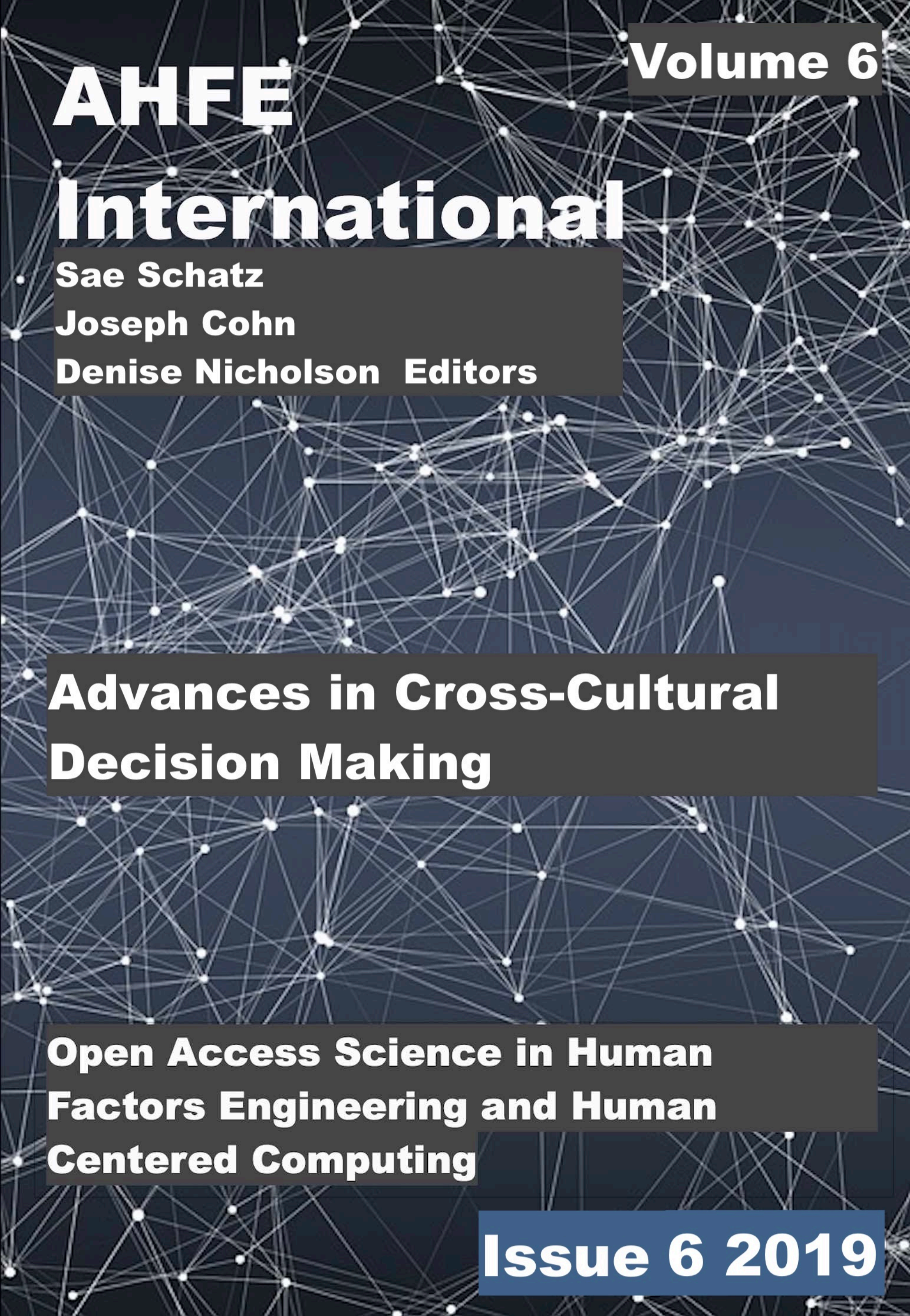


The background of the entire cover is a dark blue-grey field filled with a complex, interconnected network of thin white lines. These lines connect numerous small, light-grey circular nodes, creating a web-like or molecular structure that suggests themes of technology, communication, or human networks.

AHFE

Volume 6

International

Sae Schatz

Joseph Cohn

Denise Nicholson Editors

Advances in Cross-Cultural Decision Making

**Open Access Science in Human
Factors Engineering and Human
Centered Computing**

Issue 6 2019

*Advances in
Cross-Cultural
Decision Making*

*Advances in
Cross-Cultural
Decision Making*

*Edited By
Sae Schatz
Joseph Cohn
and
Denise Nicholson*

Published by AHFE International ©

Published by AHFE International ©

No claim to original U.S. Government works

Published in the United States of America

Version Date: 20190710

ISSN 2771-0718 (Online)

Applied Human Factors and Ergonomics International

International Standard Book Number: 978-1-4951-2095-4 (eBook)

Proceedings of AHFE International Conference (AHFE 2019), 24-28 July, 2019, Washington D.C., USA.

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the AHFE Web site at
<http://www.ahfe.org>

Table of Contents

Section 1: Modeling and Simulation to Support Cross Cultural Decision Making

Soldier narrative analysis as part of a rapid fielding program M. Jaye, P. Driscoll, D. Choe, V. Diaz and A. Keller, USA	3
Basic study on prevention of human error - Anchoring bias in relationship between objective and subjective probability A. Murata and M. Moriwaka, Japan	11
Assessing fluid rationality and its relations to cognitive styles V. Čavojova and R. Hanak, Slovakia	17
Analytical capability to better understand and anticipate extremist shifts within populations in failing states M. Bernard and A. Bier, USA	29

Section 2: Studies in Cross Cultural Decision Making

Are deliberative people more consistent in decision making? R. Hanak, Slovakia	43
Measuring the effects of cultures on website use H. C. Hsieh, Taiwan	52
Meaning making regarding threat narrative based on discourse analysis L. Fenstermacher, L. Kuznar, M. Yager and S. Shellman, USA	61
Semantic enrichment and reasoning for mobile data collection of socio- cultural data A. Caglayan, L. Cassani, T. Clark, J. Alavedra, V. Bharadwaj, N. Lee and J. Parkes, USA	71
Basic study on prevention of human error - How cognitive biases distort decision making and lead to crucial accidents A. Murara and T. Nakamura, Japan	83

Section 3: Civilization Change: Ideological, Economic, and Historical Change

The end of a civilization: What moderns might learn from Thucydides' Peloponnesian War G. McBrayer, USA	91
---	----

The roots of self-doubt (and self-loathing) in the west J. Pidluzny, USA	101
Extracting understanding from automated metaphor identification: Contrasting concepts of poverty across cultures and languages S. Taylor, L. Feldman, K. Cho, S. Shaikh, I. Cases, Y. Peshkova, G. Broadwell, T. Liu, U. Boz, K. Elliott, B. Yamrom and T. Strzalkowski, USA	113
Section 4: Advances and Challenges in Cross-Cultural Competence in the US Military, 2001-2014 and Beyond	
Socio-cultural capability requirements across all phases of military operations W. Perry, USA	125
PSYMDEV, a system intended for assisting the military analyst to construct psychological messages C. Faucher and M. Machtoune, France	137
Impact of network connectivity and agent commitment on spread of opinions in social networks D. Galehouse, T. Nguyen, S. Sreenivasan, O. Lizardo, G. Korniss and B. Szymanski, USA	149
Behavioral Influence Assessment (BIA): A multi-scale system to assess dynamic behaviors within groups and societies across time M. Bernard, G. Backus and A. Bier, USA, and W. Busch and R. Damron, Germany	161
Applying factor analysis to population surveys in Afghanistan to facilitate improved decision making J. Maddux, USA	173
Section 5: Building Models for Decision Making in Complex Cross- Cultural Domains	
Using geographic information systems analysis for mapping adverse events in an active war theater E. Cakit, Turkey, and W. Karwowski, USA	187
Developing an agent-based architecture to model population displacement J. Sokolowski and C. Banks, USA	198
Validating a hybrid cognitive-system dynamics model of team interaction A. Bier and M. Bernard, USA	209
Managing safety-related compliance of machines in global market J. Vasara and J. Kivisto-Rahnasto, Finland	221

Cross-cultural differences in preference for relationally framed decision alternatives L. Saner, A. Mathis, S. Blok, S. Glazer, I. Pavisic and S. Paletz, USA	230
Reading patterns of life: Practical tools from ethnography T. St. Benoit and C. Graffeo, USA	240
Basic study on prevention of human error - Debiasing method of cognitive biases in decision making A. Murata, T. Nakamura and S. Kubo, Japan	252

Preface

The Cross-Cultural Decision Making (CCDM) research focuses on improved decision making across a variety of cultural constructs, including geographical, historical, sociological, organizational, team, and technology interactions. This includes the research of experts and industry practitioners from multidisciplinary backgrounds, including sociology, linguistics, human-computer interaction, human factors engineering, systems engineering, military science, psychology, neuroscience, instructional design, and education, who showcase the latest advances in our understanding of the role of culture on decision making in numerous settings. Improved decision making among members of diverse teams and within organizational systems, and innovative ways to measure and assess that process, comprise the foundation for many projects discussed in these volumes. The influence of culture on decision making is pervasive, as reflected in the diverse disciplines represented by those individuals and entities involved in sociocultural research and engineering. The CCDM collection features papers that discuss emerging concepts, theories, and applications of cross-cultural decision making knowledge. The work described in these chapters reflects dedicated research by a wide range of expert academics and practitioners from around the world.

Each of the chapters of this book were either reviewed or contributed by the members of Editorial Board. For this, our sincere thanks and appreciation goes to the Board members listed below:

A. Abbe, USA	M. Kenney, USA
J. Appleget, USA	G. Klein, USA
S. Bagnara, Italy	M. Kruger, USA
M. Beaudoin, USA	J. Lincoln, USA
J. Boiney, USA	E. Lyon, USA
S. Burke, USA	N. Marmaras, Greece
A. Chan, Hong Kong	K. May, USA
C. Clegg, UK	S. Miller, USA
J. Cohn, USA	J. Morrison, USA
B. Costa, USA	D. Nicholson, USA
P. Davis, USA	M. Rauterberg, The Netherlands
E. Eldridge, USA	E. Salas, USA
I. Estabrooke, USA	D. Scapin, France
C. Fidopiastis, USA	D. Schmorow, USA
J. Fowler, USA	H. Schulze, Germany
J. George, USA	R. Stripling, USA
M. Hail, USA	R. Swisher, USA
H. Hawkins, USA	G. Zacharias, USA
J. Johnson, USA	S. Zotti, USA

We hope that this book, which is the international state-of-the-art in CCDM, will be a valuable source of theoretical and applied knowledge enabling human-centered design of variety of products, services and systems.

Sae Schatz
MESH Solutions LLC, a DSCI Company
Orlando, Florida, USA

Joseph Cohn
Office of the Secretary of the Defense
Washington, DC, USA

Denise Nicholson
MESH Solutions LLC, a DSCI Company
Orlando, Florida, USA

Editors

Section 1

***Modeling and Simulation to Support
Cross Cultural Decision Making***

Soldier Narrative Analysis as Part of a Rapid Fielding Program

Michael Jaye, Patrick J. Driscoll, Donnie Choe, Victor Diaz, and Anthony Keller*

*Department of Defense Analysis
Naval Postgraduate School
Monterey, CA 93943, USA*

**Department of Systems Engineering
United States Military Academy
West Point, NY 10996, USA*

ABSTRACT

The Warfighter Technology Tradespace Methodology (WTTM) employs technology acceptance and decision analysis theories, applying them to systems for developers to garner accurate understanding of end user expectations and concerns. WTTM applies a tradespace construct that is reliant upon three primary components as input – technical performance, logistic supportability, and human factors. The intent of this tradespace approach is two-fold: (1) to understand and assess the current state of systems as solutions to combat outpost challenges; and (2) to provide information to system developers participating in the activity as to how to tailor implemented design elements to better align with warfighter expectations, thereby maximizing the likelihood of technology acceptance upon fielding. To inform that process via a feedback mechanism, we gathered and analyzed narratives from soldier end users. We collected and tagged soldier sense-making items regarding newly fielded equipment, creating metadata for quantitative research and analysis. Here, we present the collection process as well as the results of sense-making analyses of narratives obtained from soldiers who worked with several pieces of newly fielded equipment. The intent is to highlight user dispositions as well as those factors influencing acceptance or rejection of the newly fielded equipment.

Keywords: Narratives, Equipment Fielding Process

INTRODUCTION

The Deployable Force Protection Technology Focus Team (DFP TFT) was created in 2010 at the direction of the Director, Defense Research and Engineering (DDR&E) to focus on force protection needs for troops operating in small, expeditionary bases (less than three hundred personnel), including at outposts that might be integrated with local communities. Since its inception, efforts have been directed to enhance remote operations at small bases with emphasis on capabilities that enable troops to detect, assess and defend against threats. Among the issues under investigation include holding off attackers at small bases, increasing the ability to identify shooters, return fire, and reduce response time, preventing perimeter breach, protecting critical assets, reinforce buildings to protect warfighters from blast and ballistics, assess changes to the operational environment, etc.

As a sub-organization of the DFP-TFT effort, the DFP Adaptive Red Team (ART) instituted scenario-based Technical Support and Operational Analysis (TSOA) live experiments to identify, immerse, and assess systems in operationally relevant, rigorous environments. A core component of all TSOA activities is a focused warfighter engagement with all the technology systems that participate. This warfighter engagement, which takes place on a continuous basis over several days, is intended to uncover strengths and weaknesses related to field performance and

usability as well as system vulnerabilities. TSOA events afford an early look at warfighter-defined critical capabilities at a point in the system development lifecycle where changes in design and implementation are welcome and affordable.

The ART's research and development group is the academic underpinning of TSOA. Its purpose is to develop, identify, and adapt best practices, tools, and techniques from systems engineering, logistics systems integration, human systems integration, and modeling & simulation in order to identify vulnerabilities that could threaten successful fielding, use, and operational effectiveness of potential DFP technologies. The WTTM assessment framework is a product of this group's effort. It supports comprehensive, validated, quantitative and qualitative assessments to identify new systems' vulnerabilities and risks. Structurally, it functions as an accurate cognitive mapping from observational data to quantitative scales with supporting text-based comments. What is absent in the current WTTM process, is a means of directly leveraging and analyzing the human narratives initiating this mapping sequence. The approach presented in this paper, which relies upon a sense-making construct, attempts to close this gap in a complementary fashion that exposes valuable insights currently implicit while retaining the WTTM effectiveness that has been successful over time.

HUMANS AS STORY TELLERS

Fisher's narrative paradigm offers an approach to assessing human communication which is "fundamentally stories, as interpretations of aspects of the world occurring in time and shaped by history, culture, and character" (Fischer, 1984). Fisher claims that humans are at their essence story telling beings, *homo narrans*. Others such as Franzosi suggest that stories constitute the basic form of communication through which ideas, concepts, past experience, and interpretations are conveyed (Franzosi, 1998).

Storytelling is a uniting and defining component of all individuals and communities (Snowden, 1999). Storytelling helps the individual to make sense of personal experience and individual perspective. This facilitates the transfer of knowledge and learning – sensemaking, and it is the fundamental mechanism in understanding narrative (Snowden, 1999). Sensemaking, a method in which humans integrate experiences, insights, etc., into distinct stories or narratives, often takes a complex problem or idea and deconstructs it in a way that is understandable. Sensemaking might also be thought of as a problem solving process, either collaborative or personal, that depends on perspective and interpretation (Kolko, 2010). Sensemaking has been applied to multiple fields such as: communications, intelligence, organizational studies, mathematical modeling, artificial intelligence and informatics pattern recognition (Martin & Sturmberg, 2013).

Soldier Narratives – a Novel Means to Inform the WTTM Process

We proceed by accepting that humans communicate narratives to influence one another, to capture events, and to understand and express simple and complex ideas to each other.

Yet the role that storytelling and narratives play in various aspects of warfare, especially how they shape soldier attitudes and behavior, is an understudied phenomenon. Soldiers' stories reflect their deconstruction of complex problems, and they constitute a rich source of information. The information embedded within a soldier's narrative is predominantly qualitative, yielding access to subjective experiences, thereby providing insights into new conceptions and opening up new perspectives pertinent to equipment and system assessment.

As such, soldier narratives and stories have a capacity to illuminate the sense that they have made of their experiences, including their work and dealings with new equipment. Moreover, this information naturally represents multiple perspectives, diverse experiences, and creative insights regarding a particular piece of equipment. This expands the aperture of pertinent and available information. In addition to providing equipment manufacturers specific areas for equipment improvement, narratives can serve to corroborate and validate WTTM findings and/or identify WTTM shortcomings.

Collection Methodology

Using in-depth/unstructured interviews to collect narratives or stories enables a researcher to isolate attention toward a specified population with similar life history experiences or oral history. Story collecting can be executed in a variety of ways, but the most ideal when attempting to attain sensible and self-produced data is through written stories. Story telling through writing does not adhere to the traditional pattern of a narrator telling a

story from the beginning to the end, but it enables the story teller to provide a spontaneous, uninterrupted description without external complementing or sometimes aborting influence (Czarniawska, 2004).

Consequently, we conducted a series of in-depth and unstructured interviews with soldiers from several military units (in compliance with IRB protocol and approval). We collected their written stories and then asked them to modify their stories through a series of signifying questions. This resulted in a body of empirical and qualitative data to be used for analysis and feedback to the equipment fielding process.

We interviewed 279 soldiers from three units that were chosen based on their availability within the Rest-Training-Deployment cycle. Participants were selected based on a stratified random sampling of each battalion's demographic; the main criterion for inclusion in the research was the participant's current assignment to the unit and experience with a variety of equipment. Data collection was conducted in separate engagements with three different groups from May to December 2013. There were no discriminating factors to disqualify potential test subjects from participating in the research such as education levels, rank or military experience, or military occupational specialty (MOS). As long as the test subjects were part of the units sampled and were willing to volunteer, they were free to participate. Similarly, the test subjects were informed that they could discontinue their participation at any time, even after they had provided their input. To protect the test subjects' privacy, all personal information was removed from the collected data in accordance with U.S. government institutional review board (IRB) protocol. Additionally, any information identifying specific unit, location, and mission was removed in accordance with the Department of Defense operational security policy.

The collection instrument consisted of a PowerPoint file that consisted of 25 slides, five of which were instructional. The remaining slides were designed for conveying, then signifying and amplifying the narrative. In order for the interview participant to feel more at ease and comfortable in the unscripted environment, grammatical and spelling errors were ignored. Font size, theme, and color were not prescribed (the intent being the creation of a comfortable environment that might afford free-thinking). Immediately after entering their narrative, the participant was asked to give a title of his or her story. This was intended to reinforce personal ownership of the narrative.

The following prompt was used to elicit all participant narratives:

Describe an event with the newly fielded equipment that you were a part of, that you have witnessed, or that you know about. Talk about one day, one incident, or a personal experience. What happened? What came out of this event? What would you tell others who would be using that equipment in the future? What would you tell the equipment developers that best conveys your satisfaction or dissatisfaction with the new equipment? What do they need to know? How easy was it to adapt to the new equipment? What were the particularly memorable aspects of the new equipment?

Once participants had entered their narrative and given it a title, they next answered a series of multi-dimensional signifier prompts known as triads and diads. Triads are triangulated spaces with specified vertices pertinent to a prompt that allow for amplification of the respondent's narrative in dimensions of interest. They can focus in detail on certain qualities of the service member's experience that may not have been illuminated in their narrative, and they can give the soldier the ability to reinforce certain ideas (Snowden D., 2013). Diads are similar but have one less specification vertex.

Participants were prompted with a question related to their narrative and then asked to drag a circle to a location within the triangle (or place along a scale in the case of diads). This afforded participants an opportunity to amplify his or her narrative: positioning the circle within the triad (or diad) allowed them to signify a response that best described their view relative to the prompt and the story just told. The closer to a corner, the stronger that description is taken to match the context of their story. A bubble located in the center signifies that all three aspects are equally present in their story. If the participant believed that there was no associated relationship, he or she was given the option to move the circle outside of the triangle, stating that the triad (or diad) was non-applicable to his or her narrative.

The participants' stories spanned a broad spectrum of operational system and equipment use. As such, narrative analysis offers a source of valuable information that can be used to identify potential capability gaps and technological solutions to fill those gaps. One limitation is worth noting. Because relatively few systems have been

fielded directly from the TSOA/WTTM process, we were able to identify only a small subset of participants (11) who had worked with TSOA-assessed systems. Nonetheless, as we demonstrate in what follows, those narratives are quite informative despite their small sample characteristic.

NARRATIVE ANALYSIS

We analyzed participant stories using Cognitive Edge's SenseMaker™ and SenseMaker Explorer™. From the collected narratives and their amplifications/significations (through Triads and Diads), we created a metadata file consistent with formats used by SenseMaker Explorer™. Then SenseMaker Explorer™ was used to investigate and interpret the data in a variety of ways (e.g. the individual respondent, the entire data set, a specific category, or specific item). SenseMaker Explorer™ provides various investigative options allowing for the separation of the data by filters or themes as well as their sub-components in text form. Each of the themes is associated with a Triad or Dyad, and they can indicate quantity of stories related to that specific Triad or Dyad. One can investigate specific narratives associated a particular response and we could also read each story if categorized by theme and sub-components.

It should be noted, however, that the SenseMaker suite of programs does not perform semantic analysis of the narratives. Consequently, narrative processing was not automatic - we had to read and understand participant stories in addition to whatever analysis SenseMaker might provide.

RESULTS

Figures 1 through 5 display the summary results of participant input and signification in response to narrative interview prompts. Figure 1 depicts an example of the information produced by SenseMaker Explorer™, in this case it summarizes participants' signification (or amplification) of their narratives in response to the prompt: "Based on your story, what aspect of the equipment could be improved the most?" Data displayed in Figure 1 reflect responses from over 200 participants. Each data point is associated with an individual participant's narrative and his/her interpretation of how the equipment they described in their story fit the prompt. Because the equipment described in those narratives was so diverse, one might expect the resulting signification to be spread somewhat uniformly among the three corners of the triangles vertices, as observed in Figure 1.

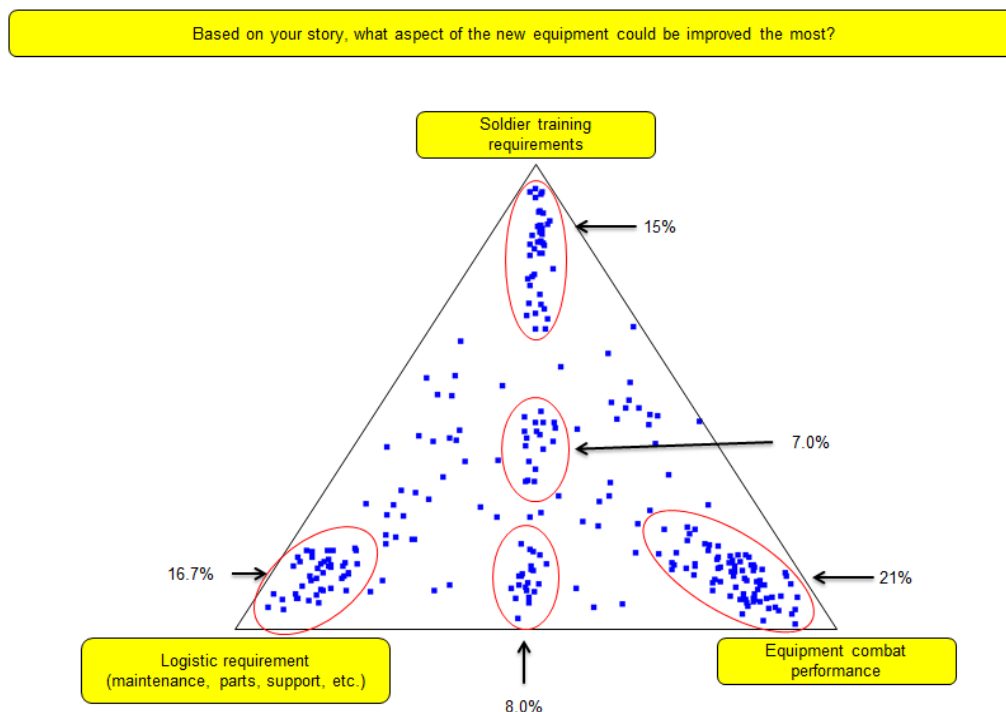


Figure 1. An example of SenseMaker Explorer output gathered from a particular narrative signification question, "What aspect of the new equipment could be improved the most?" In this example, signifiers from all narrative entries were considered in the sample space.

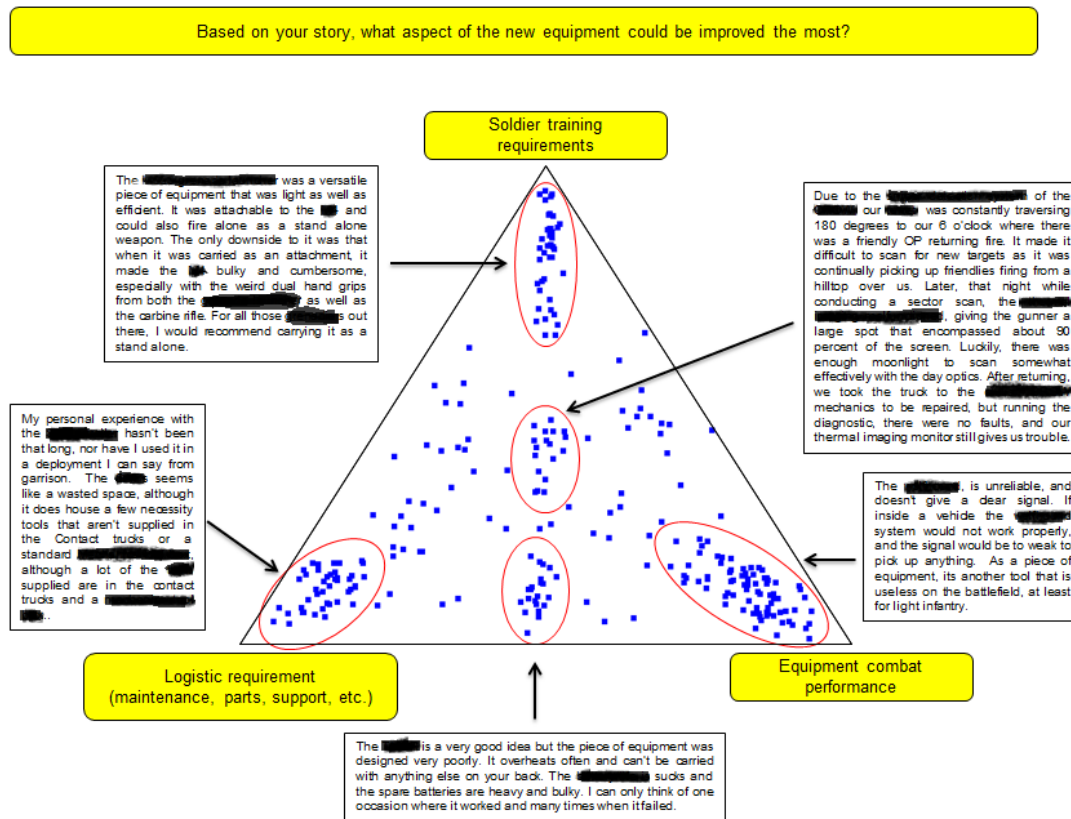
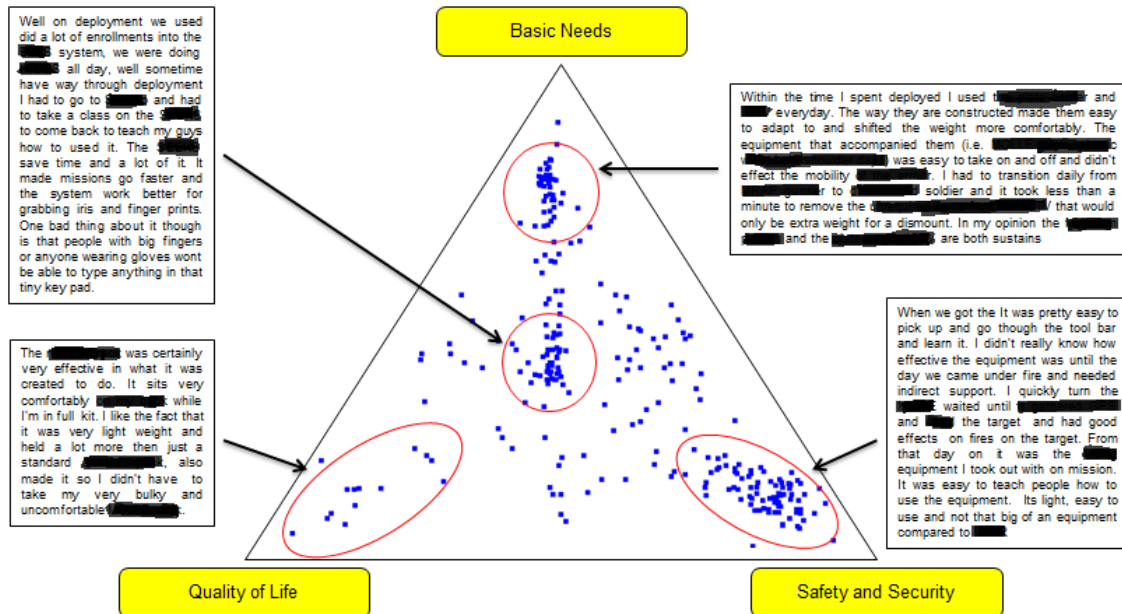


Figure 2. SenseMaker Explorer output from Figure 2, but now with narrative fragments depicting soldier commentary associated with various signification concentrations.

Figure 2 depicts the identical information as Figure 1, but now with selected or representative participant commentary identified with its associated signification (or amplification) concentrations within the triad. Portions of the narrative fragments relating to specific equipment, names, functionalities, etc., are redacted yet still convey soldier intent. The ability to mine the data set to obtain this type of representation offers insight into the data and its meaning that would not be obtained through typical surveys that employ more traditional queries such as Likert scales.

Figure 3 is another example of triad signification of participant narratives. Figure 3 displays signification and representative responses to "The benefit to your unit provided by the equipment in your story was in..." (upper figure) and "The equipment in your story failed to improve..." (lower figure).

The benefit to your unit provided by the new equipment was in...



The new equipment in your story failed to improve...

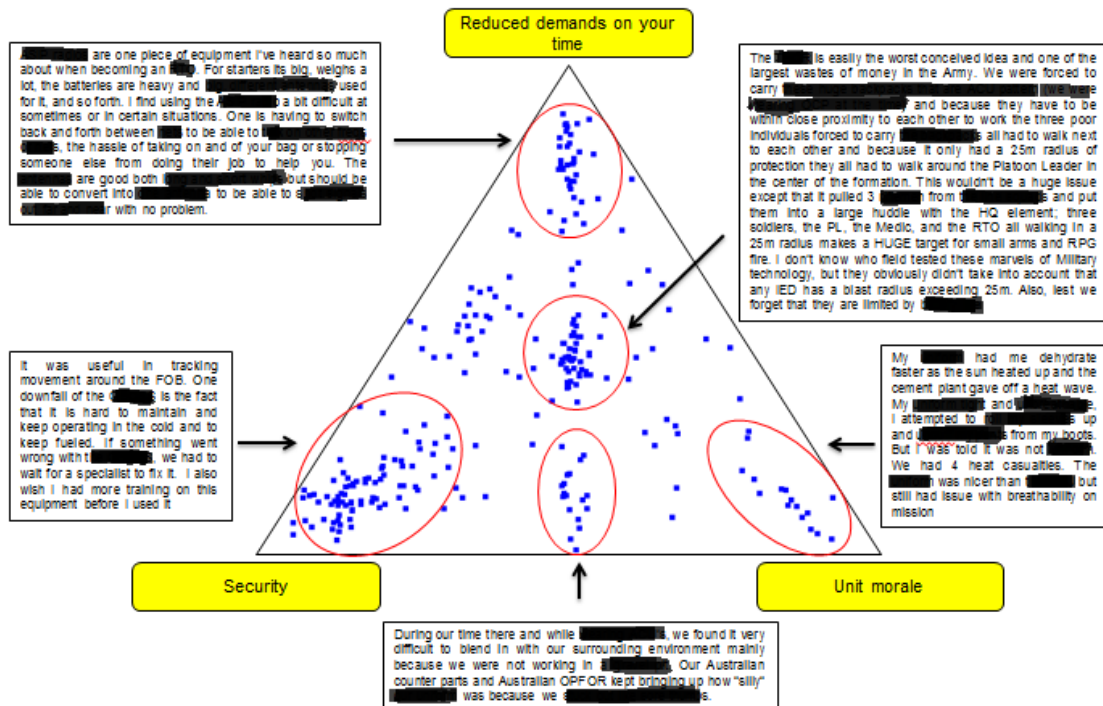


Figure 3. Participant signification and representative responses to the signifying prompts "The benefit to your unit provided by the new equipment was in..." and "The new equipment in your story failed to improve..."

Figure 4 conveys the results from the subset of participants who had worked with a piece of equipment that has been subjected to the TSOA/WTTM assessment process. The equipment's basic functionality is providing security to units engaged in combat operations, and this is corroborated by the massed location of the soldier significations as well as in their narratives.

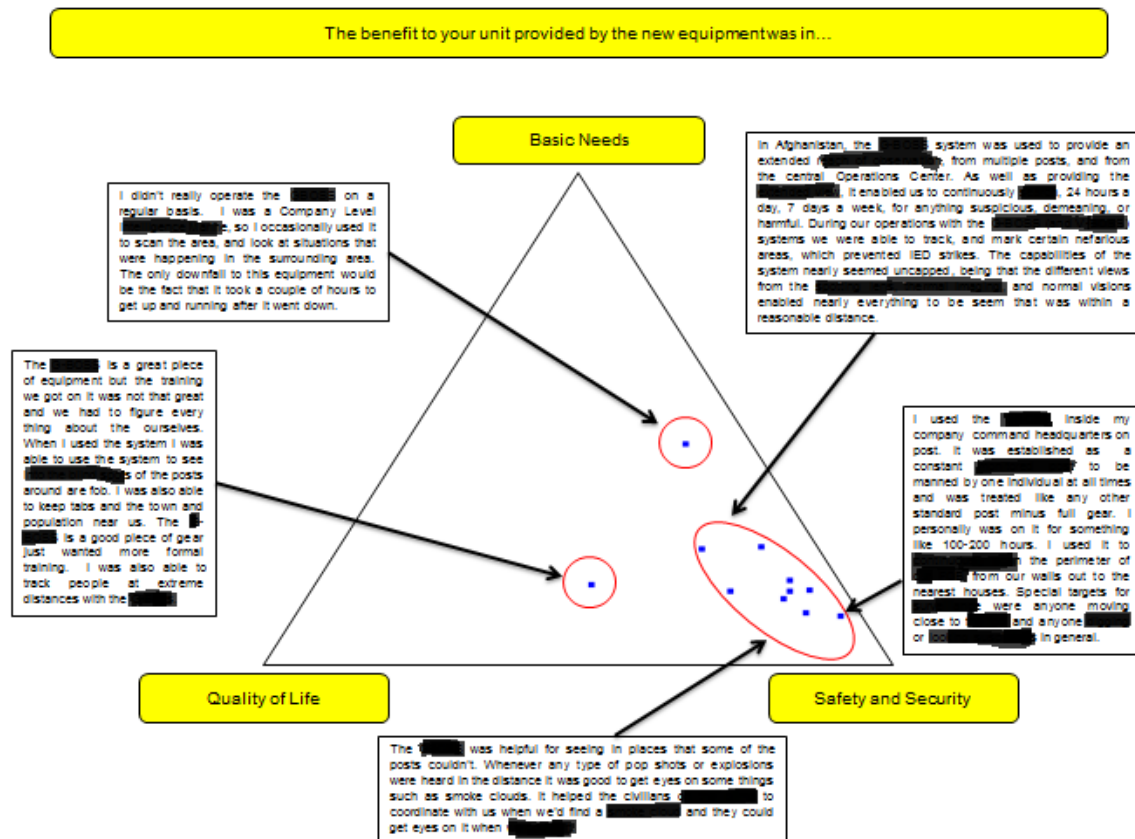
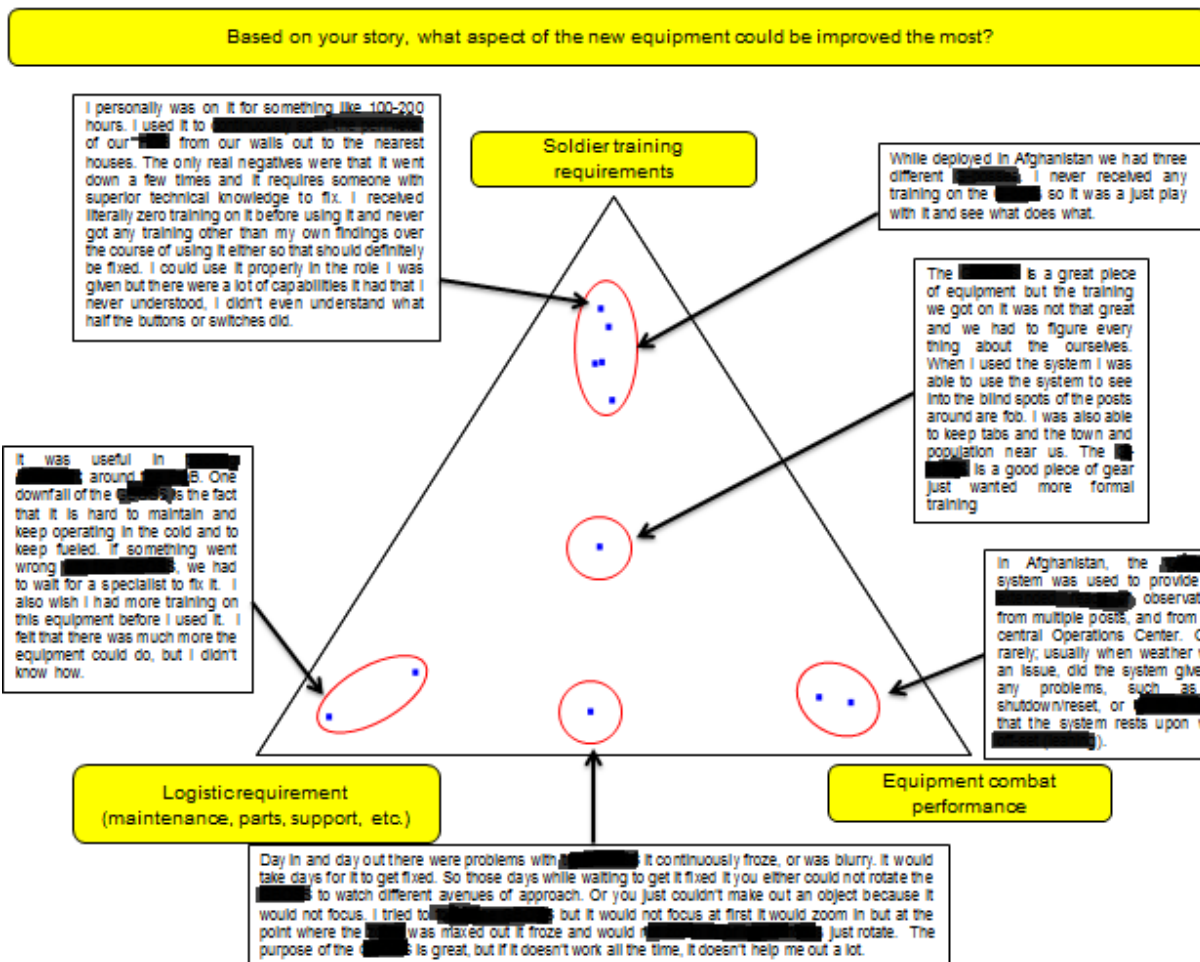


Figure 4. Participant signification and representative responses to the signifying prompts "The benefit to your unit provided by the new equipment was in..." and "The new equipment in your story failed to improve..."

Another dimension that the participants considered in amplifying their narratives regarding this piece of equipment was how it could be improved the most. Resulting signification and associated narrative fragments are depicted in Figure 5. Interestingly, not only is the relative weighting of the groupings consistent with TSOA/WTTM findings (i.e. training requirements are too burdensome, mostly due to the ill-conceived user interface; logistic and technical performance are, relatively, much better), but many of the narrative fragments are consistent with TSOA assessments and commentary as well. Moreover, while the results here are not intended to be conclusive, they do provide valuable indicators regarding potential success upon fielding – narratives appear to indicate support for acceptance based on performance if the barrier associated with burdensome training requirements could be overcome.



CONCLUSIONS

Obtaining and analyzing narratives based on participant experiences with fielded systems and equipment provides valuable insights into performance expectations that can inform both capability and design choices. In addition, significations/amplifications of these narratives combined with associated narrative fragments facilitates a deeper and richer understanding than that accessible through traditional elicitation approaches such as Likert-based surveys. Preliminary findings as reported in this paper appear to indicate that, in the case of at least one newly fielded piece of equipment, the WTTM process is capturing information that is consistent with end user experiences with that equipment as expressed in their stories and amplifications/significations. Thus, narrative analysis can serve as both a stand-alone process or as a complementary approach against which validation and consistency checking can occur.

REFERENCES

- Czarniawska, B. (2004). *Narratives in Social Science Research*. Los Angeles: SAGE Publications Inc.
- Fischer, W. R. (1984). "Narration as a human communication paradigm", in: *Communications monographs*, vol 51, pp. 1-22.
- Franzosi, R. (1998). "Narrative Analysis-Or Why (And How) Sociologists Should be Interested in Narrative". *Annual Reviews*, vol 24, pp. 517-554.
- Kolko, J. (2010). "Sensemaking and Framing: A Theoretical Reflection on Perspective in Design Synthesis", in: *Design Research Society*; Creative Commons.
- Snowden, D. (1999). "Story telling: an old skill in a new context". *Business Information Review*, vol 16 No. 1, pp. 30-37.

Basic Study on Prevention of Human Error

-Anchoring Bias in Relationship between Objective and Subjective Probability-

Atsuo MURATA, and Makoto MORIWAKA

*Dept. of Intelligent Mechanical Systems
Graduate School of Natural Science and Technology, Okayama University
Okayama, Japan*

ABSTRACT

This study empirically identified the relationship between the objective and the subjective probabilities, and confirmed whether the relationship corresponded with the hypothesized weighting function proposed in prospect theory. As a result, it was found that the experimentally identified relationship between the objective and the subjective probabilities corresponded with the hypothetical weighting function above. In other words, the estimated number of deaths was underestimated for moderate and high number of deaths, while the estimated number of deaths was overestimated for the small or rare number of deaths. Moreover, in order to examine how cognitive bias occurred in the estimation of number of deaths for a variety of deaths and accidents, it was discussed whether the different anchor biased the relationship above. The estimated numbers of deaths were remarkably affected by the anchor, and the estimated numbers of deaths for the group given the anchor A (large number of deaths) was larger than that for the group given the anchor B (small number of deaths). Some implications for risk management were given from the perspective of cognitive biases.

Keywords: Subjective Probability, Underestimation, Overestimation, Prospect Theory, Anchoring Bias

INTRODUCTION

Our cognitive information processing is conducted by System1 or System2. While System2 requires us to conduct effortful, demanding and deliberate mental activities, System1 operates quickly, automatically, without time consuming, and intuitively with little or no efforts. Although heuristic approaches that we adopt when we have no time to deliberate are based on System1, and are very simple and intuitive, such approaches are suffering from cognitive biases. Due to bounded rationality (Kahneman, 2011, Tversky and Kahneman, 1974, Kahneman and Tversky, 1984, Altman, 2012, Angner, 2012, and Bazerman and Moore, 2001), we generally cannot make decision rationally, and thus suffer from cognitive biases. Kahneman (2011), Tversky and Kahneman (1974) and Kahneman and Tversky (1984) proposed prospect theory to explain such cognitive biases in decision making under uncertainty. In prospect theory, the weighting function of objective probability such as number of deaths per year for each cause is hypothesized so that the weight is larger for the small probability range, and the weight is smaller for the large probability range. Such a hypothesis is based on only the study by Lichtenstein et al. (1978). Thus, it should be further verified whether such a hypothesis is proper or not.

After some serious accident occurred, one tends to overestimate the occurrence probability of such an accident. For example, we hesitate to use an airplane immediately after a serious aviation accident due to the overestimation of a fatal aviation accident. This means that our psychological evaluation of probability, that is, subjective probability (Wright and Ayton, 1994), changes irrespective of the constant objective probability of airplane crash. Therefore, it is important to identify the relationship between the objective and the subjective probabilities properly.

The anchoring bias is pointed out (for example, Kahneman (2011) and Bazerman and Moore (2001)). It is predicted that the difference of the anchor easily changes the relationship between the objective and the subjective probabilities above. However, such analyses have not thoroughly conducted until now. If such a hypothesis was verified, we must be very cautious in the assessment of risks (in a sense, subjective probability). It is possible that ignoring such characteristics and improperly assessing a risk lead to crucial accidents such as the Challenger space shuttle disaster (Reason, 1990 and Vaughan, 1997).

This study empirically identified the relationship between the objective and the subjective probabilities, and confirmed whether the relationship corresponded with that found by Lichtenstein et al. (1978). Moreover, it was explored whether the different anchor biased the relationship above.

METHOD

Participants

Sixty participants took part in the experiment. They were all undergraduate students at Faculty of Engineering, Okayama University.

Experimental task

The participants were distributed a paper on which 21 cause of deaths in Table 1 was randomly arranged. Their task was to estimate the number of deaths for each cause. The participants were instructed to estimate the number of deaths of 2012 in Japan.

Design and Procedure

The participants were equally divided into three groups. Two groups were given the cue for estimation. One group was given the cue (anchor) A for estimation. Another group was provided with the cue (anchor) B for estimation. The anchors A and B were as follows. The third group was not provided with the cue for estimation.

Anchor A: The number of deaths per year due to stroke is 121,602.

Anchor B: The number of deaths per year due to flood is 35.

RESEARCH HYPOTHESIS

The first hypothesis was that the estimated number of deaths was underestimated for moderate and high number of deaths, while the estimated number of deaths was overestimated for the small or rare number of deaths.

The second hypothesis was that the estimated numbers of deaths were remarkably affected by the anchor, and the estimated numbers of deaths for the group given the anchor A was larger than that for the group given the anchor B. In other words, we hypothesized that the anchor distorted the estimation of number of deaths.

RESULTS

The results of the experiment are summarized in Figure 1. In Table 1, the causes of deaths for numbers from 1 to 21 are listed. The smaller number in Table 1 shows that many numbers of deaths are induced by this cause. The larger number in Table 1 means that there are few deaths due to these causes. If the subjective probability judgments

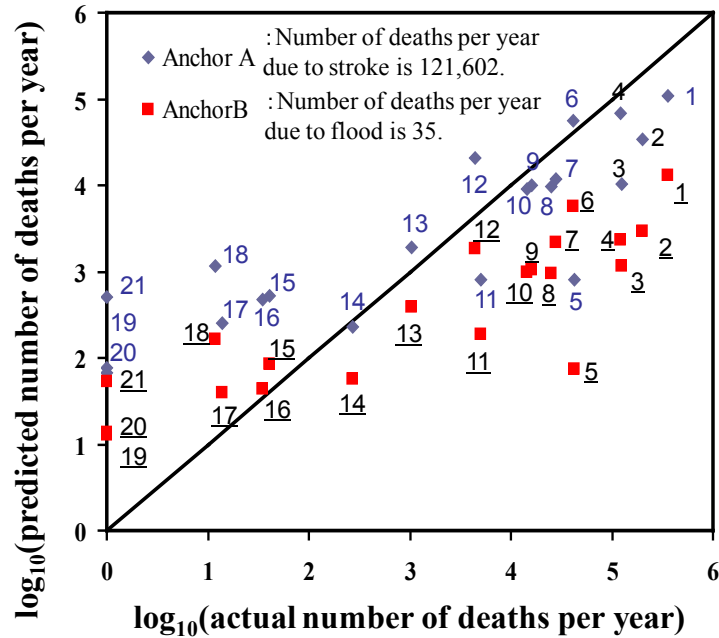


Figure 1. Relationship between estimated (predicted) number of deaths and the actual number of death (Japanese statistical data in 2012). If the estimated and the actual number of deaths are equal, the data falls on the straight line.

Table 1. Correspondence between cause of death and numbers in Figure 1.

1	all cancer	11	seasonal flu
2	heart disease	12	motor vehicle accidents
3	pneumonia	13	murder
4	stroke	14	climbing accidents
5	disease due to air pollution	15	childbirth accidents
6	all accidents	16	flood
7	suicide	17	electrocution
8	renal failure	18	food poisoning
9	liver disease	19	tornado
10	disabetes	20	helicopter crash
		21	radioactivity accidents

(estimations) were accurate, they would equal the statistical data (actual number of deaths for each cause). Although the data were dispersive to some extent, it tended that the rare causes whose actual numbers of deaths are small (for example, numbers 16, 17, and 18) were overestimated. Contrary to this tendency, the common causes of death (for example, numbers 1, 2, and 3) tended to be underestimated.

Next, the anchoring bias is discussed. The estimated numbers of deaths for the group given the anchor A tended to be larger than that for the group given the anchor B. Anchoring information was found to distort the subjective probability. The results are summarized in more detail in Figure 2. In Figure 2, the estimated number of deaths is compared among three conditions (Anchor A, Anchor B, and No anchor). In Figure 3, the difference between the mean actual number of deaths and the mean estimated number of deaths is compared between the underestimation and the overestimation areas in Figure 1 or Figure 2, and among three conditions (Anchor A, Anchor B, and No anchor). Here, the mean actual number of deaths corresponds to the mean of all number of deaths belonging to

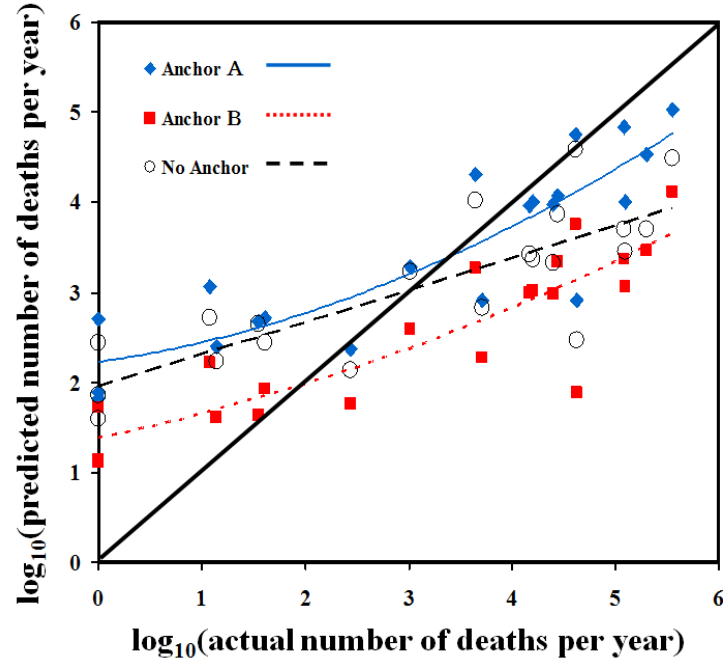


Figure 2. Relationship between estimated (predicted) number of deaths and the actual number of death (Japanese statistical data in 2012) as a function of anchoring condition (Anchor A, anchor B, and No anchor).

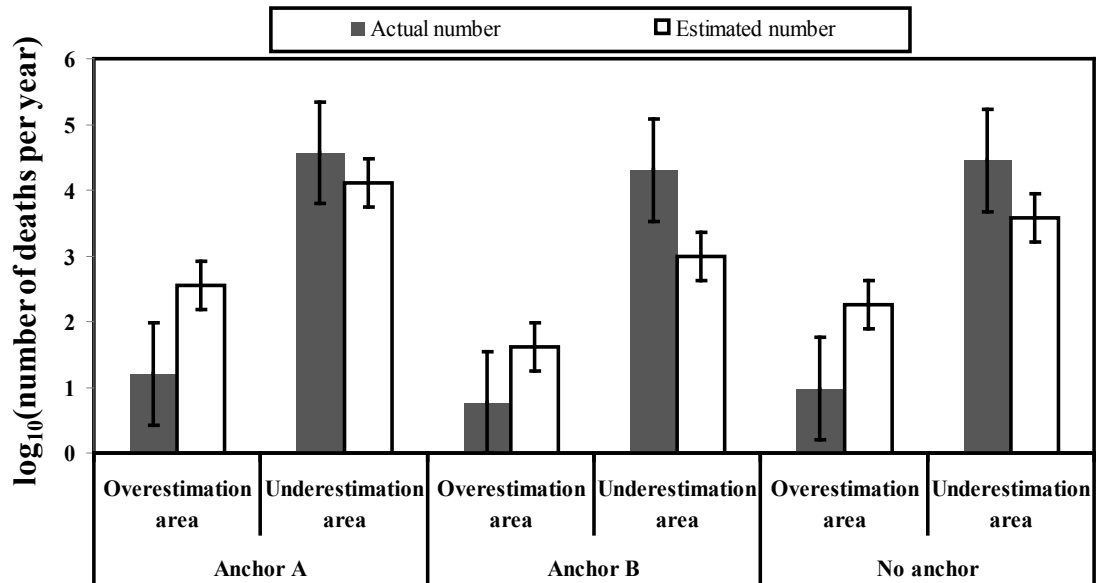


Figure 3. Difference between the mean actual number of deaths and the mean estimated number of deaths compared between the underestimation and the overestimation areas, and among three conditions (Anchor A, Anchor B, and No anchor).

underestimation or overestimation areas. The mean predicted number of deaths corresponds to the mean of all predicted number of deaths belonging to underestimation or overestimation areas.

As for the overestimation area, a statistical t -test conducted on the number of deaths revealed significant ($p < 0.01$) differences between actual and estimated number of deaths for all of three conditions (Anchor A, Anchor B, and No anchor). As hypothesized, the overestimation tendency was statistically verified for three conditions. As a result of carrying out a similar t -test for the underestimation area, significant ($p < 0.01$) differences between actual and estimated number of deaths were also detected for all of three conditions.

DISCUSSION

Relationship between objective and subjective probability

The probability weighting as shown in Figure 4 is generally assumed in prospect theory (Kahneman, 2011, Tversky and Kahneman, 1974 and Kahneman and Tversky, 1984), which aims at explaining cognitive biases such as reversal of preference and loss aversion in decision making under uncertainty. The probability weighting function (Wakker, 2011) in Figure 4 is given by

$$w(p) = \frac{p^c}{(p^c + (1-p)^c)^{1/c}}, \quad (1)$$

where p is the probability without weighting, and c is an empirically determined parameter. When c equals 1, this corresponds to the diagonal line, which means that the probability weighting (subjective probability) is equal to the objective probability. For other values of c ($=0.4, 0.6$, and 0.8), the probability weighting for the low objective probability is overestimated, while the probability weighting for the high objective probability is underestimated.

In prospect theory, the weighting function of objective probability such as number of deaths per year for each cause is hypothesized so that the weight is larger for the small probability range, and the weight is smaller for the large probability range. Such a hypothesis is based on only the study by Lichtenstein et al. (1978).

As shown in Figures 1 and 2, the number of deaths tended to be overestimated for the rare causes whose actual numbers of deaths are small (for example, numbers 16, 17, and 18). The number of deaths was underestimated for the common causes of death (for example, numbers 1, 2, and 3). This corresponded well with Lichtenstein et al. (1978).

Anchoring bias

Anchoring is related to an adjust-and-anchoring heuristic for estimating uncertain quantities like prediction of number of deaths for a variety of accidents or diseases. The adjust-and-anchoring bias seems to stem from confirmation heuristics (Bazerman and Moore, 2001). Confirmation bias can be generally explained as follows. We tend to focus on a single possible cause of an effect, neglect alternative causes of the effect, and conclude that the association between the single cause and the effect is stronger than it actually is.

The adjustment typically ends prematurely, and leads to under- or over-estimation as shown in Figures 1 and 2. As pointed out by Lichtenstein et al. (1978), the cue A or B in this study distorted the decision. When we provide participants with a rare case (small number of deaths such as “The number of deaths per year due to flood is 35”), they tended to be affected by this number and underestimate the number of deaths for a variety of deaths or accidents. When participants were provided with a common case (large number of deaths such as “The number of deaths per year due to stroke is 121,602”), they tended to be affected by this number and overestimate the number of deaths for a variety of deaths or accidents. These tendencies are demonstrated in Figure 2. When the larger number of deaths is given as a cue, the relationship between the probability and the probability weighting (subjective probability) moved upward. When the smaller number of deaths is given as a cue, the relationship between the probability and the probability weighting (subjective probability) moved downward. The cue is regarded to function as priming. The cause of such a cognitive bias might be interpreted as follows. While we frequently encounter the rare events such as tornado or helicopter crash on TV or newspaper, we seldom encounter the common events of deaths such as strokes on TV or newspaper. This difference of availability to information between two cases causes the distortion, and moves the estimation of smaller number of deaths and larger number of deaths to the overestimation and underestimation areas, respectively.

The cue definitely affected the decision making (estimation of number of deaths), and distorted the relationship between the probability and the probability weighting (subjective probability). This means that such minor information is important enough to affect our decision making. We unexpectedly tend to overestimate the occurrence probability of an airplane accident immediately after a serious aviation accident due to the overestimation of a fatal aviation accident. This means that our psychological evaluation of probability changes irrespective of the constant objective probability of airplane crash. We must bear this in mind when evaluating risks which hide behind a variety of our activities.

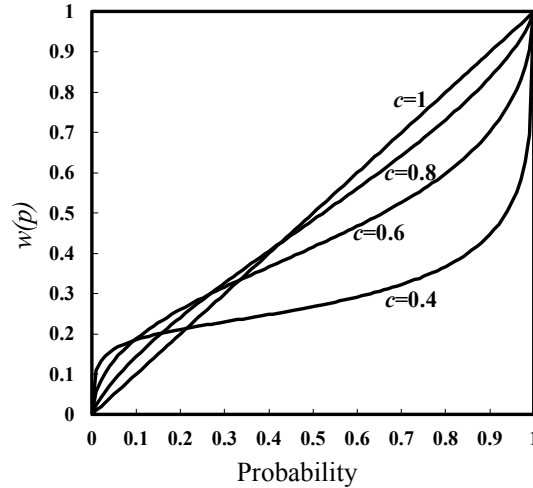


Figure 4. Typical and hypothetical probability weighting function assumed in prospect theory.

Therefore, it is important to identify the relationship between the objective and the subjective probabilities properly. Without this, our decision making unexpectedly suffers from cognitive bias, and is easily distorted. The basic source of cognitive biases is based on our cognitive characteristics, that is, bounded awareness. In risk management under uncertainty, we must bear such an anchoring bias in mind. We are easily biased by the anchor even when we engage in objective risk evaluation or assessment tasks such as estimating the number of deaths for a variety of accidents or diseases.

CONCLUSIONS

In conclusion, the relationship between the objective and the subjective probabilities on which prospect theory (Kahneman, 2011, Tversky and Kahneman, 1974, Kahneman and Tversky, 1984, Altman, 2012, Angner, 2012, and Bazerman and Moore, 2001) is based was empirically found to be valid. Moreover, although the relationship between the objective and the subjective probabilities mentioned above did not change, it was found that the anchoring biases easily occurred by changing the cue (anchor) in estimating numbers of deaths for each cause.

Future research should explore whether or not such a bias can be removed somehow. Moreover, we should develop a systematic method for checking cognitive biases in order to prevent crucial accidents or human errors stemming from cognitive biases.

REFERENCES

- Altman, M. (2012), "*Behavioral Economics for Dummies*", John Wiley & Sons Canada, Ltd.
- Angner, E. (2012), "*A Course in Behavioral Economics*", Palgrave Macmillan.
- Bazerman, M. H. and Moore, D. A. (2001), "*Judgment in Managerial Decision Making*", Harvard University Press.
- Kahneman, D. (2011), "*Thinking, Fast and Slow*", Penguin Books.
- Kahneman, D. and Tversky, A. (1984), "*Choices, Values, and Frames*", *American Psychologist*, Vol.34.
- Lichtenstein, S., Slovic, P., Fishhoff, B., Layman, M., and Combs, B. (1978), "*Judgment Frequency of Lethal Events*", *Journal of Experimental Psychology: Human Learning and Memory*, Vol.4, pp.551-578.
- Reason, J. (1990), "*Human Error*", Cambridge University Press.
- Tversky, A. and Kahneman, D. (1974), "*Judgment under Uncertainty: Heuristics and Biases*", *Science*, Vol.185.
- Vaughan, D. (1997), "*The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*", University of Chicago Press.
- Wakker, P. P. (2011), "*Prospect Theory for Risk and Ambiguity*", Cambridge University Press.
- Wright, G. and Ayton, P. (1994), "*Subjective Probability*", Wiley.

Assessing Fluid Rationality and Its Relations to Cognitive Styles

Vladimíra Čavojová and Robert Hanák

*Institute of Experimental Psychology
Slovak Academy of Sciences
Bratislava, Slovakia*

ABSTRACT

A new tri-partite model of rational thinking (Stanovich, 2011) allows for assessing various aspects of rational decision making. Because of assumed multi-faceted nature of rationality, it is necessary to establish which measuring paradigms are best suited to assess it. 531 participants (31% men), with mean age of 27.61 (SD=7.8) took part in the present study. We examined relationships between several testing paradigms from JDM literature, such as the jelly bean task (Kirkpatrick, Epstein, 1992), cognitive reflection test (Frederick, 2005) and Wason's selection task. Higher scores in these testing paradigms are hypothesized to constitute a resistance to miserly processing. We also examined two other dimensions of fluid rationality: temporal discounting of the reward and acceptance of risk (Frederick, 2005), and several measures of cognitive style containing a preference for intuition or deliberation (PID, REI, CoSi, MBTI). The lack of strong mutual relationships between these measures, together with low internal consistency for a composite score of resistance to miserly processing, rather undermined the construct of miserly processing. There were also very weak relationships with other constructs hypothesized to be other dimensions of fluid rationality besides resistance to miserly processing. Such results seem to support Stanovich's (2011) hypothesis about multifactorial fluid rationality.

Keywords: rationality, resistance to miserly processing, cognitive styles, intuition

INTRODUCTION

What is rationality and what is the best way to measure it? There has been much debate over what constitutes rationality and whether humans are rational beings (Gigerenzer, 2008, Nickerson, 2008, Sternberg, 2002, Stanovich, 2011). Rationality is hard to define, but the majority of researchers agree that it is probably multifarious – composed of a variety of cognitive abilities and thinking dispositions (Stanovich, 2011).

In dual process theories rationality has often been (wrongly) associated with Type 2 processes, which need to override the default automatic response of Type 1 processes. Evans (2010) argues that such overriding depends on cognitions as well as personality. Not only do we have to have sufficient knowledge to compute the right answer to the problem, but we also have to have a personality favoring a more reflective approach to problems. In a similar vein, Stanovich (2011) writes about individual differences in thinking dispositions, especially motivation towards rational integration (for discussion see also Saunders, Over, 2009). However, in a recent article written together by Evans and Stanovich (2013) they both emphasize the fact that rationality has to be ascribed to individual, not subpersonal processes (i.e. a person can behave rationally, but neither Type 1 nor Type 2 processes rational per se).

The view that Type 2 processes are behind all rational behavior has been reflected in naming the Type 2 processes as rational (e.g. Epstein, 2003) or by arguing that Type 1 processes must be overcome in order to avoid biases and errors (heuristic and biases program, Tversky, Kahneman, 1974). However, there are three issues to consider. Firstly, cultural studies reviewed by Buchtel and Norenzayan (2009) suggest that holistic processing can also be learned, effortful and normatively correct for solving some problems and that it is likely to reflect cultural differences in Type 2 processing. In other words, in some cultures that value holism and relationships more highly, Type 2 processes have different attributes than is typically ascribed to them (analytic, de-contextualizing, etc.). Secondly, Stanovich (2011) recently proposed a taxonomy of errors, which reflects upon a newly formed understanding that errors can be caused by not using Type 1 processes when it is adequate, not recognizing the need to engage Type 2 processing (error of comprehension, Kahneman, Tversky, 1982), or lacking the skills or knowledge necessary for successful solving the task (errors of application, Kahneman, Tversky, 1982). In this notion, rationality is identified more with a so called reflective mind (recognizing the need to intervene in default automatic processing) rather than an algorithmic mind (cognitive capacity). Thirdly, Evans and Stanovich (2013) distinguish between types (Type 1 and 2 as qualitatively distinct ways of processing information) and modes of processing, which are cognitive styles and are manifest within the domain of what is regarded as Type 2 thinking. Taking cognitive styles as only a variation in the domain of Type 2 processes largely explains the many cultural differences reviewed by Buchtel and Norenzayan (2009).

In his new tri-partite model Stanovich (2011) elaborates on the dual-process theories, but differentiates reflective and algorithmic mind at the level of Type 2 processes. It is based on extensive research of individual differences and differential association of intelligence with the two minds. Reflective mind is a source of individual differences in rational thinking dispositions and is responsible for detecting the need to override automatic Type 1 responses. Algorithmic mind is a source of individual differences in fluid intelligence and is responsible for coming up with the correct response, when reflective mind detects the need for override. Stanovich extends this model further by proposing a framework for assessment of rational thinking, where rationality can be partitioned into fluid and crystallized components by analogy to the Gf and Gc of the Cattell/Horn/Carroll fluid-crystallized theory of intelligence. Fluid rationality encompasses the *process* part of rational thought (the thinking dispositions of the reflective mind), while crystallized rationality concerns acquired knowledge (*content*) that can either help (crystallized factors) or undermine (crystallized inhibitors) rational thinking. This framework (Stanovich, 2011, p. 192) allows for assessing various aspects of rational decision making. Because of assumed multi-faceted nature of rational thinking, it is necessary to establish which measuring paradigms are best suited for assessment of rationality.

The relationship between thinking disposition (cognitive styles) and rational decision making is not straightforward. Some studies have found a positive relationship (Witteman, van den Bercken, Claes, & Godoy, 2009; Sirota, Juanchich, & Hagmayer, 2013), others have found no such relationship (Newstead, Handley, Wright, & Farrelly, 2004; Čavojská, Ballová Mikušková, Hanák, 2013). One possible explanation for these results could lie in different cultural backgrounds. For example, Buchtel and Norenzayan (2009) review several studies that suggest that not only do Eastern and Western cultures differ in dominant cognitive style (holistic vs. analytic), but that these styles are learned in the process of socialization. They assert that cultural differences are best conceptualized as differences in habits of thought, rather than differences in the actual availability of information processing strategies in the cognitive repertoire, and that holistic and analytic ways of thinking can be differently encouraged in their development and use by different cultural and situational constraints (Buchtel, Norenzayan, 2009, p. 219).

The aim of the current paper is to outline the possible relationships between some of the testing paradigms proposed by Stanovich (2011) to measure aspects of fluid rationality, such as resistance to miserly processing, temporal discounting of reward, and acceptance of uncertainty (risk). Secondly, we examine the relationships between these dimensions and cognitive styles because in Stanovich's (2011) framework faith in intuition is conceptualized as one of the rationality inhibitors belonging under crystallized rationality. Finally, we compared the results obtained in a Slovak sample with results from other cultural backgrounds.

In the rest of the paper we first introduce some concepts relevant for assessing rationality, such as resistance to miserly processing, temporal discounting of reward and risk seeking preference and then review their relationships with several cognitive styles. We then present the results of our preliminary study regarding relationships between rationality and cognitive styles, gender and cultural differences. In the final section we draw some conclusions regarding the examined model proposed by Stanovich (2011).

ASSESSMENT OF RATIONAL THOUGHT

Establishing a reliable measure for the assessment of rational thinking is of great importance because it differs from intelligence in many respects. However, due to its multifaceted nature, it is quite difficult to find just one measure that would reliably test rational thinking in all its complexity. We examined relationships between several dimensions from Stanovich's framework (described below in detail) in a sample of 531 participants (31% men), with a mean age of 27.61 years ($SD=7.8$). In this section we first introduce the three main dimensions (resistance to miserly processing, prudently discounting the future and risk preference) and tasks from different measurement paradigms. We then briefly describe cognitive styles and their measurement. Results of our analysis will be presented in the following section.

Resistance to miserly processing

Many tasks in the heuristics and biases literature (in which people fail to give the normatively correct response) reflect our tendency toward miserly processing (Stanovich, 2011). There are two ways in which the errors of faulty (not normatively correct) reasoning can occur. Earlier models presumed that miserly processing always involved defaulting to Type 1 processes (autonomous mind). However, more recent models (e.g. Stanovich, 2011) also allow for the possibility of miserly Type 2 processing. For example, recognizing the need for Type 2 processes to compute the answer, but resigning to serial associative cognition with a focal bias, resulting in a non-intuitive, but still incorrect, response.

Cognitive Reflection Test. The Cognitive Reflection Test (CRT; Frederick, 2005) is a simple task measuring our tendency towards miserly processing (or cognitive impulsiveness/laziness in defaulting to the autonomous mind) vs. the ability to postpone our judgment (cognitive reflection). It is a simple test of one type of cognitive ability that is so predictive of some preferences that it effectively functions as an expression of cognitive ability (Frederick, 2005). CRT is rooted in a dual process approach, but it also fits in the newer tripartite model proposed by Stanovich (2011) in that it directly reflects the most important aspect of the reflective mind – suppressing our tendency to give the first intuitive answer that comes to mind and recognizing the need to engage the algorithmic mind (Type 2 processes) to come to the correct answer. In contrast with classic insight problems, such as the nine-dot problem, it primes an attractive intuitive response – the (incorrect) answer comes to mind easily and the individual has to recognize that it is incorrect and calculate the right answer.

The CRT consists of three simple problems, as follows (taken from Frederick, 2005):

- 1) A bat and a ball cost \$1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost?
- 2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?
- 3) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 49 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

In an extensive study, Toplak, West and Stanovich (2011) found that CRT is a more potent predictor of a wide sample of tasks measuring rational behavior than other measures of cognitive abilities, thinking dispositions and executive functioning. Based on these results they argue that CRT is such a powerful measure because it is performance based rather than a self-report measure, and that it assesses resistance to miserly processing in a way that is not addressed by any other test of cognitive abilities or executive functioning.

Diagnostic hypothesis testing. Another important aspect of miserly processing is confirmation bias and failure to take into account information that would falsify our hypotheses. Our ability to avoid these thinking errors is measured by Wason's four card selection task. We used this problem:

"You are shown a set of four cards placed on a table, each of which has a number on one side and a colored patch on the other side. The visible faces of the cards show 3, 8, red and brown. Which card(s) must you turn over in order to test the truth of the proposition that if a card shows an even number on one face, then its opposite face is red?"

The selection task showed moderate correlations ($r=0.22$) with CRT in the study by Toplak, West and Stanovich (2013). Furthermore, CRT was a significant predictor of normatively correct performance in the selection task. Thompson, Evans and Campbell (2013) examined the selection task in a slightly different format and concluded that

people gave intuitive (incorrect) answers due to the fluency (and associated feeling of rightness) produced by Type 1 processes. Therefore, it seems reasonable to assume that the ability to suppress feelings of rightness connected with an intuitive response in the selection task will be related to an ability to resist miserly processing.

Attribute substitution refers to our tendency to lighten the cognitive load (Kahneman, Frederick, 2002). It occurs when a person needs to assess attribute A but finds that assessing attribute B (which is correlated with A) is cognitively easier, and so uses B instead. In other words, it means answering an easier question instead of a harder one. For assessing this tendency we used the Jelly bean task (Denes-Raj, Epstein, 1994), which reads as follows:

You can always win 1 EURO when you draw a red bean (not peaking) from a bowl containing white and red beans. You can choose from the two bowls from which you will draw beans: (1) A bowl containing 10 beans, 9 white and 1 red, or (2) A bowl containing 100 beans, 93 white and 7 red?

The Jelly bean task showed moderate correlation ($r=0.37$) with the CRT in the study of Toplak, West and Stanovich (2013). The Jelly bean task has often been used in studies of dual-process theories to show that people often choose according to the vividness effect – despite knowing the more unfavorable ratio of the latter bowl, the picture of 7 winning marbles seems to be more appealing than just 1 winning marble (e.g. Epstein, 2003). The normatively correct choice (bowl A) is correlated with subscales of cognitive styles that are supposed to reflect Type 2 processing (e.g. Sobyra, 2010, Witteman, van den Bercken, Claes et al., 2009; for review see Hanák, 2013).

Prudently discounting the future (Temporal discounting of reward)

Another major dimension within fluid rationality is 'prudently discounting the future' (Stanovich, 2011). It is based on the notion that more rational people should be more patient – that they “discount” future reward less (Frederick, 2005). Furthermore, they are more able to employ the algorithmic capacities of their mind to calculate the more advantageous option – the results of Frederick’s experiments show that there is no simple direct relationship between scoring high on CRT and automatically preferring future higher rewards. This was probably influenced by the fact that the high CRT group perceived themselves as both less impulsive and more concerned about inflation. Frederick (2005) concludes that “one observes considerable differences between CRT groups for choices like those in items *a* and *b*¹, where more careful deliberation or “cognitive reflection” should argue strongly in favor of the later larger reward, but negligible differences for many of the other items, for which additional reflection would not make such a strong case for the larger later reward” (p.31). Similar results showing the low CRT group as less patient than the high CRT group were found by Oechssler, Roeder and Schmitz (2009).

In the current study, we only used the items *a* to *d* of the *temporal discounting the future task* used by Frederick (2005), which covered hypothetical choices between an immediate reward and a larger delayed reward, that proved to be significant in the Frederick’s study.

Risk seeking preference

The influence of cognitive ability on risk seeking behavior is not widely studied but one can expect that people with higher rationality should be more able to choose according to the expected values of gambles (and thus employing the algorithmic capacity of their minds). Donkers, Melenberg and van Soest (2001) found that more educated participants were more tolerant of risk in hypothetical gambles. Benjamin and Shapiro (2005; in Frederick, 2005) also found a positive relationship between cognitive aptitudes tests and choice according to expected value for real decisions involving small stakes. Frederick (2005) found that the high CRT group was more willing to gamble. This was pronounced especially for higher expected value gambles, but it was clear even for middle stakes. On the other hand, for items involving losses the high CRT group was less risk seeking and more willing to accept a sure loss to avoid playing a gamble with a lower expected value. These results have a curious implication for prospect theory (Kahneman, Tversky, 1979) because they indicate that prospect theory holds for average to low cognitively reflective individuals, while highly cognitively reflective individuals are immune to such a switch of risk aversion to risk seeking when the valence of a gamble changes from positive to negative (as prospect theory suggests).

¹Frederick (2005) used 4 categories of time preferences: items *a – e* represented choices between an immediate reward and a larger delayed reward, items *f – h* represented choice between an immediate reward and a sequence of delayed rewards, item *i* represented choice between shorter more immediate message and longer more delayed message, and items *j – k* represented a choice between a smaller immediate loss or a larger delayed loss. In our study we employed only items *a – e*, because they showed as the most significant in Frederick’s studies.

Oechssler, Roeder and Schmitz (2009) similarly found that the high CRT group was more likely to choose the alternative that is compatible with risk neutrality. They argue that people with higher cognitive abilities (as reflected by CRT) might save more and receive higher expected returns, which can lead to their pronounced role in a financial markets.

Similarly, we only used the items *a* to *h* from risk seeking preferences that were significant in the Frederick's study. Concrete items are shown in Table 2.

COGNITIVE STYLES

Belief in the superiority of intuition

Although many cognitive style measures are roughly based on a distinction between Type 1 (intuitive) and Type 2 (reflective) processing, they more reflect a personal variation and preference in using Type 2 processes (Evans, Stanovich, 2013). Usually, they measure two dimensions reflecting the preference for an intuitive or reflective mode of processing. However, in Stanovich's model (2008) intuition is viewed mainly as an opposite pole from rationality and faith in intuition is hypothesized to be negatively associated with the normative response in many rationality thinking tasks. In our study, cognitive styles were measured by four questionnaires: the overview of the measures together with numbers of participants is shown in Table 1.

Preference for Intuition and Deliberation (PID, Betsch, 2004) is an instrument for measuring affective intuition, defined as being based on implicit knowledge and also as a basic mode of decision making, which uses affect as a decision criterion (high preference for intuition relates to Type 1 processes and reliance on instant impressions and emotions). On the other hand, deliberation is defined as decision making using explicit evaluation, rules of deciding, beliefs and reasons (high preference for deliberation relates to trusting rather Type 2 processes). PID consists of 18 items; 9 for PID-Intuition (e.g. "I listen carefully to my deepest feelings.") and 9 for PID-Deliberation (e.g. "Before making decisions I usually think about the goals I want to achieve."). Participants indicated their agreement with these statements on a 5-point Likert scale with 1 meaning "totally disagree" to 5 "totally agree". Our participants were in general slightly more deliberative ($M=35.4$, $SD=5.1$) than intuitive ($M=31.2$, $SD=5.1$).

Rational-Experiential Inventory (REI, Epstein, Pacini, Denes-Raj, Heier, 1996) consists of two formerly separate scales – Need for Cognition (basis for REI-Rationality) and Faith in Intuition (basis for REI-Experientiality) and this makes it compatible with Stanovich's recent (2011) framework, in which REI-R should be related to rationality and thinking tasks and REI-E should be negatively correlated with rationality. We used a 6-point scale to indicate agreements with REI items. This inventory was chosen because it has the best psychometric characteristics of inventories that measure intuitive and reflective cognitive styles (Hanák, 2013). It was also verified in a Slovak environment (Ballová Mikušková, Hanák, Čavojová, under review). Again, participants in general showed a slight inclination for rational processing ($M=80.83$, $SD=13.7$) than experiential processing ($M=75.14$, $SD=14.4$).

Cognitive Style Indicator (CoSI, Cools, Van den Broeck, 2007) is an 18-item questionnaire measuring people's preferences for perception, processing, and structuring of information. Items are scored on a 5-point scale, with 1 meaning "totally disagree" to 5 "totally agree". Authors distinguish three cognitive styles: Knowing, Planning, and Creating. The Knowing style is defined within CoSi as an emphasis on logic, precision and objectivity, which makes it similar to more traditional scales measuring reflective cognitive styles. Planning emphasizes structure, control and routine, while Creating emphasizes subjectivity, impulsivity and openness to possibilities. In CoSi, the most preferred style was Planning ($M=27.46$, $SD=4.3$), then Creativity ($M=25.87$, $SD=4.6$, and lastly Knowing ($M=14.82$, $SD=2.3$).

Myers-Briggs Type Indicator (MBTI, Myers, Briggs et al., 1998) helps to identify basic preferences for perception and judgment in terms of mental functions in combination with various attitudes, such as orientation of energy and orientation towards the external world. MBTI consists of four dichotomous categories: 1) perception: sensing (S) and intuition (N), 2) judgment: thinking (T) and feeling (F), 3) energy: extraversion (E) and introversion (I), and 4) orientation toward the world: judgment (J) and perception (P). In the present study, we used the version published by Kovacs (2011). Based on studies criticizing dichotomized categories (e.g. Arnau, Green, Rosen, Gleaves, Melancon, 2003), which is created by a forced choice between statements, we used categories as separate dimensions where each item was evaluated on a scale from 0 to 5. For the current analysis we included only subscales related to judgment (Thinking, Feeling) and perception (Intuition, Sensing). Our sample preferred more

Sensing (M=40.98, SD=6.6) than Intuition (M=31.62, SD=7.2), and Thinking (38.29, SD=8.4) over Feeling (M=34.16, SD=8.1).

Table 1: Overview of used measures within Stanovich's (2011) taxonomy of rationality assessment

Fluid rationality			Crystalized rationality
Resistance to Miserly Processing	Prudently Discounting of Future	Absence of Irrelevant Context Effect	Belief in Intuition
1. Cognitive Reflection Test (Frederick, 2005), N=531	4. Preference for Patience (Frederick, 2005), N=523	5. Preference for Risk Seeking (Frederick, 2005), N=515	1. PID (Betsch, 2004), N=446
2. Jelly bean task (Kirkpatrick, Epstein, 1992), N=531			2. REI (Pacini, Epstein, 1999), N=498
3. Wason's Selection Task, N=531			3. CoSi (Cools, van den Broeck, 2007), N=388
			4. MBTI (Myers Briggs, McCaulley, Quenk, Hammer, 1998), N=396

RELATIONSHIPS BETWEEN TESTING PARADIGMS

Assessing fluid rationality

Firstly, we investigated the possibility of creating a composite score for Resistance to Miserly Processing (as a major dimension of fluid rationality), but the five items gave Cronbach's alpha at only 0.584 level. After deleting Wason's Selection Task and the Jelly bean task, Cronbach's increased to 0.749. In Wason's Selection Task there were only 2 participants out of 531 who gave the correct answer, therefore due to a floor effect we excluded it from further analysis. The Jelly bean task shows a small, but significant correlation with CRT ($r=0.118$, $p=0.007$), but it does not bring substantial additional effect. Thus, in further analyses we treat each measure separately and not as a composite score of Resistance to Miserly Processing.

We divided participants into extreme groups according to their overall score in CRT – those giving all three correct answers were assigned into high CRT group (N=139, 26.1%) and those giving no correct answer were assigned into low CRT group (N=114, 21.4%). To analyze the associations between cognitive reflection (measured by CRT) with other dimensions of fluid rationality (time preference and risk preference) from Stanovich's (2011) taxonomy we employed chi-square analyses and the results are shown in Table 2.

Table 2: Intertemporal and risk seeking preferences for Low and High CRT Groups

	low CRT		high CRT		χ^2	df	p Value	Cramer's V
time preference	n	%	n	%				
a €3400 this month	91	33.30%	18	22.50%	3.402	1	0.042	0.098
or €3800 next month	182	66.70%	62	77.50%				
b €100 now	228	83.50%	65	81.30%	0.225	1	0.373	0.025
or €140 next year	45	16.50%	15	18.80%				
c €100 now	173	64.10%	57	71.30%	1.41	1	0.146	0.063
or €1100 in 10 years	97	35.90%	23	28.70%				
d €9 now	200	73.30%	57	70.40%	0.262	1	0.352	0.027
or €100 in 10 years	73	26.70%	24	29.60%				
risk preferences								
a €1000 for sure	116	43,10%	29	36,30%	1,199	1	0,167	0,059

	<i>or a 90% chance of €5000</i>	153	56,90%	51	63,70%				
b	<i>€100 for sure</i>	112	41,00%	23	28,40%	4,224	1	0,026	0,109
	<i>or a 90% chance of €500</i>	161	59,00%	58	71,60%				
c	<i>€1000 for sure</i>	207	75,80%	55	67,90%	2,039	1	0,101	0,076
	<i>or a 75% chance of €4000</i>	66	24,20%	26	32,10%				
d	<i>€100 for sure</i>	221	81.00%	66	81.50%	0.011	1	0.529	0.006
	<i>or a 75% chance of €200</i>	52	19.00%	15	18.50%				
e	<i>€100 for sure</i>	246	90.10%	77	95.10%	1.917	1	0.12	0.074
	<i>or a 75% chance of €150</i>	27	9.90%	4	4.90%				
f	<i>€100 for sure</i>	231	84.60%	70	86.40%	0.16	1	0.421	0.021
	<i>or a 50% chance of €300</i>	42	15.40%	11	13.60%				
g	<i>€500 for sure</i>	184	67.40%	38	46.90%	11.211	1	0.001	0.178
	<i>or a 15% chance of €1000000</i>	89	32.60%	43	53.10%				
h	<i>€100 for sure</i>	245	89.70%	64	79.00%	6.483	1	0.012	0.135
	<i>or a 3% chance of €7000</i>	28	10.30%	17	21.00%				

In contrast with Frederick (2005), our results showed that high CRT is associated with only 1 out of 4 preferences for a more patient option yielding higher reward, and 3 out of 8 preferences for a riskier option. After calculating a composite score for both time preference and risk preference (we used the same procedure as Frederick, 2005) and correlating it with total CRT scores (Table 3), there is only a slight correlation with riskier preference and no correlation with preference for the more patient option. Also Cramer's V shows only weak relationships.

Oechssler et al. (2009) similarly found more pronounced differences between low and high CRT groups in risk seeking preference than in time preferences. They employed slightly differently formulated items (and only two for risk preferences and one for time preference).

Predicting rational performance from cognitive styles

Next, we analyzed relationships between measures of fluid rationality and cognitive styles. As we stated earlier, measures of rational thinking showed only weak correlations among themselves (CRT correlating with a composite score for risk preference, $r=0.09$, $p=0.04$ and the jelly bean task, $r=0.118$, $p=0.007$; and time preference correlating with risk preference, $r=0.24$, $p<0.001$, and Wason's task, $r=0.12$, $p=0.005$).

Measures of cognitive style showed only a few weak correlations with rational thinking measures. The most useful was REI-R scale that correlated with three other measures: CRT ($r=0.14$, $p<0.01$), time preference ($r=0.105$, $p<0.05$), and risk preference ($r=0.115$, $p<0.05$). PID and CoSi subscales showed no correlations with other cognitive tasks, which is in contrast with the findings of Witteman et al. (2009). MBTI showed an interesting pattern of correlations with other cognitive tasks – intuition correlated positively with preference for the riskier option ($r=0.112$, $p<0.05$). Sensing (as opposed to intuition) correlated negatively with CRT ($r=-0.099$, $p<0.05$). However, all significant relationships were only very weak and explained only a few percent of the variance. Cognitive style measures designed to tap into individual differences in reflective and intuitive thinking show, in fact, only a poor ability to predict performance in CRT as a prime measure of resistance to miserly processing. Apart from REI-R, no subscale of PID, MBTI, or CoSi could be used as a useful predictor for CRT total score.

The REI-R consists of two parts, each having a different predictive ability. Rational engagement when used as a predictor was only close to significance ($p = 0.057$); rational ability was the better predictor ($\beta = .162$, $t(498) = 3.655$, $p = .000$), but still explained only a very small part of the variance, $R^2 = .024$, $F(1, 496) = 13.36$, $p = .000$. From four widely used cognitive style questionnaires (having 15 subscales altogether together), only Rational Ability (REI) could be used as a significant predictor for successful reflective thinking (CRT). Toplak et al. (2013) found a stronger correlation between CRT and thinking disposition ($r=0.3$), which consisted also of the Need for

Cognition Scale (which is the basis for REI-R), but probably because they calculated the correlation for a composite score of thinking disposition (Need for Cognition, Actively Open-Minded Thinking, reversed Superstitious Thinking, and Consideration for Future Consequences).

Gender differences in rationality

Because men seem to do better in CRT than women (Frederic, 2005, Toplak et al., 2013) we also analyzed gender differences in cognitive tasks (Table 3a) and cognitive styles (Table 3b).

Table 3a: Gender differences in cognitive tasks and preferences

	men		women		t	p-value	Effect size, r
	M	SD	M	SD			
CRT	1.19	1.17	0.84	1.10	3.35	0.001	0.14
Jelly bean task	0.65	0.48	0.61	0.49	0.877	0.381	0.04
composite score for Time preferences	0.44	0.29	0.34	0.28	3.527	0.000	0.15
composite score for Risk preferences	0.37	0.22	0.27	0.23	4.561	0.000	0.20

We found gender differences in CRT and almost all preferences – men seemed to be more patient in waiting for a larger reward and also generally prefer the riskier option with the more advantageous pay-off. Although differences are significant, effect sizes are close to zero or minimal. From Table 4b we can also see gender differences in cognitive style. Women scored higher in styles associated with a preference for intuition and feeling (PID_I, REI_E, MBTI_F), men scored higher in styles associated with analysis and deliberation (PID_D, REI_R, CoSi_c, MBTI_T). An exception to this was a higher score in CoSi_k and MBTI_N for men. Only in CoSi_p were there were no gender differences. As in the previous table we can see in table 4b that men differ significantly from women in almost all scales but the effect sizes are minimal or very small, explaining no more than a few percent of the variance. These results complement our previous findings (Ballová Mikušková et al., under review), where we found that combined measures of cognitive styles tapping into intuition vs. deliberation generally load to three factors, which we labeled “decision-making based on cognitions”, “decision-making based on affect and holistic processing” and “planned, structured decision-making”, and it seems likely that the preference of a particular style is affected by gender.

Table 3b: Gender differences in cognitive styles

	men		women		t	Sig. (2-tailed)	Effect size, r
	M	SD	M	SD			
PID_I	30.19	4.97	31.56	5.04	-2.618	0.009	0.12
PID_D	36.20	4.88	35.10	5.23	2.086	0.038	0.10
REI_R	86.92	12.67	78.34	13.45	6.561	0.000	0.28
REI_E	72.43	13.62	76.27	14.54	-2.732	0.007	0.12
CoSi_c	26.95	4.55	25.46	4.51	2.901	0.004	0.15
CoSi_k	15.49	2.57	14.58	2.83	2.895	0.004	0.15
CoSi_p	27.50	4.44	27.45	4.33	0.088	0.930	0.00

MBTI_N	33.84	7.76	30.78	6.74	3.871	0.000	0.19
MBTI_S	39.72	7.13	41.47	6.33	-2.361	0.019	0.12
MBTI_T	41.20	7.89	37.18	8.40	4.323	0.000	0.21
MBTI_F	31.95	7.72	35.01	8.11	-3.391	0.001	0.17

Differences between men and women in the cognitive styles are supported by CRT scores, where men differ from women significantly (Table 3a).

Gender differences, which were significant among cognitive styles and CRT, do not significantly interact together in a moderation analysis, only in combination with REI-R and PID-D. Specifically, although men and women differ in CRT score and also in PID-D, moderation analysis (with CRT score as the outcome, PID-D predictor and gender moderator) produces a model that is insignificant, and also an interaction between PID score and gender that is insignificant. This model explains minimal variance. The same results were found with REI-E as predictor.

Results are very different however, when investigating PID-I and REI-E. A moderation model with CRT score as outcome, PID-I predictor, moderated by gender, was found to be significant. The same results were found with REI-E as predictor. In both cases, the model is highly significant, as is the interaction between scales and gender and each is significant predictor alone. However, both models explain only a few percent of CRT score variance and the increase due to interaction is about 2%. Based on these results we could speculate that it is not high levels of self-reported reflective cognitive style which means that your CRT is higher. It is high levels of intuitive cognitive style in women, which makes your CRT lower. This result would provide partial support for Stanovich's (2011) claim that faith in intuition undermines rational thinking.

Comparison with other studies – comment on intercultural differences

In this section we summarize our results in terms of their comparison with findings from other cultural backgrounds. In Table 4 we can see that successful CRT answers and their distribution are comparable to Frederick's (2005) results, with our sample situated somewhere in the middle range between top colleges like MIT and Princeton and low-end colleges like University of Toledo. Results are similar to Frederick's (2005) web-based studies in terms of number of participants and also achieved level of success, possibly reflecting the heterogeneity of our sample comparable with web sampling. Our participants also achieved similar scores to participants of non-American samples (e.g. Canadian sample – Toplak et al., 2011, 2013, or Argentinian sample – Campitelli, Labollita, 2010), with the exception of the German sample in Oechssler et al. (2009) study.

It is likely that rather low levels in CRT were caused by the fact that our sample from a private college consisted of students accepted without passing entrance exams. Sirota and Juanchich (2010) on similar sample from a more elite college with hard entrance exams found a higher mean score ($M=1.3$), though their sample was smaller than ours. As Toplak et al. (2013) noted, in some populations the overall score on the three-item version of CRT might be floored, and thus the CRT is in need of supplement and extension. We agree, because our experience and results (e.g. Čavojská, Hanák, Ballová Mikušková, 2014) show the rather unflattering fact that even university students are not very successful in a variety of cognitive tasks extensively used in the JDM literature. We believe these results are caused by our unfortunate educational system rather than essential differences in rational thinking between Slovaks and other nations (e.g. Ballová Mikušková, Hanák, Čavojská, under review).

We also confirmed gender differences in the expected direction in CRT and other measures of rationality, which is in line with previous studies (Frederick, 2005, Toplak et al., 2013, Oechssler et al., 2009) and can also explain the rather low scores of our sample, as women tend to score lower in CRT and our sample consists predominately of women (69 %). On the other hand, Campitelli and Labollita (2010) did not find significant differences between men and women in their sample, though they observed a similar trend and their sample was similar to ours (73% of women).

Table 4: CRT results for several studied samples

Locations ate which data were collected	Mean CRT score	Percentage scoring 0,1,2,3				
		"Low"	" High"			
		0	1	2	3	N
<i>Our study – College of public service</i>	0.95	51.4%	17.9%	15.4%	15.3%	531
<i>Sirota, Juanchich (2010)* -- Comenius University</i>	1.3	58.51%			32.97%	94
<i>Frederick's studies (Frederick, 2005)</i>						
Massachussets Institute of Technology	2.18	7%	16%	30%	48%	61
Princeton University	1.63	18%	27%	28%	26%	121
Web-based studies	1.10	39%	25%	22%	13%	525
Bowling Green University	0.87	50%	25%	13%	12%	52
University of Toledo	0.57	64%	21%	10%	5%	138
Overall	1.24	33%	28%	23%	17%	3428
<i>Toplak, West, Stanovich (2013) study</i>	0.49					160
<i>Topalk, West, Stanovich (2011) study</i>	0.7	55.8%			6.6%	346
<i>Oechssler, Roider, Schmitz (2009)</i>	2.05	10.1%	17.7%	30.7%	41.5%	564
<i>Campitelli, Labollita (2010)</i>	0.65	59%	25%	11%	6%	157

Notes: *Percentages were calculated from the provided data, Low CRT group and high CRT groups were created by median split, where mean score for low CRT group was 0.5 and high CRT was 2.5.

We found a different pattern of connection between CRT and time and risk preferences than in the original study by Frederick (2005). A possible reason for the insignificant results in time preference could be very big differences in the possible returns. Options differ in terms of time necessary for waiting (1 month in *a* vs. 120 months in *c*, *d*), and also in return amount (€400 vs. €40). Probably, students of management made these simple financial comparisons by taking into account time, return, return ratio to average wage in Slovakia and inflation, and thus recognized that the best option for Slovak financial environment is option *a* (highest return, shortest time, high return ratio compared to the average Slovak monthly wage). Campitelli and Labollita (2010) found that the majority of their sample behaved impatiently and safely in comparison with Frederick's (2005) study. They argue that it could be caused by high inflation rates in their country (Argentina), which is not the case with Slovakia, as inflation in recent years is approaching its minimum.

CONCLUSIONS

How do our results contribute to the proposed model of fluid rationality (Stanovich, 2011)? A lack of mutual relationships together with a low internal consistency for a composite score of resistance to miserly processing (cognitive reflection, resistance to attribute substitution, and diagnostic hypothesis testing) rather undermined the unitary notion of resistance to miserly processing (measured by different testing paradigms). There were also very weak relationships between other constructs which were hypothesized by Stanovich (2011) to form other dimensions of fluid rationality together with resistance to miserly processing. On the other hand, such results seem to support Stanovich's (2011) hypothesis about multifactorial fluid rationality.

Our results in general point to the need for further refinement of the measures used to assess rational thinking, as these are typically very difficult for general (less elite) populations. This leads to either a pessimistic conclusion about the rationality of people in general or to a conclusion about the necessity of possessing specialized knowledge for solving these rationality tasks. Another issue with rational thinking tasks, which deserves attention, is that some of these tasks are beginning to be widely known and once a person acquires the solution for such tasks they become useless in further research (Toplak et al., 2013, Kostovičová, Konečný, Dudeková, 2013).

Another issue worth emphasizing is the general lack of a relationship between rationality thinking tasks

(performance based measures of rational thinking) and cognitive styles (self-reported measures of preference for rational thinking). Self-report cognitive style inventories, such as PID, CoSi or MBTI did not predict performance in the cognitive reflection test. Only rational ability (from REI) could be used as a self-reported tool useful for predicting related performance in rational thinking tasks. Our findings are in accordance with Newstead et al. (2004) who came to the similar conclusions.

However, analysis of gender differences revealed that high levels of intuition in the female group as measured by REI negatively affected miserly processing (resulting in low scores in CRT). Stanovich (2011) lists Faith in intuition as a crystallized inhibitor of rational thinking, but our results highlight the importance of analyzing the effect of faith of intuition separately for men and women. It seems to be a case that gender differences in cognitive reflection (and thus higher resistance to miserly processing) are not caused by lower cognitive abilities of women in our sample, but rather their preference for affective and holistic decision making that resulted in satisfying with the appealing intuitive (though incorrect) answers in CRT.

Generally, it seems unsatisfactory to examine rational thinking only in regard to resistance to miserly processing (or cognitive reflection) or any thinking disposition. Rather, it would seem advantageous to test the ability of such measures to predict any real-life outcomes. Many accounts of rationality embrace the notion of instrumental rationality – the rational is any behavior that leads to fulfilling personal goals, which makes its measurement rather difficult (it is difficult to establish a valid real-life outcome if it is not desirable for every person). We believe it is necessary to include into future battery tests of rational thinking measurements of personal beliefs and values, some of which probably facilitate rationality and some of them detrimental to it.

ACKNOWLEDGMENT

This work was supported by Slovak grant agency VEGA 2/0064/13 "Decision making of experts: using intuition for solution of strategic tasks" awarded to Vladimíra Čavojová.

REFERENCES

- Arnau, R. C., Green, B. A., Rosen, D. H., Gleaves, D. H., Melancon, J. G. (2003), "Are Jungian preferences really categorical?: An empirical investigation using taxometric analysis". PERSONALITY AND INDIVIDUAL DIFFERENCES, Volume 34, pp. 233–251.
- Ballová Mikušková, E., Hanák, R., Čavojová, V. (under review), "Intuition versus deliberation: Psychometric properties of two self-reported scales". JOURNAL OF PERSONALITY ASSESSMENT.
- Betsch, C. (2004), „Präferenz für Intuition und Deliberation (PID): Inventar zur Erfassung von affekt- und kognitionsbasiertem Entscheiden“. ZEITSCHRIFT FÜR DIFFERENTIELLE UND DIAGNOSTISCHE PSYCHOLOGIE, Volume 25, No. 4, 179–197.
- Buchtel, E.E., Norenzayan, A. (2009), "Thinking across cultures: Implications for dual processes." In Evans, J.St.,B.T., Frankish, K. (Eds.), IN TWO MINDS. DUAL ROCESSSES AND BEYOND. Oxford: Oxford University Press. pp. 217 – 238.
- Campitelli, G., Labollita, M. (2010), "Correlations of cognitive reflection with judgments and choices". JUDGMENT AND DECISION MAKING, Volume 5(3), pp. 182–191.
- Cools, E., Van den Broeck, H. (2007), „Development and validation of the Cognitive Style Indicator“. THE JOURNAL OF PSYCHOLOGY, Volume 141, 359–387.
- Čavojová, V., Ballová Mikušková, E., Hanák, R. (2013), "Preferencia k deliberácii: (ne)istota úspechu v kognitívnych úlohách". In KOGNITIVNÍ VĚDA A UMĚLÝ ŽIVOT. Opava: Slezská univerzita v Opavě, pp. 53-58.
- Čavojová, V., Hanák, R., Ballová Mikušková, E. (2014), "Educating rationality: Assessing Master Rationality Motive in future teachers". In INTED2014 Proceedings.
- Denes-Raj, V., Epstein, S. (1994), "Conflict between intuitive and rational processing: When people behave against their better judgment". JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY, Volume 66, 819–829.
- Donkers, B., Melenberg, B., van Soest, V. (2001), "Estimating Risk Attitudes Using Lotteries: A Large Sample Approach." JOURNAL OF RISK AND UNCERTAINTY. Volume 22, No.2, pp. 165-95.
- Epstein, S. (2003), "Cognitive-Experiential self Theory of Personality". In COMPREHENSIVE HANDBOOK OF PSYCHOLOGY (Vol. 5, pp. 159–184.).
- Epstein, S., Pacini, R., Denes-Raj, V., Heier, H. (1996), "Individual differences in Intuitive-Experiential and Analytical-Rational Thinking Styles". JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY, Volume 71, No.2, 390–405.
- Evans, J. S. B. T., Stanovich, K. E. (2013), "Dual-Process Theories of Higher Cognition: Advancing the Debate". PERSPECTIVES ON PSYCHOLOGICAL SCIENCE, Volume 8(3), pp. 223–241.

- Frederick, S. (2005), „*Cognitive Reflection and Decision Making*“. JOURNAL OF ECONOMIC PERSPECTIVES, Volume 19, No. 4. 25–42.
- Gigerenzer, G. (2008), „*Rationality for Mortals. How people cope with uncertainty*“. New York: Oxford University Press.
- Hanák, R. (2013), „*Prediktívna validita dotazníkov merajúcich intuíciu*“ [Predictive validity of PID, REI and GDMS questionnaires measuring intuition]. In R. Hanák, E. Ballová Mikušková, V. Čavojová (Eds), ROZHODOVANIE A USUDZOVANIE IV. APLIKÁCIE A LIMITY INTUÍCIE. Bratislava: "Ustav experimentálnej psychológie SAV.
- Kahneman, D., Frederick, S. (2002), „*Representativeness revisited: At-tribute substitution in intuitive judgment*.“ In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT. New York: Cambridge University Press, pp. 49–81.
- Kahneman, D., Tversky, A. (1979), "Prospect Theory: An Analysis of Decision Under Risk." ECONOMETRICA. Volume 47, No. 2, pp. 263–91.
- Kahneman, D., Tversky, A. (1982), „*On the study of statistical intuitions*“. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES. Cambridge, England: Cambridge University Press. pp. 493–508.
- Kirkpatrick, L.A., Epstein, S. (1992), „*Cognitive-experiential self-theory and subjective probability: Further evidence for two conceptual systems*“. JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY, Volume 63, 534–544.
- Kostovičová, L., Konečný, M., Dudeková, K. (2013), „*Čo to vlastne meriame? Reflexia testu kognitívnej reflexie : What do we actually measure? Reflection on the cognitive reflection test.*“ In OTÁZNIKY ROZHODOVANIA: TEÓRIA, EMPÍRIA, ŽIVOT: 31. Psychologické dni: zborník príspevkov. Editor Ivan Sarmány-Schuller. Bratislava : Stimul, pp. 166–171.
- Kovács, J. (2011), „*Manažérske dovednosti*“, Ostrava: VŠB - Technická univerzita. pp. 206.
- Myers Briggs, I., McCaulley, M. H., Quenk, N. L., Hammer, A. L. (1998), „*MBTI Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator*“, Third Edit., p. 420. Palo Alto: Consulting Psychologist Press.
- Newstead, S. E., Handley, S. J., Harley, C., Wright, H., Farrelly, D. (2004), „*Individual differences in deductive reasoning*“. THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY. A, HUMAN EXPERIMENTAL PSYCHOLOGY, Volume 57, No.1. 33–60.
- Nickerson, R.S. (2008), „*Aspects of Rationality. Reflections on what it means to be rational and whether we are.*“ New York: Psychology Press.
- Oechssler, J., Roider, A., & Schmitz, P. W. (2009), „*Cognitive abilities and behavioral biases*“. JOURNAL OF ECONOMIC BEHAVIOR & ORGANIZATION, Volume 72(1), 147–152.
- Pacini, R., Epstein, S. (1999), „*The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon*“. JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY, Volume 76, No. 6, 972–987.
- Saunders, C., Over, D.E. (2009), „*In two minds about rationality?*“ In Evans, J.St.,B.T., Frankish, K. (Eds.), IN TWO MINDS. DUAL PROCESSES AND BEYOND. Oxford: Oxford University Press. pp. 317 – 334.
- Sirota, M., Juanchich, M. (2011), „*Role of numeracy and cognitive reflection in Bayesian reasoning with natural frequencies.*“ STUDIA PSYCHOLOGICA, Volume 53(2), pp. 151–161.
- Sirota, M., Juanchich, M., Hagmayer, Y. (2013), „*Ecological rationality or nested sets? Individual differences in cognitive processing predict Bayesian reasoning.*“ PSYCHONOMIC BULLETIN & REVIEW, Volume 06, DOI:10.3758/s13423-013-0464-6
- Sobyra, J. (2010), „*The accuracy of self-reported intuitive and analytical ability.*“ HONORS PROJECT. Paper 140. Research report.
- Stanovich, K. E. (2011), „*Rationality and the Reflective Mind*“. New York: Oxford University Press.
- Sternberg, R.J., Ed., (2002), „*Why smart people can be so stupid?*“ New Haven: Yale University Press.
- Toplak, M.E., West, R.W., Stanovich, K.E. (2013), „*Assessing miserly information processing: An expansion of the Cognitive Reflection Test*“, THINKING & REASONING, <http://dx.doi.org/10.1080/13546783.2013.844729>
- Tversky, A., Kahneman, D. (1974), „*Judgment under Uncertainty: Heuristics and Biases*“. SCIENCE, Volume 185(4157), 1124–31. doi:10.1126/science.185.4157.1124
- Witteman, C., van den Bercken, J., Claes, L., Godoy, A. (2009), „*Assessing Rational and Intuitive Thinking Styles*“, EUROPEAN JOURNAL OF PSYCHOLOGICAL ASSESSMENT, Volume 25 No. 1. 39–47.

Analytical Capability to Better Understand and Anticipate Extremist Shifts Within Populations in Failing States

Michael L. Bernard & Asmeret B. Bier

*Cyber Engineering Research Institute
Sandia National Laboratories
Albuquerque, NM 87123-1327 USA*

ABSTRACT

The difficulty in adequately assessing geopolitical and sociocultural dynamics of extremist groups has led to failures in understanding, anticipating, and effectively responding to shifts in their movements and allegiances. Recent attacks in Africa highlight the need to more precisely understand and anticipate changes in societal attitudes and behaviors due to radicalization. This is particularly important as new extremist cells and affiliates have sprung up in parts of Southeast Europe, Asia, and Africa. A significant concern is their stated intent to plan and conduct attacks against populations within these regions. This paper describes an effort to build upon existing capabilities to assess the phenomena that gives rise to the support for extremism, shifts in allegiances, and active engagement in violent acts against indigenous populations. The focus of this effort is to assess how the dynamics of allegiance formation between various groups and society are impacted by conflict and by third-party interventions. We also seek to help determine how and why extremist allegiances co-evolve over time due to changing geopolitical, sociocultural, and military conditions. The aim of this paper is to discuss our initial effort to assess the dynamic interactions between an extremist group and an indigenous population over time.

Keywords: Cognitive Modeling, Behavior Influence Assessment, Extremist Groups, Social Modeling, Systems Modeling

INTRODUCTION

This paper describes an initial effort by Sandia National Laboratories to develop a capability to help assess the phenomena that give rise to active engagement in extremism and its effects on societies. The intent of this effort is to investigate underlying attitudinal and behavioral shifts over time due to the influence of extremist groups. Our test case involves extremist groups in Africa. In this case, we are considering how certain decisions affect economic and social stability in different parts of Africa and how the resulting tension may affect this society. This region was selected because of the immediate and long-term threats posed by extremist groups. For this effort we are defining an extremist group as composed of individuals that display preoccupation with an ideology, religion, or political cause to such a degree that it leads to pursuit of violence as a tactic or strategy for imposition of its members' views on mainstream groups (see also Finlay, 2010).

To accomplish this effort, we are computationally modeling the interaction dynamics within and between transnational extremist groups in response to military, social, economic, and political intercessions. We are using a data- and theory-supported assessment capability, Behavioral Influence Assessment (BIA), to better understand and anticipate (with quantifiable uncertainty) how the dynamics of allegiance formations between various groups and

society are impacted by active conflict and by third-party interventions. We are also considering how and why extremist allegiances co-evolve over time due to changing geopolitical, sociocultural, and military conditions. BIA is anchored to a theoretically consistent psychological, social, and economic foundation for analyzing behaviors of individuals and populations over time in response to military, diplomatic, and other influences (Backus et al., 2010; Bernard & Backus, 2009; Bier et al., 2011). BIA assesses multiple scales of behavior associated with the external world. The system emphasizes response/counter-response actions—the recognition of which can reduce surprises and counter-productive influences. The underlying modeling structure is based on a synthesis of expert opinion and data, including cultural and institutional constraints and conditions. Formal validation and analysis of simulation results characterize model confidence and robustness, including probability ranges of success despite uncertainty. The model uses a synthesis of data-supported, psychological and social (psychosocial) theories of group decision-making, as well as theory-based analytical studies of socioeconomic data. The psychosocial element models key cognitive processes underlying how people make decisions and express behaviors. It incorporates such elements as belief/attitude/motivation formation, framing, and strengthening by incorporating theories that are (1) mutually consistent; (2) can be integrated into a complete representation of behavior; and (3) can be instantiated, tested, and verified using accessible data. Data and SME input are used to calibrate the model, test hypotheses regarding behaviors, and quantify uncertainty. The resulting capability can delineate the temporal unfolding of societal tensions due to the social, economic, environmental, and political stresses. The models are designed to be broadly applicable (with modification), across different ethnic, political, and social groups, including regional dynamics and rest-of-the-world reactions to behaviors. For a broader discussion of BIA see Bernard, Backus, & Bier; and Bier & Bernard in this issue of the proceedings.

Dynamic modeling of extremist groups

When we study societies that are significantly influenced by extremist groups (EG), we do not model the group in isolation, since no group operates in complete isolation. In fact, many of the EGs that exist today are at least loosely affiliated with other international groups (Pillar, 2001). Moreover, local group dynamics often depend on interactions within and between people and leaders that can be greatly affected by external influences. External influences can involve a wide range of events and actions, from the imposition of new government in the area of interest or in neighboring countries to natural disasters or military incursions by foreign powers. The effect of these influences depends on the nature of the society being affected.

Extremist groups commonly seek to garner support from the local community by providing services and dispensing their brand of justice and law. This involves interacting with the community and groups outside of the community. Often, events that occur outside the purview of a local group may affect that group through its interaction with more transnational, affiliate groups. Of course, how EGs work within the culture and society of local communities can have a large effect on the support they receive. Their interaction with established cultures, tribes, clans, and other indigenous populations is often an indication of their actual support in the community (Martins, 2008; Thaler et al., 2013). This focus, however, can clash with more transnational objectives of expanding the influence of the group by focusing on broader international issues. For example, the use of a more global presence within a country that is suspicious of foreigners may actually undermine support (Hansen, 2013). An understanding of the potential effects of these types of influences is important, since they can affect the support, and thus the strength and influence, of EGs. To broadly assess these types of influences, we focused our assessment on the interactions between the entities described above. Figure 1 shows the general influences within a society that we are currently modeling. Here, the interactions between the United States, other influencing countries, international bodies, and the country or region itself are being represented. Within the country or region, the specific governmental body and various groups (both extremist and non-extremist) are also being represented. Together, these entities interact with each other in varying degrees. Exogenous influences, such as global socioeconomic and geopolitical conditions, military capabilities, ecological conditions, and the flow of information, all have some degree of influence over these interactions. These interactions can change over time due to dynamic shifts in internal and external conditions. Thus, to better understand these influences, we believe it is necessary to not only model the behaviors of EGs, but also determine how their interactions with other groups, and the external or ‘physical’ world, dynamically co-evolves in response to events and conditions over time.

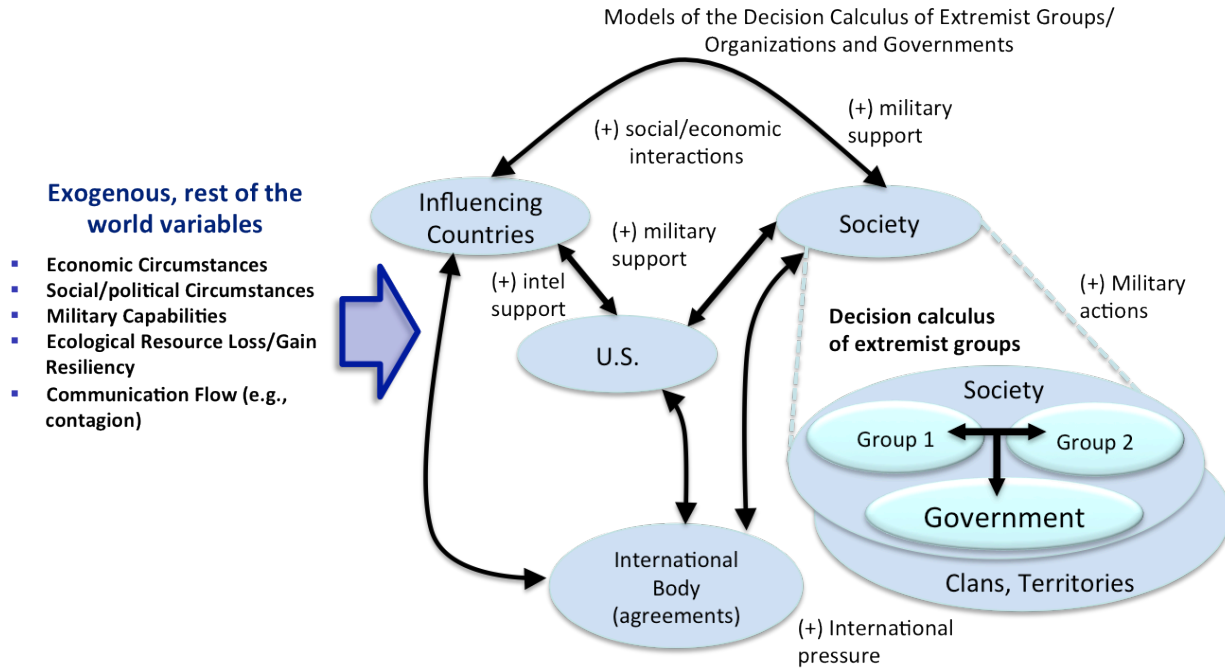


Figure 1. Conceptual view of the modeled interactions

SYSTEM MODEL DEVELOPMENT

Feedback processes among groups and the physical world unfold dynamically and can cause the outcome of an intervention to deviate from the desired direction. They can also lead to counter-responses that generate new concerns without improving the original issue. The delay between behaviors and impacts can cause secondary dynamics that make it extremely difficult to know whether the fluctuating behavioral responses and counter-responses will ultimately lead to the desired outcome. The computational modeling of these dynamic interactions needs to address this dynamic evolution, which is most readily modeled using differential equations. The differential equations used in these models not only simulate dynamics, but also causally describe why the dynamics occur (Bier et al., 2011).

The process for developing a BIA model using the system dynamics methodology starts with a description of the problem questions that are to be addressed. There is no attempt to model the entire system, but only those aspects of the system relevant to the focus problems. The next step is to develop a causal-loop diagram that relates all the interactions embodied in the theories (McFadden, 1984). The casual loop diagram is then extended (via a stock and flow diagram) to explicitly detail the flow of information and physical quantities through the system (Stermann, 2000). A key feature is the designation of stocks that represent things that accumulate—for example, the accumulation of information, experience, monetary, or physical quantities. These stocks are called “state variables” and they largely characterize the nature of the system and its responses. The difference in the value of stocks over time increments is the “differential” part of the computational modeling approach. The exact mathematical expression of theory is anchored in the flows into and out of the stocks. Only those theories that have a measurable meaning, supportable, at least in principle, by historical or experimental data, are included in the model. The data determines the parameters that control the progression of the simulated values through time. Rigorous statistical techniques determine the appropriate parameters and the uncertainty associated with their use. This uncertainty can later define the confidence in the results of an intervention analysis (Bier, 2010).

Causal loop diagram

Figure 2 shows the initial structure for one aspect of the full BIA assessment via a causal loop diagram. This initial assessment focuses on how EGs can use food as a means to obtain greater power and influence. A causal loop diagram illustrates the structure of a system using arrows to indicate causal relationships. Arrows with + symbols indicate a positive relationship (if the source variable changes, it causes the affected variable to change in the same direction), and arrows with – symbols indicate negative relationships (the two variables change in different directions). Positive and negative feedback loops can also be identified in these diagrams. Figure 2, for example, shows the complete causal loops regarding how EGs can affect the availability of food aid for communities. Here, the diagram indicates general relationships that will be discussed in detail below, shown in figures 3-7.

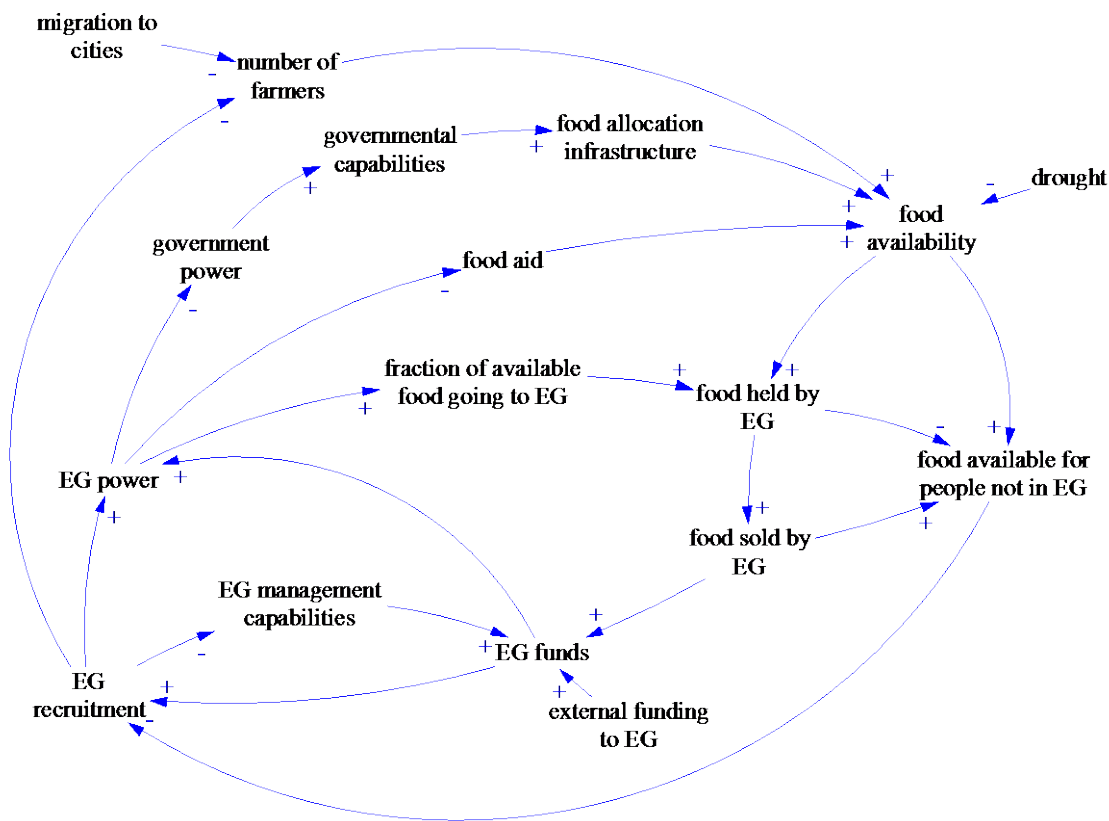


Figure 2. Causal loop diagram of EG interaction focusing on food availability and power

Figure 3 highlights a positive feedback loop pertaining to EG power and food allocation. In this scenario, the EG seeks to shift the allegiance of the population towards them. As EG recruitment increases (voluntary and/or through conscription) so does their power. This comes at the expense of governmental bodies (which can be the local government or international entities such as the United Nation, African Union, etc.). As the governmental bodies are weakened, their ability to allocate foodstuff is reduced. Thus, the availability of foodstuff for the indigenous population is reduced. This has the effect of increasing the attractiveness of EGs and the population's willingness to join an EG because of their ability to siphon off foodstuff from governmental bodies and potentially distribute it to segments of the population that are more receptive to the EG.

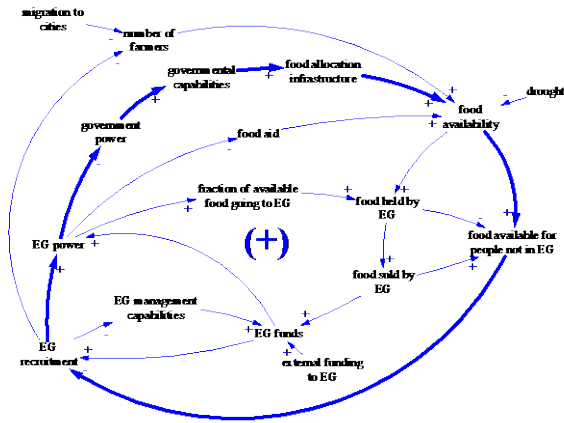


Figure 3. EG power/food allocation

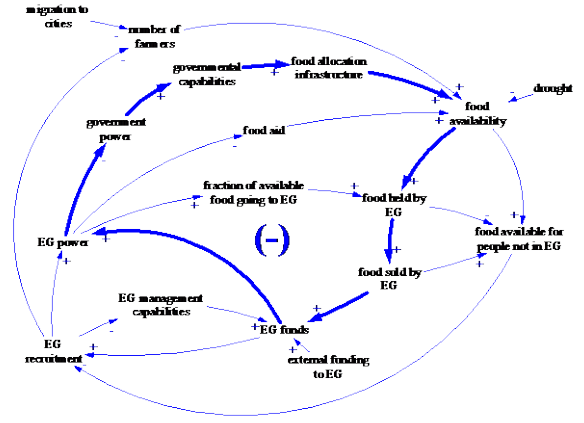


Figure 4. EG food allocation

Figure 4 shows how an EG can profit from foodstuff allocation. Here, reductions in the ability of governmental bodies to protect and properly distribute foodstuff reduces food availability for the population. A portion (sometimes as high as 50%) of the available foodstuff is taken by the EG to be used by their soldiers and sold for revenue (Harper, 2012). These funds help support EG recruitment, which increases the group's power and also increases their ability to seize more foodstuff. Other factors that affect this loop are external funds from transnational EGs, population migration away from farming areas to the cities, and droughts.

Figure 5 shows a negative feedback loop which demonstrates that as the EG grows, there is an associated greater need to manage the EG funds, foodstuff, recruitment activities, and the indigenous populations under their control. As the EG acquires greater responsibilities, they have more and more difficulty in managing these responsibilities. This causes greater inefficiencies and increases the likelihood for dissension within the population under their control. This situation can occur when an EG changes from a small protest and/or criminal organization to an organization that needs to manage larger populations. For example, when an EG assumes management of entire geographical regions, they need to administer such things as the police, schools, and infrastructure. The management of these things poses a risk to the organization if they are not well managed. The scenario where an EG has grown beyond its ability to manage itself and the population under it control has been seen in a number of cases (Besancon & Dalzell, 2013; Milton-Edwards, 2008). If it occurs, the EG may face new pressures (internal and external), which poses a risk to the survivability of that group.

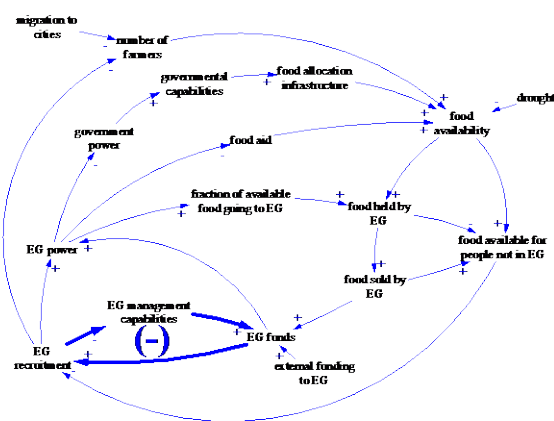


Figure 5. EG management

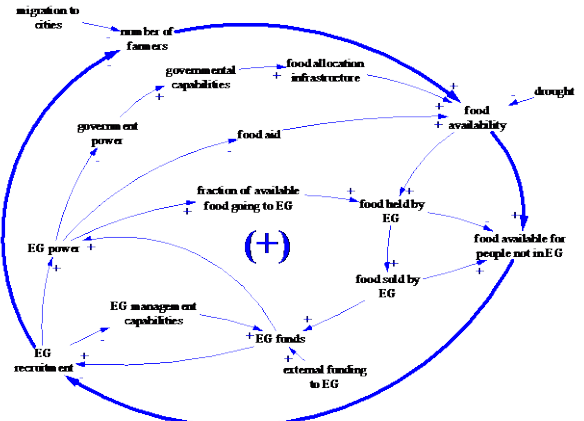


Figure 6. EG conscription

Figure 6 shows a positive feedback loop where a reduction in the availability of foodstuff causes a shortage of foodstuff for indigenous populations. This causes a greater pressure for individuals to join an EG in order to obtain greater security. Moreover, the conflict itself will make it more difficult for farmers to produce and sell foodstuff, thus, adding an incentive for them to join the EG. This produces a greater ability of the EG to recruit and conscript individuals into the EG. This produces a larger number of EG soldiers, which could cause a greater disruption in farming areas producing foodstuff. This, along with drought, would cause a reduction in food availability. This cycle continues to be reinforced until negative loops counter these cycles.

Causal loop conclusions and behavior indicators

Using a causal loop analysis, we were able to evaluate some of the major interaction effects within this dynamic EG-society system. The main conclusions of this particular assessment is the notion that the fighting between the EG and governmental bodies can create a positive feedback loop whereby the conflict is an indirect cause of recruitment, which increases the ability of the EG to participate in the conflict. While the distribution of foodstuff is often essential to the population, it can also serve to support and strengthen the EG by providing a means for food and revenue. A better understanding these causal loops could disrupt this cycle, thereby reducing the influence of EGs.

The process of creating causal loop assessments from a set of problem questions is typically iterative. Once the assessments have been sufficiently developed, more specific and measurable potential behaviors can be derived. For example, a general behavior that is at an appropriate level of granularity for a broader, systems model can be discretized into more precise behavioral elements. The behavioral elements that are generated should be at a sufficient level of granularity that is appropriate for the type of analysis that will be conducted and should conceivably be measurable. At this stage, multiple subject matter experts (SME) can expand on the behavioral elements or create their own set of behavioral elements. These behavioral elements may pertain to groups/organizations, or even to types of people (e.g., government, social, or business leaders, etc.). The process of generating these behavioral elements begins with a pre-defined conceptual idea regarding the specific domain of interest at a fairly broad, systems level. That is, one may be interested in a particular range of behaviors pertaining to a specific type of individuals as they interact within a group or country. Causal loop assessments are instrumental to help determine that. Constraining the type of behaviors one is interested in modeling makes the problem-set more tractable and helps determine the behavioral cues that are most relevant. For instance, as displayed in Figure 1, the initial conceptual model may incorporate several groups that will generate a larger and more specific set of behavior indicators to be integrated into a larger modeling structure.

EXAMPLE ASSESSMENTS OF EG SUPPORT

Within BIA, modeled humans (individuals or group of individuals) can be exposed to endogenous stimuli (i.e., efferent or “top down” sensations) and exogenous (environmental) stimuli. Stimuli can be generated within the BIA system as interactions between entities, where relevant stimuli serve as cues that can stimulate a particular belief regarding a behavior or condition that is occurring. However, the same stimuli may be interpreted differently among different groups, causing different beliefs to arise. These beliefs may stimulate pre-existing attitudes associated with norms and perceptions of the current environment. It also may stimulate levels of emotional affect (positive and/or negative) associated with the belief. Together, this has the potential for stimuli to activate motivations to perform different types of behaviors by different groups.

Group/population model

To further explore the interactions between EGs and populations within their control, we used BIA to assess the interactions between these two groups. In this assessment we modeled an EG entity that has control over a geographical region and that makes decisions regarding allowable behaviors and trade. Another entity represents the local population, whom the EG seeks to satisfy and control. We can call this the group/population scenario. This modeled scenario, like many models of human behavior, involves substantial feedback and nonlinearity. Inputs to the model are highly uncertain (especially those involving cognitive processes) and highly variable (especially

economic and social factors). An overview of the interactions within the group/population model structure is shown in Figure 7. In this scenario we are focusing on the availability of foodstuff for a population within a region under EG control. This scenario was chosen because of previous and current international crises associated with food shortages. It was also chosen because the ability to administer food is a good litmus test for the ability of the EG to generate support within the indigenous population. In this model, food demand is based on population and the availability of foodstuff. If the availability of food declines too quickly population satisfaction will decline, causing support of the EG to decline and resistance to increase. The modeled EG attempts to avoid this situation by implementing actions that could increase the availability of food for populations under their control. The EG will seek to use revenues (imposed taxes, extortion, theft, etc.) to pay for food appropriations. If these revenues are insufficient to cover the appropriations, the EG must find additional revenues to pay for it. When the EG creates additional revenues, the cost of foodstuff increases, which decreases population satisfaction and thus further reduces population support for the EG and increases resistance.

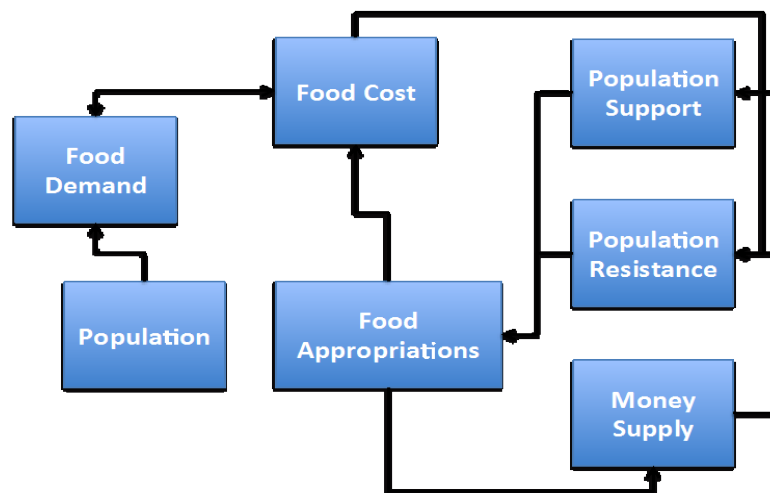


Figure 7. Overview of the EG-food model structure for the group/population model

The model works as described above, but detail is included to specify how decisions are made and how non-cognitive variables are calculated. The model simulates behaviors based on utility functions and qualitative choice theory (Ben-Akiva & Lerman 1985; Train 1986). The population in this model has three major decisions to make based on cues in the model. The population's demand for foodstuff is based on the cost (availability/price) of the foodstuff. Population resistance is determined by the cost of the foodstuff and a general cost index of goods in the society. Population support is also based on food and general cost indices, but also takes resistance activity into account. The EG has just one decision to make in this scenario: where they would like to set food cost, using the food appropriation. This decision is based on population support and population resistance, and is aimed at keeping population satisfaction with the EG high enough not to poses great resistance.

There are 12 cue inputs to the group/population model that were considered uncertain (see Table 1). The expected population resistance and support indicate levels that the EG considers desirable. Revenue is defined by a log-normal distribution. The cost adjustment describes the fraction of the indicated change in price that will actually occur. The remaining uncertain inputs are coefficients on utility functions. These inputs indicate the magnitude of the effect that a particular societal event or trend will have on a decision.

Table 1. Uncertain inputs to the population/group model

Variables	Distribution
Expected Population Resistance (EPR)	Uniform
Expected Population Support (EVS)	Uniform
Revenue	Log-normal
Price Adjustment	Uniform
Food Demand β (FD β) - How much food costs affects demand	Uniform
EG Food Allocations β (GFA β) - How much population support affects GFA	Uniform
EG Food Appropriations γ (GFA γ) - How much population resistance affects GFA	Uniform
Population Resistance β (PR β) - How much general costs affects resistance	Uniform
Population Resistance γ (PR γ) - How much food costs affects resistance	Uniform
Population Support β (PS β) - How much food costs affects support	Uniform
Population Support δ (PS δ) - How much resistance affects support	Uniform
Population Support γ (PS γ) - How much general costs affects support	Uniform

These 12 cue inputs were used to assess the degree of population support for the EG over time. Figure 8 shows a 50-run Monte Carlo simulation result of population support within the group/population model. Each line in figure 8 represents a full time series for the output population support for a simulation with random values for each of the 12 uncertain inputs in table 1. While each simulation exhibits a different pattern over time, there is a somewhat robust pattern that is shared between the simulations. At the beginning of each simulation, the EG seeks to gain support from the population by appropriating foodstuff for the allied population under their control. This causes population support to increase. However, the EG has to exact services to compensate for the appropriation, which after a time lag causes an increase in the 'cost' of the foodstuff. Thus, after an initial rise, population support for the EG can greatly decline below its initial level. What is important to note is that most situations cause the EG's support to collapse. There are, however, a small number of outlier situations where the collapse does not occur. A causal reach-back of the analysis indicates under what (few) circumstances this outcome does not occur. Secondary interventions can prevent these circumstances from occurring, or conversely, the verification of such existing circumstances would indicate that an alternative intervention strategy is needed. This process can greatly limit blind-siding and ineffective interventions. Assessing the situations where the EG's support does not collapse within the population enables us to focus on these specific variables. For example, in the group/population model, extremist group allocations of support, such as foodstuff (GFA β) is the most robust predictor for population support for EG across time. The actual cost (in money, labor, etc.) to the population for the foodstuff (PA) is the second to most robust predictor of support. The third to highest predictor of support over time is the general expected population support (EPS). This is basically the general affinity of the population towards the ideals of the EG (a positive shift towards the EG). The other variables in the model do not significantly contribute to the support of EG over time.

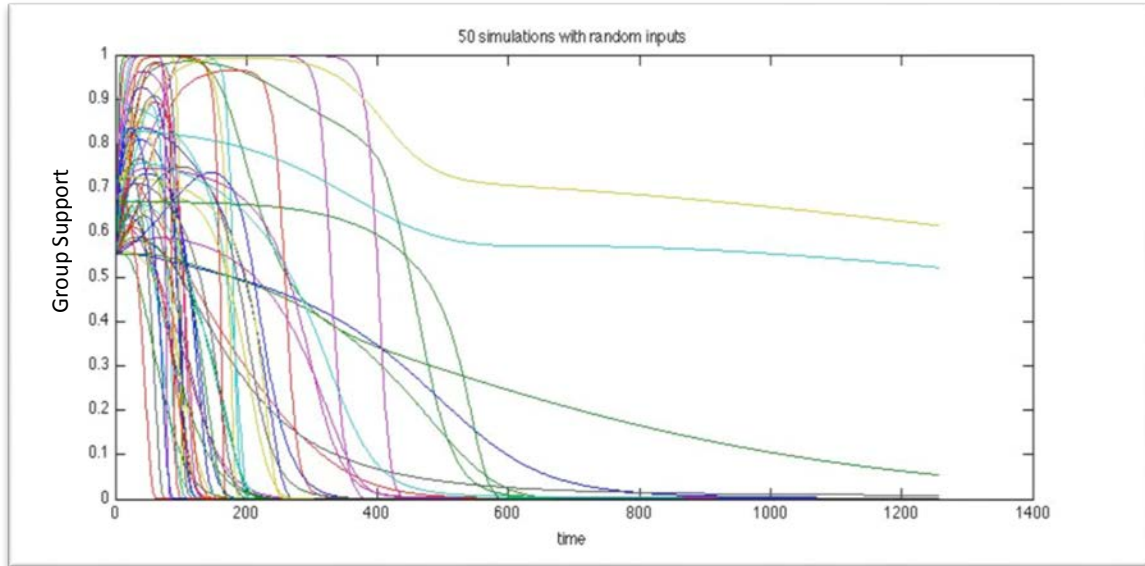
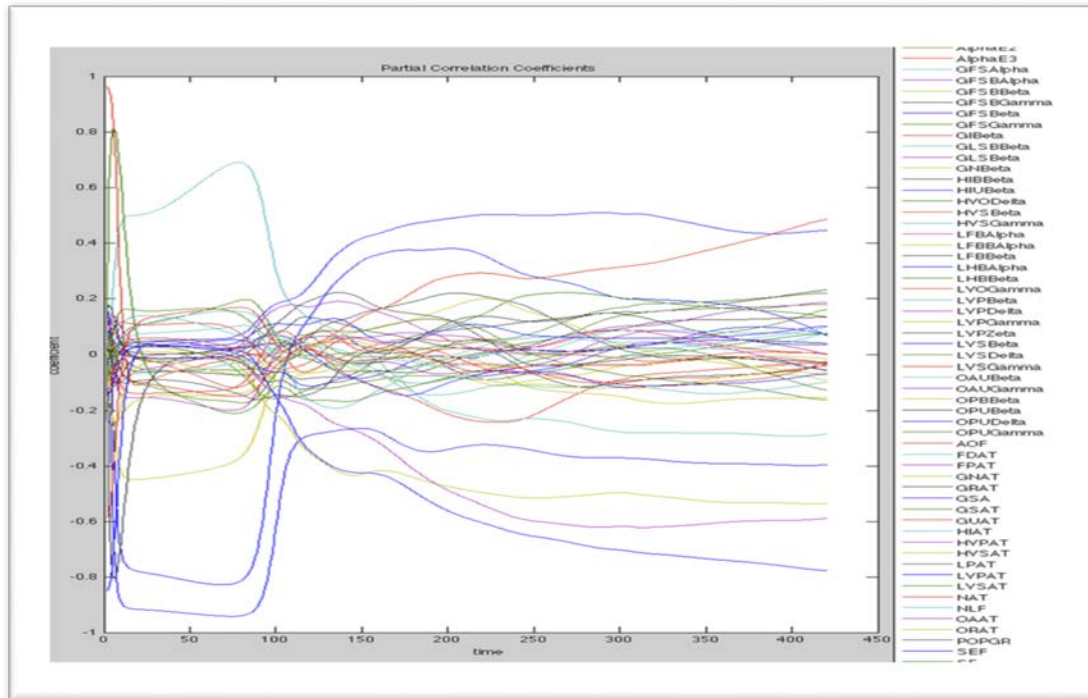


Figure 8. Population support over time (Using Monte Carlo simulation (N=50))

Partial correlation coefficients measure the strength of the linear relationship between each uncertain input and the output of interest, correcting for the linear effects of other inputs. These coefficients can vary from -1 to 1, with a value farther from 0 indicating a stronger relationship. Partial correlation coefficients over time for all of the uncertain inputs (in relation to support from populations that would typically be sympathetic to the EG) are shown in figure 9.

By plotting partial correlation coefficients over time, we can see that the relative strengths of correlations for different inputs change over time. That is, some inputs (e.g., courses of actions) may be very effective in producing certain outputs (e.g., behavior shifts), while other inputs may be ineffective. Very early in the simulation, most of the inputs are highly correlated with the output of interest. After just a few time steps, EG food allocations (GFA β) emerge as very highly correlated input. Toward the end of the simulation, GFA β reduces in influence, and other factors emerge as highly correlated with the output (behavior) of interest. This type of assessment can help determine the effectiveness of specific policies, programs, and actions over time. That is, a policy, program, or action may start off positively, but ultimately lead to negative outcomes. When this is the case, they can be eliminated or modified so that they produce positive outcomes over time.



Group/population model conclusions

The group/population model is meant to provide an example of the type of assessments being developed to evaluate the social, political, and economic conditions that both support and reduce the influence of extremist groups over time. This focus on how the availability of foodstuff can potentially affect their support is our initial effort to produce a full BIA model of the dynamic interactions between EG, governmental bodies, international governmental bodies, and other groups. The results of this assessment suggest certain scenarios that can affect a population's support of an EG over time. As discussed above, EG support for the population through distributing foodstuff can have a significant and lasting positive effect on an indigenous population's support of an EG.

MODEL CONFIDENCE

With this statistical knowledge, we can provide confidence intervals on the results of the model analyses that test interventions. By simultaneously performing uncertainty quantification for model parameters and potential interventions, BIA can determine the portfolio of interventions that have the highest (quantified) probability of success despite uncertainty. It can also quantify the risk associated with the intervention not performing as anticipated. Additionally, BIA can perform sensitivity analyses to determine what minimal additional information is needed to maximally reduce uncertainty and further assure the proposed interventions produce the desired outcome throughout the time horizon of interest. Because the model is causal, decision-makers can reach-back into detailed results of the simulation to independently evaluate the nuanced processes that caused the predicted outcomes. Moreover, the same process can determine early warning fingerprints whose measurement today or during the initial implementation of an intervention can verify or exclude the possibility of critical conditions/outcomes.

Our confidence management process involves on-going, collaborative assessment of BIA model outputs to ensure that the final product will provide useful information for the desired application. This process includes elements that take place during the modeling process, as well as some that are implemented using a completed model. This confidence management process was designed to inform the project team and end users about the level of confidence they should have in the model, as well as identifying potential improvements to the model and process that could strengthen this confidence. Confidence management consists of a suite of techniques in the categories of documentation, verification, validation, uncertainty quantification, and sensitivity analysis.

GENERAL CONCLUSIONS

This paper describes an effort to better understand how EG within failing states in Africa understand their reality; why populations choose to support or resist EG; how EG organize themselves socially and politically; and why and how their beliefs shift over time. Our initial models and assessments focused on how the distribution of foodstuff and interactions with governmental bodies, populations, and conditions (e.g., drought) dynamically affect the support of EG over time. First we performed causal loop assessments to determine the interaction effects of foodstuff allocations, governmental bodies, and EG. These assessments are necessary in framing the BIA models for more specific group and country behavioral assessments in response to specific actions or conditions. Then we developed an initial BIA model that addressed EG support across different scenarios. The assessments suggest conditions that are associated with a potential collapse in support for an EG over time. Other conditions may have little to no effect on EG support over time. Further development of BIA will include more detailed modeling and assessments of inter- and intra- governmental body, population, and EG dynamics.

REFERENCES

- Ben-Akiva, M. & Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press: Cambridge, MA.
- Backus, G., Bernard, M., Verzi, S., Asmeret, B., Glickman, M. (2010). *Foundations to the unified psycho-cognitive engine*. Sandia National Laboratories technical report SAND Report 2010.
- Bernard, M. L. Backus, G. (2009). Model the interaction between leaders and society during conflict situations. *Annual Meeting of the System Dynamics Conference*, Albuquerque, NM.
- Bernard, M. L., Backus, G. A., & Bier, A. B. (2014). Behavioral Influence Assessment (BIA): A multi-scale system to assess dynamic behaviors within groups and societies across time. In *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*.
- Besancon, M., & Dalzell, S. (2013). The soldier and the street: East African CIMIC in Somalia and beyond. *Inter-University Seminar on Armed Forces and Society*, Chicago, IL
- Bier, A. B. (2010). *Sensitivity Analysis Techniques for models of human behavior*. Sandia National Laboratories technical report SAND Report 2010.
- Bier, A. B., Bernard, M. L., Backus, G., Glickman, M., & Verzi S. (2011). Using a hybrid cognitive-system dynamics model to anticipate the influence of events and actions on human behaviors. *29th International Conference of the System Dynamics Society*, Washington, D.C.
- Bier, A. B. & Bernard, M. L. (2014). Validating a hybrid cognitive-system dynamics model of team interaction. In *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*.
- Finlay, W. M. L. (2010). The propaganda of extreme hostility: Denunciation and the regulation of the group. *British Journal of Social Psychology*, 46, 323-341.
- Hansen, S. J. (2013). *Al-Shabaab in Somalia: The history and ideology of a militant Islamist group, 2005-2012*. Oxford Press: New York.
- Harper, M. (2012). *Getting Somalia wrong? Faith, war and hope in a shattered state*. Zed Books: London.
- Martins, A. C. R. (2008). Mobility and social network effects on extremist opinions. *Physics and Society*, 1-7.
- McFadden, DL. (1984). *Econometric analysis of qualitative response models*. Handbook of Econometrics, Volume II. Chapter 24. Elsevier Science Publishers BV.

- Milton-Edwards, B. (2008). The ascendance of political Islam: Hamas and consolidation in the Gaze Strip. *Third World Quarterly*, 8, 1585-1599.
- Pillar, P. R. (2001). *Terrorism goes global: Extremist groups extend their reach*. *Brookings Review*, 19, 34-37.
- Sterman, JD. (1994). *Learning in and about complex systems*. *System Dynamics Review* 10(2-3): 291-330.
- Thaler, D. E., Brown, R. A., Gonzalez, G. C., Mobley, B. W., & Rohan, P. (2013). *Improving the U.S. military's understanding of unstable environments vulnerable to violent extremist groups*. RAND. Arroyo Center.
- Train, K. (1986). *Qualitative Choice Analysis*. MIT Press: Cambridge, Massachusetts

ACKNOWLEDGEMENTS

This research was possible in part by funding from the United States Department of Defense. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly-owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. Approved for unlimited release: SAND: 2014-1648C

Section 2

*Studies in Cross Cultural Decision
Making*

Are Deliberative People More Consistent in Decision Making?

Róbert Hanák

*Slovak Academy of Sciences
Department of Experimental Psychology
Dúbravská cesta 9, Bratislava, 841 03, Slovakia*

ABSTRACT

The preference for intuition and deliberation scale (PID) as a cognitive style measure was used to investigate whether more deliberative participants (identified by self-report PID inventory) would also show higher motivation to properly and normatively solve a task designed to measure their inconsistency and discrimination to details (CWS Index). 161 (103 women) managers and administrative workers were asked to evaluate 21 fictional job candidates. The decision task was designed so that participants could work according to their preferences – everyone had enough time to analyse the logic behind the task. Significant differences were found among all four groups (deliberative, intuitive, both below median, both above median) in levels of inconsistency. Totally consistent respondents were significantly more likely to be from the deliberative and mixed (high in deliberation and in intuition) groups.

Keywords: thinking dispositions, cognitive style, Preference for Intuition and Deliberation, PID, CWS index, motivation, consistency

ARE THINKING DISPOSITIONS MANIFEST IN THE DECISION OUTCOME?

Stanovich (2011) integrated various aspects of human thinking into a tripartite model of human mind. In this model, human thinking is divided into two types of processing. The Type 1 (Autonomous Mind) is evolutionary older and its activation is mandatory after encountering a problem. It consists of several subsystems (TASS, The Autonomous Set of Systems), which work quickly, in parallel, without intensive cognitive effort and are often associated with heuristics (Stanovich 2008, Kahneman, 2011). Relying on heuristics could be ecologically very useful in saving effort, time and cognitive capacity (Gigerenzer, 2007). On the other hand these heuristics in many (well documented) cases, tasks or problems produce non-optimal or irrational responses. There are many documented biases in the JDM literature caused by this way of thinking (Kahneman, 2011), also often referred to as miserly processing (Stanovich, 2011)

Stanovich (2008, 2011) not only sees the Autonomous mind as a set of systems, but he also differentiates Type 2 processes, responsible for our rational thinking, into two broad categories – the Algorithmic mind and the Reflective mind. The Algorithmic Mind reflects the cognitive abilities of a person and could be measured by intelligence tests. Higher cognitive abilities (intelligence) explain a large proportion of the successful solving of various logical, mathematical or “real” life tasks, problems and complex human behavior, but they fail to explain all (Pacini, Epstein, 1999; Bruine de Bruin, Parker, Fischhoff, 2007; Kokis, Macpherson, Toplak, West, Stanovich, 2002). On the other hand, the Reflective mind consists of individual differences in thinking disposition. Often labelled as

cognitive styles, some measure the functioning of the Reflective mind. In contrast with the Algorithmic mind reflecting computation power and performance, cognitive styles concern “beliefs, belief structure, attitudes toward forming and changing beliefs, goals and goals hierarchy (Stanovich, 2011, 35). Therefore, individual differences in rational (defined as optimal, normative) thinking are caused by both: individual differences in cognitive ability (intelligence) and individual differences in thinking disposition (cognitive styles). They are, in general, independent of each other and correlate only weakly (Ku, Ho, 2010, Bruine de Bruin, Parker, Fischhoff, 2007; Kokis, Macpherson, Toplak, West, Stanovich, 2002). Comparing their relative attribution to rational (optimal, normative) output of decision processes, thinking dispositions are probably more important than cognitive ability, which is only engaged when need. As Perkins and Ritchhart (2004, p. 352) explain: “How good a thinker are you? It must be answered as much in terms of people's attitudes, motivations, commitments and habits of mind as in terms of their cognitive ability”.

Among many cognitive styles, Stanovich (2011) is concerned mostly with his own Actively Open-Minded Thinking (Toplak, West, Stanovich, 2011) and Master Rationality Motive (2008), but he also uses others: Need for cognition, Need for closure, Superstitious thinking, and Dogmatism. They are self-report inventories measuring a person's willingness to think deeply, consider various aspects of problems and arguments and general motivation to engage in thinking. The Open-Minded Thinking Scale and Need for Cognition Scale have been proved as predictors for various tasks (Stanovich, West, 1998, Sirota, 2008, Toplak et. al., 2013). All of these scales are constructed on the premise that only effortful thinking – deliberation – is crucial for rational output. Intuition is viewed as the opposite of rationality; it is connected with irrational thinking, and is seen as a cause for failure in rational thinking, so it has no place in these inventories. However, many other researchers examine rationality together with intuition, not as opposing poles, but as two (Hammond, 1996, Epstein, 2003) or more (Dewberry, Juanchich, Narendran, 2013) separate dimensions where intuition is not viewed only in negative way (Glöckner, Witteman, 2010; Čavojová, 2013a, 2013b).

This view is reflected in other cognitive style inventories, to name just a few: Rational Experiential Inventory (REI, Epstein, Pacini, Denes-Raj, Heier, 1996, Pacini, Epstein, 1999), Preference for Intuition and Deliberation (PID, Betsch, 2004); General Decision Making Style Scale (GDMS, Scott, Bruce, 1995). The REI showed the highest internal reliability in comparison with other scales (for review see Koele, Dietvorst, 2010; Hanák, 2013a, 2013b). PID also works well across cultural backgrounds (Witteman, 2009, Richetin, Perugini, Adjali, Hurling 2007). Both GDMS and PID also have appropriate levels of Cronbach's alpha. REI and PID were tested in Slovak population (Hanák, Čavojová, Ballová Mikušková, 2012) and results showed their good psychometric properties (Ballová Mikušková, Hanák, Čavojová, 2014). Each of these inventories measure both rationality/deliberation and also intuition/experientiality dimensions.

The predictive validity of these and other cognitive style measures for various logical, mathematical, probabilistic reasoning and other tasks were investigated and found mostly positive results (Amstrong, Cools, Sadler- Smith, 2012, Cools, Amstrong, a Sadler- Smith, 2010, Hanák, 2013). However, there are only a few studies (e.g. Newstead, 2013) examining whether self-reported statements from cognitive style inventories are manifest connected with motivational measures. Thinking dispositions (cognitive styles) could be “translated” into specific traits closely connected with critical thinking. As such, they should be manifest in real processes and steps when solving tasks and problems. Perkins et al. (1993) introduced seven traits based on several sources of literature. They are:

1. The disposition to be broad and adventurous.
2. The disposition toward wondering, problem finding and investigating.
3. The disposition to build explanations and understanding.
4. The disposition to make plans and be strategic.
5. The disposition to be intellectually careful.
6. The disposition to seek and evaluate reasons.
7. The disposition to be metacognitive.

If we carefully and diligently apply these seven traits in combination with cognitive ability in decision process, our final decision is bound to be rational (optimal, normative). They also present motivation to solve problem. The PID cognitive style measure has nine questions in deliberative scale which fulfill these seven traits well.

PID deliberation scale (translation into English according to Betsch, 2008)

1. Before making a decision I think them through.

2. Before making decision I usually think about goals I want to achieve.
3. I consider myself.
4. I prefer making detailed plans rather than leaving things to chance.
5. I am a perfectionist.
6. I think about a decision particularly carefully if I have to justify it.
7. When I have a problem I first analyze the facts and details before I decide.
8. I think before I act.
9. I think more about my plans and goals than other people do.

Analysis of Perkins's seven traits reveals that most of them embody an intrinsic motivation to do the "job" properly and come to the best possible (normative) solution of all those available. If we offer participants a task where they have two ways to solve it, one much easier (but normatively incorrect) and one more difficult (but normatively correct), we expect that those who view themselves as deliberative should choose to resist miserly processing (cognitive laziness) and therefore, be more normatively correct.

Therefore, our aim was to examine whether more deliberative participants would act accordingly in a task designed to measure motivation for figuring out hidden principle and thus manifest high levels of consistence.

STUDY: ASSESSING MANAGERIAL STUDENTS' MOTIVATION FOR OPTIMAL OUTPUT THROUGH COGNITIVE STYLES

Participants

161 Slovak managers and administrative workers (mean age 31.5 years, SD = 9.4) participated in our study, 103 were female. All were external students of the School of Economics and Management of Public Administration in Bratislava and were in the second year of their bachelor degree. They participated for extra credit, without financial incentive. Data were collected in 2012 from respondents living in Bratislava's metropolitan area (Slovak capital). No exclusion criteria or any restriction, except age over 18, for participants were implemented. During lectures participants obtained an MS Excel file that they were asked to complete according to the presented instructions and send back via email to the teacher within one month.

Measures

Preference for Intuition and Deliberation (PID, Betsch, 2004) is a relatively new inventory designed for measuring affective intuition, defined as being based on implicit knowledge and also as a basic mode of decision making (Betsch, Kunz, 2008). This mode uses mostly affect as a decision criterion (high preference for intuition relates to Type 1 processes and reliance on instant impressions and emotions). Deliberation is defined as decision making using explicit evaluation, rules of deciding, beliefs and reasons (high preference for deliberation relates to trusting Type 2 processes). PID consists of 18 items; 9 for PID-Intuition (e.g. "I listen carefully to my deepest feelings.") and 9 for PID-Deliberation (e.g. "Before making decisions I usually think about the goals I want to achieve."). Participants indicate their agreement with these statements on a 5-point Likert scale with 1 meaning "totally disagree" and 5 "totally agree". According to recommended method (Betsch, Iannello 2010) of median split, we identified four types: pure intuitive (scoring above MDN = 30.0 in preference for intuition and below MDN = 35 in preference for deliberation), pure deliberative (vice versa), mixed (scoring above MDN in both preferences) and indifferent (scoring below MDN in both preferences). The mean for deliberation $M = 33.56$ showed a higher preference for rationality than for intuition $M = 30.17$. In this study, the internal consistency by Cronbach's alpha for PID-D was measured as .795 and for PID-I as .744, which is an acceptable level and in accordance with Slovak population ($n = 750$) where was 0.827 for PID.D and 0.738 for PID-I (Ballová Mikušková et al., 2014). Levels of Cronbach's alpha in this study are comparable to other cross- cultural backgrounds, where Betsch (2004) found .77 for PID intuition and .79 for PID deliberation for 2132 Dutch participants. Richetin, Perugini, Adjali, Hurling (2007) found in a sample of 299 British students lower levels .62 for PID intuition and .77 PID deliberation. The best results for internal consistency was found by Witteman et al. (2009) who found .87 for PID Intuition and .85 for PID

deliberation in a group of 405 Dutch students.

CWS index task (modified)

Motivation for best achievement and willingness to think thoroughly about the principles and logic behind a task was measured by a separate task based on the logic of CWS index (Shanteau, 2001, 2003). This index measures the consistency and discrimination of people by presenting them with many stimuli, some of which are identical, and measuring the ability of participants to detect differences and consistencies among stimuli. In original use of CWS index people are not allowed to know principle behind the index and do not have a time to figure it out (Weiss, Shanteau, 2003). In this work we used it exactly in opposite way and gave people a lot of time to figure it out.

The basic task was to evaluate 21 job applicants for a specific job post. All job applicants were described in one or two sentences by four cues (cognitive tests, years of experience, peer review and assessment centre). For each cue an applicant could have a positive or negative attribute. For example, for cognitive test the negative option was as follows *"The applicant achieved and average score in cognitive tests and in some tests even a below average score."* and the positive was *"The applicant achieved a highly above average score in cognitive tests."* Positive values were coded as 1 and negative values as 2. The 21 fictional job applicants were mathematically generated as all possible combinations of 1 and 2 in four cues, which made 16 combinations. Five job applicants were included twice, but participants were not informed about this fact and the applicants were labelled with different names. These 5 applicants were used for the consistency measure. Respondents were only asked to evaluate job applicants without any other instruction. For each job applicant they could assign 0 to 100 points and they could assign the same points to different candidates if they wanted. Discrimination was measured as the variance among the scores of all 21 job applicants. Applicants were different, therefore the higher variance among them, the better the participant's ability to discriminate the details and demonstrate them in their decision making. Inconsistency was measured by the variance between two applicants who with identical attributes but labelled by different names. Here variance should be minimal or zero. The procedure was as follows.

Participants had one month to fill in the task as it was part of the home assignment given during a management course. The task was designed to (indirectly) measure motivation for achieving the optimal solution. Participants received an email with an MS Excel file with 21 tables containing all 21 job applicants. Participants could spend time learning the rules and principles of the task and thus assess the job applicants in a mathematically optimal way and reach a high or absolute level of consistency. However, they needed internal motivation and desire to do this, because they were only asked to evaluate job applicants and not to figure out the principle behind the task. Solving the problem was not very difficult, but this process could be time-consuming and participants needed to work systematically, seek truth, calculate, make notes, be thorough, have the final aim in mind and compare the attributes of fictional applicants. We expected that these characteristics would be reflected in a higher score of deliberation in PID. Looking at Perkins et. al. (1993) dispositional traits in our introduction, respondents surely needed to apply almost all of them, especially disposition numbers: 1, 2,3,4,5 and partially 6 and 7, to come to the best solution (especially when not asked to do so). It would be very unlikely to reach a high score in this task purely by chance.

The opposite behaviour could be considered as indicative of miserly processing. Miserly processing is assumed to lead to mistakes because people are not expected to be able to hold in their working memory all the required combinations and facts for completing this task. PID-Intuition measures affective intuitive behaviour and we expected that those who scored high in PID-I would make their decision without deeper analysis, therefore demonstrating high levels of inconsistency.

This task was not very demanding in terms of cognitive ability (algorithmic mind) especially for college students. Rather, it reflected the principal characteristic of the Reflective mind, i.e. motivation for rational integration and overriding people's propensity for miserly processing. Therefore, we assumed that differences in consistency scores would be caused by differences in motivation for deliberative thinking or in other words abstaining from Autonomous mind rather than differences in cognitive ability.

Results

Mean discrimination to details was 428.7 (SD = 277.24) and mean inconsistency was 124.71 (SD = 210.97). Discrimination (Table 1) and inconsistency (Table 2) scores were compared between the four groups according to their PID score (*intuitive*, *deliberative*, *mixed* – both above median and *indifferent* – both below median).

Table 1. Descriptive statistics for deliberation across thinking disposition groups

Group	Mean	Standard Deviation
Intuitive (intuition above MDN)	485.77	365.23
Deliberative (deliberation above MDN)	433.79	225
Mixed (both above MDN)	402.37	271.79
Indifferent (both below MDN)	396.52	249.16

There were no significant differences in ability to discriminate details between all groups: $F(3,157) = .844$; $p = .472$. Comparing only the intuitive group to the deliberative or mixed groups also showed no significant differences. Also, the ability to discriminate details does not correlate with PID deliberation scale (motivation to solve problems), $r = -.075$, $p = .326$.

In inconsistency score, we expected that this scale to measure motivation to solve problems, therefore when solving a task where motivation plays a key role, there should be differences. Inconsistency scores ranged widely from 0 to 1370. The results for inconsistency across the four thinking dispositions groups in are in Table 2. The data were not normally distributed, Kolmogorov-Smirnov test showed not normally distributed data for each group, and therefore nonparametric tests were used. A Kruskal – Wallis test showed significant differences among all four groups ($H(3) = 9.05$, $p = 0.029$). Those in the intuitive group demonstrated a three times higher level of inconsistency compared to the deliberative group and where two times higher level compared to mixed group (both above). These three groups were investigated. Therefore, as a follow-up Mann – Withney tests were applied with Bonferroni correction, so all results are reported with 0.0167 significance. Difference between group deliberative and intuitive was not significant after Bonferroni correction ($U = 411$, $n = 69$, $p = 0.032$), but with medium effect size $r = -0.258$. Group with mixed score (above median in deliberation and also in intuition) compared to intuitive group differed significantly ($U = 620$, $n = 85$, $p = 0.016$, $r = -0.261$).

Table 2. Descriptive statistics for inconsistency (raw score) across thinking disposition groups

Group	Mean	Standard Deviation
Intuitive (intuition above MDN)	184.04	244.11
Deliberative (deliberation above MDN)	61.46	61.71
Mixed (both above MDN)	88.87	169.06
Indifferent (both below MDN)	137.78	257.33

Achieving total consistency required participants to find all five duplicate fictional applicants and to assign the same points to each of the pair. Doing this randomly was highly improbable. In fact, it suggests that these participants discovered the hidden principle behind the task and acted accordingly. There were 15 (9.3%) respondents who did this. We were interested in which group, according to PID (intuitive, deliberative, mixed, and indifferent) they belonged to. We found that only two of them were in the deliberative group, two were from the intuitive group, and one from the indifferent group (low in deliberation and intuition), and eight were from the mixed group (high intuitive, high deliberative). The remaining two had median levels of PID score, and were

therefore excluded from analysis. Differences between groups on achieving total consistency is very close to significance $\chi^2 = 7.66$, $df = 3$; $p = .054$ where Phi was .218 at $p = .054$.

Achieving total consistency was the optimal and best possible output, but we decided to also investigate those who were very close to it. These respondents also tried hard and made very small mistakes. We defined the most motivated and least prone to miserly processing as those who were under 10th percentile in inconsistency (2.5 points and less). There were 20 respondents, 3 were excluded for PID due to their mean score being equal to that which was used for dividing participants into groups. We found that four of them were in the deliberative group, two were from the intuitive group, and one from the indifferent group (low in deliberation and intuition), and ten were from the mixed group (high intuitive, high deliberative). We found significant differences between PID groups in the number of respondents achieving high consistency $\chi^2 = 10.34$, $df = 3$; $p = .016$ where Phi was .253 at $p = .016$.

Table 3. Number of respondents of total and 10th percentile group of consistency across thinking disposition groups

Group	Number of the respondents in total consistency	Number of the respondents in 10th percentile
Intuitive (intuition above MDN)	2	2
Deliberative (deliberation above MDN)	2	4
Mixed (both above MDN)	8	10
Indifferent (both below MDN)	1	1
Excluded (median levels of PID score)	2	3
Total	15	20

For further regression analyses we used a logarithmic transformation of the inconsistency score. Regression analysis showed that the PID intuitive score was a non-significant predictor for inconsistency. Only one (*I am a very intuitive person*) from nine questions correlated significantly ($r = -.154$, $p = .021$) with logarithmic transformation of the inconsistency score. It could be used as a predictor, close to significance where the standardized $\beta = -.147$, $t(174) = -1.97$, $p = .051$, but explaining in fact almost no variance $R^2 = .016$, $F(1, 174) = 3.86$, $p = .051$.

However, the PID cumulative deliberative score was found to predict inconsistency with a small effect ($\beta = -.235$, $t(174) = -3.17$, $p = .002$), explaining only very small part of variance $R^2 = .05$, $F(1, 174) = 10.08$, $p = .002$. Investigating individual items from the PID deliberative scale as predictors, only two items were significant: *I am a perfectionist* (where the standardized $\beta = -.245$, $t(174) = -2.92$, $p = .004$), and *I think before I act* (where the standardized $\beta = -.192$, $t(174) = -2.03$, $p = .044$). From 9 self-reported statements in PID deliberative scale (Table 4), two could be used as predictors (number: 5 and 8), another 3 correlate significantly with resistance to miserly processing (number: 1, 3, 7), and two were close to significance (2, 9). Only question number 4 was highly insignificant.

Table 4. PID deliberative scale questions and their relationship to inconsistency score

Item n number	PID deliberative scale items	Correlation coefficient with logarithm inconsistency	Sigma (2 tailed)
1	Before making a decision I think them through	-.123	.052
2	Before making decision I usually think about goals I want to achieve	-.105	.084
3	I consider myself	-.145	.028
4	I prefer making detailed plans rather than leaving things to chance	-.041	.296
5	I am a perfectionist	-.23	.001

6	I think about a decision particularly carefully if I have to justify it.	-.097	.101
7	When I have a problem I first analyze the facts and details before I decide	-.214	.002
8	I think before I act	-.244	.001
9	I think more about my plans and goals than other people do.	-.107	.079

CONCLUSIONS

The PID deliberation scale, as a cognitive style measure, confirmed Betsch's (2004, 2008) claims that it could be used also as a tool for measuring motivation to solve tasks or problems. Although PID deliberation does not predict correct solving of many different logical and cognitive tasks (Betsch, Renkewitz, Betsch, Ulshofer, 2010; Čavojová, Ballová Mikušková, Hanák, 2013) as expected (Betsch, 2004), it seems at least to measure the motivation to do so. Therefore, thinking disposition as measured by the PID deliberation scale could be used as a predictor for abstaining from miserly processing (cognitive laziness), thus measuring some kind of motivation to do “the job” properly and engage in thinking. Specific questions differ in their predictive validity where *I am perfectionist* and *I think before I act* perform the best. Three other scale items significantly correlated with lower levels of inconsistency but without predictive validity. The PID intuition scale was also investigated, but it has no predictive validity as a scale and only one item from nine, *I am a very intuitive person*, used as predictor was significant. It is important to note that this question is correlated and predicts same direction as all items from deliberative scale. The respondents who self-reported themselves as *very intuitive person* showed ability to behave in very rational and motivated way.

In the introduction we pointed to research tradition in cognitive styles, where intuition is not view as opposite scale of rationality or deliberation. Therefore Intuition is not view only in negative way, its attribution could be under specific conditions beneficial (Ballová Mikušková, 2013). This is in line with our own findings where we conclude that the PID intuition scale is not the opposite of the PID deliberation scale in measuring motivation to solve problems. We found the mixed group (scoring above the mean in both intuition and deliberation) dominated those achieving total consistency, being four times more likely than the other groups to figure out the principle behind the task (that the fictional applicants were identical) and fully apply it to their decision making. Little lowering criteria to get more but still highly performing respondents in terms of figuring out principle behind task we found even stronger support for our claims. The results were almost the same in the group of high achievers (up to 10th percentile) in level of consistency.

Discrimination was not expected to be in relationship with PID scales and our results support our expectations. Neither PID deliberation nor PID intuition related to the ability to seek for details (discrimination) and apply these differences to their decision making (discrimination).

ACKNOWLEDGMENT

This work was supported by Centre of Excellence at Slovak Academy of Sciences: Centre for interdisciplinary research and strategies in society – CESTA, Centre of strategic analyses.

REFERENCES

- Armstrong, S., Cools, E., Sadler-Smith, E. (2012). "Role of Cognitive Styles in Business and Management: Reviewing 40 Years of Research." *INTERNATIONAL JOURNAL OF MANAGEMENT REVIEWS*, 14, 238 - 262
- Ballová Mikušková, E., (2013). "O racionalite intuície". In. Hanák, R., Ballová Mikušková, E., Čavojová, V. (Eds.) *ROZHODOVANIE A USUDZOVANIE IV*. Bratislava: Ústav experimentálnej psychológie SAV, 34-47
- Ballová Mikušková, Hanák, Čavojová, (2014) Intuition versus Deliberation: Psychometric properties of two self-reported scales (under review)
- Betsch, C. (2004). „*Präferenz für Intuition und Deliberation (PID): Inventar zur Erfassung von affekt- und kognitionsbasiertem Entscheiden*“. *ZEITSCHRIFT FÜR DIFFERENTIELLE UND DIAGNOSTISCHE PSYCHOLOGIE*, Volume 25, No. 4, 179–197.
- Betsch, C. (2008). "Chronic preferences for intuition and deliberation in decision making". In H. Plessner, C. Betsch, T. Betsch (Eds.), *INTUITION IN JUDGMENT AND DECISION MAKING* (231–248). New York: Erlbaum.
- Betsch, C., Iannello P. (2010). "Measuring individual differences in intuitive and deliberative decision-making styles: a comparison of different measures." In Glöckner, A., Witteman, C. (Eds.), *FOUNDATIONS FOR TRACING INTUITION. CHALLENGES AND METHODS*. Hove: Psychology Press. 252 – 267.
- Betsch, C., Kunz, J. (2008). "Individual strategy preferences and decisional fit". *JOURNAL OF BEHAVIORAL DECISION MAKING*, 21, 532 -555
- Betsch, C., Renkewitz, F., Betsch, T., Ulshofer, C. (2010). "The influence of Vaccine-critical Websites on perceiving vaccination Risks". *JOURNAL OF HEALTH PSYCHOLOGY*, 15, 3, 446
- Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). "Individual differences in adult decision-making competence". *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY*, 92(5), 938.
- Čavojová, V., Ballová Mikušková, E., Hanák, R. (2013) "Preferencia k deliberácii: (ne) istota úspechu v kognitívnych úlohách". Kelemen, J. et al. (eds.) *KOGNITIVNÍ VĚDA A UMĚLÝ ŽIVOT*. Zborník príspevkov. Slezská univerzita v Opavě. Opava.
- Čavojová, V. (2013a) "O racionalite intuície". In. Hanák, R., Ballová Mikušková, E., Čavojová, V. (Eds.) *ROZHODOVANIE A USUDZOVANIE IV*. Bratislava: Ústav experimentálnej psychológie SAV, 11-33
- Čavojová, V. (2013b). "Emócie a intuícia". In D. Kusá (Ed.): *PSYCHO-LOGIKA EMÓCIÍ*. Bratislava: Veda, 87 -104.
- Dewberry, Ch., Juanchich, M., Narendran S. (2013). "The latent structure of decision styles." *PERSONALITY AND INDIVIDUAL DIFFERENCES*, 54, 566-571
- Epstein, S. (2003). "Cognitive-experiential self-theory of personality". In Millon, T., & Lerner, M. J. (Eds), *Comprehensive Handbook of Psychology*, 5: *PERSONALITY AND SOCIAL PSYCHOLOGY* (159-184). Hoboken, NJ: Wiley & Sons.
- Epstein, S., Pacini, R., Denes-Raj, V., Heier, H. (1996), "*Individual differences in Intuitive-Experiential and Analytical-Rational Thinking Styles*". *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY*, Volume 71, No.2, 390–405.
- Gigerenzer, G., (2007) "*Gut feelings*." London, Penguin Books
- Glöckner, A., Witteman, C. (2010). "Foundations for tracing intuition: models, findings, categorizations." In Glöckner, A., Witteman, C. (Eds.), *FOUNDATIONS FOR TRACING INTUITION. CHALLENGES AND METHODS*. Hove: Psychology Press. 1-23.
- Hammond, K.R. (1996). "*Human judgement and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*." New York: Oxford University Press.
- Hanák, R., Čavojová, V., Ballová Mikušková, E. (2012) "Je intuícia pri rozhodovaní v časovom strese výhodnejšia?" *SOCIÁLNE PROCESY A OSOBNOSTI 2012*. Zborník príspevkov. Ústav experimentálnej psychológie. Bratislava.
- Hanák, R. (2013a). "Prediktívna validita dotazníkov merajúcich intuíciu PID, REI a GDMS". In. R. Hanák, E. Ballová Mikušková, V. Čavojová (Eds.), *ROZHODOVANIE A USUDZOVANIE IV.*, Aplikácie a limity intuície. Bratislava: Ústav experimentálnej psychológie SAV. 101 –127
- Hanák, R. (2013b). "Nástroje na meranie individuálnych preferencií k intuitívnemu rozhodovaniu." In. R. Hanák, E. Ballová Mikušková, V. Čavojová (Eds.), *ROZHODOVANIE A USUDZOVANIE IV.*, Aplikácie a limity intuície. Bratislava: Ústav experimentálnej psychológie SAV. 71 –101
- Kahneman, D. (2011). "*Thinking, fast and slow*." London, Penguin Group.
- Koele, P., Dietvorst, G. (2010). The internal validity of self-report measures for intuitive and rational decision making. in Glöckner, A., Witteman, C. (Eds.), *FOUNDATIONS FOR TRACING INTUITION. CHALLENGES AND METHODS*. Hove: Psychology Press., 238-250.
- Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). "Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles." *JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY*, 83(1), 26-52.
- Ku, K. Y., & Ho, I. T. (2010). Dispositional factors predicting Chinese students' critical thinking performance. *PERSONALITY*

- AND INDIVIDUAL DIFFERENCES, 48(1), 54-58.
- Newstead, S. E., Handley, S. J., Harley, C., Wright, H., Farrelly, D. (2004), „*Individual differences in deductive reasoning*“. The quarterly journal of experimental psychology. A HUMAN EXPERIMENTAL PSYCHOLOGY, Volume 57, No.1. 33–60.
- Pacini, R., Epstein, S. (1999), „*The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon*“. JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY, Volume 76, No. 6, 972–987.
- Perkins, D. N., Jay, E., & Tishman, S. (1993). Beyond abilities: A dispositional theory of thinking.
- Perkins, D. N., Ritchhart, R. (2004). “When is good thinking?” In D.Y. Dai & R. J. Sternberg (Eds.), MOTIVATION, EMOTION, AND COGNITION: INTEGRATIVE PERSPECTIVES ON INTELLECTUAL FUNCTIONING AND DEVELOPMENT (pp. 351-384) Mahwah, NJ: Lawrence Erlbaum Associates.
- Scott, S. G., Bruce, R. A. (1995). “Decision-making style: The development and assessment of a new measure”. EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, 55(5), 818-831.
- Shanteau, J et al. (2003) “*How can you tell if someone is an expert? Empirical Assesment of Expertise.*” EMERGING PERSPECTIVES ON DECISION RESEARCH, Cambridge, U.K. : Cambridge University Press.
- Shanteau, J. et al. (2001) “Identifying expertise without a gold standard: Four applications.” *Paper presented at the 11th International Symposium on Aviation Psychology*. Columbus, OH.
- Sirota, M. (2008), “*Individuálne rozdiely v racionálnom usudzovaní. Kognitívne parametre a komputačné stratégie v „bayesovskom“ usudzovaní v úlohách s odlišnými číselnými reprezentáciami*”. Bratislava: Univerzita Komenského.
- Stanovich, K. E. (2011), “*Rationality and the Reflective Mind*”. New York: Oxford University Press.
- Stanovich, K. E., & West, R. F. (1997). “Reasoning independently of prior belief and individual differences in actively open-minded thinking.” JOURNAL OF EDUCATIONAL PSYCHOLOGY, 89, 342–357.
- Stanovich, K. E., (2008) “Higher – order preferences and Master Rationality Motive”. THINKING & REASONING, 14 (1), 111 - 127
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). “The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks”. MEMORY & COGNITION, 39(7), 1275-1289.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). “Assessing miserly information processing: An expansion of the Cognitive Reflection Test”. THINKING & REASONING, (ahead-of-print), 1-22.
- Witteman, C., van den Bercken, J., Claes, L., Godoy, A. (2009). “*Assessing Rational and Intuitive Thinking Styles*”, EUROPEAN JOURNAL OF PSYCHOLOGICAL SSESSMENT, Volume 25 No. 1. 39–47.
- Weiss, D., Shanteau, J. 2003 . “Empirical Assesment of Expertise”. HUMAN FACTORS, Vol 45, No. 1, Spring

Measuring the Effects of Cultures on Website Use

Hsiu Ching Laura Hsieh

Department of Creative Design
National Yunlin University of Science and Technology
123 Section3, University Rd, Douliou, Yunlin, Taiwan 64002

ABSTRACT

Under the context of globalization, it is essential to understand the cultural requirements and preferences of a target culture if a company wants to develop a website which can appropriately map that target culture. Based on the author's previous study, the culturally preferred characteristics between Taiwan and the UK have been found (Hsieh and Hong, 2013). The previous finding of the culturally preferred characteristics between Taiwan and the UK formulates the further questions: "Since different cultural preferences for web design attributes exist, how can the culturally preferred design characteristics be applied to increase usability?" Therefore, the aim of this paper is to measure the effect of cultural preferences on the use of web design. The methodologies are introduced as following. First, two cultures are selected. Taiwan and the UK are selected. Second, experimental webs are constructed. The cultural preferences for websites of Taiwan and the UK would be embedded into websites prototype, which have the same content. Finally, applying SPSS to analyze the data of the usability testing. Based on the results of the analysis and discussion, it reveals that culturally preferred design elements of the target-culture can improve the effectiveness of web communication for the users from the target-culture. This research can help web developers and designers to develop a web interface design that is culturally appropriate.

Keywords: globalization , web design attributes, cultural preferences, usability

INTRODUCTION

In the past, web developments were aligned with cognition theory and computer technology. Marcus and Gould (2000) pointed out that web designers need to do much planning, research, analysis, design, evaluation, documentation, and training to deeply comprehend the requirements of the users, markets, and businesses. Cultural diversity makes it impossible for designers depending on the instinctive knowledge or personal experiences. Under the context of globalization, Hofstede (2005) suggested that web developers could accommodate diverse cultural markets by applying localization as an alternative strategy to developing the global market. Therefore, there is a need to investigate the requirements and preferences of target-culture users. Based on the results of local website auditing in the previous research of Hsieh and Hong (2013), there are significantly different preferences for web interface design between Taiwanese and the UK cultures. Furthermore, it is questioned whether those cultural differences could be applied to improve web usability. Therefore, this research is to evaluate the effects of cultural preferences on website use.

LITERATURE REVIEW

The Meaning of Culture and Cultural Model

Trompenaars and Hampden-Turner (1997) defined culture as the way in which people in a specific area sorted out the problems and reconciled dilemmas. Furthermore, Sheriden (2003) defined culture as the means how people from certain cultural background saw and interpreted specific images and information, in terms of web globalization. Evers (2002) defined culture as formulating the way to view the world, behave themselves, express themselves and think. It was accumulated by historical experiences, value, tradition, and environments. All of the above studies suggested that the world was understood by people based on the context of culture. The way of people's behaviors, comprehension, expression and interaction was accumulated by culture. Hofstede's cultural theory is applied in this research. Referring to Hofstede's (2005) definition of culture, in which personal patterns of thinking, feeling and behaviors were accumulated in the lifetime from the childhood and the ways of feeling, thinking, and mental programming would vary the received social environments. From the investigation of IBM employees in 53 countries from 1978 to 1983, Hofstede defined the patterns of differences and similarities among the responses of employees and proposed five dimensions for the culture theory, introduced as the follows.

Uncertainty Avoidance. This refers to the degree to which people are comfortable with uncertain conditions. People from cultures with low uncertainty avoidance are more comfortable with uncertain situation. Countries (e.g., Japan, China, and Taiwan) with high uncertainty avoidance tend to be expressive, speaking with gestures and showing their emotions. People are inclined to be aggressive, emotional to stay away from ambiguous conditions, and keep an orderly structure in an organisation or society. On the other hand, people from cultures with low uncertainty avoidance (e.g., the USA and UK) tend to be less expressive, less openly anxious and people act without strongly showing their emotions.

Power Distance. This refers to the extent to which less powerful members expect and accept unequal power distribution within a culture. Cultures with high Power Distance (e.g., Malaysia and Mexico) have characterized hierarchies in organization and autocratic leadership. For example, leaders with a certain level of dictatorship, such as teachers and parents, are highly respected. Students are expected to show their respect to teachers and children have to be filial to parents. On the contrary, cultures with low Power Distance (e.g., Austria and New Zealand) tend to have characteristics such as more equal relationships between leaders and subordinates, and subordinates are more able to express their thoughts and participate in managerial decisions. Teachers and students, and parents and children regard themselves as equal relations.

Individualism vs. Collectivism. This refers to the extent to which the individuals incorporate with the group. Collectivist cultures emphasise the benefits of working in a social group. Collectivist cultures value training, skills, and physical condition, and they value harmony more than truth in family relations, using shame to achieve goals and to save face. On the other hand, individualistic cultures value personal time, freedom and truth in their family relations, and pursue the goals by guilt (Marcus and Gould, 2000). Collectivist cultures (e.g., Taiwan and China) tend to prioritize group welfare over the individual's target where the achievement of an individual is not regarded as important as the accomplishment of the group, and believe in group relationship where loyalty is dominant. Individualistic cultures (i.e. USA, Australia) are inclined to lose ties, where everyone is expected to look after themselves and their nuclear family, and are usually inclined to be independent of other people.

Long-Term Orientation vs. Short-Term Orientation. Long-term orientation presented a critical part in Asian countries influenced by Confucianism that people from such countries strongly believed in the necessity of an unequal relationship for the stability of a society. The harmony between family and society was clearly defined in the hierarchical relationship, and virtuous behaviors were identified by hard work and perseverance. On the contrary, cultures with short-term orientation revealed opposite attributes (Marcus and Gould, 2000 Hofstede, 2005).

Masculinity vs. Femininity, referring to gender roles in a culture. Countries with strong masculine cultures tended to concentrating on challenge, social recognition, and the pursuit of welfare. On the contrary, countries with less defined masculine cultures tended to collapsing gender distinction, overlapping gender roles, and emphasizing security, care of others, and the environment (Marcus and Gould, 2000 Hofstede, 2005).

The Impact of Cultures on Web Usability

Barber and Badre (1998) constructed a cultural marker approach as a kind of systematic usability method to examine hundreds of websites and defined cultural markers, such as icons, colors, metaphors, fonts, geography, sounds, motions, flags, language, preferences for text and graphics, directionality of written language, help features, and navigation tools, in order to facilitate the user's performance. Based on Baber and Badres' cultural marker approach(1998), Smith et al. (2004) figured out the culturally preferred design elements in Taiwanese and Indian cultures and defined such design elements as cultural attractors, which were the interface design attributes of the websites, reflecting their denotations matching the expectations of the local culture. Cultural design preferences could directly map into culturally appropriate design elements for a website, but were usually inclined to the stereotype. When the cultural design preferences were successfully applied to a website, they could markedly increase the usability of websites and thus address the needs of the target culture audience. As Sun (2001) mentioned, exhaustive studies could be reduced when a company realized the type of cultural preferences being used for a specific culture.

Web Components for Effective Communication

Smart et al. (2000) stated that it was necessary to identify several vital categories of web design characteristics for helping the designers convey desired meaning and the users more easily obtain the intended meaning. In this research, integrating the studies of web interface design elements from Sun (2001), Russo and Boor (1993), and Smart et al. (2000) and modifying them, the web element which would influence the web communication are defined visual representation, navigation, layout, link, and multimedia,. The detailed introduction of them is shown as below.

1) Links. Sun (2001) investigated the users from America, Germany, China and Brazil with regard to their design preferences and found the existence of different preferences. It was found that the German users preferred to the links in the navigation bar, which could be set up in alphabetical order; but, such preference was not favored by Chinese and Brazilian users.

2) Visual Presentation, including images, pictures, symbols, icons and graphics. The characteristics of visual representation can efficiently convey a message to the audience in an attracting manner. Russo and Boor (1993) state that the images, symbols, icons we recognize in our culture may have no meaning in another culture. To be successful in the global market, images must be carefully chosen and designed. Designers must sufficiently comprehend that difference among cultures to recognize images that are culturally specific.

3) Layout. This is the display structure that directs scanning information and reflects the orderly flow of tasks. It includes categories of grouping, menu placement, search functions, and orientation. It has been an important part on information retrieval system research. If the layout is properly designed, it can be easier for the viewer to access information, comprehend the information within a contextual and structural model, and facilitate the communication between the user and the system (Yu and Roh, 2002).

4) Navigation. This refers to different kinds of navigational tools, menu formats, different kinds of links, and search capabilities. Marcus and Gould (2000) declared the influences of culture on the navigation in web design. Users from cultures with a high uncertainty dimension tending to a navigation structure intended to prevent them from getting lost, while cultures with a low uncertainty dimension inclined to less control of navigation.

5) Multimedia. In web communication, the usage of multimedia in the Internet plays a significant role. Web multimedia refers to sounds, animation, moving text, and streaming video. If the multimedia can be used properly, it can enrich the experience of the user. In Web design, text and graphics can be kinetic (i.e., flash animation, gif animation), and moving text can add emphasis in a page. Integrating multimedia into Web interface design can be a very powerful means for transmitting information beyond that of text, visual representations, still images and pictures, if it is incorporated into the web appropriately. Multimedia can be used for increasing comprehension or facilitating understanding for users and enhancing the user experience. Multimedia is also an effective means to mislead, distract, and aggravate the audience (Brinck et al., 2002).

METHODOLOGIES

The methodologies in this research are introduced in this section. First, two cultures are selected. Second, hypotheses are constructed. Third, the experimental websites are constructed. Fourth, the usability test is established.

Two cultures are selected

The cultural categories used in this study are based on national culture and operationalized with websites from Taiwan and the UK that possess different cultural attributes, based on Hofstede's country cultural dimensions (2005). The UK is ranked 42/44, while Taiwan is ranked 29/30 in power distance among 53 countries; the UK is ranked 3rd, and Taiwan 44th in collectivism; the UK is ranked 9/10, and Taiwan 32/33 in femininity; the UK is ranked 47/48, and Taiwan 26th in uncertain avoidance; and finally, the UK is ranked 18th, and Taiwan 3rd in long-term orientation.

Hypotheses

In the author's previous research (Hsieh and Hong, 2013), where local website auditing was constructed, it is found that there are significantly different preferences for web interface design between Taiwanese and the UK cultures (Table 1 and Table 2). Further, the hypotheses are proposed - if the websites are embedded with culturally preferred characteristics and incorporated with their cultural dimensions, they can be more effective in communications. To testify the hypotheses, the web experiment is constructed. In the website experiment, the culturally preferred design elements are embedded in websites and the communication effectiveness (usability) would be evaluated from five aspects including learnability, efficiency, minimal errors, comprehension, and satisfaction. The details of the hypotheses are shown as below.

1. If the Nantou government website is embedded with culturally preferred design elements that reflect Taiwanese culture, it can be more effective in the communication for Taiwanese users.
2. If the Nantou government website is embedded with culturally preferred design elements that reflect Australian culture, it can be more effective in the communication for Australian users.

Table 1 Significant Taiwanese culturally preferred attributes (Hsieh and Hong, 2013)

Variables	Components
Visual representation	Iconic symbols Images of leader Photo of accomplishment Images of group Cute style illustration Photo of landscape Banner within local city image
Navigation	Accessibility icon
Links	Popup a new window Dynamic button Mouse over (color change)
Layout	Three-column
Multimedia	Stream video Flash animation Moving picture Moving text Opening

Table 2 Significant Australian culturally preferred attributes (Hsieh and Hong, 2013)

Variables	Components
Visual representation	Images of daily life Banner within color shape
Navigation	Search

Links	Accessibility on text
	Mouse over underlined
	Stay in the same window
Layout	Two-column
	Flexible width design
	Information. guide on bottom

Four experimental websites being constructed

Based on the hypotheses and integrated with the results of the local websites audit, the culturally preferred features are embedded into four websites, which have the same content from Nantou county government in Taiwan, the Chinese version of Nantou government website reflecting Taiwanese culture, the English version of Nantou government website reflecting Taiwanese culture, the Chinese version of Nantou government website reflecting the UK culture, and the English version of Nantou government website reflecting the UK culture .

Evaluation criteria for usability test

Based on Nielsen (1993), four evaluation criteria are interpreted as the following. 1) Efficiency can be the assessment of the time needed for carrying out a task. The process of more rapidly executing a task implies the greater efficiency. 2) Learnability points out the easiness to learn the system. 3) Minimal errors indicate the evaluation criteria of errors being defined as users making as few errors as possible when using the system. 4) Satisfaction refers to the extent of the users preferring the system. 5) Comprehension: It is easy to understand and readable.

CONSTRUCTION OF WEB EXPERIMENTAL AND USABILITY TEST

Participants and experiment procedure

Fifteen UK participants were recruited in this experiment, and these participants were native English speakers, averagely aged 30 years old, stayed in Taiwan less than 3 months, taught English, and were born and brought up in the UK. Taiwanese participants were native Chinese speakers, the PhD and Master students from National Yunlin University of Science and Technology, and aged 30 years old averagely. Each participant was invited to Laboratory DA405 in National Yunlin University of Science and Technology to participate in the web experiment. The Taiwanese participants were asked to carry out the tasks assigned on Chinese version of Nantou website embedded with Taiwanese preferences and Chinese version of Nantou website embedded with the UK preferences, and the UK participants were asked to carry out the tasks assigned on English version of Nantou website embedded with Taiwanese preferences and English version of Nantou website embedded with the UK preferences according to the instruction and task assigned. When the participants executed the tasks, they interacted and navigated with the websites, during which the time and clicks for carrying out each task were captured and recorded. Finally, the usability questionnaires were used for their responses, subjective opinions, and satisfaction from the participants.

Task assigned

Using Spool et al.'s method (1999), the questions were designed so that the answer comprised a single fact and only one correct answer. Four sets of tasks were constructed for each website. The tasks were equivalent for the Nantou website which was embedded with Taiwanese culturally preferred attributes and the Nantou website which was embedded with the UK culturally preferred attributes. Taiwanese participants needed to locate requested pages on the Nantou website which reflect Taiwanese culture and the Nantou website which reflect the UK culture, and the UK participants needed to locate requested pages on the Nantou website which reflect Taiwanese culture and the Nantou website which reflect the UK culture.

Table 3 Task assigned for Taiwanese and the UK participant

	The Chinese and English version Nantou web reflecting Taiwanese culture	The Chinese and English version Nantou web reflecting the UK culture
Task01	When is the Beauty Leg Princess Election in Nantou held and write it down?	Where is the Beauty Leg Princess Election in Nantou held and write it down?
Task02	Can you find out Government Ethics Bureau's free call for prosecution and write it down?	Can you find out Government Administrative Bureau's call and write it down?.
Task03	When is the Job Fair held by Nanto County Government?	How many positions are available in the job fair held by Nanto County Government?
Task04	Can you tell me Nantou County Government located in the left side or right side of Highway?	Can you find out the address of Nantou County Government and write it down.

In the Chinese and English version Nantou web reflecting Taiwanese culture website. The first set of tasks are designed to test the efficiency; when the participants carry out this set of tasks, they experience the cultural preferred attributes including dynamic button, Mouse over (color change), and parallel structure. The second set of tasks is designed to measure the minimal errors; when the participants execute this set of tasks, they interact with the cultural preferences including icon symbols, cute style illustration, photo of landscape banner within local city image, deep hierarchy information structure. The third set of tasks is designed to evaluate the learnability; when the participants carry out this set of tasks, they experience the cultural preferred features, such as popup a new window and images. The fourth set of tasks is designed to measure the comprehension; when the participants execute this set of tasks, they experience the cultural preferred characteristics, such as iconic symbols and cute illustration. In The Chinese and English version Nantou web reflecting the UK culture. The first set of tasks are designed to test the efficiency; when the participants execute this set of tasks, they interact with the cultural preferences including logical layout, search function. The second set of tasks is designed to measure the minimal errors; when the participants execute this set of tasks, they interact with the cultural preferences including images of daily life, banner within color shape. The third set of tasks is designed to evaluate the learnability; when the participants execute this set of tasks, they interact with the cultural preferences including stay in the same window. The fourth set of tasks is designed to measure the comprehension; when the participants execute this set of tasks, they interact with the cultural preferences including no use of flash.

RESULTS

Using One-way ANOVA to compare the performance between Taiwan and the UK cultures within each version of web prototype, the significant differences are shown as below.

Table 4 Performance of two culture participants within each version of web prototype

		Sum of Square	df	Mean Square	F	Sig
prototype reflecting	Within groups	1952.133	1	1952.133	79.064	.000
Taiwanese culture	Between groups	691.333	28	24.690		
efficiency	Total	2643.467	29			
prototype reflecting	Within groups	20.833	1	20.833	16.892	.000
Taiwanese culture	Between groups	34.533	28	1.233		
minimal errors	Total	55.367	29			
prototype reflecting	Within groups	4992.300	1	4992.300	45.438	.000
the UK culture	Between groups	3076.400	28	109.871		
efficiency	Total	8068.700	29			
prototype reflecting	Within groups	43.200	1	43.200	5.492	.026
the UK culture	Between groups	220.267	28	7.867		
minimal errors	Total	263.467	29			
prototype reflecting	Within groups	2323.200	1	2323.200	6.397	.017
the UK culture	Between groups	10168.667	28	363.167		
learnability	Total	12491.867	29			
prototype reflecting	Within groups	4.033	1	4.033	38.500	.000
the UK culture	Between groups	2.933	28	.105		
comprehension	Total	6.967	29			

Within the web prototype reflecting Taiwanese culture, with regard to efficiency, $F(1, 28)=79.064$, $p<0.01$; related to minimal errors, $F(1, 28)=16.892$, $p<0.01$. Within the web prototype reflecting the UK culture, in term of efficiency, $F(1, 28)=45.438$, $p<0.01$; related to minimal errors, $F(1, 28)=5.492$, $p<0.05$; with related to learnability, $F(1, 28)=6.397$, $p<0.05$; in term of comprehension, $F(1, 28)=88.5000$, $p<0.01$. Based on the results of One-Way ANOVA, there is significantly different performance between Taiwan and the UK cultures, particularly, the results of efficiency, minimal errors within the sites which reflect Taiwanese culture, and the results of efficiency, learnability, minimal errors, and comprehension within the sites which reflect the UK culture. Further, the Independent samples t-test is applied to comparing the performance of the participants of the same culture between two version prototypes.

Table 5 Taiwanese's response between web reflecting Taiwanese culture and web reflecting UK culture

Evaluation criteria	The web reflecting Taiwanese culture	The web reflecting UK culture
Efficiency (sec)	27.93	35.80
Learnability (sec)	19.00	30.53
Errors (clicks)	5.00	5.67
Comprehension (correct percentage)	80%	40%
Satisfaction (mean of questionnaire)	3.47	3.07

Table 6 The UK users' response between web reflecting Taiwanese culture and web reflecting UK culture

Evaluation criteria	The web reflecting Taiwanese culture	The Web reflecting UK culture
Efficiency (sec)	14.80	10.00
Learnability (sec)	16.00	12.93
Errors (clicks)	4.60	3.27
Comprehension (correct percentage)	33%	93%
Satisfaction (mean of questionnaire)	3.00	4.00

Based on the results in Table 5, comparing the performance of Taiwanese participants between the web reflecting Taiwanese culture and the web reflecting UK culture, it indicates that Taiwanese participants take a shorter time for executing task01 and task02 and click less in carrying task03, and 80% of Taiwanese users get the correct answers when they interact with the sites which reflect Taiwanese culture. Whilst Taiwanese takes a longer time for carrying out task01 and task02 and clicks more in carrying task 03, and only 40% of Taiwanese users get the correct answers when they interact with the sites which reflect the UK culture. Comparing the mean of Questionnaire for satisfaction, the mean from the prototype reflecting Taiwanese culture is higher than the other one.

According to the results in Table 6, comparing the performance of the UK participants between the web reflecting Taiwanese culture and the web reflecting UK culture, it indicates that the UK participants take very short time for executing task01 and task02 and click less in carrying task03, and 93% of the UK users get the correct answers when they interact with the sites which reflect the UK culture. Whilst the UK users take a longer time for carrying out task01 and task02 and click more in carrying task03, and only 93% of the UK users get the correct answers when they interact with the sites which reflect Taiwanese culture. Comparing the mean of Questionnaire for satisfaction, the mean from the web reflecting the UK culture is higher than the other one.

DISCUSSION AND CONCLUSION

Efficiency. The results of Table 4 indicate that the UK users have a significantly different performance between the websites which reflect Taiwanese culture and the websites which reflect the UK culture. Table 5 and Table 6 shows that Taiwanese users perform better in the sites which reflect Taiwanese culture, compared to the sites which reflect the UK culture. This indicates that Taiwanese users work efficiently when they interact on the sites which reflect Taiwanese culture, and UK sites which reflect the UK culture can not benefit them, whilst the UK users do not work efficiently in the sites which reflect Taiwanese culture because it takes them the longer time to perform the tasks. Based on the data for time performance, it is more effective when the UK users use UK style websites.

Minimal errors. Based on the results of the user testing and from reviewing Table 5, it is found that Taiwanese participants make fewer errors in the websites reflecting Taiwanese culture and more errors in the websites reflecting the UK culture. The UK participants make fewer errors in the websites which reflect the UK culture and more errors in the websites reflecting Taiwanese culture.

Learnability. The results of user testing in Table 5 indicate that Taiwan users have a higher learnability in the sites which reflect the Taiwanese culture, whilst the UK users has the highest learnability in the sites which reflect the UK culture.

Comprehension. With regard to the comprehension, 80% Taiwanese users get the correct answers within the sites which reflect Taiwanese culture, whilst only 46% English users get the correct answers. 33% Taiwanese users get the correct answers within the sites which reflect the UK culture, whilst 93% English users get the correct answers in such sites.

Satisfaction. Based on the results of the questionnaire, the UK participants are more satisfied with the websites which reflect the UK culture (see Table 5), whilst Taiwanese are less satisfied with these. Taiwanese users are more satisfied with the sites which reflect Taiwanese culture (see Table 6).

Overall, the results are consistent with the hypotheses. According to the above findings, it indicates that incorporating cultural dimension and study local cultural preferences should be considered for web developers and designers in order to develop an effective website. The UK culture has individualism, short-term time orientation, lower power distance, and higher masculinity. Overall the UK culturally preferred elements improve the effectiveness for the UK users, such as, images of daily life, banner with color, two column, search function , more mouse over underlined, stay in the same window, and no use of flash. Taiwanese culture has collectivism, long-term time orientation, high context dimension, higher power distance, and lower masculinity. Overall, the Taiwanese

culturally preferred elements improve the effectiveness for Taiwanese users, such as, iconic symbols, cute style illustration, flash animation, moving picture, moving text, popup a new window, dynamic button, mouse over(color change) and three columns. It reveals that if the web is embedded with culturally preferred design elements of the target-culture, it can benefit the users from the target-culture. This research has the potential to help web developers and designers to develop a web interface design that is culturally appropriate.

Acknowledgement. The author gratefully acknowledge the support for this research provided by Taiwan National Science Council under grants no. NSC 102-2410-H-224-035-.

REFERENCES

- Barber, W., and Bardre, A.: (1998), *Culturability: The Merging of Culture and Usability*. In: Proceedings of the 4th Conference on Human Factors and Usability. <http://zing.ncsl.nist.gov/hfweb/att4/proceedings/barber>
- Brinck, T., Gergle, D., and Wood, S. D. (2002), *Designing web sites that work: usability for the web*. Morgan Kaufmann Publishers, San Francisco.
- Evers, V. (2002), "Cultural aspects of user interface understanding: an empirical evaluation of an E-learning Website by international user group". Ph.D thesis, Open University.
- Hofstede, N. (1999), "Global Issues, local Concerns". Technical COMMUNICATION, Second Quarter 1999, pp. 145-148.
- Hofstede, G. (2005), *Cultures and Organizations: Software of the Mind*. McGraw-Hill, London.
- Holland, R., and Young, M. (2009), "A Theoretical Model for Cross-Cultural Web Design", LNCS, Vol.5619, pp. 712-721.
- Hsieh, and Hong, S.D.(2013), "Localization of Web Design: An Investigation of Culturally Preferred Web Attributes in Taiwan and the UK", LNCS, Vol.8024, pp. 404-413.
- Marcus, A., Gould, E.W. (2000), "Cultural Dimensions and Global Web User-Interface Design". Interactions, pp. 33-46 (July/August)
- Nielson, J. (1993), *Usability Engineering*. San Francisco: Academic Press.
- Russo, P., Boor, S. (1993), "How fluent is your interface? Designing for international users. Human Factors and Computer Systems". Proceedings of 4th conference on Human factors in computing systems. Addison-Wesley, Boston, pp.342-347.
- Sheriden, E. F. (2003), Cross-cultural Web Site Design. Considerations for developing and strategies for validating locale appropriate on-line content, *MultiLingual Computing & Technology* #43,12 (7). <http://www.multilingual.com/FMPro?-db=archives&-format=ad%2fselected%5fresults.html&-lay=cgi&-sortfield=magazine%20number&-sortorder=descend&-op=eq&Ad%20Type=reprint&-max=1817529521&-recid=33326&-token=%5bFMP-currenttoken%5d&-find=>
- Smart, K. L., Rice, J. C., and Wood, L. E. (2000), "Meeting the Needs of Users: Toward a Semiotics of the Web Technology & Teamwork", IEEE, pp.593-605.
- Smith, A., Dunckley, L., French, T., Minocha, S., and Chang, Y. (2004), "A process model for developing usable cross-cultural websites", *Interacting with Computers*, Vol16, No1, pp.63-91.
- Spool, J.M., Scanlon, S., Schroeder, W., Snyder, C., and De Angelo, T. (1999), *Web Site Usability: A Designer's Guide*. Morgan Kaufman Publishers Inc.
- Sun, H. (2001), "Building a culturally-competent corporate web site: an exploratory study of cultural markers in multilingual web design". In: Proceedings of the 19th annual international conference on Computer documentation, pp. 95-102.
- Trompenaars, F. and Hampden-Turner, C. (1997), *Riding the Waves of Culture: Understanding Cultural Diversity in Business*. London: Nicholas Brealey.
- Yu, B. and Roh, S. (2002), "The effects of menu design on information-seeking performance and user's attitude on the Web". *Journal of the American Society for Information Science and Technology*, Vol53, No11, pp. 923-933.

Meaning Making Regarding Threat Narrative Based on Discourse Analysis

Laurie Fenstermacher^a, Lawrence Kuznar^b, Mariah Yager^c, and Steve Shellman^d

*^a Air Force Research Laboratory
Wright Patterson Air Force Base, OH 45433 USA*

*^b Indiana University Purdue University Fort Wayne
Fort Wayne, IN, USA*

*^c National Security Innovations, Inc.
Fort Wayne, IN, USA*

*^d Strategic Analytic Enterprises
Williamsburg, VA, USA*

ABSTRACT

Often after an act of violence, a forensic analysis of what the responsible individual(s) or group(s) said or wrote would reveal “signals” that would have foreshadowed the event. Although these signals frequently occur well in advance; they are often nuanced, requiring a different lens to find and interpret discursive patterns and practices related to social identity, affect, integrative cognitive complexity, trustworthiness, and worldview. Threat narrative is the behavioral (actions/words) manifestation of subjective reality regarding threat. These lenses help an analyst reason about how an individual or group sees themselves and others, their perception of threat and propensity to negotiate, cooperate or engage in violence. The result is a tomographic view, albeit imperfect one, of the threat narrative.

The Air Force Research Laboratory (AFRL) has been engaged in research aimed at enabling meaning making from discourse regarding threat narratives for several years. Previous research developed multi-lingual methodologies (Arabic and Pashto), documented in primers transitioned to operational customers, including the National Air and Space Intelligence Center (NASIIC), which enable the detection and interpretation of discourse related to social identity (in-group/out-group) (Fenstermacher et. al. 2012). This paper will focus on two projects designed to enable meaning making from the analysis of discourse, one employing a systematic approach to creating codebooks for automated analysis, and another employing taxonomies for automated analysis of identity and intent.

A grounded theory approach, using human coders, was used to identify relevant discursive practices and patterns (themes and rhetorical devices), including intensifiers used to express trust, trustworthiness or distrust in Farsi. Key themes were identified such as Islam, positive virtues, and advanced age and/or experience. Association with a trusted individual, expert citation, language related to intimacy and poetry were typically associated with trust. Conversely, distrust was conveyed in themes related to negative virtues and government agendas and by use of figurative language such as metaphors and allusions.

An automated approach focused on understanding the link between affect and behaviors using quantitative models of the effects of emotions (eight classes coded: trust, fear, surprise, sadness, disgust, anger, anticipation and joy) on behaviors of competing actors in Syria, Egypt and the Philippines (e.g., a dissident group, government and population). This approach highlights similarities and differences in resulting behaviors. For example, in both Egypt and the Philippines, societal fear, anger and disgust toward dissidents resulted in increases in dissident hostility. Conversely, in Egypt, government hostility increased in response to societal disgust whereas in Philippines it decreased.

This research effort identified several apparently independent features: idea density and vocabulary diversity (proxies for integrative cognitive complexity) and affect expressed regarding in-group and out-group. Preliminary results indicate that the combination of these features would enable accurate forecasting of Naxalite bombings (.92 in sample, .8 out of sample correlation between model and actual bombings). These results are promising but preliminary; the generalization and robustness of these factors relative to different groups and languages will be assessed in a newly started research effort.

The coding methodologies and the text analytic algorithms are a significant step forward in assisting analysts to systematically interpret threat narrative related language, characterize sources and reason about future behaviors and influence as well as helping to mitigate information overload by cueing analyst attention to potentially relevant documents and important events.

Keywords: Discourse Analysis, Sentiment, Affect, Text Analytics, Forecasting, Intent, Cognitive Complexity, Thematic Analysis, Grounded Theory, Social Identity, Threat Narrative

INTRODUCTION

When an act of violence occurs, forensic analyses invariably identify precursor “signals”. These signals are in a variety of forms, including physical and discursive behaviors. Each event, seen in hindsight, was preceded by a pattern of other events or behaviors. These behaviors have been described and studied in previous research. Bandura (1990) outlined a number of mechanisms used to disengage the moral sanctions to violence, including dehumanization (e.g., the Nazi cartoons depicting Jewish people as rats during the lead up to the Second World War). Numerous studies on the relationship between Integrative Complexity (IC) and subsequent behaviors demonstrated a link between a drop in IC and subsequent hostility/violence (Suedfeld, 1977, 2010). Social identity (in-group/out-group) has been identified as a factor in motivating violence. As threats from out-groups increase, in-group identification increases and, as in-group identification increases, out-group derogation increases, as does the likelihood of violence against that group (Pyszczynski, 2013). Previous work developed methodologies to identify discursive patterns related to social identity in order to forecast hostility (Toman et al., 2010).

New research led by the Air Force Research Laboratory seeks to leverage previous research on discourse analysis and threat behaviors in order to develop methodologies and algorithms to enable analysts to make meaning about threat and threat narratives. Threat narrative is the externalized behavioral “storyline” in which knowledge and subjective reality regarding threat is expressed (Moon, 2009). It includes the answer to questions as perceived by the object of study: Is there a threat? What’s motivating the threat? What’s the intent? What are the capabilities of the threat? How can the threat be deterred or defended? Where did the threat come from? Why does it exist? What are the “right” actions regarding the threat? What are preferred future outcomes? The focus on threat narrative implicitly requires an understanding of not only the “what”, “who” and “how” related to threats, but also enables the development of an understanding of “why” – providing key clues regarding potential nonkinetic courses of action.

DEVELOPING DISCOURSE ANALYSIS METHODOLOGIES TO ENABLE MEANING MAKING ABOUT THREAT

Previous AFRL discourse analysis research resulted in methodologies that enabled analysts to have a “lens” to identify and interpret the language associated with social identity in Arabic (Toman, et. al. 2010) and Pashto (Kuznar and Yager, 2012). This proved to be useful not only for forecasting future hostility but also for discerning the difference in rhetorical strategies and language used by the mujahedeen in Afghanistan (Soviet era) and the Taliban (current era). A new research effort was initiated to use the same approach to identify the language associated with trust and trustworthiness in Farsi. The goal was to be able to improve source characterization (e.g., is the source representative, trustworthy, etc.), an important step in analysis.

The primary objectives of this study included:

- Identifying themes and rhetorical devices associated with trust in the Persian language,
- Determining valence, or positive/negative connotation, of rhetorical devices,
- Providing diachronic analysis of narrative theme and linguistic device stability,

- Exploring the effects of genre, author's profession, and purpose of communication,
- Generating a codebook of these dimensions, and
- Creating a structured database for reuse by analysts and researchers.

Methods

In consultation with native speaking Persian scholars, a corpus of 71 Iranian Persian (Farsi) texts that in some manner concerned trust was compiled, including works from Medieval sources, Classical Persian literature, contemporary literature and political discourse. They came from news articles (47%), prose literature (24%), websites/blogs (19%), Opinion Editorials (7%) and poetry (3%).

Focus groups of Persian-speakers with experience in language analysis provided an initial assessment of key rhetorical devices and themes used to identify trust and distrust in Persian. The information from these focus groups was used to produce an initial codebook of trust themes and rhetorical devices.

In order to sample the range of interpretations and reactions an audience may exhibit, a larger group of Persian-speakers read and then answered questions about a set of Persian documents. Readers rated the importance of trust in each document and provided their commentary on what uses of language expressed trust or distrust. It is important to note that the readers in this study were well-educated, urban ex-patriot Iranians, and so probably best represent how well educated, urban Iranians would react to rhetoric surrounding trust issues.

Feedback from readers and focus groups was coded using a typology of 122 themes and rhetorical devices, along with measures of how important readers claimed the themes were for expressing trust or distrust and how strongly these sentiments were expressed. NSI researchers then created a structured database that contained each instance of each type of theme and rhetorical device. The database (both in Excel and SPSS format) is searchable, and provides for further analysis to answer questions about the stability of themes and rhetorical devices across multiple variables.

Key Findings – Most Common Themes Associated with Trust/Distrust

In parallel with trust definitions developed by social psychologists studying Western populations (Mayer, Davis, and Schoorman, 1995), the key to trust expressed by our Persian consultants is the ability to be vulnerable and not have others take advantage of one's vulnerability. Native Persian speakers also reported that the ability to believe what one is told, or for it to necessarily be true, were secondary aspects to defining trust.

Many of the themes associated with trust could easily be categorized according to the Western Aristotelian Pillars of Persuasive argumentation: *ethos* (credibility of a source), *pathos* (emotional appeals), and *logos* (appeals to logic and evidence) (Hovland, Janis, and Kelley, 1953; McCrosky and Teven, 1999). Each of these aspects of persuasive argument was important to the Persian participants, and in roughly equal measure.

The most common themes associated with trust/distrust include:

- **Religion (Islam)** generally establishes trust.
- Suspicion that there is a hidden **government agenda** generates suspicion and distrust.
- Perceived positive **virtues** establish trust, perceived **vices** undermine it.
- Advanced **age and experience** of a messenger establishes trust.
- Being perceived as an **enemy** or **foreigner (*farangi*)** creates distrust.

Common themes that undermine trust include **anger, lying, fear, conspiracy, secrecy, betrayal, and violence**.

Rhetorical devices are ways of using language. The following rhetorical devices that most commonly establish or undermine trust include:

- **Grouping**, which is the association of one party with another that is trusted or distrusted, therefore conveying the same quality to that party.
- **Citing experts** helps to establish trust.
- Citing relevant **examples** helps to establish trust.

- Establishing social closeness, or **intimacy**, increases trust; creating social distance undermines trust.
- The use of **poetry** is generally associated with establishing trust.
- **Figurative language** (allusions, metaphors) and symbols are often used to either establish trust or distrust, depending on the comparison being made.

Intensifiers are rhetorical devices that do not in themselves make references to anything, but either increase or decrease the effect of another theme or rhetorical device.

- The only common intensifier is the use of **hyperbolic**, or extreme, language, which strongly discredits a statement and creates distrust.

Key Findings – Valence of Themes / Rhetorical Devices

Not all themes and rhetorical devices convey the same degree of trust or distrust. A trust valence metric, which varies from -2 (extreme distrust) to 2 (extreme trust), measures how strongly a theme or rhetorical device conveys trust or distrust. The themes and rhetorical devices whose valences were statistically significantly different from zero (no valence) demonstrate the strongest degrees of trust or distrust.

Only one theme conveyed trust in a strong way that is statistically significant: **provide evidence**. Authors who provide evidence for their arguments are more likely to be trusted. Themes that convey distrust were more frequently identified. They include: emotive themes such as **hyperbole**, **intensifiers**, **anger**, **fear**, **loss of trust**, **secrecy**, **conspiracy**, **divisiveness**, **enemy**, **threat**, **virtue**, **betrayal**, and **lying** (

Table 1).

Table 1. Valence of Statistically Significant Themes Related to Trust

Theme/Rhetorical Device	Valence (+ Trust / - Distrust)	p value < or =	Description
Provide Evidence	1.60	.001	Hyperbolic language undermines trust
Hyperbole	-1.60	.002	Use of intensifiers tends to back-fire and undermine trust
Intensifiers	-1.50	.002	Use of intensifiers tends to back-fire and undermine trust
Anger	-1.60	.002	Expressions of anger undermines trust
Fear	-1.61	.002	Creating fear undermines trust
Loss of Trust	-1.61	.003	Erosion of trust over time understandably undermines trust
Secrecy	-2.00	.0001	Being secretive undermines trust
Conspiracy	-1.80	.0002	Acting in collusion against others undermines trust
Divisive	-2.00	.0001	Being socially disruptive undermines trust
Enemy	-1.50	.001	Enemies are not trusted
Government Agenda	-1.79	.0001	Fearing that the government has ulterior motives undermines trust
Opposing Sides	-2.00	.0006	Focusing only on the weaknesses of an opponent's point of view undermines trust
Threat	-1.83	.001	Making threats undermines trust
Virtue	-1.81	.0001	Most of the time, when an individual's virtue is mentioned, it is to focus on vices that undermines that individual's credibility
Betrayal	-1.70	.003	Betrayal undermines trust
Lying	-1.79	.0001	Lying undermines trust

Effects of Genre and Purpose

Several independent variables were hypothesized to vary with trust themes and rhetorical devices including: Time Period, Genre, Author's Profession, and Purpose of a document. Differential use of themes is measured by examining the percentage of text segments in the corpus associated with a particular category within these variables (e.g. percentage of times Clerics used a theme versus other professions). The strength of these following relationships were measured with Cramer's V and were found to be statistically significant at the $p < .05$ level.

- **Figurative language** (Allusions, Symbols, Metaphors) is more often used on **websites** (5.5%, Cramer's $V = .136$, $p = .031$).
- **Clerics** (16.3% Cramer's $V = .191$, $p = .004$) and documents sympathetic to the **Green Movement** (42.9% Cramer's $V = .328$, $p < .0001$) made most use of **expert citation**.
- Documents sympathetic to the **Green Movement** (21.4% Cramer's $V = .224$, $p = .0001$) rely more on **providing evidence**.
- **Clerics** use **poetry** about as much as literary writers (8% and 9.9% respectively, Cramer's $V = .218$, $p < .0001$), and poetry is also common in **pro-democracy** (8%) and **anti-regime** (5.6%) writing (Cramer's $V = .205$, $p = .001$).
- **Objectivity** is most valued in **Green Movement** writing (14.3% Cramer's $V = .19$, $p = .002$).
- **Students** (18.2% Cramer's $V = .183$, $p = .008$) and documents sympathetic to the **Green Movement** (14.3% Cramer's $V = .166$, $p = .015$) are most likely to invoke the personal characteristics of the **messenger** to establish trust or distrust.
- **Students** (18.2% Cramer's $V = .219$, $p < .0001$) and **pro-democracy** (8% Cramer's $V = .151$, $p = .043$) writing are more likely to invoke **fear**, which undermines trust.
- Concern over a **government agenda** peaks in 2009 (33.3% Cramer's $V = .197$, $p = .035$), and is most common among **pro-democracy** advocates (19% Cramer's $V = .159$, $p = .024$) and **students** (83.3% Cramer's $V = .316$, $p < .0001$).
- **Students** are more likely to cite **violence** (18.2% Cramer's $V = .174$, $p = .015$), **betrayal** (18.2% Cramer's $V = .175$, $p = .014$), **conspiracy** (36.4% Cramer's $V = .307$, $p < .0001$) and **loss of trust** (18.2% Cramer's $V = .175$, $p = .014$) as reasons to distrust.
- **Secrecy** is cited to establish distrust on **websites** (4% Cramer's $V = .144$, $p = .019$), by a **cleric-politicians** (16% Cramer's $V = .231$, $p < .0001$), and in **pro-democracy** writings (16% Cramer's $V = .233$, $p < .0001$).
- **Religion**, in particular appeals to Islam, are most common on **websites** (25.7% Cramer's $V = .262$, $p < .0001$), most commonly used by **politicians** (60% Cramer's $V = .381$, $p < .0001$), and to some extent by clerics and scholars, and are prominent in both **pro-** (22.6%) and **anti-regime writings** (22.6%) (Cramer's $V = .240$, $p < .0001$).
- **Pro-regime** writings are most likely to invoke the specter of **enemies** to establish distrust (7.6% Cramer's $V = .152$, $p = .039$).
- **Pro-regime** writings are also the ones that invoked the specter of invading or interfering **foreigners** (5.7% Cramer's $V = .157$, $p = .028$), although all of these are very recent (2013).
- Concern over **threat** is most common on **websites** (5.5% Cramer's $V = .133$, $p = .037$) and in **pro-democracy** writings (16% Cramer's $V = .215$, $p < .0001$), and used by **politicians** (22.7% Cramer's $V = .227$, $p < .0001$).

DISCOURSE ANALYSIS ALGORITHMS FOR FORECASTING

Affect and Behaviors

Many standard approaches to forecast behaviors have focused largely on static factors (e.g., poverty, grievances) at the expense of dynamic factors (e.g., who is doing what to whom, attitudes and/or sentiment of key actors/groups). Recent research explored the impact of affect on forecasting behaviors, focusing on the dynamics between the government, dissidents (a non-state actor who uses non-traditional means (protest, violence) to meet their political needs, goals) and population (Shellman 2006a, b, 2008). The population is considered a key actor because both dissidents and (most) governments vie for their support. Text analytic algorithms extracted events as well as affect from data gathered on the Philippines and Egypt from 2001 to 2012. Emotions data were derived from extracted sentiment expressions by mapping the expressions to an affect category; for example anger or fear (Plutchik, 1980). For example, “The Swiss government lauded Suu Kyi for her courage” would be coded as “positive” sentiment and under the “joy” affect class; whereas, for “The King is appalled” the sentiment would be coded as “negative” and the affect class as “disgust”. The relationship models of the government, dissidents and population were estimated using Auto Regressive Conditional Heteroskedasticity (ARCH) due to the non-constant variance in the time series data. The focus was on understanding the relationship both the impact of societal emotions on dissident and government behaviors as well as the impact of government and dissident behaviors on societal emotions.

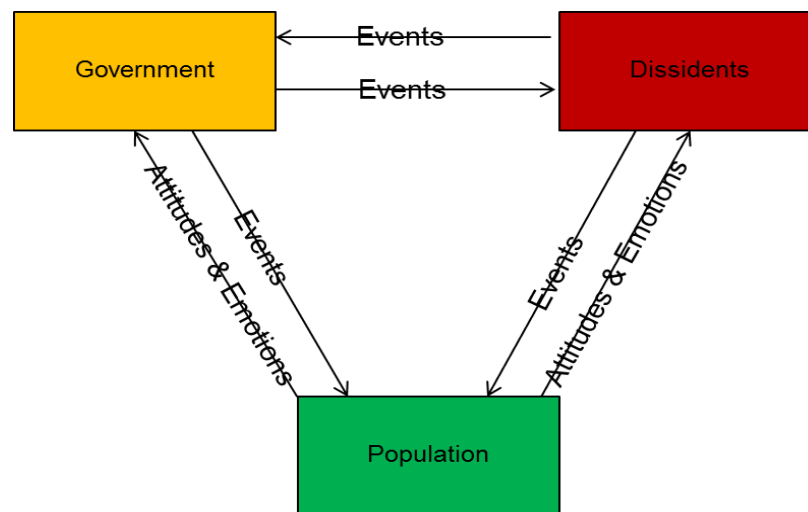


Figure 1: Conceptualization of Conflict as Strategic Interactions

Across the two cases considered, Egypt and the Philippines, there were several similarities in relationships between the dissidents and government and population, but also several differences. For example, in Egypt and in the Philippines, intense societal fear of dissidents and societal disgust toward the government was associated with increases in dissident hostility (See Figure 2) and, conversely, societal anger towards dissidents was associated with a reduction in dissident hostility. However, differences arose in response to disgust expressed by the population toward the government. In the democratic Philippines, government appeared to view negative social attitudes toward dissidents, principally anger, as a pretext to justify cracking down on dissidents through escalation of repressive behaviors; however, these behaviors eased in cases where society began to fear the government or display disgust toward it. This change in behavior makes sense to the extent that democratic leaders require the support of the people to attain or retain political office. Alternatively, the authoritarian Egyptian government reacted to similar forms of disgust by intensifying repression (see Figure 3). In time-series analysis of the emotions variables, only societal anger could be explained by our government and dissident behavior indicators across both cases. Dissident hostility tended to reduce societal anger while government hostility tended to inflame it.

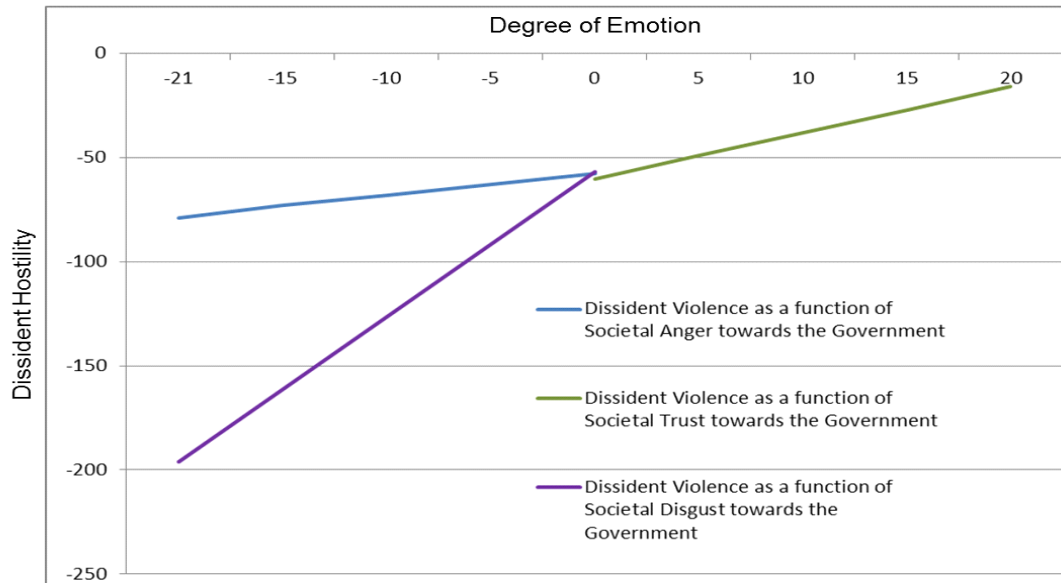


Figure 2. Impacts of Societal Fear and Anger Directed Toward the Government on Egyptian Dissident Hostility, 2001-2012.

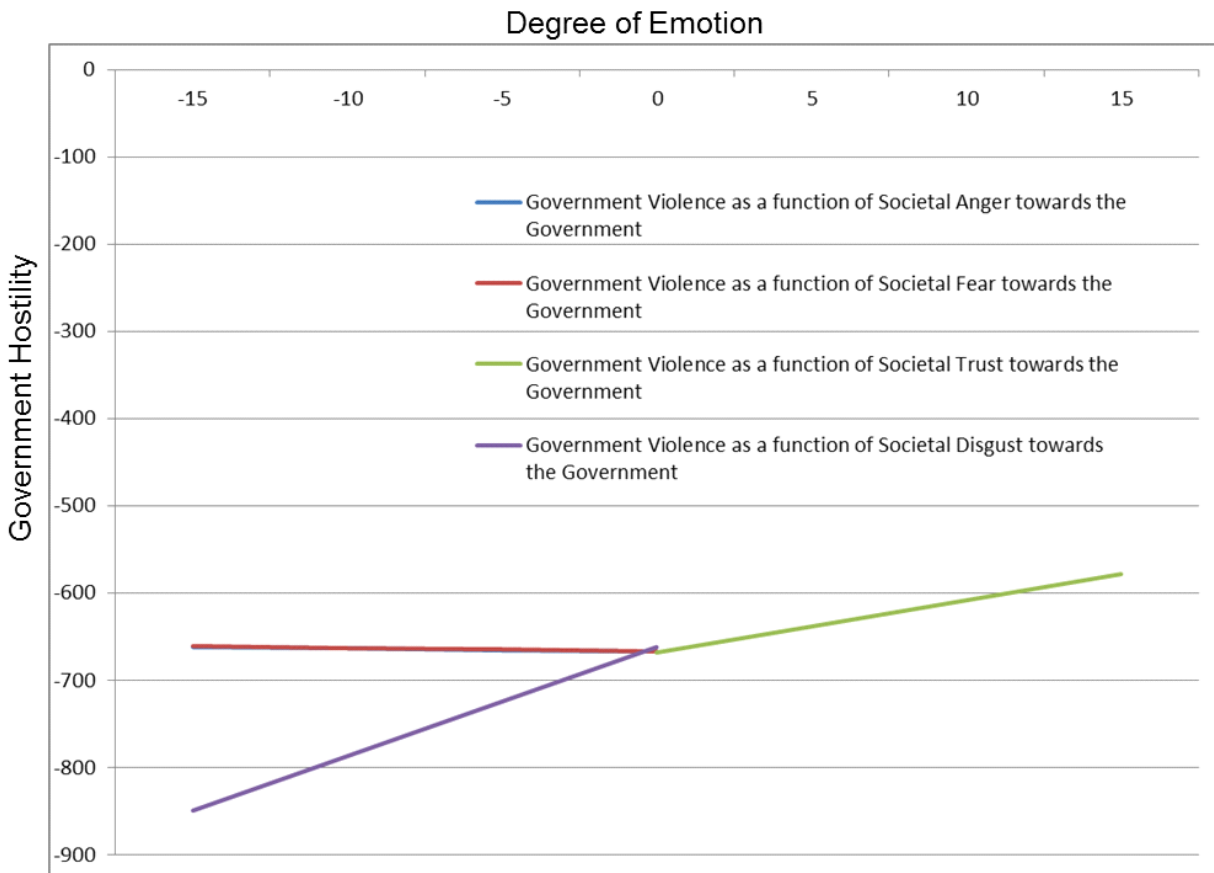


Figure 3. Impacts of Societal Emotions Directed Toward Government on Egyptian Government Hostility, 2001-2012.

Forecasting Group Intent

Research was initiated to investigate ways to automate the processing of discourse in order to extract discourse features in order to identify intent and/or forecast behaviors. One of the resulting efforts focused on extracting, measuring, and modeling five discourse features: cognitive complexity, sentiment expressed by in-group toward out groups (social identity), idea density, vocabulary diversity and keyness. Cognitive complexity is the extent to which a person is open to new ideas/ perspectives and is able to integrate them), idea density is the amount of information contained in a certain number of words, vocabulary diversity is the range of vocabulary used in a text, and keyness is the frequency of the use of certain words in a text relative to comparative texts (a way to track the salience of topics over time). The notion was that either or both idea density and vocabulary diversity would be good proxies for cognitive complexity. Cognitive complexity is a measure intended to approximate the one used by Baker-Brown et. al. (n.d.) in which the sum of the density of high (>2) cognitive complexity (language) indicators is divided by the number of words in the document (and then multiplied by 100 for easier visualization). The sentiment of the in-group (Naxalites) toward various out-groups (Muslim dissident organizations, pro-democracy groups, the government, police, military and other state officials) was measured by extracting language related to anger, fear and disgust (negative sentiment). The expectation was that expression of anger or especially disgust toward an out-group would be associated with increased violence (Matsumoto et al., 2012) and expression of fear would be associated with decreased violence. These features were extracted from 35-50 page newsletters produced regularly by the People's War Group (PWG) from 1998 to the present. The PWG is a violent organization associated with the Naxalite communist movement in India and the newsletter serves to justify opinions, euphemize the group and derogate their opponents, among other objectives.

It is no surprise that there was clear statistical evidence showing that a significant decline in cognitive complexity on the part of the Naxalites was associated with increased levels of violence, consistent with previous findings (Toman et al., 2010, Satterfield, 1998). Idea density and vocabulary diversity also correlated with Naxalite violence, and may serve as orthogonal measures of cognitive complexity (see Figure 4). Keyness is similar to topic tracking, useful for tracking the use of discourse over time and potentially useful for identifying early indicators of violence including certain uses of imagery and propaganda.

Figure 5 below shows the power of using these discourse markers together to predict violence in the case of the PWG. The blue line shows a monthly frequency count of PWG bombings throughout India. The red line shows the predicted level of bombings using only measures of vocabulary diversity, idea density, cognitive complexity, and in-to out-group displays of *fear*, *anger*, and *disgust*. Each variable is statistically significant and there is a 91% correlation between the model-fitted values and the actual number of monthly bombings.

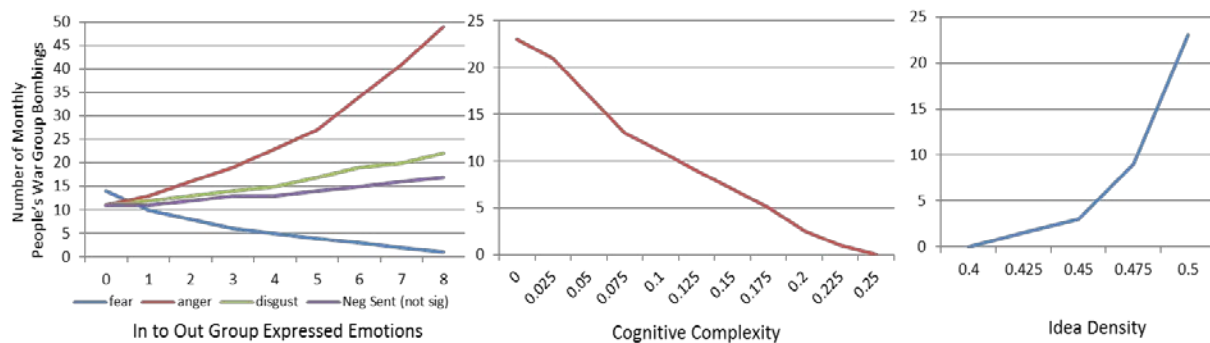


Figure 4: Effects of Various Discourse Indicators on PWG Bombings in India

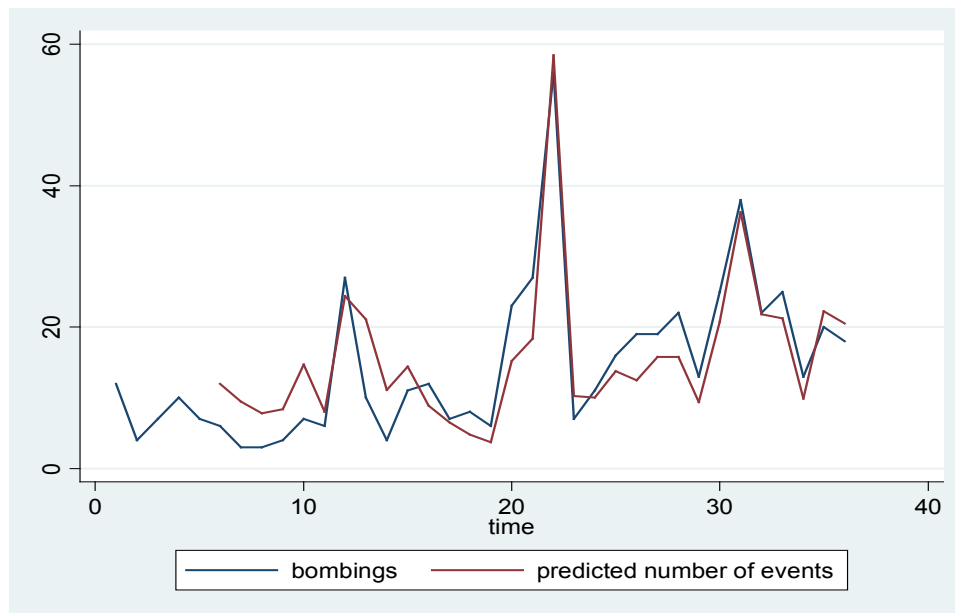


Figure 5: Actual v. Predicted PWG Bombings as a Function of Several Autogenerated Discourse Variables
(6 month lagged IVs; No lagged DV, no event or repression variables – all discourse indicators)

The model only included automated discourse indicators (cognitive complexity, sentiment/affect toward out-group, idea density and vocabulary diversity). There were no lagged dependent variables, structural indicators, event based or repression based measures included in the model. And yet, discourse markers alone yielded an incredibly, perhaps surprisingly, powerful model of group-level violence.

PIECING TOGETHER THE THREAT NARRATIVE

Advanced methodologies for extracting meaning and algorithms for automated text analytics are described in this paper. Ultimately, neither result in a complete description or understanding of the threat narrative. This understanding must be built by an analyst based on cues and clues from these tools. Existing methodologies (Fenstermacher, 2012) focused on the language of social identity and the methodology described previously can provide clues about how a group sees their capability, how group members see each other (e.g., are they trustworthy?) and the capability of the threat and support forecasting. The automated algorithms described in this paper, once matured, can provide important clues to an analyst about the existence and intent of the threat. New methodologies and algorithms will have to be developed to support the analyst's understanding about what is motivating the threat, what the "right" actions regarding the threat are and what the preferred outcomes are. These are part of their worldview and are explicitly and implicitly woven into and manifest in discourse and physical behaviors.

CONCLUSIONS

The AFRL discourse analysis research program and other related research programs have developed methodologies and automated algorithms to cue an analyst's attention and to help them identify the language characteristic of the threat narrative. There is no existing silver bullet available to do this. Ultimately to enable an analyst to make meaning about threat narrative, it will take the integration of existing (deployed) methodologies and algorithms in addition to the maturation and transition of new methodologies and algorithms to foster a truly nuanced understanding of adversaries and their "emic" perspective of threat. This integration must be done in a way to enable the piecing together and inferential reasoning about the threat narrative, providing appropriate cues about impending action, clues into worldview and potentially information about vulnerabilities and potential courses of action to shape or influence behaviors.

REFERENCES

- Baker-Brown, G., Ballard, E. J., Bluck, S., Vries, B. D., Suedfeld, P., & Tetlock, P. E. (2004). Coding manual for conceptual/integrative complexity. Retrieved from <http://www2.psych.ubc.ca/~psuedfeld/MANUAL.pdf>
- Bandura, A. (1990), "*Mechanisms of moral disengagement*", in: *Origins of terrorism: Psychologies, ideologies, theologies, states of mind*, W. Reich (Ed.). Cambridge: Cambridge University Press.
- Covington, M.A. (2007), CPIDR 3 User Manual. CASPR Research Report 2007-03 (<http://www.ai.uga.edu/caspr>), Artificial Intelligence Center, The University of Georgia.
- Fenstermacher, L., Kuznar, L., Yager, M. (2012), "*Analysis of Discourse for Indications and Warnings*," in: *Proceedings 2nd International Conference on Cross-Cultural Decision Making*, San Francisco, CA, 21-25 July 2012, 7307-7316.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953), *Communication and persuasion: Psychological studies of opinion change*. New Haven, CT: Yale University Press.
- Kuznar, L. A., & Yager, M. C. (2012), *Analysis of Pashtun narratives: Report on results. What cues do Pashto speakers use in understanding how to draw in- and out-group distinctions?* Unpublished manuscript, Report prepared for Air Force Research Lab (AFRL), Contract FA8650-10-C-6106, Wright-Patterson Air force Base, Dayton, OH
- Matsumoto, D., Hwang, H.C., Frank, M.G. (2012), "*Emotions expressed in speeches by leaders of ideologically motivated groups predict aggression*", *Behavioral Sciences of Terrorism and Political Aggression*, DOI: [10.1080/19434472.2012.716449](https://doi.org/10.1080/19434472.2012.716449)
- Mayer, R. G., Davis, J. H., & Schoorman, F. D. (1995), "*An integrative model of organizational trust*", *Academy of Management Review*, 20(3), 709-734.
- McCroskey, J. C., & Teven, J. J. (1999), "*Goodwill: A reexamination of the construct and its measurement*", *Communication Monographs*, 66, 90-103.
- Moon, W. (2009). "*Worldview – Subjective Reality – Threat Narrative: The Macro View*". Working paper.
- Plutchik, R. (1980), *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion* 1, New York: Academic Press.
- Pyszczynski, T. (2013), "*Appendix B: Group Identity, Intergroup Conflict and Support for Violent Political Action*" in: *Mixed Initiative Discourse Analysis System* (Ed. Patten, T., Romero, V., Koelle, D., et. al.), AFRL-RH-WP-TR-2013-0029.
- Satterfield, J. M. (1998), "*Cognitive-Affective States Predict Military and Political Aggression and Risk Taking: A Content Analysis of Churchill, Hitler, Roosevelt, and Stalin*." *Journal of Conflict Resolution* 42 (6): 667-690.
- Shellman, S. M. (2006a), "*Leaders & their motivations: Explaining Government-Dissident Conflict-Cooperation Processes*", *Conflict Management & Peace Science* 23(1):73–90.
- Shellman, S. M. (2006b), "*Process Matters: Conflict & Cooperation in Sequential Government-Dissident Interactions*", *Security Studies* 15(4):563–99.
- Shellman, S. M. (2008). "Coding Disaggregated Intrastate Conflict: Machine Processing the Behavior of Substate Actors Over Time and Space", *Political Analysis* 16(4): 464-77.
- Suedfeld, P., Tetlock, P. (1977), "*Integrative Complexity of Communications in International Crises*", *The Journal of Conflict Resolution*, 21 (1): 169-184.
- Suedfeld, P. (2010), "*The Scoring of Integrative Complexity as a Tool for Forecasting Adversary Intentions*", DRDC Report DRDC Toronto CR 2010-039.
- Toman, P., Kuznar, L., Baker, T., et. al. (2010), "*Analysis of Discursive Accent and Discursive Practices I&W*", AFRL-RH-WP-TR-2010-0128.

Semantic Enrichment and Reasoning for Mobile Data Collection of Socio-Cultural Data

Alper Caglayan, Laura Cassani, Tim Clark, Jose Alavedra, Vinay Bharadwaj, Nick Lee, John Parkes

*Milcord
303 Wyman Street, Suite 300
Waltham, MA 02169, USA*

ABSTRACT

Leveraging the widespread adoption of mobile devices, we have developed a semantic knowledge management, mobile data collection, and situational awareness capability that enables the automatic semantic annotation and fusion of mobile collected data. The major components of our semantic knowledge management system consist of form management; data collection, aggregation, and storage; and spatially aware semantic analysis. Our research transforms the analysis of large amounts of data collected from mobile devices from an unstructured process to a structured approach leveraging semantic technology and automated tagging. Semantic technology is used to author, publish, and distribute forms to mobile devices. When submitted from the mobile device, the structured form is mapped to a semantic form to automatically generate semantic annotations from the mobile collected data for semantic enrichment. The semantic representation of the data enables users to organize, analyze, visualize, and create custom information products from the aggregated data.

Keywords: Semantic Knowledge Management, Geospatial Analysis, Mobile Data Collection, Socio-Cultural Data

INTRODUCTION

Estimates suggest that there are now more than five billion mobile phone subscribers in the world, with this trend increasing in relevance particularly in the developing world (Melanson, 2011). In support of the widespread adoption of mobile phones, we have developed a semantic knowledge management cloud service and mobile data collection and situational awareness capability called *SemLayer* that enables users to automatically semantically enrich mobile collected data, fusing information with contextual knowledge for analysis, visualization, and sharing purposes. Semantic enrichment is the process of augmenting unstructured data with semantic relations, defining concepts and the relationships between entities in an ontology. The benefits of harnessing semantic technology for information fusion and decision support are well represented in the literature (Caglayan, 2013; Caglayan et al, 2012; Stroh et al, 2010; Uren et al, 2006). By semantically enriching data, we are able to support semantic queries, allowing for analysis across multiple dimensions within the data. Semantic queries specify some conditions that describe what is asked for, such as the subject, category, properties, and values. Semantically enriched data also enables semantic reasoners to generate new relationships from the existing data (Horridge et al, 2008).

In support of this effort, we developed an integrated solution using a semantic wiki for knowledge management and an Open Data Kit (ODK) inspired mobile data management system. Semantic technology is used to author, publish and transmit structured forms to mobile devices for data collection. When submitted from the mobile device, the structured form is automatically mapped to a semantic form to generate semantic annotations from the mobile collected data for semantic enrichment. The semantic representation of the data enables users to organize, analyze, and create custom information products using semantic queries from the aggregated data, visualizing results with charts, tables, calendars, maps and timelines. We have also extended the semantic knowledge management constructs to represent and reason using geospatial methods. Mobile users can receive geospatially-enabled notifications and alerts to improve situational awareness and to refine the data collection plan.

We use semantic annotations as a semantic enrichment representation for knowledge management, leveraging machine automation to capture semantic enrichment instances such as form based data collection to capture semantic annotations; user defined semantic annotations in the wiki text; semantic inferencing to infer semantic enrichments from collected data; and semantic forms in the wiki to capture semantic annotations. The knowledge base automatically links categories, properties, and values in the data collected with contextual knowledge captured in the knowledge base ontology ingested from other data repositories. Twitter style markup enables the knowledge base to link user defined tags with existing categories, properties, and values, or creates these as new categories, properties, or values if they don't currently exist in the ontology.

Using a Description Logics (DL) formalism, semantic fusion consists of semantic queries that search across different category and property dimensions, with the semantic query results available in a variety of visualizations. Semantic queries fuse data across different data sources (contextual knowledge, mobile collected data, user defined data, inferred data, etc.), across different data dimensions (e.g. class (i.e. meta-data, schema) vs. instance (i.e. data)), and across different dimension instances (e.g. categories, properties). Using this format, we are able to support Level 2 and Level 3 data fusion by representing the operational environment and the relationships among the entities and the events in it for situation assessment and through semantic reasoning we are able to infer about the impact of the assessed situation and various vulnerabilities for the threat assessment level (Das, 2009).

Additionally, a DL based semantic solution is scalable to high density data streams. For example, several commercial solutions use triple stores that have been deployed to handle billions of facts. A DL based semantic solution is also flexible in that it is able to handle all types of data by incorporating the knowledge into a common ontology within the semantic store.

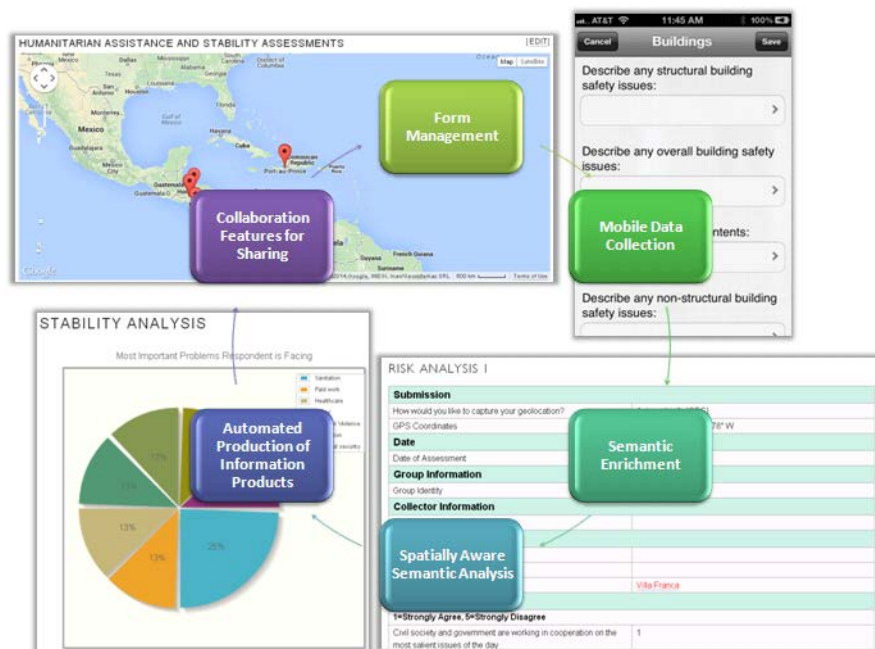


Figure 1. Semantic enrichment and fusion for mobile data collection conceptual framework.

SEMANTIC KNOWLEDGE REPRESENTATION

Description Logics

The underlying logical foundation for our research to leverage machine automation to generate semantic annotations as a semantic enrichment representation for knowledge management is Description Logics (DL), which are subsets of first-order logic. DL separates knowledge representation into: Terminological Box (TBox) for representing ontology classes and axioms about classes and Assertional Box (ABox) for describing specific objects and facts about individual objects. The TBox represents ontology classes and attributes, the hierarchy of concepts and the relationships between concepts. The ABox describes specific entities and facts about individual entities (Nardi and Brachman, 2003). The combination of TBox and ABox defines the full ontology, which is commonly referred to as the knowledge base. The DL representation enables semantic fusion which consists of semantic queries to search across different category and property dimensions.

We transform the structured knowledge collected on forms in mobile devices into semantic annotations using Resource Description Framework (RDF) triples. Formally, an annotation is a tuple consisting of the format “subject, property, object”. RDF and its extension OWL (Web Ontology Language) together provide an explicit and intuitive representation of common knowledge representation syntax for data transfer in the form of RDF triples.

Semantic Wikis

We use the Semantic MediaWiki platform as the basis of our knowledge management system. A semantic wiki differs from a regular wiki by enabling users to annotate wiki content by defining domain concepts, and relations between concepts, in a manner that can be queried like a database. Each wiki page represents an entity (e.g. person, organization, event, location) with the semantic wiki organizing the knowledge into categories and properties so users can easily search and discover relationships within the data. In contrast to a traditional wiki where thousands of users curate knowledge, querying the underlying structured knowledge base dynamically generates most of the pages in a semantic wiki.

Below we present a high level overview of semantic knowledge representation. Description Logics generates a hierarchy of concepts through ‘roles’. Such a hierarchy enables the children nodes to inherit the parent relationships. For instance, Boston has a population property because city has a population and Boston is an instance of the category city.

Table 1: Semantic knowledge representation

Knowledge Representation	Description Logics	Semantic MediaWiki	Example
TBox	Concept	Category	City
ABox	Individual Name	Is a	Boston
Relations	Role	Property	City [[has::population]]
Values		Object	Boston [[has population::625,000]]

Semantic wikis are the ideal platform for knowledge management because of the ontological organization of knowledge, the discovery of knowledge using faceted and semantic search, and linked data storage. For instance, adding semantics to wikis enables semantic search and navigation, provides a more intuitive interface, supports tagging and folksonomy management, and links the wiki content to external resources (El Ghali et al, 2007). There are also a number of benefits semantic wikis provide to support the learning process (Schaffert et al, 2006). For example, semantic annotations lead to reflection about knowledge. The analyst reflects on the content while performing a faceted search on the material yielding to the discovery of relations in the topic of interest. Second, semantic wikis enable fellow analysts to share formal models and build a common model collaboratively. Finally, reasoning and inference capabilities of semantic technologies can lead to discovery of knowledge without active user search thanks to the machine maintenance of organized knowledge (Horridge et al, 2008). The structured format of the semantic annotations enables the automatic generation of knowledge through dynamic tables, charts, graphs, and maps for analysis and visualization.

Semantic Forms

One of the mechanisms we use to capture the semantic annotations easily within the wiki is semantic forms, which enables the rapid input of structured content. Semantic forms are part of an extension to MediaWiki that allows users to use forms for adding, editing, and querying data without requiring any programming. Semantic forms enforce the use of templates by enabling a user to populate a pre-defined set of templates for a page. The data is turned into semantic properties once the page is saved. Forms can be used to edit the data in an existing page, with an “edit with form” tab showing on the top of any page for easy access. Semantic forms also support auto-completion of fields, so users can see previously entered values for a given field. This greatly helps to avoid issues of naming ambiguity and spelling errors.

MOBILE DATA COLLECTION

Semantic Authoring, Publishing, and Distribution of Forms

With an understanding of semantic knowledge representation concepts in our semantic wiki knowledge base, we turn to our mobile data collection capabilities that implements form based structured data collection for semantic annotation in support of semantic enrichment. To semantically enrich and fuse mobile collected data, we have developed an automated system using semantic technology to author, publish, and distribute structured forms to mobile devices and the semantic wiki knowledge base. We built an app that allows users to download forms to their mobile device, and fill out the form with the submissions automatically creating pages for each form instance within the semantic wiki. As the forms are ingested into the semantic wiki, the form fields map onto semantic forms so that data is automatically semantically annotated in the “subject, property, object” format, where subject is an instance of a category. By tagging data in this format, we are able to semantically fuse the data through the generation of semantic queries.

Every data collection form that exists within the mobile app resides as an Excel spreadsheet, or XLS Form, on the semantic wiki. These XLS forms adhere to a standardized format that allows them to be converted to XForms upon upload to the wiki. XForms are XML files that provide a data processing model for XML data and specify how the forms are presented to the user. The XForm contains the structure and content of the form; the full list of questions and information about how these questions should be presented. Most rows represent a question; the rest of the rows specify control structures such as groups. The spreadsheet also is used to specify the options for multiple choice questions, which helps to control the ontology. For example, mobile data collectors interacting with the public can record the sentiment of an interviewee using positive, negative, or neutral annotations, as opposed to having an open text field. Having the data collected in a structured format such as this enables a semantic query to aggregate data across dimensions such as understanding demographic data for those individuals that reported a negative view.

Users are able to download the XLS Form from the wiki, apply any necessary changes to the fields, and re-upload the spreadsheet into the wiki. Upon uploading the spreadsheet, the newest version of the form is represented with the automatic propagation of changes to the form, template, and XForm pages within the semantic wiki, as well as to forms within the mobile devices. This simplifies form management, as users are able to automatically update the mobile and semantic forms simultaneously, editing the ontology, while creating the necessary categories, properties and templates in the background, empowering real-time oversight of data collection requirements.

To demonstrate the types of properties associated with a particular category, consider the below example of a form that can be used for conducting a context analysis assessment. Here we can see the example of how the XLS Form specifies the field names, property names, and datatypes for semantically representing the data within the knowledge base. The field name is how the question appears in the form to the user on the mobile app and within the semantic wiki semantic form. The property name is the actual name of the property from the ontology. The datatype documents the property’s datatype, allowing the semantic wiki to properly evaluate the property value.

Table 2: Context analysis form field names, property names, and datatypes

Field Name	Property Name	Datatype
Identify the primary drivers of the conflict using the list of multiple selections	Primary driver of conflict	String
Hazards in the region have been nominal or non-existent	Level of hazard risk in the region	Number
Survey start date and time	Date and time of report	Date
List the NGOs operating in the area	Has NGO	Page
Primary purpose of the building	Has purpose	String
GPS Coordinates	Coordinates	Geographic coordinates

As the spreadsheet with the above outlined field names, property names, and datatypes is uploaded to the wiki, the form will be automatically distributed to the mobile app and the semantic wiki. Below we can see an example of how the form appears within the mobile app for data collection.

3.1 Primary purpose

☐ Hospital

☐ School

☐ Government

☐ Housing

☒ Other

☐ NA

Specify other.

Figure 2. Mobile app with question from context analysis form.

Semantic Annotations

As data is collected and the form is submitted from the mobile device, it automatically maps onto a semantic form with the associated semantically enriched data available in the semantic wiki. Each of the fields of the form are represented as standardized properties to create a common ontology. The specific assessment or form that is being completed becomes the subject, with the field represented as the property, and the inputted data as the object. As an example of a semantic annotation from mobile collected data represented in the “subject, property, object” format, the semantic wiki would store information about a particular building assessment using the name of the building unit as the subject such as “Clark Elementary, Has purpose, School” or “Clark Elementary, Year built, 1984”. Since each individual assessment is captured in a form on an individual wiki page, a user can go in manually and edit the semantic form at any time to add to or correct the field collected data. Using the mobile device’s GPS locator, we also automatically capture the coordinates of where the assessment took place to enable geospatial analysis and visualization. Below we can see an example of the semantic form within the semantic wiki.

Submission	
How would you like to capture your geolocation?	☐ No answer ☒ Automatically (GPS) ☐ Manually (MGRS/Lat Long)
GPS Coordinates	geopoint
Date	
Date of Assessment	7 May 2013
Group Information	
Group Identity	☒ No answer ☐ Volunteer ☐ Community ☐ Staff
Collector Information	
Assessment Collector	
Peace Building and Conflict Sensitivity	
What are the key historical turning points (international, national, and NGO sector) that have shaped the current conflict?	
Identify the major actor-groups (4 maximum) who currently exert a positive influence on the conflict?	
1)	
2)	
3)	
4)	

Figure 3. Context assessment form in the semantic wiki.

We present a summary of all of the queryable annotations in a factbox at the bottom of each page. Below we can see an example of some of the annotations captured in a context analysis assessment form submission.

<u>Accelerators of conflict</u>	Degraded local water source has exacerbated the social conflict +
<u>Action ensuring that there is no discrimination</u>	Yes +
<u>Assessment team</u>	Community +
<u>Community capability rating</u>	Excellent +
<u>Coordinates</u>	14.11262° N, 87.23308° W +
<u>Country has National Disaster Management systems</u>	No +
<u>Country has disaster threats</u>	Flooding + and Tsunami +
<u>Country in disaster prone area</u>	Yes +
<u>Criminal groups view INGOs as desirable target</u>	Yes +
<u>Date / time of report</u>	7 May 2013 +
<u>Description of violent crime circumstances</u>	Violent crime is specifically targeted at American aid workers operating within five miles of the village. +
<u>Essential services infrastructure capability rating</u>	Average +
<u>Hazards caused loss of life</u>	Yes +
<u>Healthcare capability rating</u>	Below Average +
<u>Hostility to INGOs in country</u>	Yes +
<u>How capture geolocation</u>	Manually (MGRS/Lat Long) +
<u>Human induced vulnerabilities increase risk</u>	Yes +
<u>Intended action occurring to advance humanitarian imperative</u>	No +
<u>Local security forces sufficient and free of corruption</u>	No +
<u>Political violence exist in country</u>	Yes +
<u>Primary drivers of causes of conflict</u>	Social + and Environmental +

Figure 4. Context assessment form factbox presenting a summary of semantic annotations.

Semantic Internal Objects

The advantages of employing a spreadsheet-driven solution to form authoring are manifold. In particular the structure of the XLS Form allows us to add augmented functionality. For example, our solution allows users to place groups of properties and their values inside semantic internal objects within the wiki. Semantic internal objects allow us to create “bundled annotations” which represent a collection of multiple values in a single annotation on a page. The value of leveraging these internal objects can be seen in the following example. Suppose a question within a form asks the user, “List this person’s siblings and their dates of birth” It is important to create distinct instances for each individual person and associate their birth date with the correct individual, if we are to treat each person as a unique instance and query it within semantic search. In this case, if we are to add multiple instances of people, we want to ensure that each additional property instance such as birth date is associated with the correct individual, as opposed to being associated with the overall form. We therefore use subobjects to represent these individuals within the factbox, and maintain their annotations separately. Leveraging semantic internal objects allows us to semantically store and query subgroups of compound values.

MOBILE COLLECTED DATA ANALYSIS AND FUSION

Semantic Queries

A semantic annotation format facilitates data aggregation, and enables semantic inference engines to semantically fuse the data and process information for further analysis. In contrast to database queries where terminological knowledge plays no part in the query, our DL based representation enables the combined query of both terminological and assertional knowledge for greater situational understanding. Because data is captured about each

assessment in the form of semantic annotations, those annotations can be leveraged to delve into multiple dimensions of the data to examine particular subsets of the assessments.

For example, we can search by:

- `[[Category:Medical Survey]]` which shows all pages in a category such as Medical Survey.
- `[[Has disease::Cholera]]` which gives all pages annotated with the ‘Has disease’ property that has a property value of Cholera.
- `[[City::Saint Marc]]` which gives all pages annotated as being about the city Saint Marc.

Combining these annotations into one search would enable a user to identify Medical Surveys that have taken place in Saint Marc, where the individual interviewed was diagnosed with Cholera. Once we had this subset of data of interest, we could cross analyze the various dimensions of the data asked in the survey such as, where did the individual access their water. In this way, we are able to analyze across semantic dimensions within our data.

Semantic queries have several different options for result formats, such as maps, charts, graphs, timelines, calendars, and tables. Semantic queries are embedded within the pages of the semantic wiki, so that they are dynamically generated automatically with the most relevant and recent data incoming from the mobile collected forms, without requiring the user to actually understand and populate the semantic query syntax.

In the below example, we can see the result of a query about data collected about non-governmental organizations (NGOs). In this example, we query all NGOs in the knowledge base that are located in Thailand, viewing the associated URL for the organization, with the results displayed in a table format.

	URL
AACCU - Association of Asian Confederation of Credit Union	http://www.aaccu.coop/
AAT - Alliance Anti Trafic	http://aatthai.org/
ABCA - Asian Business Coalition on AIDS	http://www.abconids.org/
ABV - Australian Business Volunteers	http://www.abv.org.au/
ACET International - AIDS Care Education and Training Ltd	http://www.acet-international.org/
ACFOD - Asian Cultural Forum on Development	
ACHR - Asian Coalition for Housing Rights	http://www.achr.net/
ActionAid International	http://www.actionaid.org/
ACT - Organic Agriculture Certification Thailand	http://eng.actorganic-cert.or.th/
ADF - Asian Disaster Foundation	http://www.adf-thailand.org/
ADPC - Asian Disaster Preparedness Center	http://www.adpc.net/
ADRA - Adventist Development and Relief Agency International	http://www.adra.org/
	http://www.adrathailand.org/
ADRA (Thailand Asia Regional Office) - Adventist Development and Relief Agency International	http://www.adra.org/
	http://www.adraasia.org/
AFC - Association of Finance Companies	http://www.afc.or.th/
AFD - Agence Française de Développement	
AFFECT - Association for Akha Education and Culture in Thailand	http://www.akhathai.org/
AFET - Agricultural Futures Exchange of Thailand	http://www.afet.or.th/
AFI - Alliance for Financial Inclusion	http://www.afi-global.org/
AFIC - Asian Food Information Centre	http://www.afic.org/
AFPPD - Asian Forum of Parliamentarians on Population and Development	http://www.afppd.org/

Figure 5. Dynamically generated table of NGOs and the property URL.

In constructing more complex visualizations, we encounter the need to conduct searches for properties that are semantic queries in and of themselves. In such visualizations, the information we need to query is not tagged within the pages, but by nesting a semantic query as a property value we can infer knowledge from the other semantic relationships that exist. For example, say there is a page for “Cholera” where we define a property whose value is assigned by a semantic query that returns a total for the number of cholera diagnoses aggregated from several locations. Because the property value is defined by a semantic query, it is a dynamic value. Thus, we can create an additional page within the semantic wiki to report an aggregation of various types of disease data for reporting, where the size of the bar on the pie chart may increase or decrease depending on the number of cholera diagnoses reported to the wiki at any given time, and where we can break out the analysis by the age of the patient.

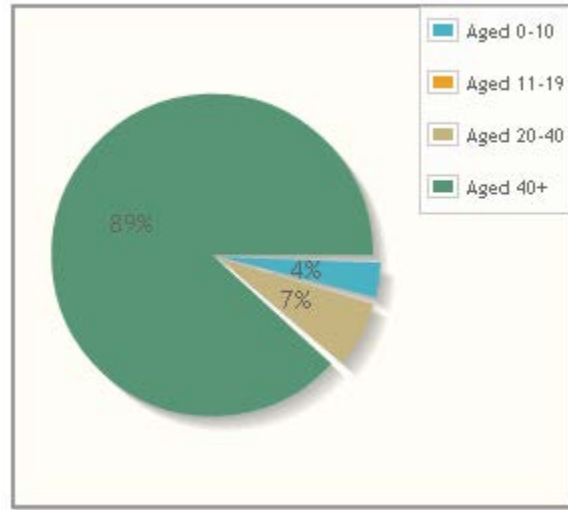


Figure 6. Dynamically generated pie chart of cholera diagnoses aggregated from a medical operation and broken out by patient age.

This is a query within a query, as we are delving into the multidimensional semantic relationships that exist, rather than the tags within a page, to deliver information. This innovative feature is a force multiplier as it demonstrates that visualizations and reports can be enhanced to drill down into multiple dimensions of data, querying for relationships nested among other relationships, to derive insight and produce refined visualizations in support of situational awareness.

Semantic Templates

By leveraging parser functions alongside semantic queries, we can also conduct statistical analyses. For example, we can create a template that embeds hundreds of individual semantic queries within the page that are all dynamically updated with the incoming mobile collected tagged data for data aggregation and analysis requirements. For example, perhaps we are interested in aggregating patient registration data for a medical support operation. These tables provide medical planners with on the ground analysis of how many patients have passed through medical registration throughout the day, the demographic information of these patients, as well as the most common treatments sought. Medical planners can then leverage this real-time information to plan medical relief accordingly.

As an example, we can create a template that allows a user to create a one-line table that depicts the most recent assessment submitted in any given category. To embed this template within a page, we can include the name of the template 'Most Recent Assessment' within double braces. When using the template, we include the name of the category of interest to view the most recent submission.

```
{{Most Recent Assessment | Category of Assessment}}
```

For instance, perhaps we are interested in viewing the most recent submitted assessment for a school assessment form.

```
{{Most Recent Assessment | School Assessment}}
```

A user can list additional qualifiers; for example, perhaps a user wants to view the most recent submitted school assessment, but only for those submitted for a particular location.

```
{{Most Recent Assessment | School Assessment | Location::Port-au-Prince}}}
```

For a more advanced analysis, templates with parameters produce different contents or have different behaviors depending on specifications provided by the user. For example, we can utilize a template to allow users to embed charts containing statistics about patient registrations completed by a particular team in a specific location as part of a medical assistance operation. Here we can see a chart that provides a patient breakdown by age.

Age	Total Patient Registrations	% of Patient Registrations
0-10 years	36	21.6%
11-19 years	7	4.2%
20-40 years	7	4.2%
41+ years	117	70.1%

Figure 7. Dynamically generated analysis of medical data from an operation using a template embedded on a page for automatic population of data as it is ingested from mobile device forms.

This template contains wiki syntax that constructs a formatted table with twelve semantic queries, however, a user would just need to transclude the template on a page rather than having to format the table and define each semantic query. To embed this template within a page, we include the name of the template we created such as ‘Medical Summary’ within curly double braces. When using this template, a user must pass parameters for location and medical team so that knowledge base understands which medical team and location the user is interested in.

{{Medical Summary|medical team=Team 1|location=Port-au-Prince}}

Geospatial Analysis

We utilize an interactive map for our geospatial analysis functionality to construct geospatial semantic queries. The user can write a query and see the results on a map, then refine the query and see the new results immediately, with no page refresh required. For example, we have defined the geospatial property “Within”, to translate to a semantic query, which provides all of the results whose coordinates fall within the polygon. The user does not have to understand this syntax however; they are able to select “draw polygon” from the map in the semantic wiki and draw a bounding area. From there the query is automatically generated representing the “Within” semantics. For example, here we can see the resulting forms that have been submitted from within the defined circle. Other geospatial operators that we have incorporated into our semantic search interface include “Contains”, “Covers”, “CoveredBy”, “Overlaps”, “Crosses”, “Disjoints”, “Equals”, “Intersects”, and “Touches”. This automated knowledge growth empowers users to monitor team activities, visualize impact, and geospatially analyze operations.

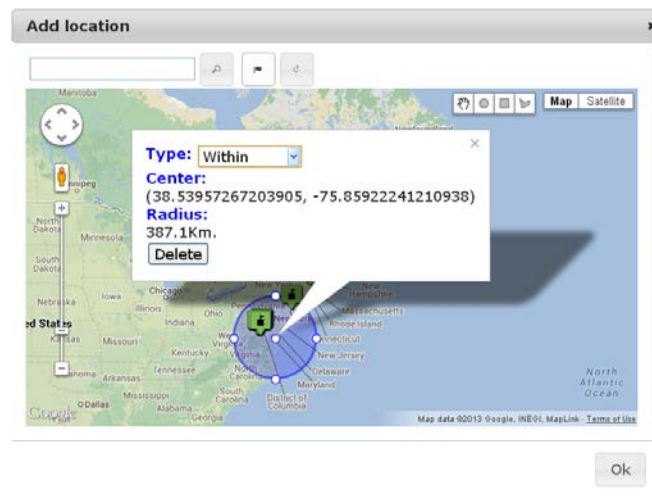


Figure 8. Semantic query incorporating geospatial semantics using the "within" operator.

Linked Data

The linked data within the semantic wiki also allows users to browse associations among the mobile collected data. For example, the datatype listed in the XForm can specify that a form entry ought to be a page within the semantic wiki. That field will therefore contain an automatically generated link that directs the user to a page for that entity

containing existing contextual knowledge. We have leveraged a number of contextual knowledge sources for thousands of pages in the wiki, automatically ingesting this information and producing the relevant annotations. For example, we semi-automatically ingested data pertaining to administrative divisions for particular countries, and annotated information within these pages about key leaders and social network data; cultural knowledge; and infrastructure information, and linked these locations to surveys completed by mobile data collectors. In this way, a user can submit an assessment in a particular area and then be able to go to that submission page on the semantic wiki and see a link to a page about that location with pertinent socio-cultural data. Here we can see a page referencing socio-cultural data that was automatically ingested into the semantic wiki.

Aheu Language	
Primary Language Name	Aheu
Alternate Language Name(s)	Kha Tong Luang, Phon Sung, Phonung, So, Thanun
ISO 639-3 Code	lhm
Population of Speakers	750
Writing System(s)	
Region(s) Where Spoken	

Aheu is a language. Approximately 750 people speak Aheu as a first language^[1]

References

1. Lewis, M. Paul (ed.), 2009. Ethnologue: Languages of the World, Sixteenth edition. Dallas, Tex.: SIL International. Online version: <http://www.ethnologue.com/>

Category: Language

Facts about Aheu Language

Alternate Language Name: Kha Tong Luang → Phon Sung → Phonung → So → Thanun

ISO 639-3 Code: lhm

Population of Speakers: 750

Primary Language Name: Aheu

Figure 9. Aheu Language page automatically ingested from Ethnologue Website (<https://www.ethnologue.com>) for linking socio-cultural information to mobile collected data.

It is important to note that when items are blue within the semantic wiki, it means that there is a page about that entity and that there is content populating that page. When items are red, it denotes that the wiki has indicated that this entity is a page, but there is no content on that page. This type of demarcation can assist operators with data collection requirements. For example, if a user sees that a particular entity is referenced numerous times in the wiki, but that there is no content supporting the understanding of this entity in the knowledge base, meaning it is marked as red, the user can prioritize data collection requirements against this entity. The semantic wiki provides a list of the most referenced red entities to assist in prioritizing data collection requirements.

Developing methods and means for automatically ingesting relevant socio-cultural data from websites and databases, tagging the information, and representing it semantically within our semantic wiki was one of the most significant developments for our solution as it enables the augmentation of mobile collected data by automatically linking to relevant socio-cultural data for situational and threat awareness.

CONCLUSIONS

Our semantic knowledge management solution for the semantic enrichment and fusion of mobile collected data, *SemLayer*, eliminates the challenge of form management for mobile data collection campaigns, providing a cloud service for collection, aggregation, analysis, and collaboration in support of big data analysis. We use semantic annotations as a semantic enrichment representation for knowledge management, leveraging machine automation to capture semantic enrichment through form based data collection and user defined semantic annotations in the wiki text. Semantic forms in the wiki capture semantic annotations for semantic enrichment allowing the knowledge base to automatically link categories, properties, and values in the data collected with contextual knowledge captured in the knowledge base ontology ingested from other data repositories. Semantic inferencing can be effectively used to infer semantic enrichments for greater data aggregation and analysis.

What is new about our approach is that we enable authoring of form semantics; automate form publishing and distribution using semantic technology; and automate semantic enrichment and the semantic fusion of mobile collected data. Semantic technology enables the use of spreadsheets to define the semantics of a form, and upload it to the knowledge base for automated generation of semantic forms. Upon upload, the form is automatically distributed to mobile devices while any changes to the forms are automatically propagated. Form submissions are

automatically ingested into the knowledge base, and linked with contextual knowledge. The semantic representation enables the use of semantic queries to analyze and gain insight in the collected data, while spatial semantics are fully integrated with semantic queries, which can be specified and visualized using maps. Semantic queries fuse data across different data sources, across different data dimension types, and across different dimension instances.

ACKNOWLEDGMENTS

This research was supported by the Office of Naval Research under the direction of Dr. Ivy Estabrooke and Gary Kollmorgen.

REFERENCES

- Caglayan, A. (2013, June). “*Mobile Enabled Spatially Aware Semantic Wiki for Information Management using Open-Source Software Platforms*”, Presentation at Semantic Technology and Business Conference, San Francisco, CA.
- Caglayan, A., Cassani, L., Mooney, L., Morgan, W., Boyles, A. (2012, November). Semantic Technologies for Civil Information Management during Complex Emergencies. In *2012 IEEE Conference on Technologies for Homeland Security (HST)* (pp.523-528). Boston, MA.
- Das, S. (2009), *High-Level Data Fusion* Norwood, MA: Artech House.
- El Ghali, A., Tifous, A., Buffa, M., Giboin, A., Dieng-Kuntz, R. (2007, September). Using a Semantic Wiki in Communities of Practice. In *Proceedings of the 2nd International Workshop on Building Technology Enhanced Learning Solutions for Communities of Practice* (pp. 22-31), Crete, Greece.
- Horridge, M., Bauer, J., Parsia, B., Sattler, U. (2008, October). Understanding Entailments in OWL. In *Proceedings of the Fifth OWLED Workshop on OWL: Experiences and Directions*. Karlsruhe, Germany.
- Melanson, D. (2011), *UN: worldwide internet users hit two billion, cellphone subscriptions top five billion*. Engadget Website: <http://www.engadget.com/2011/01/28/un-worldwide-internet-users-hit-two-billion-cellphone-subscript/>
- Nardi, D., Brachman, R.J. (2003), “*An introduction to description logics*”, in: The Description Logic Handbook, Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (Eds.). pp. 1-40
- Schaffert, S., Bischof, D., Bürger, T., Gruber, A., Hilzensauer, W. (2006, June). Learning with Semantic Wikis. In *Proceedings of the First Workshop on Semantic Wikis: From Wiki to Semantics (SemWiki2006)*. Budva, Montenegro.
- Stroh, L., Caglayan, A., Rashed, T., Burke, D., Eaton, G. (2010), “*Geospatial Campaign Management for Complex Operations*”, in: *Advances in Cross-Cultural Decision Making*, Schmorow, D., Nicholson, D. (Eds.). pp. 460-469
- Uren, V., Cimiano, P., Iria, J., Handschuh, S., Vargas-Vera, M., Motta, E., Ciravegna, F. (2006). “*Semantic annotation for knowledge management: Requirements and a survey of the state of the art*”. Web Semantics Volume 4 No. 1.

Basic Study on Prevention of Human Error

- How Cognitive Biases Distort Decision Making and Lead to Crucial Accidents -

Atsuo MURATA and Tomoko NAKAMURA

*Dept. of Intelligent Mechanical Systems
Graduate School of Natural Science and Technology, Okayama University
Okayama, Japan*

ABSTRACT

On the basis of the analysis of past case studies of accidents, it was examined how cognitive biases were ubiquitous in the process of accident outbreak, distorted decision making, and led to crucial accidents. We made an attempt to point out that cognitive biases distort decision making and are potentially related to crucial accidents. Using a few cases of crucial accidents (Challenger space shuttle disaster, collision accident between the Japanese Aegis-equipped destroyer “Atago” and fishing boat, and Three Mile Island nuclear power plant accident), it was demonstrated how cognitive biases are related to these accidents. It was demonstrated that heuristic-based biases such as confirmation bias, groupthink, and social loafing surely appeared in the process of accident breakout. Overconfidence-based biases such as illusion of control, fallacy of plan, and optimistic bias were also ubiquitous in the route to a crucial accident. Moreover, framing biases was found to contribute to the distorted decision making, and eventually turn into the main cause of crucial accident. In such a way, as well as human factors or ergonomics approaches, the prevention and the deletion of cognitive biases were indispensable for preventing crucial accidents from occurring.

Keywords: Cognitive Bias, Confirmation Bias, Groupthink, Fallacy of Control, Fallacy of Plan, Framing.

INTRODUCTION

Different from the traditional economics, the bounded rationality is commonly assumed in behavioral economics (Kahneman, 2011, Tversky and Kahneman, 1974, Kahneman and Tversky, 1984, Altman, 2012, Angner, 2012, Bazerman and Moore, 2001, Ariery, 2009, Ariery, 2010, and Ariery, 2012). Due to bounded rationality, we generally cannot make decision rationally, and thus suffer from cognitive biases pointed out by Kahneman (2011), Tversky and Kahneman (1974), and Kahneman and Tversky (1984). Kahneman and Tversky (1984) states that our cognitive information processing is conducted by System1 or System2. While System2 requires us to conduct effortful, demanding and deliberate mental activities, System1 operates quickly, automatically, without time consuming, and intuitively with little or no efforts. Although heuristic approaches that we adopt when we have no time to deliberate are based on System1, and are very simple and intuitive, such approaches constantly suffer from cognitive biases.

One of the major causes of the Challenger space shuttle disaster (Reason, 1990 and Vaughan, 1997) is regarded to

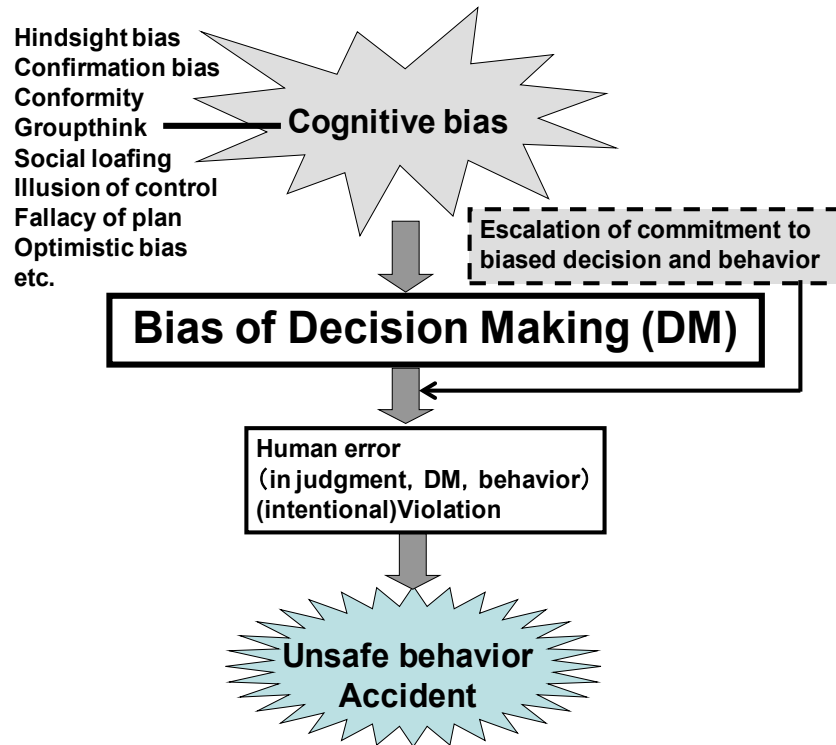


Figure 1. Relational model between cognitive biases and unsafe behaviors or accidents.

be due to groupthink, especially illusion of unanimity (Janis, 1982). In this case, although the manufacturer of O-ring recognized the risk of malfunction of O-ring under the severely cold temperature, the manufacturer agreed with the launch of the Challenger space shuttle because of illusion of unanimity. After some serious accident occurred, one tends to overestimate the occurrence probability of such an accident. For example, we hesitate to use an airplane immediately after a serious aviation accident due to the overestimation of a fatal aviation accident. This property is one type of cognitive biases called hindsight bias. It is suggested that such a bias becomes an obstacle to the objective survey of an accident.

Therefore, it is important to get insights into how the cognitive biases are related to, and lead to crucial accidents. In this study, it was discussed how cognitive biases induce crucial accidents by distorting decision making. We made an attempt to point out that cognitive biases distort decision making and are potentially related to crucial accidents. Using a few example of crucial accidents, it was demonstrated how cognitive biases are related to these accidents.

HOW COGNITIVE BIASES LEAD TO UNSAFE BEHAVIORS OR ACCIDENTS

As shown in Figure 1, it is hypothesized that cognitive biases cause the distortion of decision making, and thus that this leads to human errors in judgment, decision making, and behavior and eventually (at the worst case) triggers crucial accidents if the commitment to the biased judgment, decision making, and behavior is escalated.

According to Bazerman and Moore (2001), we summarized how cognitive biases are induced. It is hypothesized that the heuristics such as availability, representativeness, confirmation, or affect cause the biases such as confirmation biases, and anchoring and adjustment. In Figure 2, not only heuristics but also overconfidence and framing are shown as causes of biases. Moreover, it is assumed that our bounded awareness and uncertain (risk) situations form the basis of heuristics, overconfidence, and framing. Due to such bounded ability, it is valid that we humans cannot behave rationally but irrationally.

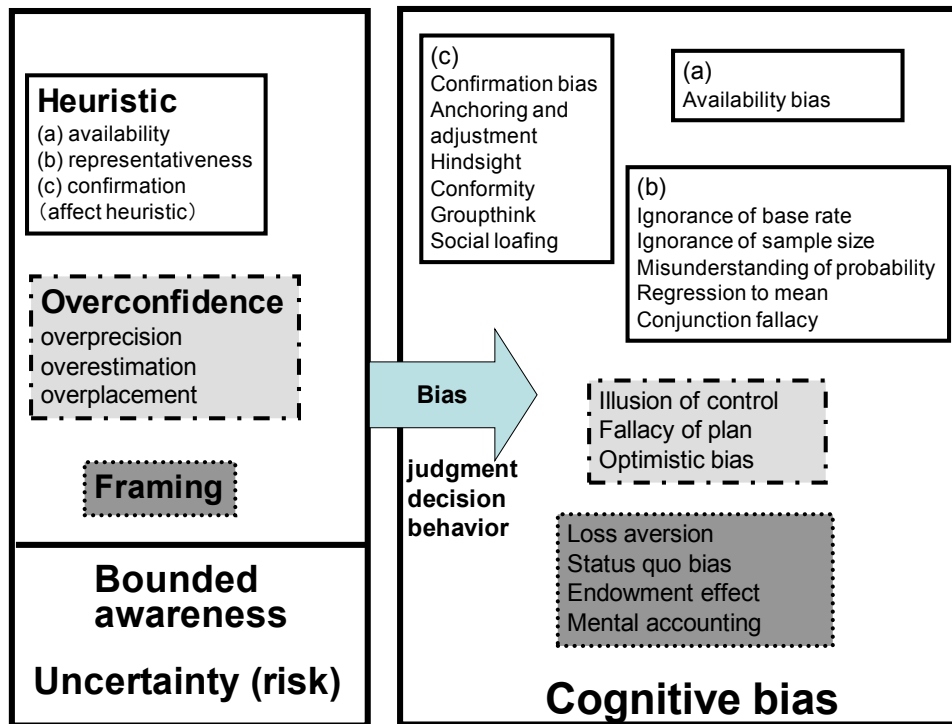


Figure 2. Mechanism of cognitive biases due to heuristics, overconfidence, and framing.

As mentioned above, we frequently tend to behave irrationally, and are in most cases unaware of how and to what extent these irrational behaviors influence us. Such irrational tendencies are sure to distort our decisions, and in the worst cases this leads to crucial accidents according to the model in Figures 1 and 2. Without consideration of our bounded rationality (irrationality), we cannot approach the prevention of crucial accidents and analyze the true (genuine) cause (source) of accidents.

In the next chapter (Chapter 3), it is demonstrated how the cognitive biases form the main cause of crucial accidents using three case studies of crucial accidents.

ANALYSES OF ACCIDENTAL CAUSE FROM THE PERSPECTIVE OF COGNITIVE BIASES

Challenger space shuttle disaster (Reason, 1990 and Vaughan, 1997)

One of the major causes of the Challenger space shuttle disaster is regarded to be due to groupthink, especially illusion of unanimity. As shown in Figure 2, groupthink stems from confirmation heuristic. In this case, although the manufacturer of O-ring recognized the risk of malfunction of O-ring under the severely cold temperature, the manufacturer agreed with the launch of the Challenger space shuttle because of illusion of unanimity (groupthink (Janis, 1982)). This eventually led to the crucial accident. The summarized analysis of this accident is shown in Figure 3.

Collision accident between the Japanese Aegis-equipped destroyer “Atago” and fishing boat

It is possible that social loafing, which stems from confirmation bias or optimistic bias, leads to a crucial accident. On 19, February in 2008, the Japanese Aegis-equipped destroyer “Atago” that belonged to Japanese Ministry of Defense collided with the fishing boat “Seitokumaru.” Consequently, two crews of the fishing boat were missing

The manufacturer of O-ring recognized the risk of malfunction of O-ring under the severely cold temperature.

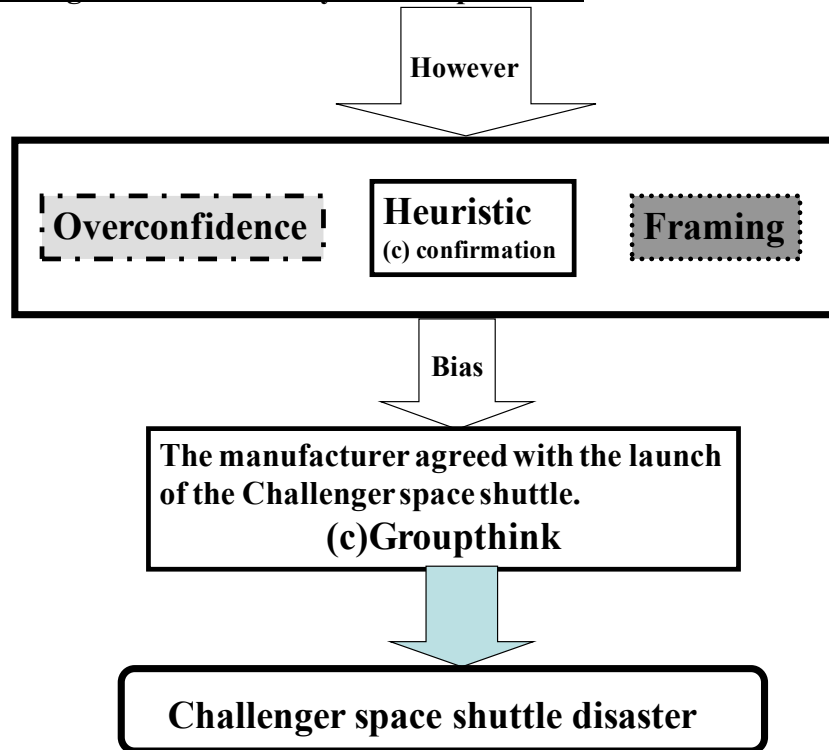


Figure 3. Summarized analysis of Challenger space shuttle disaster.

and thereafter identified as dead. One of the main causes of this accident is inferred as follows. Although twenty-four crews were working (on duty) on the Aegis-equipped destroyer when the accident occurred at about 4 a.m., nobody properly noticed the fishing boat, and thus could not take a proper countermeasure against the collision. In spite of many crews on duty, every crew must optimistically reckon that someone would notice the fishing boat, if any. This corresponds to the social loafing phenomenon which stems from cognitive bias (confirmation or optimistic bias). The summarized analysis of this accident is shown in Figure 4.

Three Mile Island Nuclear Power Plant Accident

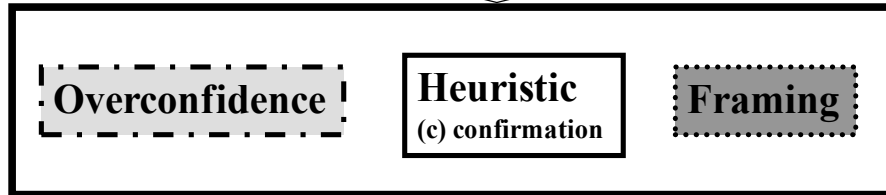
The main cause of this accident is that the operators forget to open the valve of the auxiliary (secondary) water feeding pump after the maintenance task of the auxiliary water feeding (emergency feedwater) pump, and did not notice this error for longer hours. This is related to the confirmation bias (see Figure 2) to make operators believe that such a subtle error actually should not be the cause of criticality accident. In spite of automatic operating of ECCS (Emergency Core Cooling System), the operators could not notice the malfunction of the nuclear reactor due to the availability heuristic (especially Halo effect).

On the basis of the past report on the malfunction of the pilot operated relief valve and the past accident related to this malfunction, it should be expected that the pilot operated relief valve cannot be closed in emergent situations. The confirmation bias that such a trouble would not occur in Three Mile Island nuclear power plant also prevented operators from noticing and identifying the malfunction of pilot operated relief valve.

It took long time to identify the cause of the meltdown of the reactor core due to the rapid and abrupt increase of the reactor core temperature. This must be due to the framing effect (bias) that the operators generally can not identify the cause from the multiple perspectives especially under emergent situations. The operators adhered to the narrow frame which they usually have, and could not apply another frame to the problem solving (identification of the cause).

The preparedness or provision for the accident must be enough due to the engagement of job by many (24) crews.

However



Bias

The crew was trapped into a bias that someone will notice the risk (existence of fishing boat).

(c)Social loafing

Collision accident between the Japanese Aegis-equipped destroyer “Atago” and fishing boat

Figure 4. Summarized analysis of collision accident between the Japanese Aegis-equipped destroyer “Atago” and fishing boat.

The design of the central control room must be also a main cause of the criticality accident. The display system of the central control room was not designed so that the states of the nuclear power plant such as the pilot operated relief valve, the emergency feedwater pump, and the drain tank were easy to recognize and notice, and eventually the cause of the rapid and abrupt increase of the increase of the reactor core temperature could not be identified until the meltdown of the reactor core occurred. The optimistic bias, the fallacy of plan, and the fallacy control (in more detail, see Figure 2) lie here. The designers of the control room and the plant operator must optimistically predict that any situations can be noticed and recognized by the system in status quo. Therefore, they did not take the worst situations into account in designing the central control room due to both fallacies of plan and control. The summarized analysis of this accident is shown in Figure 5.

CONCLUSIONS

As seen and recognized from three examples, many crucial accidents include cognitive biases as one of main causes of the accident. The correction or modification of bias in decision making must be one of the promising measures for preventing crucial accidents.

When the designers, the engineers, and the managers of modern technologies such as transportation systems, nuclear power plants, and social inflation systems don't understand humans' fallibility (error-prone properties) related to our irrational mind, we tend to design new and improved systems that don't take our limitation (irrationality) into account.

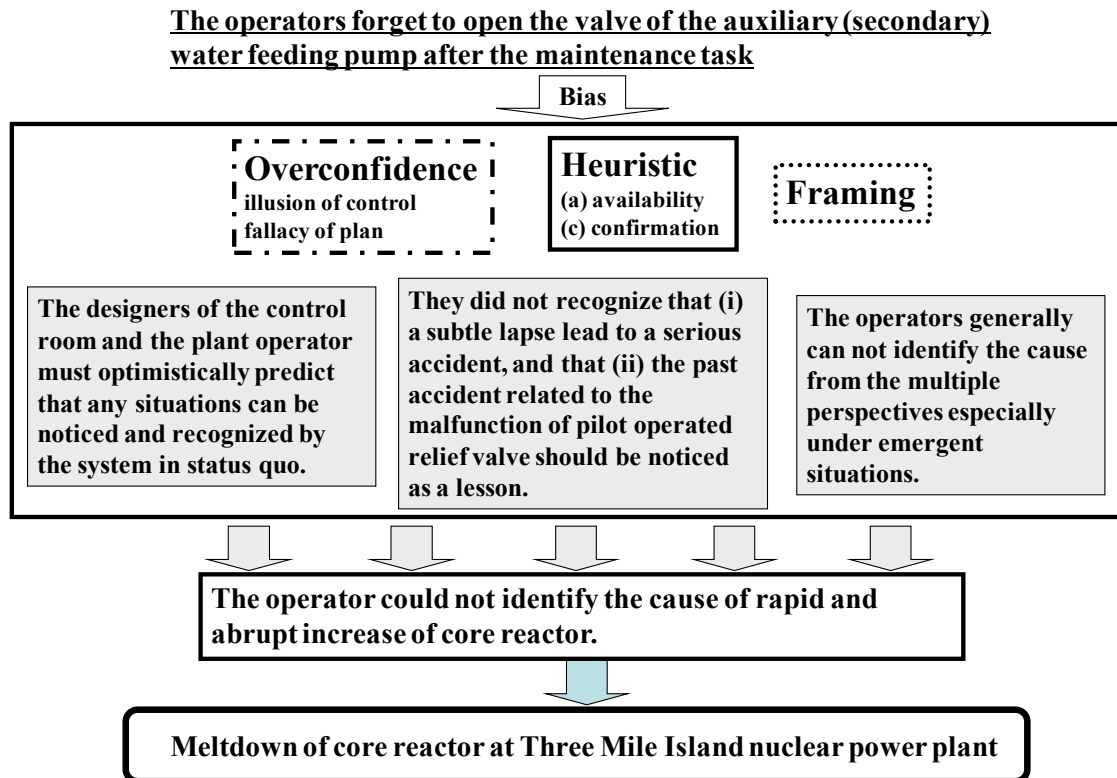


Figure 5. Summarized analysis of Three Mile Island nuclear power plant accident.

Consequently, we inevitably distort our decisions, and make our errors or mistakes serious, and these distortions or errors lead to magnificent and crucial accidents such as analyzed in the previous section. Without such an understanding to our irrationality, we unwillingly repeat crucial accidents, and cannot get out of vicious circles of similar crucial accidents. The understanding of how cognitive biases distort decision making and lead to crucial accidents is essential in order to avoid such vicious circles as pointed out by Dekker (2006).

REFERENCES

- Altman, M. (2012), *“Behavioral Economics for Dummies”*, John Wiley & Sons Canada, Ltd.
- Angner, E. (2012), *“A Course in Behavioral Economics”*, Palgrave Macmillan.
- Ariery, D. (2009), *“Predictably Irrational -The Hidden Forces that Shape our Decisions-“*, Harper.
- Ariery, D. (2010), *“The Upside of irrationality -The Unexpected Benefits of Defying Logic at Work and at Home-“*, Harper.
- Ariery, D. (2012), *“The (Honest) Truth about Dishonesty -How We Lie to Everyone-Especially Ourselves”*, Harper.
- Bazerman, M. H. and Moore, D. A. (2001), *“Judgment in Managerial Decision Making”*, Harvard University Press.
- Dekker, D. (2006), *“The Field Guide to Understanding Human Error”*, Ashgate Publishing.
- Janis, I. L. (1982), *“Groupthink: Psychological Studies of Policy Decisions and Fiascoes”*, Cengage Learning.
- Kahneman, D. (2011), *“Thinking, Fast and Slow”*, Penguin Books
- Kahneman, D. and Tversky, A. (1984), *“Choices, Values, and Frames”*, American Psychologist, Vol.34.
- Reason, J (1990), *“Human Error”*, Cambridge University Press.
- Tversky, A. and Kahneman, D. (1974), *“Judgment under Uncertainty: Heuristics and Biases”*, Science, Vol.185.
- Vaughan, D. (1997), *“The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA”*, University of Chicago Press.

Section 3

*Civilization Change: Ideological,
Economic, and Historical Change*

The End of a Civilization: What Moderns Might Learn from Thucydides' Peloponnesian War

G. A. McBrayer

*School of Public Affairs
Morehead State University
Morehead, KY 40351, USA*

ABSTRACT

Thucydides self-consciously presents the Peloponnesian War as the greatest war the world had ever seen to that point in history, insofar as it was a contest between the two greatest Greek powers—Athens and Sparta—at the peak of Greek Civilization. The war, however, would mark the beginning of the end of this great civilization. Although Thucydides does not unequivocally blame Athens for the war that ultimately leads to the destruction of Greece, it is clear that he thinks Athenian devotion to motion, or to the perpetual pursuit of progress, spurred it on. Thucydides appears to lament the great expansion of education, in particular the sophistic education that became prevalent in Greece and contributed heavily to the theoretical justification behind the Athenian Empire. Even or especially education at its highest—Socratic philosophy—seems to bear some culpability for, or is at least symptomatic of, Athens' decline, and ultimately Greece's decline as well, in Thucydides' view. This paper will examine Thucydides' teaching regarding the decline of civilization to see if it can offer any guidance to the current crisis of civilization in the West.

Keywords: Thucydides, Pericles, Diodotus, Alcibiades, Civilization, Athens, Empire, Sophistry, Philosophy

INTRODUCTION

The modern project has long held that peace and progress in the arts and sciences go hand in hand, that there is a kind of natural harmony between progress in the sciences and arts, on the one hand, and politics, on the other. Peace and progress arise concomitantly, each mutually reinforcing the other in a sort of virtuous circle. Indeed, from its inception, modern political rationalism has promised peace to those who would follow the precepts of its science. As Hobbes says, if political scientists had discharged their duty as well as geometers have, “mankind should enjoy such an immortal peace” but for scarcity (Hobbes, *Epistle Dedicatory*). As the famous critic of modernity, Carl Schmitt, says of the modern project, “Technology appeared to be a domain of peace, understanding, and reconciliation” (Schmitt 91). Modernity was going to pave the way for peace, not just through progress in the sciences and arts, but also through republicanism (or liberalism) and increased secularism. Perhaps the most famous proponent for the possibility of peace is Immanuel Kant, who elaborated how it could be achieved in his tract *Perpetual Peace: A Philosophical Sketch*.

Moreover, progress in the arts and sciences was supposed to render man less afraid of powers invisible, less credulous regarding religion. As man's situation became safer and more prosperous, his need for religion would gradually fade away. Hear Montesquieu on the matter:

“Therefore, one does not succeed in detaching the soul from religion by filling it with this great object, by bringing it closer to the moment when it should find religion of greater importance; a more certain way to attack religion is by favor, by the comforts of life, by the hope of fortune, not by what reminds one of it, but by what makes one forget it; not by what makes one indignant, but by what leads one to indifference when other passions act on our souls and when those that religion inspires are silent. General rule: in the matter of changing religion, invitations are stronger than penalties” (Montesquieu, Bk. 25, Ch. 12).

As a result of his newfound secularism, modern man was thought to be rendered more pacific, especially since religion had been the cause of so many conflicts throughout Europe in the centuries prior to modernity. With the goal of minimizing conflict, modern thinkers sought to decrease the public role of piety. Republicanism, progress in the sciences and arts, and a depreciation of the role of religion were going to usher in a new golden age.

But such was not to be the case. Modernity did not culminate in a new age of perpetual peace. In fact, the 20th century was likely the bloodiest in history, and much of the crisis of modernity can be traced to a sudden and violent rejection of faith in the power of modern rationalism to bring about a peaceful and just world political order. Indeed, World War II has been taken as evidence that enlightenment rationalism, modern rationalism had self-destructed. Instead of achieving a new civilized era for humanity, modernity plunged us into a new kind of barbarism (Horkheimer and Adorno xi).

It appears as though modern rationalism has failed to understand adequately the relationship between progress in the arts and sciences and civilizational flourishing. Accordingly, one might be inclined to reject rationalism out of hand, as have many contemporary political theorists. Perhaps, however, there is an alternative to modern rationalism that can help us to understand more clearly the relationship between progress in politics, science, and the arts, on the one hand, and the prospects of civilizational decline, on the other. There are other, earlier forms of rationalism, forms that holds out more modest hopes for what reason can hope to achieve in the world, forms that recognize the fragile status of justice among nations.

Thucydides’ *War of the Peloponnesians and the Athenians* offers such an alternative understanding of politics, one that still seeks to understand politics rationally. Given the crisis of modernity, or the crisis of modern political thought, one is justified in turning to alternative forms of rationalism in an attempt to understand politics. After all, if modern rationalism rested on a rejection of earlier forms of thought, and now modern rationalism is itself called into question, earlier forms of thought reemerge as serious candidates for understanding the world around us. Thucydides considered his book, moreover, to be a possession for all time (Thucydides, I.22.4; hereafter, all unmodified numbers refer to Thucydides’ work). Thucydides thought that his book contained knowledge, or even wisdom, and that he had much to teach future generations, including ours, about politics, the nature of war, the role of chance or fortune in human affairs, and, of course, human nature. Perhaps there is much we could stand to learn from Thucydides. Thucydides self-consciously presents the Peloponnesian War as the greatest war the world had ever seen to that point in history, insofar as it was a contest between the two greatest Greek powers—Athens and Sparta—at the peak of Greek Civilization, the greatest civilization the world had seen to that point. The war, however, would mark the beginning of the end of this great civilization. If Greece was at the peak of its power, a descent must of necessity follow. Precisely because he recounts the greatest war that ever occurred among the Greeks, Thucydides shows us the downfall of a civilization.

Thucydides challenges, moreover, the modern view regarding civilizational progress and war. Rather than showing civilizational decline as a radical departure from the virtues and habits that led to civilization in the first place, Thucydides prods his reader to consider whether progress in the arts and sciences coupled with the increased political power that results from such progress—in a word, civilization—fashion the very tools that will destroy civilization. The end of civilization is not decadence or decline or a departure from the principles that allowed for the growth of civilization in the first place; rather, the decline is the culmination or inevitable conclusion of the self-same sources that allowed for civilizational progress in the first place.

THE GREATEST “MOTION”

The Peloponnesian War was the “greatest motion” ever known—among the Greeks and among the barbarians, and my well be the greatest motion ever in the history of mankind to that point (1.1). This is because the Greeks were at the height of their power and prosperity. Such was not always the case. Originally, the Greeks were a primitive

people who wandered around, and there were no real cities to speak of. Athens, in particular, occupied relatively infertile land and was accordingly more or less left alone. Indeed, Thucydides says that primitive Greeks were not even Greeks—the term Hellenes was not yet applied to them until well after the Trojan War (1.3.3). Because the Greeks were so weak and primitive, their wars were also not very great. This is particularly the case, Thucydides argues, with respect to the Trojan War (1.11). After the Trojan War there was a series of upheavals in Greece, but finally Greece returned to peace, quiet and tranquility (1.12). During this time of peace and tranquility, Greece grew in power and wealth. Corinth especially grew in wealth and power as it developed commerce and trade; most importantly was the growth of the Corinthian navy. The navy served the purpose largely of promoting trade, which included protecting merchants and cities from pirates. The Greeks increasingly engaged in commerce and navigation; the result was growth and the acquisition of wealth throughout Greece (1.13.1). One cannot underestimate the role that such power and wealth eventually played in making the war greater; progress in the sciences and arts rendered the Greeks more capable of waging war on a grander scale—it made large-scale killing and destruction easier.

At this point of relative prosperity, Thucydides recounts the events of the Persian Wars. Twice the Greeks were able to repel the Persian invasion, and the result was even greater power among the Greeks, although after the second failed invasion power was concentrated, above all, in two principal cities: Athens and Sparta. The war served the purpose of showing the Greeks their relative strength and establishing two poles of power within Greece. But it also showed the cities in Greece the need to prepare militarily in case the Persians ever tried to invade again. Indeed, one could say with a measure of justice that the impetus for the growth of Greek military power arose out of fear of the threat posed by Persia. This is perhaps also the origin of the Athenian desire for empire (1.73).

All of these factors that led to the growth of Greek civilization contributed also to its downfall. In what follows, I will focus above all on Athens and the role it played in bringing about the downfall of Greece. While the wealth and power no doubt fashioned formidable tools of destruction, I will focus instead on the intellectual changes that made such a great war possible, and I will focus above all on such changes in Athens. In the first place, the limits of the paper compel me to narrow its scope, but there is also an argument for why one should focus on Athens. Athens, much more than Sparta, seems to have been the cause of the war (1.23.5). Athens' rise, Thucydides reports, was particularly problematic; it was the city most devoted to "motion" and its motion inspired fear in other Greek cities, principal among which was Sparta. Sparta, however, was devoted to rest and, to some degree, to moderation; it seems to have been pulled reluctantly into war and would have preferred rest. Not many cities are capable of faring well through prosperity. Indeed, Thucydides admits that the Spartans and the Chians are the only people he knew of who knew how to be moderate while flourishing—the general rule for those who are flourishing is immoderation—perpetual movement (8.24.4). Since immoderation is the general rule, and since the immoderation of Athens is a cause of the war, we are permitted to focus on Athens. Athens is the more typical case. We are not focusing on Athens because it was the more splendid or glorious city; we are focusing on Athens because of its immoderation, but the two are connected: "Thucydides cannot have been blind to the glory which was Athens, a glory which is inseparable from the spirit of daring innovation and from that madness, that mania, which rises far above moderation" (Strauss, 1989).

THE DARING INNOVATION OF ATHENS

Athens was characterized by boldness, daring, and innovation; the Athenians are quick to act and adventurous. Even or precisely its enemies recognized these traits in Athens (1.70.2). One could perhaps point to a singular event as the source of Athenian daring innovation. To return to the Persian Wars, in the face of the Persian threat, Athens undertook a bold, innovative measure, and its decision to do so ensured its very survival. At the direction of Themistocles, the Athenians came to hold that the best chance they had of survival, the best chance they had to repel the barbarian invasion, was to abandon their city and break up their homes and use the material to fashion a navy. This innovative measure proved to be the key to their success. "The Athenian character—in particular its daring side—was formed, or at any rate came into its own, only at the time of the Persian Wars" (Orwin, 1986). To repeat, this daring act was also impious, and revealed to other Greek cities that Athens would transgress ordinary boundaries of propriety. Accordingly, other Greek cities grew to fear Athens (1.90.1). While we moderns may marvel at the ingenuity of this design, and surely it is something at which to marvel, one cannot but help to think about the lessons Athens drew from it. In the first place, war, that violent teacher, taught Athens the value of innovation. But one must also recognize the other side of this daring undertaking. For in abandoning the city, the Athenians abandoned temples, shrines, sacred spaces, and the graves of their ancestors. To put it as plainly as possible, abandoning the city was an act of gross impiety, and "it would be a real question within the conventions of Greek piety whether a

city could have any being at all under such circumstances” (Forde, 1986). Importantly, this act of impiety did not meet with divine retribution. In fact, to the contrary, it seems that abandoning the city was instrumental if not decisive in securing the city’s well being. The lesson that the Athenians, or at least some Athenians, drew from this particular instance was that the gods do not punish the wicked and reward the just and pious, though the Athenians refrain, initially, from making this conclusion explicit. The gods are not providential, and believing this compels one to call into question the traditional teaching regarding the gods, the teachings of the poets. This initial act of impiety thus gave way to a period, the Athens of Pericles’ time, of even more widespread heterodoxy, agnosticism, and even outright atheism. “Thucydides belongs in a sense to Periclean Athens—to the Athens in which Anaxagoras and Protagoras taught and were persecuted on the ground of impiety” (Strauss, 1964). This is not to say that everyone throughout Athens was impious, quite the contrary. As the very idea of these thinkers being persecuted shows, there was still a strong pious sentiment among many Athenians. Nevertheless, “At the peak of the biggest rest, which partly extends into the biggest unrest, there is power, wealth, the arts, refinement, order, daring, and even the over-coming of poetic magnification by the sober quest for truth” (Strauss, 1989).

Thucydides seems to go to lengths to show that the spirit of sophistry and the spirit of philosophy that permeated Athens insinuated themselves into the foreign policy decisions of the Athenian elites. Moreover, he reveals that these precarious sources of thinking about world politics often had pernicious effects. The novel and innovative character of thinking led to novel and innovative policies that sometimes succeeded but were often immoderate and unrestrained from conventional limits. While such flexibility can in fact be useful from time to time, and while perhaps the *best* foreign policy decisions are informed by philosophy, there is a grave danger in sophisticated opinions that fall short of philosophy guiding foreign affairs. Using the examples of Pericles, Alcibiades, and Diodotus—three men in Thucydides’ work who clearly received something like a philosophic education—I will show that Thucydides harbors deep suspicions regarding the sophisticated education prevalent in the Athens of his time. The daring innovation of Athens is simultaneously the source of the glory of Athens and a cause of its ruin.

The Case of Pericles

Thucydides reports three speeches from Pericles, but by far the most revealing for our purposes is his funeral oration. Pericles’ famous funeral oration arguably captures better than any other speech the spirit of Athens, revealing the connection between Athens’ greatness and its own inevitable demise. While it might seem strange to turn to a eulogy to lay bare Pericles’ understanding of Athens, very little of Pericles’ funeral oration is in fact a eulogy proper: only one of the twelve chapters (2.42) is devoted to praising the dead. The central, and by far longest, part of the eulogy is in fact a praise of the Athenian regime (*politeia*, 2.37-41). Leading up to this praise of the Athenian regime, Pericles suggests that Athens is now at the height of its greatness. The previous generations, especially the Marathon generation, were impressive, but Pericles implies that the current ones who have added to the empire are no less impressive (2.36.3).

These impressive men contributed to the development of a regime that has become a pattern for others. The beauty or nobility of Athens is “inseparable from the regime and the way of life” (Orwin, 1994). As Pericles makes clear, Athens’ daring character, so central to its way of life, made its empire possible (2.43.1). The Athenian regime, instilling a daring spirit in its citizens, paved the road by which the city reached its empire (2.36.4). Indeed, the Athenian empire seems to be the natural consequence of the Athenian regime. This regime is devoted to freedom, and equal justice for all. It subscribes to unwritten laws, and has plenty of leisure activities. Athens is engaged in commerce and enjoys products from all over the known world into its harbors as a result. Athenians alone are educated to live as they please, unlike especially their rivals the Spartans. Freedom here again seems to be central: the Athenians alone are free to live as they please. The regime is permissive and freedom to live as one pleases entails a kind of liberation from ordinary or conventional restraints. As individuals, Athenians could pursue whatever they pleased, and the variety of tastes and pursuits contributed to the myriad beautiful things to be enjoyed in Athens. This private freedom, however, manifests itself publicly as well: Pericles views the Athenian decision to hold an empire as an undertaking freely chosen (2.40.4-5, Cf. 3.45.6). Pursuing empire seems to be the only way for Athens to express its freedom on a world level. Athens alone is not bound by the yoke of necessity like other cities; it can transgress the ordinary boundaries of propriety in regards to foreign affairs.

But there is a problem: Athens needs its citizens to support the empire, in many ways but most decisively with their lives if need be. Freedom to live as one pleases, however, means traditional restraints no longer bind. Morality, at least as it is conventionally or traditionally understood, no longer holds, and appeals to moral grounds will fall on the deaf ears of those who are inclined, thanks to the blessings of liberty, to live as they please. This is why Pericles

does not appeal to a sense of duty when he urges his fellow citizens to admire and imitate the fallen and to defend the city and its empire. While Pericles exhorts his audience to put their lives in the service of Athens, he can no longer do so by reference to a conception of morality that is bound up with self-sacrifice, for the permissiveness of living as one pleases has taken its toll on appeals to duty and sacrifice. "Tradition," Clifford Orwin says in this context, "had exhorted men to risk their loves for parents, wives, children, and ancestral gods" (Orwin, 1994). Since the Athenians are now free to live as they please, it is difficult if not impossible to make painful or even potentially life threatening demands of them. Their devotion to Athens must get them something; it must pay, so to speak. Accordingly, Pericles appeals to glory as the motivating factor for Athenians to support the empire, no longer being able to appeal to a sense of duty or justice that would weigh heavy on the conscience. The Athenians are now free to live as they please, thanks to their regime, and choosing to pursue glory would be a noble pursuit that citizens could freely choose. But it is not clear, indeed it is highly doubtful, that love of glory can be safely relied upon for defense of the city; appeals to duty or morality would seem to resonate more with those whose education has focused on doing what one ought to do. But, to repeat, such appeals are no longer persuasive to an Athenian audience thanks to the liberal, daring character of the way of life that the Athenian regime instills in its citizens. The spirit which leads to the founding or origin of Athenian Empire thus devours itself—freedom from traditional pursuits encourages Athens to pursue the perilous task of running empire but cannot adequately motivate its citizens to take up arms on its behalf.

Sophistic and philosophic education were increasingly in the air, further undercutting, and in many cases providing a theoretical justification for, abandoning traditional conceptions of justice and piety. Philosophy, too, is a kind of liberal or free endeavor. Its existence relies on a kind of freedom or permissiveness and its goal is further liberation, insofar as it aims to liberate one from conventional opinions, especially conventional opinions about justice, the law, piety, and the gods. It is therefore unsurprising that a funeral oration that praises philosophy is almost completely silent about the gods. Pericles clearly links the pursuit of glory liberated from traditional constraints with the prevailing educational currents. He proclaims of the Athenians: "We love the beautiful (*philokaloumen*) without extravagance and we love wisdom (or, we philosophize "*philosophoumen*") without softness." As Clifford Orwin notes, Pericles means by philosophizing without softness deliberating combined with daring (2.40.3; Orwin 1994). Pericles thus traces the cause of Athenian imperialism to their love of beauty combined with their love of wisdom (or literally, philosophy). Let us not forget, after all, that the philosopher Anaxagoras was Pericles' teacher, and, according to Plutarch, Anaxagoras more than anyone else was responsible for Pericles' majesty and the worthiness of his moral character (See Plutarch, *Life of Pericles* 4.4). So this chief architect or most enthusiastic proponent of the Athenian Empire was philosophically educated. The rise of sophistry and philosophy in Athens provided its outstanding youths with first-rate educations, educations that severed them from traditional thinking and allowed them to pursue novel, bold policies. The Athenian Empire itself was the boldest policy, and, in another speech, Pericles is willing to admit how bold it is: "Besides, to recede is no longer possible, if indeed any of us in the alarm of the moment has become enamored of the honesty of such an unambitious part. For what you hold is, to speak somewhat plainly, a tyranny; to take perhaps was unjust, but to let it go is unsafe" (2.63.3 with minor changes to the translation).

Pericles is often cited as an excellent example of a statesman who put his philosophic education to good use in service of Athens; he recognized, however, the tyrannical character of empire and nevertheless supported it. He clearly possessed excellent rhetorical skills, and the funeral oration is a prime example of rhetorical acuity. And Athens' empire flourished with Pericles at the helm. Once he was gone, however, his successors misguided foreign affairs and too often conducted affairs according to the whims of the multitude, producing a host of blunders. While Pericles put his philosophic education to use successfully guiding the empire, he was an extraordinary man. Thucydides, by emphasizing Pericles' unique and impressive character (2.65), leads one to question the desirability of the widespread diffusion of a liberal, philosophic education. The daring or innovative character, however, allowed these types of education to flourish, and sophists from all over Greece came to Athens.

Pericles was liberated him from concern with the ancestral gods as well as with the traditional concern with justice; this liberty allowed or enabled Pericles to put support for the empire on a new, non-traditional ground. Citizens are now to be devoted to the common good because they stand to gain something. But perhaps, "[Pericles] did not realize that the unjust understanding of the common good is bound to undermine dedication to the common good however understood. He had not given sufficient thought to the precarious character of the harmony between private interest and public interest; he had taken that harmony too much for granted" (Strauss, 1964). Strauss continues, "It is in the long run impossible to encourage the city's desire for 'having more' at the expense of other cities without encouraging the desire of the individual for 'having more' at the expense of his fellow citizens" (Strauss, 1964).

Post-Periclean Athens lacked proper devotion to the common good—Athens became increasingly unjust, if we can call devotion to the common good justice—and we can see Pericles’ contribution to that problem.

The Melian Dialogue

Although many Athenian speeches reflect a position that is liberated from ordinary constraints of morality, none is more famous or infamous than the so-called Melian Dialogue. This dialogue arguably contains the clearest expression of Athenian foreign policy as it came to be divorced from questions of morality—or to speak more clearly, justice. The Melian dialogue, above all, exemplifies the novel, daring character of Athenian foreign policy. And it also, above all, reveals itself to be the product of a kind of philosophic education. This dialogue contains the bold, sophistic assertion that “the strong do what they can and the weak suffer what they must” (5.89), so reminiscent of the rhetoricians and sophists from the Platonic dialogues like Thrasymachus, Protagoras, and Callicles.

Directly before recounting the Melian dialogue, Thucydides mentions that the Athenian general Alcibiades happens to have been sailing around the islands of the Athenian empire (5.84). Alcibiades is thus fresh in the reader’s mind as Thucydides turns to recount the conversation between the Athenian envoys and Melian representatives. Although one could by no means with certainty show it to be the case, the implication seems to be that Alcibiades, or at the very least someone very much like him, speaks for Athens in Melos. At the very least, one could say that Alcibiades “might have said what the Athenians on Melos say... while Pericles might never have said what the Athenians said on Melos” (Strauss, 1964).

The Melian dialogue stands apart from the other speeches in Thucydides’ work; all of the other speeches are rhetorical speeches: there are exhortatory speeches, warning speeches, display speeches, and apologies or defense speeches. The Melian dialogue stands alone as a dialogue, i.e., not a rhetorical speech at all in the precise sense, but rather a conversation. The unique character of this speech draws attention to its special character. It is, as it is commonly referred to, a dialogue, the form of speech preferred by Socrates. Thucydides has just mentioned that Alcibiades, widely known to be a student of Socrates, was in the area, so to speak. And while Plato presents Alcibiades as a rather poor student of Socrates, Xenophon shows that Alcibiades in fact picked up a great deal of Socrates’ manner (See Xenophon, *Memorabilia* 1.2.20-46).

The dialogue in Melos takes place under the following circumstances. Athens sailed to Melos, a small Spartan colony that had striven to remain neutral in the war, and asked the Melians to submit to their rule. The Athenians sent ambassadors to Melos to conduct negotiations, and the Melian few have urged that they meet privately, i.e., not in front of the people (5.84.3). Upon immediately meeting, the Athenians propose that the negotiations be conducted dialectically, criticizing long rhetorical speeches as seductive and deceitful (5.85). “The Athenian ambassadors talk as if they had been listening to Socrates’ censure of Protagoras or Gorgias” (Strauss, 1964). Conversation is safer and more conducive to teaching (5.86, 5.98) than rhetoric, and the Athenians set out to educate the Melians regarding their true interest. Regarding safety, the Athenians seem to presume that the Melian few are elites like themselves, free of the wrongheaded religious and moral concerns of the many. They think that they can focus solely on self-interest, relieved of the compulsion of making beautiful speeches about justice: they can get down to brass tacks, so to speak. The Athenians set out to show the Melians that it is in their interest to surrender; the only alternative is annihilation. One should note from the outset how coolly the Athenians begin: they urge surrender strictly from the vantage point of cold calculating interest.

The Melians, however, appeal immediately to equity or common decency (*epeikeia*), an oblique reference to justice; this is an early indication that the Athenian ambassadors’ assumption that the Melian few share their view of the world may have been wrong. But the Athenians will have none of this discussion of decency. They recognize in the Melian appeal to decency a latent appeal to justice. If the Melians are here to talk about anything other than their own safety, they assert, they can simply stop talking. They go on to say that they will not trouble the Melians by speaking beautifully about how they rule their empire with justice, for they know that justice is only in question among equals, as far as the world goes, “while the strong do what they can and the weak suffer what they must” (5.89). The Athenians in Melos boldly and coolly dismiss any pretense of justice; they have been loosed from these conventional demands and limits on proper behavior.

The Melians, compelled to dismiss claims of justice, now make an appeal to self-interest—though their argument is that it would be in Athens’ self-interest to pay attention to justice! They argue that one day the Athenian Empire will fall, and when it does, as it inevitably will, the Athenians would be well served if they could point to the fact that they observed justice when they ruled (5.90). In response, the Athenians say that they are not worried about

being conquered by another rival empire, for any other empire would recognize the law, so to speak, that the strong do what they can and the weak suffer what they must. Any future conqueror, therefore, would be free of the indignation that arises as a result of the belief in justice and would lack the motivation, then, to treat Athens in an overly harsh manner. The Athenians return to the task at hand, advising the Melians that it is in their interest to submit; they point out that the Melians will suffer the worst possible fate if they do not submit, revealing at the very least the Athenians' suspicion that there is no afterlife in which mortals would be rewarded or punished. This is our first indication, as it were, of the connection between the Athenians' immorality and their disbelief in the traditional ancestral gods.

It is these very gods to whom the Melians appeal. The Melians refuse to submit, and they give two reasons to hope that they will not suffer the fate that Athens has said in no uncertain terms they will suffer: the Spartans and the gods will come to their aid. These two factors, with a little luck, give the Melians reason to hope for a favorable outcome (5.102).

In response to these suggestions, the Athenian ambassadors make very bold speeches about the gods, speeches that betray beliefs that may have always been present, for example in Pericles' characterization of the Athenian regime and in his general silence regarding the divine in his funeral oration, but have never been articulated clearly and openly to one's enemies as it is now. The Athenians assert here, boldly, that the gods do not come to the aid of human beings; there are not providential gods. Not only do the gods not come to the aid of the just and not punish the unjust, it would be nonsensical for them to do so.

Human beings are compelled to behave as they do, hence there is no room for praise and blame regarding human action. We have seen the Athenians consistently make this argument or something like it among human beings, but we now see how it informs their theology.

In chapter 105, the Athenians offer a new theology, one that is based on the so-called Athenian Thesis. The Athenians flatly deny the notion that gods will punish them for their empire. Human beings rule when they can; this is a law of human nature (5.105.2). The strong will do what they will, the weak suffer what they must, and that is human nature. An account of the gods that says they punish human beings is false and bad—it makes no sense to punish human beings for being human beings—just as it would make no sense to punish a lion for eating a gazelle. Athens and Melos are the same in the eyes of gods—each does what it must. Human “injustice” if you will, is natural—it is innate. And a god that punished human beings for being what human beings are by nature would be unreasonable. Worse, it would show that the god does not understand human nature. The Athenian Thesis regarding the nonsensical character of praise and blame extends to the gods. It is very clear that the new Athenian morality, the new Athenian view of justice, is inseparable from abandoning the opinions and worship of the ancestral gods. The Athenians are now freed from any traditional understanding of justice and piety, and this enables them to guide their empire ruthlessly.

Ultimately, the Athenians are incapable of persuading the Melians to surrender. And, as Athens had predicted, the Melians fall at the hands of the Athenians. Neither the gods nor the Spartans came to the aid of the Melians. Of course, the actions that follow are absolutely horrific, and it seems clear that Thucydides intends to leave the reader with a very foul taste in his mouth. The Athenians kill all of the adult male inhabitants of Melos and enslave all of the women and children (5.116.3-4). While it is unknowable whether the ambassadors would have agreed with this course of action, it is undeniable that Thucydides leaves the impression that the speeches in Melos are clearly connected to the deed. Increasingly we see that the spirit of daring innovation conduces to transgressing ordinary boundaries of decency.

The Case of Theodotus

This is not to say that Thucydides thinks that all philosophical thought is pernicious, or that all innovative thought is harmful. Perhaps the clearest case of a philosophically inclined speaker putting his philosophic and rhetorical skills in the service of justice, advocating moderate measures, occurs earlier in the war in response to the revolt in Mytilene. Here, like in the case of Melos, the Athenians have decided to kill the entire adult male population of a city and to enslave the women and children. This time, however, there is a different outcome, and the favorable outcome seems to come about as a result of the philosophic speech of one Diodotus. The Mytilenians, a colony of the Athenian Empire, have revolted, unsuccessfully. The Athenians have successfully squashed the rebellion thanks in large part to the capitulation of the Mytilenian people. The Athenians are pretty upset because unlike their treatment of other colonies, they had actually treated Mytilene pretty well. They were not, strictly speaking, under the Empire

like others. Mytilene was an oligarchy, and it was the oligarchs who decided to revolt against Athens. But their attempt to revolt did not fare very well. The troops were not winning and they were running out of food. The Spartan commander, then, tried to arm the population in Mytilene against Athens. This sort of backfired—the many decided that it would be better to threaten to turn the city over to Athens unless the many gets the food. The Oligarchs decide to turn things over to the Athenians themselves. The Athenians are nonetheless upset with the Mytilenians, and have decided to execute all the adult males of the city and enslave all of the women and children. The subsequent debate we see in Athens is over punishing the People, the *Demos*. But it should be remembered that it was the oligarchs who started the revolt, and we never get an indication where the Many stood. The debate, then, is a reconsideration of the previous assembly's decision, as the Athenians have already decided to kill them and enslave the women and children, and they have already sent a ship to Mytilene to do the deed.

The Mytilenian Debate, it seems to me, may in fact be the most important debate in the entire work—more important than even the much more famous Melian Dialogue which we have already discussed. The reason is that this debate explores more deeply, I think, than any other series of speeches, the so-called Athenian Thesis that the Athenians first offer in Sparta. Let me briefly state that, as I understand it. In Sparta, when Corinth was trying to convince Sparta to go to war against Athens they charge that Athens behaves unjustly in conducting its empire. Against these charges, Athens says, more or less, “we do not apologize; we will not defend ourselves.” Why do these Athenians refrain from apologizing for their empire? Because there is no room for praise and blame in international politics, because there is no room for justice. Justice does not matter in international relations. All cities are compelled to pursue what they believe to be best in the light of three considerations: fear, honor, and self-interest (1.76.2). The compelled part clinches the argument for them: because they are compelled, praise and blame do not make sense. We only blame or praise actions that we think are the product of volitional action. The Athenians deny that cities are free from compulsion, hence, no praise or blame.

Now, Cleon comes forward to advise destroying the Mytilenians and standing by the decision they have already made. His speech appears cold, calculated, and ruthless. It would seem to represent the Athenian Thesis. It is possible, however, that Cleon's speech represents either a confused Athenian thesis, or it makes clear that the Athenian thesis is itself confused for the following reason. The Athenians in Sparta excuse their own actions by recourse to compulsion, as I just explained. But if that is correct, the same compulsion led the Mytilenians to do what they did. In other words, compulsion explains the actions of not only the strong but also the weak; it explains the actions of those who are seeking empire and those who are seeking to get out from under empire. The anger with which Cleon speaks about the Mytilenians, assuming it is not feigned, is misplaced. Cleon's speech is tinged with justice: How dare those Mytilenians revolt after we treated them so well? Whether we see it or not on first glance, his angry speech betrays a sense of justice.

Then we hear from Diodotus, who begins his speech with the words, “I do not blame.” In other words, he begins by recognizing a key part of the Athenian Thesis, that praise and blame have no place in the conduct of cities. Diodotus jettisons concerns for justice in the deliberation at hand, remarking that justice is appropriate in a court of law, but has no business in deciding matters between cities (3.44.4). Diodotus says the only consideration for Athens with respect to what to do about the Mytilenians is advantage: if it is advantageous to kill all the Mytilenians, fine, and if not, also fine. Who cares about justice? Diodotus thus presents himself as even less concerned with morality and justice than his rival Cleon. Interestingly, though, Athens' self-interest is served, he argues, by sparing the Mytilenians. He insists that in advocating for sparing the Mytilenians, he is only considering self-interest. Ultimately, Diodotus' speech carries the day—barely. By the slimmest of margins the Athenians vote to undo the previous decision and spare the Mytilenians. Diodotus, the hard-nosed advocate for profit and self-interest, has worked to spare the lives and freedom of an entire city.

There is a problem with this interpretation of the speech: Diodotus, who makes every effort to sound hard-nosed, admits that he is compelled to use a disguise in addressing the democratic assembly, that he will in fact lie (3.43.2). This raises an interesting question about prudence, truth-telling, demagoguery, and democracy. But I think there is a question of greater importance: What is Diodotus lying about? Toward the very end of his speech, he quickly mentions that butchering the people of Mytilene would be unjust (3.47.3). I contend that Diodotus' lie is in his denial that he is concerned with justice. He is in fact concerned with justice, but he recognizes the fragility of justice in the world (3.45, see also Orwin, 1997 and Strauss, 1964). Justice can only be achieved with great effort, and its achievement is always insecure.

Diodotus, like so many Athenian speakers before and after him, evinces a certain liberation from justice traditionally understood—he is willing to use duplicity, after all, in a democratic assembly—but he nonetheless shows himself to care deeply about justice. He opposes the butchering of the Mytilenian people, and despite his claims to oppose this

only on the grounds of self-interest, his open admission of the need to use disguises and deceit compels the reader to discern the character of his deceit and disguise. And his disguise is a brilliant one. Diodotus thus reveals himself to be a philosophically educated man who resorts to lies and trickery precisely because of the prevalence of demagoguery or rhetorical sophistication of less than competent, less than decent men. The only way for justice to succeed in circumstances such as these is by foul means.

Diodotus seems to believe more consistently in the Athenian Thesis than the other Athenians. He recognizes that cities and humans are compelled to behave the way they do, but does not use this insight to justify the satisfaction of the basest longings. He appears to put his knowledge in the service of justice. So while the vast majority of Athenian speakers' liberation from conventional morality leads them to the heights of immorality, Diodotus shows that this liberation from conventional morality can in fact be the best defense of morality in immoral times. One should recognize immediately that appeals to conventional morality would have fallen on deaf ears in the assembly. Moreover, the Diodotuses of the world are rare; in fact, there is doubt that this man even existed. In other words, it is difficult to find men still devoted to the common good who have been exposed to and even adopted a sophisticated education. More often than not, such an exposure will often be taken as an excuse to pursue one's own advantage. We have not answered the question of why men like Diodotus would be concerned with the common good once liberated from traditional reasons to be so devoted. Indeed, this is puzzling, so let us leave it at suggesting that there may be a rational account of why one should support the common good, one that relies on reason, not tradition.

CONCLUSIONS

I have tried to show that Athens' boldness or daring links the tyrannical empire to the peak of progress in the sciences and the arts, namely philosophy. In other words, the heights of injustice and philosophy coincide in Athens, but really it is no coincidence. Both are the products of a liberated bold, daring innovative spirit. Both within the city and without, Athens and Athenians were no longer devoted to the common good. "Post-Periclean Athens lacked the singular public-spiritedness which was the honor of Athens from the time of the Persian war until the age of Pericles" (Strauss, 1964). Athens was more public spirited in previous generations, and that public-spiritedness was linked to the divine. Athens' first successful repulsion of the Persians was at the Battle of Marathon, when "belief in the ancestral gods was in its greatest vigor" (Strauss, 1964). Even during the time of Pericles, the Athenians recognized that the virtue of these men was extraordinary (2.34.5). As civilization progresses in the arts and sciences, with concomitant growth or expansion in politics, this entails a rejection of the traditional or the ancestral. And one wonders whether devotion to the common good, to justice, can exist without the support offered by belief in the ancestral gods. "It is precisely for this reason that human beings seek support for right in the gods or that the question of right cannot be considered entirely apart from the question concerning the gods" (Strauss, 1964). Athens' daring character meant that Athens dared to reject the ancestral gods and the ancestral conception of justice. A city no longer devoted to piety and justice cannot expect to survive long, as it transgresses boundaries of ordinary decency that foreign cities still observe, and it will increasingly have a difficult time persuading its citizens to sacrifice on the city's behalf. The intellectual currents undermined meaningful devotion to the city and its empire—one supported it provided it made no serious demands.

It should be recalled that we have no external evidence that Diodotus exists, unlike the other persons represented in Thucydides' work. It is possible, then, that Diodotus, whose name means "Gift of God," stands in for Thucydides, representing the kind of speech Thucydides would have offered himself in the situation in which Diodotus found himself. That is, Thucydides, the author of the work, was surely familiar with what I have been calling the Athenian Thesis. Familiarity with or understanding even of this doctrine takes a mind that has been exposed to a kind of philosophic education. No doubt Thucydides himself was familiar with the intellectual currents of his day. Thucydides, after all, considers himself a wise man, and he believes that his book is a possession for all time. No doubt a philosophic education can provide us with a work comparable to Thucydides' great book, and it can also guide politics in a corrupt time in the manner of a Diodotus. That is, philosophy can be a good guide for theory and practice, but in limited cases. Through the deeds and speeches he presents, however, as well as his own concealed way of writing, Thucydides leaves us with the impression that the diffusion of a somehow philosophic education is pernicious, eroding the source of morality and the common good without which a city, or a civilization, inevitably declines. Perhaps sophistry is nothing other than diffused or vulgarized philosophy. I have not sufficiently distinguished the two in this paper.

Thucydides is much less certain than we are that there is a harmony between progress and peace. In fact, his work seems to evince doubt that what we call progress is possible at all. Quite to the contrary, history seems to be some-

what cyclical, not in the sense that there is a clear pattern, but in the sense that civilizations rise and fall. Reading Thucydides, we can have a greater appreciation for the fragility of civilization, and take a more cautious view of unlimited progress. We can come to see that science liberated from prudence can have potentially catastrophic consequences, and we can have a more realistic view regarding the prospects for enlightenment and peace.

REFERENCES

- Forde, Steven. (1991), "Thucydides on the Causes of Athenian Imperialism." *American Political Science Review*, Volume 80, No. 2. pp. 433-448.
- Hobbes, Thomas. (1991), *De Cive*. Indianapolis: Hackett Publishing.
- Horkheimer, Max and Theodor W. Adorno. (1969), *Dialectic of Enlightenment*. New York: Continuum Publishing.
- Kant, Immanuel. *Kant: Political Writings*. (1991), Ed. H.S. Reiss. Trans. H.B. Nisbett. Cambridge: Cambridge.
- Montesquieu. *The Spirit of the Laws*. (1989), Trans. Anne Cohler. Cambridge: Cambridge University Press.
- Orwin, Clifford. *The Humanity of Thucydides*. (1994), Princeton: Princeton University Press.
- Plutarch. (1916), *Lives*, Vol. VIII. Trans. B. Perrin. Cambridge: Harvard University Press.
- Schmitt, Carl. *The Concept of the Political*. (2007), Trans. George Schwab. Expanded Edition. Chicago: University of Chicago Press.
- Strauss, Leo. *City and Man*. (1964), Chicago: University of Chicago Press.
- Strauss, Leo. "Thucydides: The Meaning of Political History." (1989), *The Rebirth of Classical Political Rationalism*. Chicago: University of Chicago Press.
- Thucydides. *The Landmark Thucydides: A Comprehensive Guide to the Peloponnesian War*. (1996), Ed. Robert B. Strassler. Trans. Richard Crawley. New York: Free Press.
- Xenophon. *Memorabilia*. (1994), Trans. Amy L. Bonnette. Ithaca: Cornell University Press.
- . *The Landmark Xenophon's Hellenika*. (2009), New York: Random House.

The Roots of Self-Doubt (and Self-Loathing) in the West

Jonathan W. Pidluzny

*School of Public Affairs
Morehead State University
Morehead, KY 40351, USA*

ABSTRACT

Democratic self-doubt is manifest in many ways, some of them a threat to the perpetuation of the very political orders responsible for the exceptional appeal of life in Western societies. This paper begins by defining the phenomenon and proceeds to examine its roots. Five contributing factors are identified, all of them in some way a constituent element of the very way of life, our confidence in which they threaten to erode. The factors discussed are the following: first, the value neutral nature of experimental natural science and its effect on the liberal arts; second, European romanticism and its suffusion through the entire cultural life of the West; third, a *mania* for equality which, although a fundamental principle of healthy republicanism, can (when pushed to the extreme) erode citizens' pride in their own regime; fourth, materialism and material abundance, which distract citizens from the innately human longings and cut them off from participation in public life; fifth, egalitarianism desiccates the liberal arts—one of the greatest legacies of the West—and enervates the democrat's capacity to be moved by them. The paper ends by considering a more insidious form of self-hatred that emerged from the thought of Friedrich Nietzsche and Karl Marx.

Keywords: Western Civilization, Crisis of the West, Democratic Self Doubt, Alexis de Tocqueville, Egalitarianism

INTRODUCTION

The modern West suffers from a remarkable lack of self-confidence. Never in the annals of history have so many stable, prosperous and powerful states—each devoted to the protection of individual liberties, ruled according to the consent of the governed, internally stable with abundant economic opportunity for all, at peace with one another and content with established national boundaries—existed simultaneously. And yet, even as an almost unlimited number of emigrants and would-be emigrants for the privilege of building a life in the developed countries of the Western world, Westerners themselves doubt the superiority, even the choice-worthiness, of the very way of life they enjoy. This paper examines the crisis of the West. Why are Westerners, especially Western elites, so harshly self-critical? What are the dangers of this pervasive self-doubt? And what are its roots?

WESTERN CIVILIZATION

It is impossible to talk about Western self-doubt or a crisis in Western civilization without first defining two pregnant terms: Western and civilization. When we say “The West,” we collapse the terms together, not to indicate a geographical reference, but rather, to denote a confluence of a particular culture and community (Berman 1983, 2-

3). Let us say, preliminarily, that an individual's most deeply ingrained opinions, convictions and *mores*—his habits of heart and mind—are in some sense an artifact of, and constitute his membership within, a community that is generally broader than a single state or nation. As Samuel Huntington has put it, “a civilization is thus the highest cultural grouping of a people and the broadest level of a cultural identity people have short of that which distinguishes humans from other species” (Huntington 1993, 24). Angelo Codevilla gets at the definition from the angle of the constituent member: a civilization “is a package of habits and precepts that not only affects the way people live but to some extent defines what it means to be happy” (Codevilla, 50). Put another way, a civilization is the authoritative moral, ethical and aesthetic convictions its members share: communal dedication to notions of right and wrong, good and evil, noble and base, decent and obscene, permitted and impermissible, beautiful and ugly—judgments in the context of which an individual conceives of his happiness and the sort of life it is worth aspiring to. Civilizations thus set boundaries for statecraft: the kinds of political regimes and laws that will succeed in a given place and time must suit the spirit of the people, or at least not make spiritually implausible demands of them. (Matching regime type to civic character is especially important in polities that employ majoritarian institutions. If the people are, in effect, required to animate a institutions in a certain way to maintain the constitutional arrangement—pushing back against an encroaching executive, for instance, by electing guardians of their rights to the legislative branch—it is vital that the people understand their political role thus or they are unlikely to play it.)

Western civilization is more or less coterminous with a particular cluster of evolving moral and intellectual commitments that have laid the foundation for the way of life enjoyed by inhabitants of the North Atlantic states. In defining “Western,” Leo Strauss emphasizes our Greek and Judaic roots: “Western man became what he is, and is what he is, through the coming together of biblical faith and Greek thought” (Strauss 1967, 45). At his best, Western man believes that it is possible *to know* and worthwhile *to be good*. If a few perceive a tension between the Greek and biblical conceptions of wisdom and goodness—and even the bases upon which they are apprehended and realized—they nonetheless remain “open to both and willing to listen to each” (Strauss 1967, 46). Where a civilization remains moored to both propositions—the possibility of science (or philosophy) *and* the choice-worthiness of moral action—remarkable polities can be built. The best among them open a way to, and sustain, the heights of human achievement, while they simultaneously provide fertile ground for a broad range of decent and fulfilling lives. If such a state becomes powerful, it exercises its power with restraint and even benevolence; it might even sacrifice blood and treasure to build, or try to build, a better world order—to spread science and the blessings of liberty.

And yet, the pursuit of truth has always been dangerous. Vainglorious men who believe they have apprehended it, whether by reason or through their reading of revealed scripture, have not hesitated to try to instantiate it. The history of tyranny is almost identical to the history of government; the worst tyrannies are those which join force to ideology, with the most horrifying examples—*totalitarian* examples—closest in time to the present. In ushering in the era of modern science, the Enlightenment made human beings more powerful than ever before, but not, it appears, any better nor any wiser.

Sensing, perhaps, that there is something else about Western civilization than can, in principle, limit the scope of government, other commentators have emphasized the distinctiveness of Western law—a body of legal strictures separable from tribal custom or religious command. In this context, Berman notes the significance of the development of a “body of legal principles and procedures clearly differentiated from other processes of social organization” (Berman, 50). Long before Westphalia, the establishment in Western countries of professional courts and judiciaries, legislative bodies, legal professions, and legal literatures began to sunder the moral traditions and religious inheritance of a place from the laws enforced by the coercive power of the state. This embryonic separation of church and state—and the concomitant establishment of a private sphere—is a second important feature of Western civilization, something which distinguishes the West from both pre-modern societies as well as the Eastern (Talmudic and Islamic) alternatives.

The new conception of law permitted a sundering of the merely legal from that which a political community believes to be moral, right, and just. The communal acknowledgment that not every immoral act need be proscribed by law and punished by the community's coercive apparatus constitutes an important root of limited constitutional government, that hallmark of the *modern* West. After all, individual freedoms are only secure where government remains limited; there can be no durable check on governmental authority where the governed themselves believe *everything* is properly within the purview of the law and, therefore (today), the state.

Western civilization is roughly coterminous with Christendom because the ideas that lie at its foundation—the

legacies of Greece and Jerusalem, along with the example of Roman law (which could exist simultaneously with Christianity, even as its moral empire grew)—were spread by the Roman Empire to the edges of the known world. Rome's political primacy disintegrated, but its moral and intellectual legacy endured, in a variety of shifting forms thanks to Christianity and the Church. With the rebirth of Greek philosophy during the Renaissance and Enlightenment, the tension between reason and revelation could be restored in the West, its garden larger than ever before insofar as Christianity had homogenized *mores* through much of the known world.

On these bases, a bloc of countries emerged that have, for centuries, dominated the globe. At their best, the North Atlantic states establish and protect an enviable way of life, never before so durably, or so widely, instantiated. Within them, individual liberties are secure; science is pushed forward, more quickly than at any other point in man's history; the humanistic arts are cultivated, or at least celebrated, studied, and enjoyed; and the stable political environments they provide is conducive to the generation of levels of abundance and comfort unimaginable outside of the West. It is not for nothing that the political form generally associated with all of this, constitutional democracy, is today presumed to be the only legitimate political regime. It is, on the face of it, quite strange that we are not willing to do more to protect it. Stranger still, Western elites tend, today, not only to depreciate the West's tremendous accomplishments; they frequently blame the West for problems that exist, problems that have always existed, in distant hemispheres.

THE *CRISIS* OF THE WEST

It has become commonplace to speak of “the ‘crisis’ or ‘decline’ of the West” (Hail, 377). Crisis and decline are, of course, separable (if ultimately related) problems. The concern of this paper is not so much “decline”—in the sense of the diminution of power and influence of the Western states, their *exhaustion* in the face of a rising East—but a crisis emanating from Westerners' shaken confidence in (even their utter ambivalence to) the purposes or aims of their political regimes. As Strauss puts it, “The West was once certain of its purpose... it had a clear vision of its future as the future of mankind. We no longer have that certainty and that clarity” (Strauss 1964, 44).

A crisis of this nature is dangerous for a number of reasons, not least because it fuels and accelerates decline; timid Westerners, especially Western elites, will neglect to take steps, difficult though within their power, which might slow the West's decline. They would behave differently if they not doubt the righteousness of the Western way of life. There are other reasons for concern as well: prudent political judgment, rare in the most favorable situations, will be rarer still where policymakers do not know what they are aiming at. In matters of war and diplomacy, especially, it becomes difficult to know when and how to deploy the state's resources, military and otherwise, when no one is sure what the state ultimately stands for. When is war worth it? Or to bring the problem closer to home, at what point do our politicians decline to sacrifice further American liberties and privacy for one more increment of security from a potential terrorist threat? Security is *a* good. But it is not the highest good. When is the cost of peace or domestic security too high? There can be no principled answers to these questions where Americans themselves are unsure of their own regime's ends and principles, questions bound up with the country's civilizational inheritance. Strauss' cogent encapsulation is illustrative on this point too: “a society which was accustomed to understand itself in terms of a universal purpose cannot lose faith in that purpose without becoming completely bewildered” (Strauss 1964, 44).

It is appropriate to consider a distinctly American example of the transformation insofar as America holds a special space among Western nations by virtue of its preponderance of power and its uniquely liberal founding. The old clarity of purpose is beautifully captured in Thomas Jefferson's last letter, prepared on the occasion of the fiftieth anniversary of the Declaration of Independence. Writing to Roger Weightman, Jefferson's pride in the country's founding work of political poetry is inspiring.

May it be to the world, what I believe it will be, (to some parts sooner, to others later, but finally to all), the signal of arousing men to burst the chains under which monkish ignorance and superstition had persuaded them to bind themselves, and to assume the blessings and security of self-government. That form which we have substituted, restores the free right to the unbounded exercise of reason and freedom of opinion. All eyes are opened, or opening, to the rights of man (Jefferson, 1517).

It is hard to imagine any such celebration of American principles today, not least from American politicians. President Obama himself publicly dismissed the notion that the United States is exceptional in any determinative respect: “I believe in American exceptionalism, just as I suspect that the Brits believe in British exceptionalism and the Greeks believe in Greek exceptionalism.” His point: national attachments aside, there is nothing distinctively praiseworthy about America, about the civilizational legacy the U.S. inherited and had helped to perpetuate in the world. It is unusual for an American president to minimize the significance of his country’s history, in particular, the remarkable success of America’s republican experiment. As Alexander Hamilton hoped it would, the framing of this country proved an important point to the world—that government “by reflection and choice” *is* possible, that there *is* an alternative to government by those who dominate through force or by accident. As James Ceaser has suggested, the Founding generation believed America had a “mission” to help defend and disseminate polities devoted to the advancement of science and the protection of man’s natural rights in the modern world—both offspring of Western civilization—though they had not worked out a particular policy to achieve it (Ceaser 2012, 13-16).

The 2008 election is illustrative for a number of reasons. For the first time, Americans elected a truly cosmopolitan president, even, in important respects, a post-American president. Far from praising Americans’ bold sacrifices over the course of the twentieth century, which did more to spread freedom and democracy and any other force in human history, President Obama has instead personally apologized for its actions abroad—as though the United States has done more harm than good in the world. On the campaign trail, similarly, candidate Obama promised nothing less than a wholesale transformation of the country. All criticism of President Obama’s policy and rhetoric aside, however, his election is significant insofar as the tremendous public support he amassed reveals a pervasive wobbliness and confusion in the citizenry. A people with a healthy attachment to their guiding ideals—a confidence that they are just and good—would not be moved to support a president promising to eschew them in favor of a grandiose transformation of society and state.

To dwell on the foibles of the current president is not to imply that this problem falls along the ordinary Republican-Democrat divide. That President Bush came across as a stronger partisan of America and the West does not mean he understood the critical features of Western civilization terribly well. If anything, his enthusiasm to disseminate constitutional democracy to the far reaches of the Middle East reveals a misunderstanding of its character and prerequisites (cf. Pidluzny 2012). It is possible, after all, to sense or intuit the value of a thing without quite understanding it. Enthusiasm itself is usually immature, even childish. It is possible to arrive at a mature appreciation for the virtues of Western civilization without feeling compelled to devote the resources of one’s state to spreading them willy-nilly. Take John Quincy Adams’ proud endorsement of liberal democracy, which he coupled with a prudent foreign policy realism.

America, with the same voice which spoke herself into existence as a nation, proclaimed to mankind the inextinguishable rights of human nature, and the only lawful foundations of government. America, in the assembly of nations... has uniformly spoken among them, though often to heedless and often to disdainful ears, the language of equal liberty, of equal justice, and of equal rights. She has, in the lapse of nearly half a century, without a single exception, respected the independence of other nations while asserting and maintaining her own. She has abstained from interference in the concerns of others, even when conflict has been for principles to which she clings, as to the last vital drop that visits the heart. Wherever the standard of freedom and Independence has been or shall be unfurled, there will her heart, her benedictions and her prayers be. But she goes not abroad, in search of monsters to destroy.

It is hard to imagine a similar union of prudence and principle today. One of the main reasons for this is the state of civic and liberal education. One need not look any further than the state of the country’s leading colleges and universities today for confirmation that the West is indeed suffering from a crisis of confidence. The priorities of higher education are one of the best indicators of the health and direction of a civilization not simply because the academy molds a polity’s future leaders. More than this, educational institutions are paramount because they have the power to perpetuate or transform entire cultures—by educating teachers, journalists, government administrators, novelists and filmmakers. The academy effectively *shapes* the character and concerns of subsequent generations.

A recent report by The National Association of Scholars, entitled *The Vanishing West*, is thus an urgent cause for alarm. The study of 50 leading colleges and universities chronicles a “decline and near extinction of the Western

Civilization history survey course” (iv). In 1964, every one of the elite curricula investigated “made students familiar with Western civilization in some form” (whether by requiring classes in Western civilization, or through a general education or great books curriculum; 1). In 2010, not one of the same fifty institutions required students to complete a class in Western civilization in order to graduate; only 32% made one available as an optional part of the general education curriculum. The study found, furthermore, that even history majors are seldom required to study Western civilization today.

What explains this startling, and counterintuitive, neglect? Why are colleges and universities, which have done so much to build and sustain Western civilization, so apparently unconcerned with teaching it today? In what follows, I argue that Western self-doubt has five distinct roots.

MODERN SCIENCE

The depreciation of the liberal arts, of genuinely liberal education, has coincided with the ascendancy of the natural sciences. The ascendancy of the latter is one of the reasons for the West’s current crisis. Whereas a healthy liberal arts program introduces students to alternative ways of living and organizing political communities through the world’s great works of literature, philosophy, religion and art—all with the aim of evaluating them, of determining which best instantiate justice—the sciences treat a more limited subject matter with clinical detachment. A civilization that loses sight of its place in the annals of history—in favor of petri dishes and electron microscopes—is destined to lose sight of its distinguishing features and their value. Somewhat perversely, modern science has undermined the West’s dedication to science.

One important distinguishing feature of Western civilization is its consciousness, dating to ancient Greece, of two important dichotomies: the distinction between nature and convention, and the distinction between knowledge and opinion. The twin recognition that some beliefs are true and others mistaken (however deeply held, however rooted in tradition), and that the touchstone by which knowledge is distinguished from opinion is *nature* (understood broadly as the *cosmos* or ruling order), constitute the very root of science. Indeed, science would be rare, not to say impossible, in a political community that considers established beliefs to be inviolable—true for being old or inherited—or one which denied the existence of a reliable standard by which to evaluate a polity’s opinions. To do their work, scientists and philosophers *must be* cognizant that most opinions are conventional (the product of human art, culture, history, etc.), and they must have confidence in their rational faculties; they must believe themselves intellectually competent to separate truth from mistaken and incomplete opinion. Thus, they (slowly) add to what is actually known—either by an individual knowledge-seeker, or collectively by a collective group of inquirers.

Until the seventeenth century, the terms science and philosophy were interchangeable: both terms described the endeavor to *know* the natural order. The hard distinction we draw between natural science and philosophy today is one reason for the West’s shaken confidence (even while simultaneously responsible for its tremendous economic and scientific prowess). Put simply, experimental science has, by its technical success, led us to doubt the possibility of solid answers to normative questions. Its new methods were first proposed as a coherent alternative to the classical (contemplative) approach by philosophers with an agenda in the sixteenth and seventeenth centuries. Francis Bacon and Rene Descartes sought nothing less than to render man “master and possessor of nature” with the ultimate goal of “[relieving] man’s estate” by their proposed reorientation.

They were successful. And there is certainly a difference between science and philosophy today. The species of knowledge ascertained by the investigation of the permanent natural order by way of repeatable and falsifiable experiments in a laboratory *is* distinguishable from the kind of understanding Socrates and the great moral philosophers could claim. It is even natural to suspect that the findings of the natural sciences are *firmer* than the conclusions of social scientists. This aspect of the difference is clear: the knowledge yielded by physics permits engineers to build airplanes and nuclear reactors; political scientists have trouble forecasting an election result. The hard line has been drawn between hard science and soft science in recent centuries, with the more useful sciences occupying the place of pride in the eyes of Westerners, is thus hardly a surprise. We benefit from the advancement of learning in innumerable obvious ways. The philosophy of science implicitly teaches only questions amenable to investigation by the positivistic methods of the natural science are answerable with any degree of scientific certainty.

This presumption—that only the hard sciences can discern the hard truth—is politically pernicious, however, no

matter how glorious its technical fruit. Where people doubt that knowledge of good and evil, noble and base, decent and obscene, beautiful and ugly is even possible, it becomes difficult to defend the choice-worthiness of one's way of life, his regime, even his civilization. Indeed, what reason is there to prefer freedom to despotism if one presumes social scientists cannot discover *truths* about justice and the principles of political legitimacy? What reason is there to prefer a political community that permits scientific inquiry? Natural science is silent on value-laden political questions because its methods are inapplicable to subject matter (human action and their motivating beliefs) that does not behave in predictable, unchanging and law-governed ways. The rigid assumption that the methods of natural science are nonetheless the *only* way to *know* thus abets the emergence of "value relativism"—the lazy presumption that opinions on normative questions *cannot* be refined but are irredeemably subjective, the product of this or that time, culture, or socio-economic situation, or personality (Malcolmson, Myers & O'Connell 1996, 20; cf. Lange 2011, 121).

The separation of fact and value, and its implicit rejection that knowledge in the sphere of "value judgments" is possible, has a dangerous twin. To doubt there are compelling *reasons* to live a good or just life, to doubt the possibility of *proving* the superiority of one's way of life or regime—(say) limited government (built on egalitarian principles and devoted to the protection of individual rights)—inevitably spawns a pervasive wishy-washiness in the face of the most important human concerns. It can even lead to outright nihilism. For if one presumes that there are established facts in the arena of the hard sciences—but that when one moves from the behavior of electrons to the behavior of human beings one enters a subjective realm of values, where it is possible to opine, to emote, to will, to feel, *but never to know*—it is impossible to justify the choices one makes. A paralyzing helplessness, encapsulated in a question, follows fast upon that realization: *Why bother?* What a schizophrenic endeavor!—to seek to live well at once believing it is impossible to justify one's way of life. Ambivalence is the natural result. But ambivalence to what have long been believed to be the perennial human questions is dangerous—not only personally, but from the perspective of the regime as well. If a civilization is properly defined as the highest cultural grouping—participation in which requires that one share the moral, ethical and aesthetic judgments that separate one's own from alternative civilizations—the relativist's outlook would seem to make genuine participation in one's culture or civilization impossible.

EUROPEAN ROMANTICISM

It is commonly held that the modern West—a strange fusion of Greek, Jerusalem and Rome—emerged in the latter part of the twentieth century as a result of momentous revolutions in thought aimed squarely at the intellectual dominance of Christianity and the Church. And yet, the opinions put forth by Enlightenment and Reformation thinkers—though sufficient to depoliticize Christianity—did not alone suffice to build the social consciousness of the contemporary West. In particular, the Enlightenment and the Reformation do not account for the West's ethic of toleration. After all, the presumption that certainty on the most difficult questions is possible—whether it is achieved by reason or revelation—leads inexorably to attempts by the ambitious to instantiate those truths by force.

Enlightenment thinkers believed all of nature—its natural laws and its moral laws—could be discovered by rational investigation, that human beings were sufficiently endowed to discover those laws without divine help, and that all of nature's truths were consistent. What they did not sufficiently appreciate is how little room they had left for disagreement among those who were absolutely certain their opinions were the right ones. Or put another way, the assumption that an ordered nature can be known entirely by unaided human reason opens the way to fanatical attachments to *rational* accounts of what is best for mankind, while at an instant undermining centuries of tradition—the moral convictions built over time on the authority of revealed teachings—that might otherwise have stood in the way of this new species of certitude.

The aspiration to certainty on the most important moral and political questions, even when it is sought by reason, can inspire utopian political treatises and dangerous political ambition—quite as well, in fact, as professed religious certainty can lead to fanaticism when political authority is intermingled with it. Enlightenment ideas helped to fuel the French Revolution and its terrors. The terrible totalitarian movements of the twentieth century in the West—National Socialism and Soviet Communism—were, likewise, inspired by atheistic ideologies which sought to engineer totally new societies on the basis of theories and the theorists who believed they had discovered truths worth instantiating.

To move from the presumption that there is single unalterable truth according to which all of society should be organized, whether derived from revelation or discovered by human reason, to the acceptance of the possibility that a variety of ways of living may be worthwhile, and that the political arrangement should protect individuals' freedom to pursue happiness as they themselves conceive of it within very generous boundaries protected by law, a further revolution in social consciousness was necessary. Isaiah Berlin calls European Romanticism "the greatest single shift in the consciousness of the West that has occurred" (Berlin 2001, 1). Its new emphasis on sentimentality and emotion, the charms of aesthetic life, love, all the things men and women feel—in a word, the inner (and not always fully rational) life of the individual—represents a dramatic step away from the hitherto accepted assumption that public life should, so far as possible, be organized according to a single coherent system of morality. This privileging of the sentimental side of man, so long suppressed by authoritative conceptions of good and evil, noble and base, decent and obscene, would progressively replace what had for a long time been the North Star of European civilization: the search for truth, rational or religious, and the belief that the community should be organized according to it.

European Romanticism—ironically made possible by the Enlightenment's assault on authority—amplified it from a new, and arguably more powerful, direction. If there is not a single truth to know, other excellences can be elevated: creativity, genius, beauty, passion, the sublime, depth of feeling, etc. These were depicted beautifully, movingly, in the poetry, art and literature of the Romantic era. Thus, from the middle of the Eighteenth Century, worthy action—the meaning of "noble"—comes increasingly to be interpreted in light of a loose constellation of often incompatible standards, and less and less, measured against a man's fidelity to truth or his righteousness (Berlin 2001, 10). Suddenly grand action—(insert your own definition)—meant more than the pursuit of wisdom or virtue, both of these things, irreparably rooted in thought. It became possible, perhaps for the first time, to admire another's dedication to ideals one did not share, and in some cases, ideals one emphatically rejected. Action and feeling were glorified; contemplation and the pursuit of truth, meanwhile, were depreciated.

The resultant toleration of significant disagreements on the most fundamental questions of justice and morality is a truly impressive achievement from the perspective of limited and liberal government. To accept with indifference that fellow citizens, family members, clansmen, etc.—people to whom one is united by powerful bonds of affection—will disagree profoundly on the most fundamental questions, and should nonetheless be tolerated and even embraced, is not in any way an automatic disposition. Where individuals care deeply for one another, one would rather expect great efforts to be made to save those who stray from the right path. It is not unusual to take an active interest in the morality or propriety of the decisions the people we care about make; it is far more unusual to expect that those who care most for us will muster a well-meaning indifference in the face of our choices when they believe them to be profoundly misguided.

This new ethic helped make liberal democracy possible by making citizenries in the West *gentle*—restrained and relatively easygoing in their use of their considerable freedoms in the private sphere. The birth of tolerance has an unhappy twin, however. The prior disintegration of authoritative moral guidance (even the belief in the possibility thereof) made possible, as the logic of Romantic ideas unfolded into value relativism, a dramatic radicalization of individualism and redefinition of liberty. Romanticism, by progressively divorcing the definition of worthy action from any single conception of virtue or excellence apprehended by the intellect, appears in practice to legitimate the freer and freer and freer pursuit of happiness. When the American Founders spoke of the pursuit of happiness, happiness bore an essential connection to virtue understood along classical or Christian lines. By disconnecting the idea of happiness from notions of excellence (whether intellectual or moral), which depend for their continuation across generations on confident espousal by the regime's respected voices, liberty too came to take on a new, entire self-indulgent, meaning.

Originally, liberty meant freedom from interference by the state, and implied a right to help guide the polity by patriotic participation in government; most important, perhaps, for the American founders it meant the freedom to pursue one's own salvation free of government interference. As Romantic ideas took hold, it came to mean the freedom to freely will and live a law apprehended by reason (or some other way) for oneself. For progressive political reformers, this justified active state support for individuals in the name of helping them to realize their potentials. In the 1960s and 1970s, however, liberty took on an altogether new meaning again: the unlimited and unbridled license to indulge whatever passions, whatever inclinations, whatever whims, one prefers, with the expectation that no one will judge the resultant behavior according to anything resembling an exacting standard—moral, aesthetic or otherwise.

In sum, then, European Romanticism struck a great blow for liberty and tolerance in its infancy but grew corrosive with age. The movement—built on sentiment, sensuality and willfulness—exerts two separate pernicious effects, each of which strike important pillars of Western civilization. In the first place, the romantic ethic strengthened

value relativism and its denial that knowledge is possible in the realm of values. Secondly, it transformed Westerners' conception of one of the very things—liberty—our civilization exists to protect. Third, in glorifying man's sensual nature and depreciating the West's Christian inheritance, Romanticism levied a powerful challenge to the notion that it is worthwhile to be good, that goodness and virtue have a direct connection to human happiness. It is much harder to defend the righteousness of a civilization where its very *raison d'être* are understood in such a way as to condone and legitimate the antics of a Paris Hilton, a Lady Gaga or a Miley Cyrus.

DEMOCRATIC EGALITARIANISM AND WITHDRAWAL

The immortal words of the Declaration of Independence, "all men are created equal and endowed with certain unalienable rights" announced America's dedication to principles discovered by the Enlightenment's leading political thinkers. The idea spread throughout the West like wildfire, perhaps because Christianity, which taught that all men are equal in the eyes of God whatever one's earthly station, make it spiritually plausible to believe in the equal dignity of all human beings. Once again, however, a constituent element of Western civilization—the very keystone of the exceptional polities built in the modern era—ultimately contributes to its unraveling.

No thinker better understands the double-edged nature of egalitarianism and democracy than Alexis de Tocqueville. To be sure, Tocqueville appreciated the attendant benefits: class mobility is a feature of a just regime; family life is sweeter where inheritance and hierarchy do not interfere with parental and filial love; religious worship can bring deep and lasting satisfaction when it is divorced from political ambition; in unleashing the productive energy of the entire citizenry, democratic states have the potential to become more powerful than their aristocratic ancestors. These are all happy consequences of the emergence of an egalitarian social state, which Tocqueville believed to be inevitable—at least in the Christian world. Nonetheless, it is the threat equality represents—to the potential happiness and greatness of democratic individuals and nations—which occupies the vast majority of Tocqueville's attention. He believed the democratic social state to be inevitable; but he did not believe healthy republican governments would necessarily or inevitably issue from it. Egalitarianism can just as easily prepare the way for insidious new forms of tyranny and despotism. Tocqueville wrote to teach democrats how to avoid the perils of egalitarianism in the hope of rendering democracy "profitable to mankind." Nor are Tocqueville's concerns confined to the political dangers of egalitarianism. He also feared that a *mania* for equality might transform human longing altogether, robbing democrats of "several attributes of their humanity."

Tocqueville's key insight is that equality of conditions transforms ordinary selfishness into what he calls "individualism." Democratic citizens are free and equal, but they are politically impotent: "As conditions are equalized, one finds a greater number of individuals who, not being wealthy enough or powerful enough to exert great influence over the fates of those like them, have nevertheless acquired or preserved enough enlightenment and goods to be able to be self-sufficient" (Tocqueville 2002, 483-4). Private life thus comes to occupy the whole of an individual's attention. Sensing that he will never distinguish himself among the entire body of the citizenry, the democrat seeks to isolate himself, retreating into coteries (small groups) of like-minded individuals the better to pursue his narrow interests and inclinations. As Tocqueville explains, "Only with difficulty does one draw a man out of himself to interest him in the destiny of the whole state, because he understands poorly the influence that the destiny of the state can exert on his lot" (Tocqueville, 487). If the country is at war, or in fiscal crisis, he hardly notices so distant do the affairs of state seem to him to be from his own. On the one hand, America is exceptional in part for Americans' active participation in their government; Tocqueville calls the township the "seed" of free institutions. On the other hand, however, by way of a pervasive individualism, egalitarianism destroys citizenship: "there is no need to tear from such citizens the rights they possess; they willingly allow them to escape. The exercise of their political duties appears to them a distressing contretemps that distracts them from their industry" (Tocqueville, 515).

This retreat from public life is pernicious for a number of reasons, not least because it abets the rise of Soft Despotism. For present purposes, its effect on the citizen's self-understanding, the democratic psychology, is more important still. Individualism and withdrawal erodes citizenship by weakening citizens' sense of belonging, their sense of duty to the broader political community. Such a man is hardly even a citizen anymore—highly unlikely to sacrifice narrow private interest from love of country, not even when the public good urgently requires it. This effectively transforms democratic self-government into a kind of sum of narrow preferences will of the majority. What is more, for *feeling* less attachment to his polity, less bound to for his identity, this new kind of citizen-subject

feels less and less impetus to learn about his country, its principles and purposes, less and less reason to celebrate its legacy. Withdrawal from country—in America’s case especially—means disengagement from the civilizational legacy as well. Is it any surprise that young men, lost in their video games and Twitter feeds, take next to no interest in the historical and philosophical roots of Western civilization? Where democrats are drawn out of themselves, they join with those like them—who share and interest or a passion—and derive their identity from the associations nearest to them rather than the political community as a whole.

DEMOCRATIC MATERIALISM AND THE DESICCATION OF HUMAN LONGING

One’s sense of his own impotence, joined to habitual indulgence in narrowly private pursuits, has a further damaging effect: it distracts individuals from worthwhile endeavors and wraps them up entirely in making money. For Tocqueville, greatness generally *requires* leisure and education, which is to say greatness requires inherited wealth. Where moneymaking is an object of scorn (by those born into it), the ardor to be admired—pride—spurs the gentleman to develop a reputation for excellences that extend well beyond the marketplace. Aristocratic brilliance flows from the way unequal social conditions mold and shape human longing. By channeling the energy of men and women toward noble, worthwhile pursuits—achievement in politics, literature, art, science and philosophy, some other great contribution to country—*aristocratic* inequality inspires high culture and political greatness. This is the reason Tocqueville laments its passing. Democratic egalitarianism, by contrast, is obsessively pecuniary in focus and orientation, and almost necessarily so; the many will always have to work—to survive, to live. That work inevitably dominates life where there is unlimited hope of improving one’s station is only a part of the problem. Some of that work (the tedious task of placing pinheads on pins is Tocqueville’s example) is so narrow as to exert a dehumanizing effect: it shrinks man’s conception of himself (Tocqueville, 530-2). Others devote every last bit of energy to ameliorating their fortune without improving themselves or achieving anything truly worthwhile. Democracy channels energy bit of human energy in the same direction and, in so doing, raises a herd of hardworking but otherwise mediocre businessmen: “What above all turns men of democracies from great ambition is not the smallness of their fortune, but the violent effort they make every day to better it. They compel the soul to employ all its strength in doing mediocre things” (Tocqueville, 601).

When excellence at moneymaking becomes honorable, the spur that drives capacious individuals ceases to raise their aspirations and begins, instead, to desiccate them. Human passion can itself be “appeased and debased” (Tocqueville, 604). Tocqueville’s prose betrays a particularly poignant dejection when he laments the soul-deadening effect of this reorientation of life toward money and utility: “What I reproach equality for is not that it carries men away in the pursuit of forbidden enjoyments; it is for absorbing them entirely in the search for permitted enjoyments...” Materialism, Tocqueville goes on to explain turns democrats to the pursuit of “material enjoyments... with an insane ardor” and in the process “softens [souls] and in the end quietly loosens all their tensions” (Tocqueville, 509, 519). The tensions—between duty and inclination, present situation and imagined glories, love of country and religious devotion, family and high romantic attachment—can drive human beings to exert energy, sometimes tremendous energy, in an effort find peace in resolution. Democratic centuries loosen those tensions in two separate ways. Where the marketplace rules, money-loving men paint easier horizons, which give safe—read: mundane—shape and tenor to the longings and aspirations that motivate members of the community. At the same time, distractions are multiplied and brought within reach of everyone: either by the furious busyness of the marketplace or the distractions that come from the successes it lays open, democrats grow ambivalent to the longings of soul that, in differentiating man from beast, once provoked him to strive to perfect the higher features of his nature. We live in an environment that all but *compels* “the soul to employ all its strength in doing mediocre things” (Tocqueville, 601). One of the unforgettable images of Allan Bloom’s *Closing of the American Mind* is his arresting portrait of pubescent male youth: utterly absorbed in a private mood cocoon, withdrawn from the world, completely content to strive for nothing so long as he has his Walkman (Bloom, 77-81). It is not for nothing that Bloom wanted to call his siren work *Souls Without Longing*.

Democratic man, as Tocqueville found in embryonic form when he visited America, is a pitiable and pathetic creature—blind to the higher longings that might bring deep and lasting fulfillment, it is as though he “despises himself to the point that he believes himself made only to taste vulgar pleasures” (Tocqueville, 604). At the extreme, he becomes viciously self-indulgent insofar as his ubiquitous focus on material concerns tends to morph

into a dangerous theoretical materialism. Thus distracted from his mortality, from all of higher longing, bourgeois man is apt to seize upon the mechanistic understanding of the cosmos put forth by modern science to soul-desiccating effect. It is a short step from denying the possibility of God and an immortal soul to thoroughgoing nihilism. Every consistent nihilist, meanwhile, lives life as a hedonist, unable to find meaning in the world outside of himself. The democrat's self-indulgent retreat saps life of its deepest potential satisfactions; it is hard to feel fulfilled or happy lacking something larger than oneself to which to contribute. Hedonists are politically pernicious, too—unlikely to sacrifice for the greater good, unmoved by its claims to noble purpose, unconcerned with the endurance of their political community past their own time in it. The community feels the damage keenly; this new kind of man is ignorant, even contemptuous, of the legacy he has inherited and his role in helping to continue it.

THE EROSION OF THE LIBERAL ARTS

The liberal arts—poetry, philosophy and history—stand proud, as prospective remedies for the ills thrust upon us by egalitarianism and capitalism. They have always reminded the best among those who engage in them of the brilliant heights mankind *can* reach, inspiring some to *want* to equal and exceed the pinnacles of human achievement, thought and experience. Western civilization's openness to the great works from every period—and, indeed, from every civilization—is one of its distinguishing features. Tocqueville expresses a hope that teaching the great works of antiquity, in particular, can “counterbalance our particular defects” and “prop us up on the side where we lean” (Tocqueville, 452).

Tocqueville also foresaw that the liberal arts themselves, and the way they are engaged, were likely to be affected by democratic *mores*, however. On the one hand, democrats are apt to emphasize the useful and the practical, depreciating the purely theoretical and the beautiful (Tocqueville, 439). Artists themselves are forced to please an ever larger and ill-educated market, bustling obsessed with practical and material concerns. This inevitably debases the “art” they produce: “in aristocratic centuries, enjoyments of the mind are particularly demanded of the sciences; in democratic, those of the body” (Tocqueville, 437, 448). What is more, the impotent, equal and lonely democrat, alone among a crowd of similarly unimpressive individuals, naturally doubts the possibility of great human achievement simply. Tocqueville keenly observes that there is “nothing so small, so dull, so filled with miserable interests, in a word, so antipoetic, as the life of a man in the United States” (Tocqueville, 461). Thus, there is no subject matter for the poet; “equality... dries up most of the old sources of poetry” (Tocqueville, 460).

Egalitarianism and materialism also transforms democrats' manner of engagement with the works of antiquity. Who imagines that the ambition of an Alexander or a Caesar is possible today? This naturally gives a fatalistic character to democratic history (Tocqueville, 471). Individuals do not bend the arc of history; democrats perceive powerful historical forces and trends, beyond the control of any individuals or society, sweeping entire populations along in spite of themselves. The democrat in effect loses sight of his place in history, his country's place in history, even the distinguishing features of his civilization. A population that detaches itself from the great achievements of the Western tradition is unlikely to trumpet and celebrate that tradition. It is as though he believes himself to be beneath the dignity of such lofty concerns. As Irving Kristol has observed, the entire political community is lowered where the citizenry does not believe itself worthy of its rich inheritance or meaningful collective pursuits:

a society needs more than sensible men and women if it is to prosper: It needs the energies of the creative imagination as expressed in religion and the arts. It is crucial to the lives of all our citizens, as it is to all human being at all times, that they encounter a world that possesses a transcendent meaning, a world in which the human experience makes sense. Nothing is more dehumanizing, more certain to generate a crisis, than to experience one's life as a meaningless event in a meaningless world (Kristol, 134).

CONCLUSION

Unfortunately, the West's elite universities are unlikely to combat the crisis of meaning so apparent today. They have not only abandoned their responsibility to teach future generations about the West, its merit, and the

importance of its endurance. Far from buttressing and improving Western civilization, they contribute mightily to its further erosion today. In fact, the most impactful geniuses of the late nineteenth century, their influence still felt and amplified by the very universities their thought quickly and thoroughly corrupted, sought deliberately to undermine the foundations of Western civilization today.

The barbarians within are often more dangerous than the barbarians at the gates. For us, the seductive ideas of Karl Marx and Friedrich Nietzsche are the biggest danger, turning democratic self-doubt into a pervasive kind of self-loathing. Nietzsche's proud depreciation of the possibility of knowledge—his affirmation of a kind of blind, animalistic, will to power—provided a powerful theoretical reason to dismiss the principles which redeem and justify the West. His claim that the West's foundational assumptions are in fact a source of weakness—likely to usher in the reign of the last man—served only to reinforce the dangerous relativistic assumptions promulgated by modern science. Equally pernicious, an important tenet of Marxist ideology, which sprung from Romanticism to criticize the failings of capitalism, is alive and well today. While the failed Soviet experiment has cured Westerners of their desire for communal ownership of the means of production, the underlying assumption of Marxism, that capitalism (not human nature) is the cause of human greed, continues to fuel a pervasive self-hatred among many Western elites. This erroneous assumption is Pavlovian truth for Western elites today, and it leads them to blame their own regimes for the injustice of the world—as though *we are* the cause of human greed, oppression and war, both at home and abroad.

If the liberal arts exacerbate the problem today, they may also be the solution. In spite of all of the societal forces that sap lofty ambition and eradicate refinement of taste, Tocqueville is unusually optimistic through his discussion of the liberal arts. He is confident that great minds will nonetheless emerge, from time to time, even in spite of the thoroughly enervating effect of egalitarian *mores* (Tocqueville, 429, 437). He seems to hope these rare lofty geniuses—lovers of truth and greatness—will work to refine and ennoble their societies, helping to pull against the desiccating tendencies of egalitarianism. He hopes, in a word, that a kind of natural aristocracy—proud of its heritage and public spirited—will be cultivated somehow, even if the spirit of the times altogether rejects the notion. Culture can change quickly in democratic centuries thanks to its destruction of hierarchy and tradition. It will be left to such men and women, “perfect master[s] of ancient literature and wholly steeped in its spirit,” to save the West from the barbarians within (Tocqueville, 452).

REFERENCES

- Berlin, Isaiah (2001). *The Roots of Romanticism*, Henry Hardy, ed. New Jersey: Princeton University Press.
- Berman, Harold (1983). *Law and Revolution: The Formation of the Western Legal Tradition*, Cambridge: Harvard University Press.
- Bloom, Allan (1987). *The Closing of the American Mind*. New York: Simon & Schuster.
- Brague, Remi (2010). “Liberty and the Judaeo-Christian Inheritance,” in Roger Scruton, ed., *Liberty and Civilization*. New York: Encounter Books.
- Ceaser, James W. (2012). “The Origins and Character of American Exceptionalism, *American Political Thought* 1:1 (Spring).
- Codevilla, Angelo (2009). *The Character of Nations*. New York: Basic Books.
- Hail, Michael (2013), “The Philosophic Foundations of the Crisis of Western Civilization,” in *Advances in Cross-Cultural Activities, Part II*, Denise M. Nicholson and Dylan D. Schmorow, editors. New York: Taylor & Francis, CRC Press.
- Hallaq, Wael (2005). *The Origins and Evolution of Islamic law*. New York: Cambridge University Press.
- Huntington, Samuel (1991). *The Third Wave: Democratization in the Late Twentieth Century*. Norman: University of Oklahoma Press.
- Huntington, Samuel (1993). “The Clash of Civilizations?” *Foreign Affairs* (Summer, 1993).
- Jefferson, Thomas (1984). *Writings*, Merrill Peterson, ed. New York, The Library of America.
- Lange, Stephen (2011), “Relativism and its Consequences for Western Civilization” in *Advances in Cross-Cultural Decision Making*, Denise M. Nicholson and Dylan D. Schmorow, editors. New York: Taylor & Francis, CRC Press.
- Kristol, Irving (1995), *Neoconservatism: The Autobiography of an Idea*, New York: The Free Press.
- Ricketts, Glenn, Peter Wood, Stephen Balch and Ashley Thorne (2011). “The Vanishing West, 1964-2010,” A Report by The National Association of Scholars. http://www.nas.org/articles/The_Vanishing_West_1964-2010. Last accessed March 1, 2014.
- Pidluzny, Jonathan W. (2011). “Democracy's Sacred Opinions and the Radicalization of Islam in the Twentieth Century” in *Advances in Cross-Cultural Decision Making*, Denise M. Nicholson and Dylan D. Schmorow, editors. New York: Taylor & Francis, CRC Press.
- Scruton, Roger (2010). *Liberty and Civilization : the western heritage*. New York: Encounter Books.

Strauss, Leo (1964). "The Crisis of Our Time" in *The Predicament of Modern Politics*, Harold J. Spaeth, ed. Detroit: University of Detroit Press.

Strauss, Leo (1967). "Jerusalem and Athens," *Commentary* (June).

Tocqueville, Alexis de (2000). *Democracy in America*, Harvey Mansfield and Delba Winthrop trans. Chicago: University of Chicago Press.

Extracting Understanding from Automated Metaphor Identification: Contrasting Concepts of Poverty Across Cultures and Languages

Sarah Taylor^a, Laurie Beth Feldman^b, Kit W Cho^b, Samira Shaikh^b, Ignacio Cases^b, Yuliya Peshkova^b, George Aaron Broadwell^b, Ting Liu^b, Umit Boz^b, Kyle Elliott^c, Boris Yamrom^b, and Tomek Strzalkowski^b

*^a Sarah M. Taylor Consulting, LLC
Orlando, FL 32816-2993, USA*

*^b University at Albany, State University of New York
Albany, NY 12222, USA*

^c Plessas Experts Network

ABSTRACT

We compare several conceptual metaphors for POVERTY – LOCATION, ENEMY, DISEASE, PLANT, BURDEN - as they appear in English, Russian and Spanish using two different corpora. We additionally explore the English POVERTY IS A LOCATION metaphor in detail, looking at the relative importance of its different dimensions and at the picture of POVERTY IS A LOCATION that these dimensions suggest. These analyses provide a prototype for future cross language analyses of metaphor to provide cultural insight.

Keywords: Metaphor, Cross-culture, Poverty, Natural Language Processing

INTRODUCTION

Metaphors describe or explain one concept (the target) using words or phrases that are also closely associated with some other thing or idea (the source concept). In the metaphorical sentence “Syrian refugees in Lebanon [are] drowning in poverty” (TopNewsToday, 2013) “poverty” is the target and “drowning in” is a phrase clearly associated with a body of water. This metaphor thus likens what is happening to these people in poverty to what might happen to them in a body of water. To paraphrase the sentence and make the comparison explicit, it means something akin to: ‘poverty like a body of water is overwhelming, smothering, incapacitating and possibly killing Syrian refugees in Lebanon’.

Typically, as metaphors are often explained, the target concept is an abstraction, and the source concept is something concrete and familiar. The language and ideas associated with the source supply terms that indicate to the reader salient properties of the target, frequently carrying with them a sensory and emotional component. Thus many, but not all, metaphors depend on sources from the material world, either natural or man-made. However, it also happens that abstractions, even though targets of metaphors in one context, can serve as source concepts for other metaphors. So while poverty is the target in the sentence cited above, it is the source in the passage “The other poverty is the poverty of ideas. Let us ask each of our leaders and politicians for just one relatively new and interesting idea to solve a pressing problem. Just one.” (Schwartz, 2014).

In our work we call the metaphorical words or phrases (‘drowning in’ and ‘poverty of’ in the two cited sentences) the *relation*, because these words relate, or link, the target to the source concept. A specific instance of a metaphor,

“drowning in poverty”, is called a *linguistic metaphor*. A *conceptual metaphor* is defined by a number of related linguistic metaphors with the same target and source concepts (Lakoff & Johnson, 2003; Deignan, 2005). The presence of a conceptual metaphor indicates that the people using its various inter-related descriptive phrases regularly discuss one idea in terms of the properties of the other. For example, the concept of poverty (hereafter POVERTY, when meant as a concept) is frequently likened to the concept of place or location. Many phrases suggesting POVERTY is like a LOCATION are common in English: “live in poverty”, “climb out of poverty”, “road to poverty”. In standard metaphor terminology, we say the conceptual metaphor comparing poverty and location, in English, is POVERTY IS A LOCATION.

The existence of the POVERTY IS A LOCATION conceptual metaphor in a culture in no way implies that POVERTY is always conceived of as a LOCATION. There are many conceptual metaphors used with POVERTY and they appear to vary with language. We address several – POVERTY IS A DISEASE, POVERTY IS AN ENEMY, POVERTY IS A PLANT, and POVERTY IS A BURDEN. However, when POVERTY is described as a LOCATION, and a LOCATION with a number of inter-related properties then certain aspects of POVERTY are emphasized.

Dimensions of the Source Concept

Not all metaphors employing the same source domain highlight the same features of it. Two common POVERTY IS A LOCATION phrases, “fallen into poverty” and “trapped in poverty”, refer to different aspects, or *dimensions*, of a protagonist’s possible relationship to a LOCATION – that a person can enter it, and that a person can remain there. These different dimensions carry very different import. The linguistic metaphors associated with them evoke different kinds of reality, imply different mental imagery, and suggest different affective judgments.

One way of dealing with this variety of source dimensions is to consider each a different conceptual metaphor, with obvious relationships analyzed as hierarchical relationships among them, or other combinations of conceptual metaphors operating in concert (Barcelona, 2001). Our team’s general approach is ‘bottom up’ or corpus-driven (Deignan, 2005). We believe the jury is still out on what level of aggregation or abstraction is useful for naming and defining a conceptual metaphor source (Ritchie, 2003, Krennmayr, 2013). At the same time, we need a method of organizing related linguistic metaphors that can capture the similarity of their semantics. We call the aspects or features of a source *dimensions*. The dimensions of the source LOCATION are: (i) getting into, (ii) getting out of, (iii) existing in, and (iv) properties of. We begin to develop this concept to be able to compare metaphors at a corpus scale on the basis of meaning. We employ the dimensions for POVERTY IS A LOCATION to support one of the comparative studies described below.

THE PRESENT STUDY

In this study, we compare several POVERTY conceptual metaphors – LOCATION, ENEMY, DISEASE, PLANT, BURDEN - as they appear in English, Russian and Spanish. We additionally explore the English POVERTY IS A LOCATION metaphor in detail, looking at the relative importance of its different dimensions and at the picture of POVERTY IS A LOCATION that these dimensions suggest. This analysis of dimensions suggests a prototype for future cross language analyses of metaphor to provide cultural insight.

The genesis of the project was to compare the *importance* of a metaphorical source across cultures on the basis of corpus evidence. The work is ongoing so we now only provide a preliminary report. Since conceptual metaphors tell us something about the way that people think, and they have both shared and differing behavior in different cultures (Deignan, 2003), source domains is one obvious point of comparison. If one culture group relies heavily on certain sources for discussing POVERTY, it could tell us something interesting about how poverty is viewed and dealt with in that culture (Thibodeau & Boroditsky, 2011). The broader importance of certain metaphor sources across many domains could also imply more general insights into a culture. Do US Americans, for example, rely as heavily on conflict and competitive metaphors as it sometimes seems they do? Is this a characteristic of the culture? Is it present in some cultures and not others, and so forth. But to answer such questions we need to have a way of defining and measuring *importance*.

The validity of our corpus-based comparison across cultures rests on the resolution of two core problems: how to quantify importance and how to make appropriate comparisons using our available corpora. Our Analyses section rests on our solutions to these problems. However, we make our working assumptions explicit here as they

determine the potential theoretical impact of our findings.

Importance

We now rely on two quantitative measures of importance, *frequency of use* and *elaboration*. Frequency of use is the number of times a metaphor appears in a corpus. For example, we compare the number of POVERTY IS A LOCATION metaphors in an English corpus, expressed as a percentage of all the POVERTY metaphors in that corpus, to the number of POVERTY IS A LOCATION metaphors in a Russian corpus, also as a percentage of all the Russian poverty metaphors.

Our second measure of importance is *elaboration* or variation of expression across a dimension of the conceptual metaphor. It is a measure of semantic richness not unlike a word's contextual diversity (Adelman, Brown, & Quesada, 2006) or number of senses or features (Pexman et al., 2008; Yap et al., 2012). Our justification for *elaboration* is that the number of ways or varieties of expressions that a language uses to express fundamentally the same idea reflects the centrality of this metaphor to that culture. At a minimum, the degree of variation in how a core metaphorical idea is expressed reflects a breadth of language users, registers, genres, document and spoken language contexts for the metaphor.

The dimensions of the source domain help us categorize linguistic metaphors into groups with similar meanings. At this point we have only operationalized the *elaboration* measure for the English POVERTY IS A LOCATION conceptual metaphor and use it to compare the importance of the different dimensions of this metaphor in English. However, the method is theoretically possible to extend to other comparisons of source importance.

To be clear, *frequency* and *elaboration* are only two of the possible measures of metaphor importance. As developed here they are preliminary, as they are limited to the one strongest relation and would fail to detect the importance in a situation of several weaker but associated relations. Other possible measures of metaphor importance, for example, include (a) composite measures of sensory vividness and emotional arousal, (b) usage patterns over extended time periods, (c) association of particular metaphors with foundational religious, literary, and intellectual texts. These latter measures are more complex to develop than frequency and elaboration, and their full development awaits a later phase of this work.

Comparability

We employed two rather different corpora in our comparisons, to see how sensitive our conclusions were to the nature of the corpus on which they were based. The characteristics of each corpus required the use of rather different methods for searching for metaphors to calculate frequencies. Later phases of this work will endeavor to compare and coordinate the results from different corpora more closely.

Our team's metaphor database (Shaikh et al., 2014) is based on a four-language corpus (English, Russian, Spanish and Farsi) drawn from Internet sources, heavily weighted toward a variety of newspapers, online news sources and blogs, primarily from 2008 to the present. The database consists of linguistic and conceptual metaphors automatically discovered for a limited number of target domains focused on government and social issues, of which POVERTY is one (Strzalkowski et al., 2013, Broadwell et al., 2013). For the work reported in this paper, the POVERTY linguistic metaphors were checked for accuracy and manually assigned to conceptual metaphors by senior linguists. The process for discovery and proposal of the linguistic metaphors, however, was entirely automatic. An identical method of linguistic metaphor identification is employed across all four languages, but our systems for all four languages are not yet of equal maturity. Searches and frequency counts in this database are easy to perform and take into account a wide range of lexical and syntactic variants in each language.

The second corpus we used is the Google books corpus, readily searchable in the Ngram Viewer (Michel et al., 2010). We searched Google books from 2005 to 2008, using the American English, Russian and Spanish corpora. The data for these corpora are the most recent available (version 20120701). Google books collection is larger than our collection and contains a wider range of genre and subject matter. The books also employ a more formal language than the shorter Internet news texts and blogs of our team developed corpus. Neither corpus incorporates extensive examples of every day speech, but both represent language understood by a broad audience, although our own corpus includes fewer examples of specialized language (such as from the academy or professions) than does Google books.

To use the ngram index we searched for frequencies of known metaphorical phrases. We know from the results of our own automated metaphor discovery, as well as from the work of other researchers (Cameron & Deignan, 2006), that many linguistic metaphors employ fixed lexical sequences. We used such fixed expressions to obtain counts of known linguistic metaphors. Many of the fixed expressions -- for example, “mired in poverty” -- employ *relations* that correspond to those metaphorical relations stored in our team database. Therefore, we used the POVERTY metaphors discovered in the team developed database as a guide for linguists to develop the lists of English, Russian, and Spanish phrases to be searched in the Google books index.

Relations: Strong Versus Weak Links to the Metaphorical Source

Crucially, the fixed expressions (for ngram searching) must be reliably indicative of the source domains we wish to compare. Thus, we introduced a distinction between relations that are tightly linked to the source concept and relations that are only compatible but not exclusively linked with a particular source concept. For example, the sentence “Syrian refugees in Lebanon [are] drowning in poverty” already cited tightly links poverty to a body of water, likely of some depth. However, the sentence “They were rescued from poverty and given a job” (Watts, 2012) could be considered only compatible with the notion of drowning. It indicates poverty as a danger, but implies no single source of danger. In our comparison of sources “rescue from” could easily apply to four of the five source domains we examined (LOCATION, DISEASE, ENEMY, and BURDEN, but not PLANT).

We intend our frequency counts to indicate the relative importance of the sources (Figure 2), therefore we used only phrases where the relations were strongly tied to the specified source domains. Finally our measures within and across languages and cultures capture relative proportion. The metaphoricity of the fixed phrases as well as their utility for cross language comparison was concluded from their presence in our database. The strength of their ties to the source domains was judged by members of the linguistic team.

RESULTS

We have performed three comparative studies to throw light on the utility of our approach as well as to discover differences and similarities between POVERTY metaphors using the sources LOCATION, DISEASE, ENEMY, PLANT and BURDEN among three languages. The first uses the database we are developing. The second makes a similar comparison based on Googlebooks ngram searches. The third addresses the elaboration measure. We focus on English and Russian for the first comparison, on English, Russian and Spanish for the second, and on English POVERTY IS A LOCATION metaphors for the third analysis.

1. Frequency of POVERTY Metaphor Source Domains: English and Russian

Using frequency as a measure of importance, we compare the frequency of POVERTY metaphors in our team metaphor database in several domains for English and Russian. The results are shown in Figure 1. For this comparison we count the number of instances of POVERTY metaphors associated with each of five sources – DISEASE, LOCATION, ENEMY, PLANT, and BURDEN. These five cover the most frequent sources for English and Russian in our database. BURDEN is also included as we use it in the second comparison below. In the database, the metaphors were automatically discovered in text, manually reviewed for accuracy and manually assigned to source domains. We consider POVERTY a state of being, and POVERTY IS A LOCATION to belong to the metaphor A STATE OF BEING IS A LOCATION (Lakoff, Espenson, & Schwartz, 1994) and thus to include relations indicating many types of locations such as high and low places, bodies of water, confined spaces.

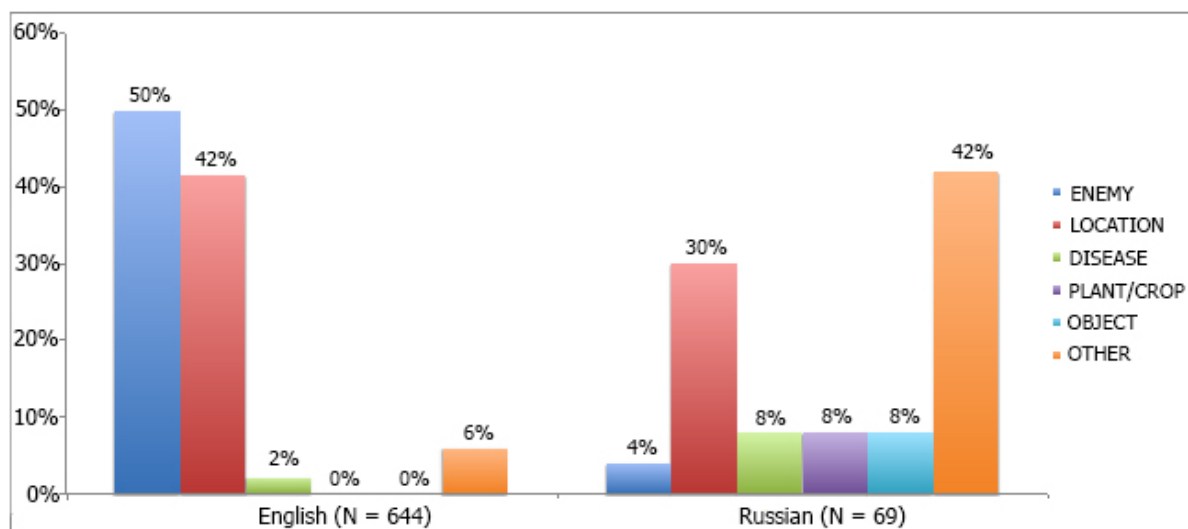


Figure 1: Initial comparison of source frequencies for poverty metaphors

We caution that the English and Russian results are incomplete and reflect the status of a database continuing under development. However, we present them to indicate the potential utility of such a comparison as a method based on metaphor suggesting insights into culture. Additionally, we used the more common relations and associated sources in our database as a basis for developing the ngram searches used in our second comparison.

2. Frequency of Poverty Metaphor Sources Based on Ngram Searches: English, Russian, Spanish

Relying again on frequency as a measure of importance, we compare the frequency of POVERTY metaphors in English, Spanish and Russian using source domains of DISEASE, LOCATION, ENEMY, PLANT AND BURDEN. The frequencies were obtained from ngram searches of Google books. For expediency one relation was selected to stand for the source in each language. These relations were selected based on two criteria: strong and unambiguous links to a single source and high frequency levels for that relation within the conceptual metaphor (target + source). Generally this resulted also in their being frequent relations for that conceptual metaphor in the team database. These relations were selected independently for each language.¹ While this use of a single relation to represent each source may not be ideal, searches in English suggest that one relation will tend to overwhelm others in frequency for any particular conceptual metaphor. For example, almost 60% of all occurrences of the 59 identified POVERTY IS A LOCATION relations, used in our third comparison study, were variants of “live in poverty”. For purposes of estimating importance such a comparison can thus be indicative.

Each relation can occur in multiple syntactic forms, such as “curing poverty” and “cure for poverty”. The relative frequencies of these forms vary. A linguist enumerated the possible forms for each relation for each language, disregarding their likelihood as to actual occurrence. We then searched for every form and summed the results. Table 1 shows the relations used in each language for each source. Note that we determined the relation separately for each language and did not translate. For example, страдать от means suffer from, not cure. The syntactic forms that provided the most frequent results are shown in Table 2.

¹ A separate attempt, intended for comparison, to use translated terms based on high frequency English relations yielded few or no ngram search results from the other languages.

Table 1. Key relations for the five source domains

Source	English relation	Russian relation	Spanish relation
DISEASE	cure	страдать от	erradicar
ENEMY	fight	борьба с	luchar
LOCATION	live in	жить в	vivir en
PLANT	grow	расти	crecer
BURDEN	burden	бремя	se reducir ²

Table 2. Top most productive syntactic form by relation.

English	Frequency	Russian	Frequency	Spanish	Frequency
cure for	43%	страдают от бедности	28%	erradicación de	61%
fight against	37%	борьба с бедностью	83%	lucha contra	90%
living in	53%	n/a		viven en	28%
growing	60%	плодить нищету	61%	crecimiento de	88%
burden of	90%	n/a		se redujo	52%

The search of Ngram strings in English, Russian and Spanish results in about 73000 hits in English, 5400 in Russian and 49000 in Spanish. Figure 2 depicts the distribution of source domains for POVERTY, expressed as a percent of total hits in each language. In English and in Russian, the most common source domain associated with poverty is LOCATION with percentages at 70% and 79% respectively. The second most common is ENEMY with percentages at about 20%. These two constitute more than 90% of all POVERTY metaphors that we detected in English and in Russian. The distribution looks very different in Spanish, where relations of ENEMY are most prevalent and constitute 55% of all metaphors that we detected. Spanish also differs from the other two with respect to the proportion of DISEASE metaphors- 22% in Spanish and fewer than 1% in English.

² This relation requires additional testing in Spanish for the ambiguity of its relation to BURDEN.

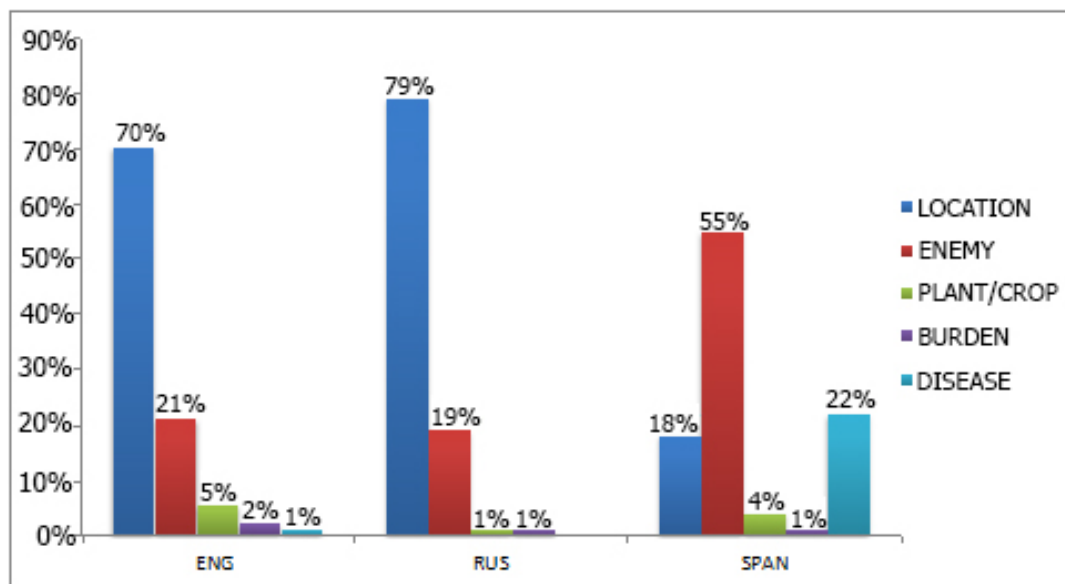


Figure 2: Comparison of frequency of POVERTY metaphor sources based on Google books Ngram searches

3a. Elaboration as Possible Measure of Importance: Comparing the Importance of Dimensions in American English POVERTY IS A LOCATION

An extensive search of English POVERTY IS A LOCATION metaphors in Google books (2004-2008) enabled us to compare the importance of the different dimensions of POVERTY IS A LOCATION in contemporary American English. We then compare the elaboration measure to the frequency measure of importance. We exhaustively catalogued the list of POVERTY IS A LOCATION relations for English, emphasizing those associated with the three dimensions Get into, Get out of and Be in. As these are focused on verbs, most employing prepositions indicating location, they constitute a more tractable list to prepare than relations for Features of, which can include a wide variety of descriptive relations.

Table 3 shows the relations searched for in the ngram corpus and their association with each dimension. All were extensively tested by the linguist, using manual review of returns from the Google internet index, to determine that 1) they were used metaphorically with poverty in an estimated 90% or more of cases and for 2) close association with the LOCATION source. Some apparently obvious phrases, for example “place of poverty”, were not used because they frequently occur in literal as well as some metaphorical contexts. Certain of the relations have a significantly different meaning depending upon their syntactic form, such as “sink into poverty” (implying the process of sinking) and “sunk in poverty” (implying the process is complete and now the condition of being pertains) and thus may be associated with different dimensions. All relations were tested for natural occurrence using the Google internet corpus (see column 4). However, some relations occurring in that corpus did not appear in the ngram books corpus (those in *italics* in column 4).

Table 3. Percent of relation occurrences for each POVERTY IS A LOCATION dimension, based on ngram corpus (2004-2008)

Dimension	Elaboration: # of identified relations (internet corpus)	Elaboration: # identified relations (ngram corpus)	Relations	% of all relation occurrences (ngram corpus)
Get into	24	18 (36%)	drive, drop, enter, fall, lower, slide, slip, cast, land, plunge, push, sink, throw, <i>stumble</i> , <i>tumble</i> , <i>plummet</i> , <i>topple</i> , <i>dive</i> , way, path, <i>pathway</i> , road, route, down	10%
Be in	11	9 (18%)	live, drown, engulf, lost, mire, stick, submerge, sink, entrench, <i>swamp</i> , <i>swim</i>	71%
Get out of	13	12 (24%)	emanate, emerge, lift, climb, pull, <i>drag</i> , rise, path, pathway, route, way, road, up	12%
Features of	12	11 (22%)	brink, edge, deep, <i>shallow</i> , endless, vast, vast sea, quagmire, quicksand, morass, mire, swamp	7%

This initial comparison of the importance of each dimension based on elaboration and on frequency does not suggest a close correlation between these two possible measures. Even restricting the analysis to the first three dimensions as potentially being more reliably enumerated than the fourth suggest that importance based on frequency and importance based on elaboration are distinct properties of metaphors.

3b. What kind of place is poverty?

States of being are treated as LOCATIONS. However, different states of being can be viewed as different kinds of locations. The semantics of the relations, categorized within each dimension, tell us what kind of location poverty is. We have the following in English, based on possible expressions developed and tested for each POVERTY IS A LOCATION dimension. Each expression was examined and categorized depending upon whether its meaning suggested a low place, a wet place or a confined place, or none of these qualities. For example, “fall into” means going downwards, and thus implies a low place. A “mire” is a swampy place, and so “mired in”, frequently used with “mud” in its literal use, implies a wet place, as well as confinement, since it means one cannot get away.

Table 4 shows that fully two thirds of the relations identified for talking about POVERTY IS A LOCATION in English are associated with at least one characteristic of lowness, wetness or confinement. Fourteen, or roughly 25% of these terms are associated with at least 2 of these characteristics. The remaining terms are mostly neutral to the type of location poverty is, being expressions such as “live in”, “route to” or “pathway from”.

Table 4: Relations for POVERTY IS A LOCATION metaphors in English

Dimension	Relations	Low Place	Wet Place	Confined Place	Other associations
Get into	24	12	5	2	9
Be in	11	4	7	6	2
Get out of	13	4	0	4	6
Features of	12	4	5	5	3

CONCLUSIONS

The analyses we have performed suggest that both frequency and elaboration can be employed to explore the importance of a metaphor or of metaphor dimensions within and across languages. However, the relationship between these two measures requires further exploration because they appear to reflect different properties of metaphors. Employing the Google ngram corpus provides a significant extension of our capability to investigate metaphorical usage within and across languages. We were somewhat surprised by the differences across our database and the Google ngram corpus and will continue to compare corpora as our database expands. The difference between relative frequencies of various metaphorical phrases and expressions in the Google Books and the Internet database suggests that more detailed comparisons are warranted.

ACKNOWLEDGEMENT

This work was supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Defense US Army Research Laboratory contract number W911NF-12-C-0024. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoD/ARL, or the U.S. Government.

REFERENCES

- Adelman, J. S., Brown, G. D. A., and Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science*, 17, 814-823.
- Barcelona, A. (2001). On the systematic contrastive analysis of conceptual metaphors: case studies and proposed methodology. In M. Pütz, Niemeier, S., Dirven, René (Eds.), *Applied Cognitive Linguistics II: Language Pedagogy*, 117-146.
- Broadwell, A. B., Boz, U., Cases, I., Strzalkowski, T., Feldman, L., Taylor, S., Shaikh, S., Liu, T., Cho, K., & Webb, N. (2013). *Using imageability and topic chaining to locate metaphors in linguistic corpora*. Proceedings from SBP'13 Proceedings of the 6th international conference on Social Computing, Behavioral-Cultural Modeling and Prediction. (pp. 102-110). Springer-Verlag Berlin, Heidelberg: Springer.
- Cameron, L., & Deignan, A. (2006). The emergence of metaphor in discourse. *Applied Linguistics*, 27, 671-690.
- Deignan, A. (2003). Metaphorical expressions and culture: An indirect link. *Metaphor and Symbol*, 18, 255-271.
- Deignan, A. (2005). *Metaphor and corpus linguistics*. Philadelphia, PA: John Benjamins.
- Krennmayr, T. (2013). Adding transparency to the identification of cross-domain mappings in real language data. *Review of cognitive linguistics*, 11, 163-184.
- Lakoff, G., Espenson, J., Schwartz, A. (1994). The master metaphor list. Retrieved from <http://cogsci.berkeley.edu/>.
- Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. Chicago: University of Chicago Press.
- Michel, J-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Brockman, W. The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011) Quantitative analysis of culture using millions of digitized books. *Science*, 331, 176-182.
- Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: effects of

- three measures of semantic richness on visual word recognition. *Psychonomic Bulletin, & Review*, 15, 161–167.
- Ritchie, D. (2003). “ARGUMENT IS WAR” – Or is it a game of chess? Multiple meanings in the analysis of implicit metaphors. *Metaphor and Symbol*, 18, 125-146.
- Schwartz, B. (2014). The other poverty: The poverty of ideas. Retrieved from <http://bobmschwartz.com/2014/01/07/the-other-poverty-the-poverty-of-ideas/>.
- Shaikh, S., Strzalkowski, T., Liu, T., Broadwell, G. A., Yamrom, B., Taylor, S., Feldman, L., Cho, K., Boz, U., Cases, I., Peshkova, Y., & Lin, C-S. (2014). A Multi-cultural repository of automatically discovered linguistic and conceptual metaphors. The 9th Edition of the Language Resources and Evaluation Conference.
- Strzalkowski, T., Broadwell, G. A., Taylor, S., Feldman, L., Yamrom, B., Shaikh, S., Liu, T., Cho, K., Boz, U., Cases, I., & Elliott, K. (2013). Robust extraction of metaphors from novel data. Proceedings of the First Workshop on Metaphor in Natural Language Processing. (pp. 67-76). Meta4NLP, The 1st Workshop on Metaphor in NLP. Atlanta, GA.
- Thibodeau, P. H., Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLoS ONE*, 6, 1-11.
- TopNewsToday (2013). *Syria refugees in Lebanon drowning in poverty, says Oxfam*. Retrieved from <http://world.topnewstoday.org/world/article/8742619/>.
- Watts, J. (2012). Violence in the vineyard. Retrieved from <http://nampaumc.org/sunday-april-1-2012/>.
- Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, 6, 1-10.

Section 4

Advances and Challenges in Cross-Cultural Competence in the U.S. Military, 2001-2014 and Beyond

Socio-Cultural Capability Requirements across All Phases of Military Operations

Walter Perry

*RAND Corporation
1200 S. Hayes Street
Arlington, VA 22202-5050, USA*

ABSTRACT

In modern military operations, it is critical that commanders and their staffs be sensitive to the social and cultural norms of populations with whom they must interact. Most military operations have the mission of promoting stability and reconstruction, or eventually evolve to such after major combat. Gaining the cooperation of the local population is critical to a secure environment. This paper identifies the socio-cultural capabilities needed to execute military missions, identifies which of these the military has access to, and proposes mitigating measures to close requirement gaps. We identified 27 missions across all phases of military operations along with 12 socio-cultural capability requirements organized in five categories: planning, communication and coordination, gaining detailed cultural awareness, engaging the population and assessing the impact of this engagement. We identified several gaps but in general, we found that substantial progress has been made in terms of closing these gaps. Personal initiative accounted for a large fraction of these efforts and attests to the growing appreciation among military personnel for the important role of socio-cultural capabilities. Although we provide recommendations to fill specific gaps, finding ways to encourage personal initiatives will nurture the development and testing of original ideas that could be used to address socio-cultural needs more broadly.

Keywords: Military operations, socio-cultural capabilities, human terrain, cultural awareness.

INTRODUCTION

In modern warfare it is critical that Commanders and their staffs be sensitive to the social and cultural norms of host nations. Most military operations have the mission of promoting stability and reconstruction, or eventually evolve to such after major combat as in Iraq. Gaining the cooperation and active assistance of the local population is critical to ensuring a secure environment. The Army and Marine Corps doctrine manual on counterinsurgency operations emphasizes the importance of understanding the local culture and social structure (FM 3-24, 2006):

...staffs should identify and analyze the culture of the society as a whole and of each major group within the society. Social structure comprises the relationships among groups, institutions, and individuals within a society; in contrast, culture (ideas, norms, rituals, codes of behavior) provides meaning to individuals within the society.

Although this addresses counterinsurgency operations, the Army and Marine Corps recognize that understanding a society's culture and social structure is necessary in all phases of military operations.

This paper identifies the socio-cultural capabilities needed to execute missions and tasks across the full spectrum of

military operations, identifies which of these the military has access to, and proposes mitigating measures to close requirement gaps. The analysis draws on the capability-based assessment (CBA) process. A CBA is designed to identify solutions to provide the warfighter with capabilities needed to accomplish missions. A CBA consists of three parts: the Functional Area Analysis (FAA) resulting in the identification of needed capabilities; the Functional Needs Analysis (FNA) resulting in the identification of gaps in the required capabilities identified in the FAA process; and the Functional Solutions Analysis (FSA) resulting in the identification of means to fill the capabilities gaps.

THE MISSIONS

The full spectrum of military operations is described in terms of operational *phases*. Joint Publication 3-0, *Joint Operations*, describes the military operations phasing model. This model is designed to organize operations by helping Commanders and their staffs to think through an entire campaign from the early *shaping* phase through the *combat phases* to *reestablishing civil authority* in the host nation. There are six phases identified. Although presented as a sequence, it is clear that there is considerable overlap of the phases of an operation, thus precluding a distinct separation between phases. This reality highlights the need to integrate socio-cultural factors throughout a military operation. (Joint Publication 3-0, 2010)

We describe the activities associated with each of the operational phases along with the missions drawn from the sources cited above. For each phase, we summarize the mission set by citing the major objectives associated with the phase. The full list of missions is included in Table 1.

Phase 0: Shaping: This phase consists of benign military operations designed to enhance international legitimacy and gain international support for U.S. Military and national objectives. The focus is on building local capacity, developing host nation military capabilities, and dissuading potential adversaries. The objective is to create a favorable opinion of the United States and to establish or reinforce relations with friendly countries.

Phase I: Deter: In this phase, the objective is to demonstrate resolve by engaging in preparatory actions that clearly demonstrate that a joint force is both capable and committed to deterring enemy activity. Some of the activities in this phase include demonstrating capability, deterring criminals and violent extremists, and gaining support for the command's operations concept. In general, U.S. forces move pro-actively to prevent a potential attack or crisis.

Phase II: Seize the Initiative: If adversaries cannot be deterred, the U.S. may move quickly to launch combat operations to seize the initiative, regardless of whether it is a hostile nation state or a non-state group posing the threat. In non-combat operations the objective is to apply appropriate joint force capabilities to gain the initiative. Phase II activities can include offensive combat operations designed to destroy an adversary, kill or capture its senior leaders, reduce its freedom of movement, dislodge its forces from their positions, and degrade an enemy's capability to threaten U.S. or allied interests, or undertake acts of violence.

Phase III: Dominate: This phase focuses on breaking the enemy's will for organized resistance or, in non-combat situations, control of the operational environment. If the United States succeeds in seizing the initiative from an adversary, it can attempt to dominate the environment and prepare to transition from offensive operations to longer-term stability operations. In Phase III, the United States might conduct continuous operations aimed at suppressing threat networks and isolating threat groups, including violent extremist organizations, operating within a partner country.

Phase IV: Stabilize: This phase applies when there is not a functioning government in the host nation, or only a weakly functioning government in place. This scenario may arise from the collapse of a regime. The phase is characterized by the transition from combat operations to stability operations. In this phase the United States may be compelled to step in and use military forces to provide basic security, re-build critical infrastructure, and ensure that the essential needs of the host country population are met, including medical needs.

Phase V: Enable Civil Authority: In Phase IV, the joint force has assisted the re-establishment of governance by a host country. In Phase V the joint force focuses on providing support to a legitimate, functioning government, one

that it may have helped establish. Support to a functioning government is provided in accordance with some agreement. U.S. Military forces will typically work in coordination with other U.S. government agencies, foreign government agencies, and NGOs and IGOs.

We identified 27 missions across the full spectrum of military operations (Table 1). These missions were culled from several Joint Staff, Major Command and Service documents. They are, in essence, summaries of more theater-specific missions and tasks. We reviewed Joint Publication 3-0, which describes activities that Commanders might anticipate in various phases of a theater campaign. We consulted the planning guidance and tasks assigned to combatant Commanders in the *Joint Strategic Capabilities Plan* (Joint Strategic Capabilities Plan, 2008). We reviewed Joint Warfighting Center materials to identify tasks organized by type of operation developed for mission analysis and published in the *Universal Joint Task List*, a Chairman of the Joint Chiefs of Staff manual (CJCSM 3500.04C, 2002). Finally, we took future missions into consideration by reviewing scenario data published for each of the selected combatant commands in the Multi-Service Force Deployment documents promulgated by the Joint Staff.

Table 1: Military Missions by Phase

Phase	Missions
Phase 0: Shaping	Build local capacity and support infrastructure projects,
	Develop host nation military capabilities and solidify relationships with friends and allies
	Dissuade or deter potential adversaries
	Improve information exchange and intelligence sharing, and provide U.S. forces with peacetime and contingency access
Phase I: Deter	Employ the joint force to demonstrate capability and resolve
	Deter criminal and violent extremist activities
	Mobilize and tailor the force to address the crisis in coordination with OGAs, IGOs, and NGOs
	Support the command's CONOPS by obtaining over-flight permissions and planning for force protection and logistics
Phase II: Seize the initiative	Disable or destroy the enemy through kinetic means
	Limit collateral damage to infrastructure and partner country populations
	Provide assistance to relieve conditions that precipitated the crisis
	Influence and inform the local population
Phase III: Dominate	Prevent instability from spreading to adjacent areas,
	Influence population to counter anti-U.S. messages
	Disrupt activities of violent extremist organizations
	Control the situation
	Overmatch the enemy by employing the full joint force
	In irregular warfare, dominate the environment using conventional and unconventional, information and stability operations
Phase IV: Stabilize	Improve local governance
	Assist in the provision of basic services and re-build, improve, and protect infrastructure
	Build capacity of local law enforcement and justice system, and partner with local authorities to increase security and restore order
	Ensure that conditions leading to the original crisis do not recur
	Begin redeployment operations as soon as possible
Phase V: Enable civil authority	Build, support, improve, and secure legitimate civil government and enable civil authority's provision of services
	Coordinate actions with multinational OGAs, IGOs, and NGOs
	Persuade locals to regard U.S. and local civil authority favorably
	Disrupt and deter human/drugs/weapons trafficking and erode support for criminal and violent extremist organizations

THE SOCIO-CULTURAL CAPABILITIES

Across the globe, in commands, theaters, and campaigns, there is an increasing recognition that socio-cultural knowledge and expertise is critical for the success of military missions. The U.S. Military and its allies are increasingly dealing with internal conflicts within states, responding to complex humanitarian emergencies, and other situations far outside the scope of “traditional” state-to-state warfare (TRADOC OE White Paper, 2009). In these new operational environments, the enemy is often not a clearly discernible group, and may change rapidly from day to day. Such environments call for new capabilities that can assess and make productive use of operationally relevant information regarding the social and cultural landscape of the area of operations. As recent campaigns have shown, this is no small feat. Human terrain is inherently local and by its nature rapidly evolving. Our next task then was to develop a list of socio-cultural capabilities necessary to meet this challenge.

The first step in developing a list of required socio-cultural capabilities (SCCs) required our developing a working definition of the term itself. We assembled a team of military experts (most of them former U.S. Military officers) and social science experts with applied research backgrounds in Anthropology, Psychology, Political Science, and Public Health. The research team conducted a literature search for descriptions of how socio-cultural knowledge and research had been used to inform military operations, as well as calls for improvements in the use of socio-cultural knowledge for future military operations. This involved consulting doctrine, military documents such as reports from Human Terrain Teams, academic journals and books, and the existing RAND literature on the topic. In addition, we vetted the definition of SCCs with military and non-military experts during one-on-one and group interviews we conducted over the course of the study. This process produced the following definition:

*A socio-cultural capability is the ability to employ **human and material resources** to gather, synthesize, analyze, interpret and share **operationally relevant** information about a population’s **social and cultural landscape**, and use this to inform and assist the planning, preparation, execution and assessment of a military operation*

We further defined the bolded elements of the above definition. Table 2 summarizes these additional definitions.

Table 2: Socio-Cultural Component Definitions

Component	Definition
Operationally relevant	Pertaining to the requirements of the supported unit and mission
Social landscape	A particular group’s political, economic, institutional, kinship and class structure and organization
Cultural landscape	Meanings, beliefs, attitudes, behavioral norms, values, and perceptions shared by and distributed among a group
Human resources	Resources that support an individual or unit’s socio-cultural capability development. Examples include interpreters/linguists, subject matter experts, tribal leaders, key figures in local governance, NGOs, social scientists, etc.
Material resources	Information storage and sharing systems; software for data analysis; local media for dissemination and influence, etc.

Our SCC definition includes a ‘know’ element (knowing the social and cultural landscape) and a ‘do’ element (using information to plan, prepare, execute and assess a mission). Modern social science views the social and cultural as causally intertwined; thus, the term “socio-cultural” is a reflection of the myriad causal forces and processes that influence patterns of thought and behavior in a society or group.

Based on interviews with approximately 30 military and non-military experts and an extensive literature review we reduced a rather long list of socio-cultural capabilities to a manageable set of 12 which we divided into five categories. The categories were organized to approximate a chronological progression of how socio-cultural capabilities might be used throughout the planning and execution of military missions. They consist of planning, communication and coordination, gaining detailed cultural awareness, engaging the population and assessing the impact of this engagement. We describe the categories below and Table 3 lists the associated required socio-cultural capabilities.

Planning: For the purposes of mission planning, the Commander must have the ability to discover the features of the human terrain that are most operationally relevant—in other words, to know what needs to be known. Furthermore, he must have the ability to make use of socio-cultural information to inform mission objectives. This requires the ability to identify relevant socio-cultural information and use it to plan operations.

Coordination and Communication: One critical socio-cultural capability is the capacity to communicate information and coordinate productively with other parties operating in the area of interest (AOI). This requires the ability to communicate findings and information to both military and non-military organizations and to work well with organizations in the region to plan collaborative activities.

Gaining Detailed Cultural Awareness: The most important and arguably most difficult required capability is the capacity to obtain valid and reliable operationally relevant information regarding the social and cultural landscape of an AOI. Some of the requirements in this category include the ability to acquire knowledge of the culture, social structure, political dynamics and economics of the area and how they interconnect, design and administer efforts aimed at obtaining this information, and verify the accuracy and reliability of the information.

Engaging the Population: Increasingly, through the use of provincial reconstruction teams, female engagement teams, and the like, the U.S. Military is recognizing the need for and difficulty of engaging in sensitive interactions with the local population in a culturally appropriate manner. This requires an ability to communicate with the local population, build trust, and productively interact with them. A number of initiatives attempting to address these challenges have been undertaken. “Smart cards” with basic phrases in the local language and information about the culture have been issued to deploying personnel (Davis, 2010).

Impact Assessment: After socio-cultural capabilities have been employed in the pursuit of mission success, it is important to assess the impact of these efforts in order to inform future mission planning. This requires an ability to obtain reliable information on post-engagement conditions and the ability to understand the implications of local changes for successive engagements. For example, effectiveness of language training courses is routinely assessed. Within the Army, “language tests are mandatory for soldiers who have received foreign language training at government expense,” excluding survival-level courses (Army Regulation 11-6, 2009). However, annual retesting of speaking skills is not currently required for all language-dependent military occupational skills. The Defense Language Transformation Roadmap sets proficiency level goals for language professionals (Defense Language Transformation Roadmap, 2005).

Table 3: Socio-Cultural Capabilities by Category

Category	Required Socio-Cultural Capabilities
Planning	The ability to identify operationally relevant socio-cultural information categories and requirements (Salmoni and Eber, 2008), (Medby and Glenn, 2002).
	The ability to use socio-cultural information to design missions and activities (FM 3-24, 2006) (Salmoni and Ebert, 2008). Such as: possess access to best practices for using socio-cultural information to inform or alter mission objectives (Flynn, Pottinger, and Batchelor, 2009), (Salmoni and Eber, 2008).
Coordination and communication	The ability to communicate findings and information effectively to both military and non-military organizations (Shaley, 2007).
	The ability to store and establish access to information for multiple agencies and organizations (3-24, 2006), (Kjeldsen, 2006), (Defense Science Board, 2009), (Byman, 2001).
	The ability to work with organizations, both U.S. and foreign, in the region to plan collaborative activities (Medby and Glenn, 2002), (Byman, 2001), (FM 3-24, 2006).
Gaining Detailed cultural awareness	The ability to acquire knowledge of the culture, social structure, political dynamics and economics of the area and how they interconnect. Such as: (1) understand where to go to in order to get information; (2) have access to documents describing the local culture; (Medby and Glenn, 2002) and (3) possess access to experts (HTS, RRCs, SMEs, CAs, etc.).
	The ability to design and administer appropriate efforts to gain detailed situational awareness. Such as: (1) possess the ability to design and administer surveys, interviews, focus groups (Morgan, 1993), and conduct ethnographic observations; and (2) possess the ability to conduct quantitative and qualitative analyses and make appropriate inferences.
	The ability to verify the accuracy and reliability of the information gained (Jandora, 2006), (Legree, 2010), (FM 3-24, 2006). Such as: (1) have procedures in place to crosscheck the accuracy and reliability of information; and (2) to update with new information as it becomes available.
Engaging the population	The ability to communicate with the local population. Such as: (1) possess survival language skills for limited interactions; (2) possess ability to communicate fluently for in-depth interactions; (3) have access to translators/interpreters if needed; and (4) if translators/interpreters are needed, understand how to use them appropriately.
	The ability to productively interact with the local community (Jandora, 2006), (LeGree, 2010), (FM 3-24, 2006). Such as: (1) understand the meaning and significance of local culture (Abbe and Halpin, 2009-2010), (Jager, 2007), (Salmoni and Eber, 2008), (Wunderle, 2006), (Abbe, Gulick, and Herman, 2008), (FM 3-24, 2006); (2) possess awareness of appropriate social etiquette, with people and groups

Category	Required Socio-Cultural Capabilities
	across the social hierarchy (Abbe and Halpin, 2009-2010), (Cushner, and Brislin, 1995), (Cushner and Landis, 1996), (Salmoni and Eber, 2008); (3) possess the ability to build effective relationships with local institutions and persons of influence or high status (Jandora, 2006), (LeGree, 2010), (Salmoni and Holmes-Eber, 2008), (Stenmark, 2006), (FM 3-24, 2006); and (4) possess locally appropriate influence, negotiation, and conflict resolution skills (Black and Porter, 1991), (Brett and Okumura, 1998), (Gelgand, Erez, and Aycan, 2007), (Gelfand and Dyer, 2000), (LeGree, 2010), (Morris, Williams, Leung, Larrick, et al, 1998), (Ohbuchi and Takahashi, 1994), (Sheer and Chen, 2003), (Sullivan, Peterson, Kameda, and Shimada, 1981), (Tyler, Lind, and Huo, 2000).
Impact assessment	The ability to obtain reliable and relevant information on post-engagement conditions to assess the efficacy of employing SCCs (Language Skills, Regional Expertise, and Cultural Capabilities, 2011). Such as: (1) the ability to develop a credible socio-cultural impact assessment plan; (2) possess knowledge of various assessment methods; (3) the ability to access trained personnel and resources required to conduct proper assessments; (4) the ability to assess the impact of U.S. operations on the local socio-cultural environment (may focus on balance of power, traditional practices, etc.); and (5) the ability to assess the impact of the socio-cultural milieu on U.S. operations.
	The ability to understand the implications of local changes for successive engagements. Such as: the ability to integrate relevant information and conclusions from impact assessment efforts into planning processes for successive missions.

SOCIO-CULTURAL CAPABILITIES BY OPERATIONAL PHASE

With the missions and the required socio-cultural capabilities identified, the next step was to identify which socio-cultural capabilities were needed to accomplish the missions across all phases of military operations. We proposed to develop a matrix with the 27 missions as rows and the 12 socio-cultural capabilities as columns. The result was to be a matrix with 314 cells containing the manner in which the column socio-cultural capability can assist in accomplishing the row mission. Figure 1 illustrates the matrix structure. We refer to the cell entries as “intersections.”



Figure 1. Intersections of SCCs by Mission Activities

While useful for validation purposes and to produce a large base of “raw data” for further work, this level of resolution was found to be excessively specific and therefore not likely to be useful for military planners or other operatives across the diversity of conditions found in different COCOMs and operations. Thus, once we created the set of mission activities and socio-cultural capabilities described above, we created a diminished matrix consisting of the six phases as rows and the five categories as columns. This resulted in a more manageable matrix consisting of 30 cells. The socio-cultural capability categories and phases intersect at a more general level. Table 4 illustrates the process by depicting two of the five category columns and all six of the operational phases.

Table 4: Sample Intersections of Phases by SCCs

SCC Categories	Planning	Coordinating and Communicating
Operational Phases		
Phase 0: Shape	Develop compelling messages aimed at both potential adversaries and supporters that decreases negative and increases positive perceptions of U.S. actions in the region, minimizing the use of threats or punitive language if possible.	Gain access to appropriate communications media to target specific audiences in ways that reach them in time to influence their response.
Phase I: Deter	Develop messages that inform the population of: (1) how the U.S. military's demonstration of resolve deters enemy activity; and (2) how the preparatory actions they are observing contributes to a capable joint force.	Disseminate messages developed using a wide range of media to reach a diverse audience to include the local government, the military and the local population.
Phase II: Seize the initiative	Develop CCIRs for the commander aimed at understanding the local population demographics and cultural sites to help avoid unnecessary collateral damage. Develop messages to inform the population of the military operations to the extent possible.	Disseminate messages that inform the local population of the operations underway. Aim at convincing the populace that the military will not deliberately destroy homes and culturally significant sites.
Phase III: Dominate	Advise the commander to avoid collateral damage when conducting combat operations by identifying culturally significant sites as well as major population centers. Develop messages to inform the population of military operations to the extent possible.	Disseminate messages that inform the local population of the operations underway. Aim at convincing the populace that the military will not deliberately destroy homes and culturally significant sites.
Phase IV: Stabilize	Advise the commander on ways to engage the population that account for cultural sensitivities. Provide humanitarian assistance and engage in nation-building.	Develop messages that convince the local population that the U.S. humanitarian assistance and nation building efforts are in their best interests.
Phase V: Enable civil authority	Partner with and provide support to the legitimate government while continuing to provide humanitarian assistance.	Disseminate messages developed using a wide range of media to reach a diverse audience to include the local government and military and the local population.

GAPS AND MITIGATION MEASURES

Our research consisted of reviewing the pertinent socio-cultural literature and military documents in addition to conducting several interviews of military and non-military personnel. In the process, we identified several gaps in the provision of SCCs, organized along our five major categories. We noted that not all gaps constituted a complete lack of capability. Some were due to limited access to required capabilities or limited knowledge that such capabilities existed. In addition, competing priorities resulted in precluding access. The following is a summary of our major findings and recommendations. Table 6 lists all the gaps we uncovered in this research along with the measures we recommended to mitigate the gaps arranged by Doctrine, Organization, Training and education, Materiel, Leadership, Personnel, and Facilities (DOTMLPF). We have omitted Facilities from the list as it had not been mentioned in any of our interviews. The numbers at the head of each column refers to the coded gap mitigation measure listed by DOTMLPF category in Table 5.

Table 5: Gap Mitigation Measures

DOTMLPF Category	No.	Gap mitigation measures
Doctrine	1	Incorporate SCC information into long-term strategic planning
	2	Develop doctrinal publication that provides guidance on how to do the above, and incorporates lessons learned from employment of SCCs over the last decade
Organization	1	Repeat deployment to same local area
	2	Utilize resident SCC expertise
	3	Expand and systematize cooperation with NGOs, OGAs, IGOs
	4	Systematize collaborative links with colleges, universities, research centers / “think tanks.”
	5	Increase pass-through of socio-cultural information during RIP/TOAs
Training and education	1	Pre-deployment training in socio-cultural aspects of local AOI
	2	Training in proper use / caveats of using translators and interpreters
	3	Develop training in “culture-general” skill sets (non-verbal communication, cultural relativism, empathy, etc.)
	4	Training track in data acquisition, verification, analysis, interpretation.
Materiel		Develop standardized field-based platforms for data acquisition (geospatial, survey, social network, etc.), as well as both field-based and centralized standardized analysis and interpretation tools with broad accessibility
Leadership	1	Education tracks for senior military leaders in social science and related fields
	2	Bolster existing education efforts in SCC areas in academies and civilian schools
	3	Support Commanders’ individual initiatives for pre-deployment training in socio-cultural arena
Personnel	1	Hiring tracks to aid recruitment of military personnel with SCC skills
	2	Career tracks that incentivize development of SCCs and balance rewards for kinetic and non-kinetic skills
	3	Include assessments of SCC-related skills in officers’ performance evaluations
	4	Develop review boards to assess quality of socio-cultural information, education, and training
	5	Develop reference list of regional and thematic experts for consultation on socio-cultural issues

Table 6: Gaps and Mitigation Measures

	Doctrine		Organization					Training and education				Materiel	Leadership			Personnel				
	1	2	1	2	3	4	5	1	2	3	4		1	2	3	1	2	3	4	5

	Doctrine		Organization					Training and education				Materiel	Leadership			Personnel				
	1	2	1	2	3	4	5	1	2	3	4		1	2	3	1	2	3	4	5
Planning																				
<i>Lack of appropriate personnel</i>				X		X							X	X		X	X	X		X
<i>Plans not sensitive to local context</i>			X		X	X	X	X			X	X								
<i>Plans based on hunches</i>			X		X	X	X	X			X	X								
<i>Kinetic bias</i>	X	X															X			
Coordination and Communication																				
<i>Problems w/in military</i>								X				X								
<i>Ineffective outreach</i>			X		X		X	X		X								X		
<i>Military vs. non-mil</i>					X	X				X								X		
Gaining Detailed Cultural Awareness																				
<i>Unsure where to access</i>				X		X	X	X				X	X							X
<i>Validity checking</i>						X						X							X	
<i>Unclear roles</i>													X							
<i>Leveraging experts</i>				X	X	X		X							X					X
<i>Lack of culture-general tools</i>										X										
<i>Local AOI information</i>			X		X		X	X			X	X			X					X
<i>Methodological training and expertise</i>						X				X		X	X	X		X	X	X		
<i>Lack of appropriate personnel</i>				X		X							X			X	X	X		X
Engaging the Population																				
<i>Basic language skills</i>			X					X						X	X	X	X	X		
<i>Fluency</i>			X										X	X		X	X	X		
<i>Interpreters/translators</i>									X									X		
<i>Etiquette</i>			X					X						X				X		
<i>Relationship-building</i>			X		X		X	X		X								X		
<i>Conflict resolution</i>			X				X	X		X								X		
<i>Conflict resolution</i>								X		X				X	X	X		X		
<i>Negative perceptions of military</i>			X		X		X	X		X								X		
Impact Assessment																				
<i>Not incorporated in SOP</i>	X	X										X								
<i>Lack of appropriate personnel</i>				X		X				X			X			X	X	X		X
<i>Poor/biased assessments</i>						X				X		X							X	
<i>No assessments of SCC training</i>						X													X	

CONCLUSIONS

In **Planning**, we identified challenges in finding the appropriate personnel to use socio-cultural knowledge in the service of strategic planning. Our research also indicated that the continuing use of “one size fits all” directives or plans based on hunches rather than data was a problem in the use of SCCs in planning missions and activities. In **Coordination and Communication**, we identified several gaps related to SCCs, many related to cross-cultural

interactions between military and non-military organizations as well as difficulties coordinating among different military entities. Negative perceptions of the U.S. Military by international government organizations (IGOs) and non-governmental organizations (NGOs) present a significant challenge to prospects for supplementing SCCs through collaborative effort with other organizations operating in the area of interest.

In **Gaining Detailed Cultural Awareness**, we found there to be a gap in the ability to access up-to-date, verified information about specific areas of interest, as well as a lack of systematic approaches and systems for obtaining, analyzing, storing, and sharing socio-cultural information. Similarly, we discovered a lack of sufficient access to military and non-military personnel with the methodological and area expertise to cover the wide range of areas of interest covered by the military. With respect to **Engaging the Population**, our research identified the need for increased training in languages and/or the use of interpreters and translators, as well as a better capacity to build productive relationships with local populations, engage local leaders, and resolve conflicts in culturally appropriate ways. We identified significant gaps in **Impact Assessments**, most notably the fact that impact assessment for the provision of SCCs is not systematically incorporated in standard operating procedures. Additionally, military socio-cultural education and training has not been systematically assessed.

Our research team used the DOTMLPF framework to provide recommendations to mitigate these gaps (Table 6). We recommend a doctrinal or doctrine-like publication that focuses on lessons learned in the provision of SCCs during the past decade, as well as the use of SCCs in long-range strategic planning. With respect to **Organization**, we recommend organizing repeated deployments to the same local areas as well as improving processes during Relief in Place/Transfer of Authority operations to transfer socio-cultural knowledge and ensure the proper hand-off of relationships with key leaders. As well as building greater socio-cultural capacities within the armed forces, we recommend expanding and systematizing relationships with NGOs and IGOs, as well as universities and think tanks that provide access to SCCs.

With respect to **Training**, we recommend expanding the provision of socio-cultural knowledge specific to the precise area pre-deployment (rather than regional or national training), expanding training in the use of interpreters and translators, and also expanding training in “culture general” skill sets, including empathy, suspension of judgment, and the ability to adapt to new cultural environments. For the category of **Materiel**, we recommend establishing systems for the standardized acquisition, analysis, and interpretation/validation of both existing sources of information and new streams of data from units deployed and collecting data in the field. As a model, we point to the SERENGETI system developed at AFRICOM. SERENGETI is an Africa-centric data repository that is populated with open source data according to a specific architecture and taxonomy. It was developed in AFRICOM’s knowledge development division (Lee, undated). We make several recommendations pertaining to **Personnel**. First, we suggest instituting specialized hiring efforts and promotional tracks to both attract personnel with socio-cultural skill sets and to reward the expansion of these skills during service members’ careers. We also recommend making SCC-related parameters a systematic part of officers’ performance evaluations and implementing efforts to monitor and assess socio-cultural training and education.

On the whole, we found that substantial progress has been made in terms of building socio-cultural capabilities, and that when gaps emerged there were at least limited attempts to address unmet needs. Personal initiative accounted for a large percentage of these ad-hoc efforts and attests to the growing appreciation among military personnel for the important role of socio-cultural capabilities in achieving mission goals. Although we provide recommendations for how to fill specific gaps, finding ways to encourage and support personal initiatives will not only help meet SC needs as they arise in specific locales, but will nurture the development and testing of original ideas that could be used to address SC needs more broadly. Issues that on the surface appeared unrelated to socio-cultural issues were often the main impediments to building greater SC knowledge and skill sets: a position’s promotion potential, the length and repetition of deployments to an area, the commissioning of assessments, and the quality of the rotation process. Resolving these over-arching issues would likely produce marked improvements over multiple SCC categories and as such, might be cost-effective areas for focusing future SCC reform efforts.

REFERENCES

- Abbe, Alison, Lisa M. Gulick, and Jeffrey L. Herman, *Cross-Cultural Competence in Army Leaders: A Conceptual and Empirical Foundation*, ARI Study Report No. 2008-01. Arlington, Va.: U.S. Army Research Institute for the Behavioral and Social Sciences, 2008.
- Abbe, Allison, and Stanley M. Halpin, "The Cultural Imperative for Professional Military Education and Leader Development," *Parameters*, Vol. 39, No. 4, Winter 2009-2010, pp. 20-31.
- Black, J.S. and L. W. Porter, "Managerial Behaviors and Job Performance: A Successful Manager in Los Angeles May Not Succeed in Hong Kong," *Journal of International Business Studies*, Vol. 22, 1991, pp. 99-113.
- Brett, J.M., and S. Okumura, "Inter- and Intracultural Negotiation: U.S. and Japanese Negotiators," *Academy of Management Journal*, Vol. 41, 1998, pp. 495-510.
- Byman, Daniel, "Uncertain Partners: NGOs and the Military," *Survival*, Vol. 43, No. 2, Summer 2001, pp. 87-114.
- Chief of Naval Operations, *Language Skills, Regional Expertise and Cultural Awareness Strategy*, Washington, D.C.: U.S. Navy, January 2008.
- Department of Defense, *The Defense Language Transformation Roadmap*, January 2005.
- Flynn, Michael T., Matt Potting, and Paul D. Batchelor, *Fixing Intel: A Blueprint for Making Intelligence Relevant in Afghanistan*, Washington, D.C.: Center for a New American Security, 2009.
- Gelfand, M.J., and N. Dyer, "A Cultural Perspective on Negotiation: Progress, Pitfalls, and Prospects," *Applied Psychology: An International Review*, Vol. 49, 2000, pp. 62-100.
- Gelfand, M.J., M. Erez, and Z. Aycan, "Cross-Cultural Organizational Behavior," *Annual Review of Psychology*, Vol. 58, 2007, pp. 479-514.
- Headquarters, Department of the Army, *Counterinsurgency*, FM 3-24 / MCWP 3-33.5, Washington, D.C., December 2006.
- Headquarters, Department of the Army, *Army Foreign Language Program*, Army Regulation (AR) 11-6, Washington, D.C., 31 August 2009.
- Headquarters, Department of the Army, Training and Doctrine Command (TRADOC), *Operational Environment 2009-2025*, Ft. Monroe, Va., TRADOC OE White Paper V-6, August 2009.
- Jager, S.M., "On The Uses of Cultural Knowledge," *Strategic Studies Institute of the U.S. Army War College: Monograph*, 2007.
- Jandora, J., "Military Cultural Awareness: From Anthropology to Application," *Landpower Essay No. 06-3*, The Institute of Land Warfare, November 2006.
- Joint Chiefs of Staff, *Joint Strategic Capabilities Plan*, Washington, D.C.: Department of Defense, 2008.
- Joint Chiefs of Staff, *Joint Operations*, Joint Publication 3-0, Department of Defense, Washington, D.C.: March, 2010.
- Joint Chiefs of Staff, *Universal Joint Task List*, CJCSM 3500.04C, Washington, D.C.: Department of Defense, 2002.
- Kjeldsen, Stig, *Training Impact Assessment of the United Nations Civil-Military Coordination (UN-CMCoord) Course*, Geneva, Switzerland: United Nations Office for the Coordination of Humanitarian Affairs, 2006.
- Lee, Bari A., IKD Knowledge Development Division, SERENGETI, U.S. Africa Command, "US Africa Command: Delivering Cultural Competence," Briefing, Stuttgart, Germany, Not Dated.
- LeGree, Larry, *The Battle for the Minds: IO and COIN in the Pashtun Belt*, *Military Review*, September–October 2010, pp. 21-31.

- Medby, Jamison Jo and Russell W. Glenn, *Street Smart: Intelligence Preparation of the Battlefield for Urban Operations*, MR-1287-A, Santa Monica, Calif.: RAND Corporation, 2002. As of February 25, 2011: http://www.rand.org/pubs/monograph_reports/MR1287.html.
- Morgan, David L., ed., *Successful Focus Groups: Advancing the State of the Art*, Thousand Oaks, California: Sage Publications, 1993.
- Morris, M.W., K. Y. Williams, K. Leung, R. Larrick, M. T. Mendoza, D. Bhatnagar, J. Li, M. Kondo, J. Luo, and J. Hu, "Conflict Management Style: Accounting for Cross-National Differences," *Journal of International Business Studies*, Vol. 29, 1998, pp. 729-747.
- Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, *Report of the Defense Science Board Task Force on Understanding Human Dynamics*, Arlington, VA, 2009.
- Ohbuchi, K., and Y. Takahashi, "Cultural Styles of Conflict Management in Japanese and Americans: Passivity, Covertness, and Effectiveness of Strategies," *Journal of Applied Social Psychology*, Vol. 24, 1994, pp. 1345-1366.
- Rochelle Davis, "Culture as a Weapon System," *Middle East Report*, 225, Summer 2010, <http://www.merip.org/mer/mer255/davis.html>.
- Salmoni, Barak A. and Paul Holmes-Eber, *Operational Culture for the Warfighter: Principles and Applications*, Quantico: Marine Corps University Press, 2008.
- Shaley, Orit, "Constructing Multi-Dimensional Perspective of Multiculturalism in Armed Forces," in Cees M. Coops & Tibor Szvircsev Tresch, eds., *NATO Defense College: Cultural Challenges in Military Operations*, Rome, Italy: NATO Defense College, 2007, pp. 130-140.
- Sheer, V.C., and L. Chen, "Successful Sino-Western Business Negotiation: Participant's Accounts of National and Professional Cultures," *Journal of Business Communication*, Vol. 40, 2003, pp. 50-85.
- Stenmark, T., *Looking for Gold Nuggets in the Melting Pot: Language, Cultural Awareness, and the Fourth Generation Warrior*, Maxwell Air Force Base, AL: Air Command and Staff College, April 2006.
- Sullivan, J., R. B. Peterson, N. Kameda, and J. Shimada, "The Relationship Between Conflict Resolution Approaches and Trust—A Cross Cultural Study," *The Academy of Management Journal*, Vol. 24, 1981, pp. 803-815.
- Tyler, T.R., E. A. Lind, and Y. J. Huo, "Cultural Values and Authority Relations: The Psychology of Conflict Resolution Across Cultures," *Psychology, Public Policy, and Law*, Vol. 6, 2000, pp. 1138-1163.

PSYMDEV, a System Intended for Assisting the Military Analyst to Construct Psychological Messages

Colette Faucher and Malika Machtoune

Laboratoire des Sciences de l'Information et des Systèmes (LSIS)
Aix-Marseille University
Marseille, FRANCE

ABSTRACT

In asymmetric conflicts, the Armed Forces generally have to intervene in countries where the internal peace is in danger. They must make the local population an ally in order for them to be able to deploy the necessary military actions with its support. For this purpose, psychological operations (PSYOPS) are used to shape people's behaviors and feelings by spreading out messages thanks to different media (tracts, loudspeakers, video clips, etc.). In this paper, we present PSYMDEV (PSYchological Message DEViser), a system that helps the military analyst to construct messages that trigger specific feelings in members of the population selected by social criteria like age or political opinion and called the *info-targets*. Given such a sociocultural group and a feeling that the latter must feel, the system provides a *twofold-situation* that consists of, on the one hand, a *categorization-situation* meant to induce a positive or negative initial state of mind in the info-targets depending on the type of feeling to be triggered through a psychological mechanism inspired by theories stemming from Social Psychology and an *action-situation* aiming at effectively triggering the specific feeling through a psychological process explained by the Intergroup Emotion Theory, an extension of the Appraisal Theory of Emotions. These situations are illustrated by means of images or a film or some auditive elements, thanks to adapted media generally used by the military like tracts or video clips, for example. Therefore, the twofold-situation gives birth to a psychological message intended to trigger a feeling. After presenting the theories underlying the system and its overall structure and functioning, we more specifically focus on the conception of a categorization-situation.

Keywords: Psychological Operations, Appraisal Theory, Intergroup Emotion Theory, Social Categorization Approach, Common Ingroup Identity Model, Sociocultural Identity, Situation, Sociocultural knowledge.

INTRODUCTION

In asymmetric conflicts, the Armed Forces generally have to intervene in countries where the internal peace is in danger. They must make the local population an ally in order for them to be able to deploy the necessary military actions with its support. For this purpose, psychological operations (PSYOPS) are used to shape people's behaviors and feelings by spreading out messages thanks to different media (tracts, loudspeakers, video clips, etc.). In this paper, we present PSYMDEV (PSYchological Message DEViser), a system that helps the military analyst to construct messages that trigger specific feelings in members of the population selected by social criteria like age or political opinion and called the *info-targets*. Given such a sociocultural group and a feeling that the latter must feel, the system provides a *twofold-situation* that consists of a *categorization-situation* meant to induce a positive or negative initial state of mind in the info-targets depending on the type of feeling to be triggered, through a

psychological mechanism inspired by theories stemming from Social Psychology and an *action-situation* aiming at effectively triggering the specific feeling, through a psychological process explained by the Intergroup Emotion Theory, an extension of the Appraisal Theory of Emotions. These situations are illustrated by means of images or a film or some auditive elements, thanks to adapted media generally used by the military like tracts or video clips, for example. Therefore, the twofold-situation gives birth to a psychological message intended to trigger a feeling. In this paper, we present the theories stemming from the Psychology of Emotions and Social Psychology that underlie PSYMDEV. We then explain the notion of salience of a social identity that is essential to our system and we report a few computational works addressing the notion of social identity. We explain the purpose of PSYMDEV, its functioning and its overall structure, before presenting the modeling of the main components involved in the system. We then focus on the conception of a categorization-situation that we illustrate thanks to the presentation of an example. Our conclusion ends the paper.

THE PSYCHOLOGICAL THEORIES UNDERLYING PSYMDEV

The Appraisal Theory

The basic question that motivated appraisal theories is why people react to the same things differently. Even if they live the same, or a similar situation, all people will react in slightly different ways depending on their perception of the situation. These perceptions elicit various emotions that are specific to each person.

Dating back to the 1940s and 1950s, Arnold wanted to “introduce the idea of emotion differentiation by postulating that emotions such as fear, anger, and excitement could be distinguished by different excitatory phenomena” (Arnold, 1954). For this purpose, she developed her “cognitive theory” in the 1960s, according to which, in an individual, an emotion starts with their appraisal of the current situation that arouses both the appropriate actions and the emotional experience itself, so that the physiological changes accompany, but do not initiate, the actions and experiences (Arnold, 1960).

Lazarus pursued research in the same direction as Arnold and tried to answer the following questions: what is the nature of appraisals which underlie separate emotional reactions and what are the determining antecedent conditions of these appraisals (Lazarus, Averill and Opton, 1970), (Lazarus, 1991). He distinguished two types of appraisals: the *primary appraisal*, directed at the establishment of the significance or meaning of the event to the individual in terms of its motivational relevance, and the *secondary appraisal*, directed at the assessment of the individual’s ability to cope with the consequences of the event.

On the basis of those pioneering works, Scherer finally succeeded in developing a consensual theory of emotion. The Appraisal Theory of Emotions (Scherer, Schorr, and Johnstone, 2001) postulates that the emotion that a human being (or any living organism) experiences in a given situation, whether the latter is lived, imagined or remembered, results from the assessment of a few cognitive criteria that can be categorized into four families that answer the following questions: Is the situation relevant to me, does it affect my well-being? (*Relevancy*). What are the implications of the situation and how do they affect my well-being and my short-term and long-term goals? (*Implications*). To what extent can I face the situation or adjust to its consequences? (*Coping*). What is the significance of the situation as regards my social norms and my personal values? (*Normative Significance*). Scherer’s version of the appraisal theory includes 16 specific criteria that belong to the previous categories (Scherer, Schorr, and Johnstone, 2001). A combination of values of the criteria determines in a unique way a specific feeling, but the assessment of the different criteria is subjective. Thus, the same situation can trigger different emotions in people with different traits and coming from different cultures. Only the correspondence between a combination of values and a specific feeling is universal.

The Social Identity Approach

The Social Identity Approach comprises both the Social Identity Theory and the Self-Categorization Theory (Tajfel and Turner, 2004). It addresses the ways in which people perceive and categorize the others and themselves. Individuals form self-conceptions that are based on a twofold identity: a personal or self identity, and a collective identity. Personal or self identity refers to our unique, personal qualities such as our personal beliefs, our abilities and skills, etc. The collective self includes all the qualities that arise from being part of a society, culture, family, groups, clubs, etc. A distinction is made between people belonging to the same group(s) as us (the ingroup(s)) and people not belonging to it (them) (they belong to our outgroup(s)). Besides, individuals actually have multiple social identities that can be activated (that is, can become salient) and prevail over the others depending on the context within which they stand. For example, if you are a male professor, a father of three children and a supporter of the

Lakers, in the context of a game of your favorite team, your identity as a supporter will be salient, whereas, in a professional context, your identity as a professor will prevail over the others and in a family context, your identity as a father will be the most important.

The Intergroup Emotion Theory

The Intergroup Emotion Theory (Mackie, Devos, and Smith, 2000) suggests that the emotional experience of a person as a member of a group is identical to the experience they live as an individual, as it is described in the Appraisal Theory. The only difference is that the Intergroup Emotion Theory implies the cognitive evaluation of a situation that concerns the social identity of an individual (traits that connect the person to social groups) instead of involving their personal identity (the aspects that make the person unique). According to Scherer and Garcia-Prieto (Garcia-Prieto and Scherer, 2006), the criteria that are sensitive to the social identity of a person are the ones that have a social connotation like the conduciveness/obstructiveness, the causal agency, the responsibility, the control, power and adaptability and the compatibility with the social norms of the group(s) the person belongs to (Sander and Scherer, 2009). Moreover, the feeling triggered depends on the strength of the identification of a person with their social group(s). Table 1 shows the extension of Scherer's appraisal dimensions to the intergroup level.

The different assessment criteria mentioned in this table answer the following questions :

- **Goal conduciveness/obstructiveness:** Does this situation constitute an obstacle to my goal as a member of my ingroup(s) or does it facilitate it?
- **Causal agency:** Who is responsible for the failure or the success of this situation?
- **Responsibility:** What is the intentionality?
- **Control:** Does(do) my ingroup(s) control the consequences of the situation?
- **Power:** Does(do) my ingroup(s) have any power in the context of this situation?
- **Adaptability:** Will my ingroup(s) adapt to the consequences of this situation?
- **Normative significance:** Is this situation in accordance with the standards of my ingroup(s) (norms and values)?

Table 1 : Extension of Scherer's appraisal dimensions to the intergroup level (Garcia-Prieto and Scherer, 2006)

	Propositions for the interpersonal context	Propositions for the intergroup context
Appraisal dimension	Self vs. other	When a social identity is salient Ingroup vs. outgroup
<u>Implication assessment</u>		
Goal conduciveness/obstructiveness	Likely to be appraised in terms of personal goals	Likely to be appraised in terms of salient ingroup's goals
Causal agency	Likely to be appraised in terms of	Likely to be appraised in terms of
Responsibility	the self in relation to other(s) interpersonally	the salient ingroup in relation to the target outgroup
<u>Coping potential</u>	Likely to be appraised in terms of personal	Likely to be appraised in terms of
<u>Determination</u>	control, power and adaptability in regards to	the salient ingroup's control, power
Control	the consequences of the event	and adaptability in regards to the
Power		consequences of the event
Adaptability		
<u>Normative significance</u>	Likely to be appraised in terms of general	Likely to be appraised in terms of
<u>Evaluation</u>	social norms	salient ingroup's norms
External standards		

The Common Ingroup Identity Model

Within the framework of the Social Identity Theory, individuals are characterized by their social and cultural identity involving a set of external and internal traits, behaviors, beliefs, values, goals that are crucial for the assessments they make about a given situation. The Social Categorization Theory affirms that people show favoritism towards members of their ingroup(s) compared to those of their outgroup(s). Conversely, they are inclined to prejudice and discrimination towards their outgroup members. The goal of the Common Ingroup Identity Model is to reduce intergroup conflicts by reducing this bias. The idea is to modify in people the perception of their ingroup(s) and outgroup(s) by modifying their categorization. That way, people that were classified as outgroup

members will be reclassified as ingroup members (Dovidio et al., 2006), (Crisp and Hewstone, 2007). The solution that has been proposed is to induce the perception of a common ingroup which is achieved by increasing the salience of an existing common superordinate membership (e.g. a school, a company, a nation) or by making salient new elements that redefine group relations (e.g. common goals or shared fate). Identities are thus assumed to be structured within a hierarchy of specialization.

More specifically, according to Crisp (Crisp and Hewstone, 2007), the three best possible strategies to reduce the bias between two groups with their respective identities, are:

- If both identities have a common superordinate identity, the stress must be put on the salience of this identity, but without erasing the salience of the subordinate identities. As a matter of fact, for some type of strong identities like race or for minority groups, the bias may be increased instead of decreased due to the fact that the subordinate groups may feel that their own identity is threatened (Sander and Scherer, 2009).
- If there is no superordinate identity, an alternative is to have both groups collaborating on a common task, with an equal status and towards a common goal (strategy inspired by Allport (Allport, 1954)).
- Yet another option can be to decategorize people and then to avoid intergroup bias, by having members of both groups interact individually with each other and no longer as members of their respective groups.

NOTION OF SALIENCE OF A SOCIAL IDENTITY

The salience of a social identity is determined by its *accessibility* and its *situational fit* (Turner et al., 1987). The *accessibility* of an identity is defined as the easiness or the spontaneity with which this identity is recalled. The *situational fit* is the extent to which a particular context activates an identity (Kopecky, Bos, and Greenberg, 2010). The *accessibility* of an identity for an individual partially depends on the past experiences of the person, on their expectations and present motives, on their values, goals and needs, but some social identities are naturally more accessible whatever the individuals (e.g. professional status as regards marital status). Besides, some identities are permanent like ethnic group, others are flexible like political opinion. On the other hand, the subjective identification of an individual to their ingroup identity plays a crucial role in determining the accessibility of this individual to their identity.

The *situational fit* has two aspects: the *comparative (structural) fit* and the *normative fit*.

- The *comparative fit* is based on the MetaContrast Principle that defines the fit in terms of the emergence of the sharpness of a category (an identity) on a contrasting background. Let's imagine the following situation: you attend a basket-ball game, Boston Celtics vs. Chicago Bulls. In the audience that stands on the bleachers of the stadium, you catch sight of several people that wear Boston Celtics' supporter T-shirts, while most of the others wear Chicago Bulls' supporter T-shirt or a neutral jersey. Given that you wear a Boston Celtics' supporter T-shirt, you self-categorize yourself as such and you see the people wearing a Chicago Bulls' jersey as an outgroup.

Among a group of individuals, people who look alike tend to stand out among the other people, thus allowing their categorization and the activation of the corresponding self-categorization.

- The *normative fit* refers to the fact that for an individual to categorize spontaneously people into a social group, the people must behave in accordance with the typical way the members of the group are expected to behave. They must conform to the expectations an individual has concerning their external appearance and their behavior, the latter being even more important than the external appearance. Let's take the following example. You are in a railway station. Among the crowd, you catch sight of several people wearing very formal clothes, while most of the others wear informal clothes. Wearing yourself used jeans and a T-shirt, you categorize yourself among the careless people and view the strict ones as an outgroup. All of a sudden, there's an argument between the ticket seller, who is a black person, and a female customer. The latter pronounces a racist insult. Among the witnesses, some, regardless of their clothing, take sides for the ticket seller and others for the lady. Your anti-racist identity then becomes salient and you take sides for the ticket seller. It's then this last identity that dictates your behaviors. It has become salient all the more easily that it is important in your eyes (much more than an identity in terms of clothing). You then consider the members of the racist group as similar to each other, although they wear different clothes and this all the more easily

that they conform to the image you have about the members of this social category.

SOCIAL IDENTITY MODELING

Few models focus on the computational representation of social identities, (Kopecky, Bos, and Greenberg, 2010). The *Simulate Cultural Identities for Predicting Reactions to events* (SCIPR) model, concentrates on political opinions as flexible identities. It uses a simple social network of influence. The Salzarulo's *MetaContrast* model illustrates the phenomenon of discrimination as defined in Turner's theoretical approach to self-categorization. This model shows how polarization and extremism can occur due to the combination of attraction to ingroups and repulsion of outgroups. The *Political-Science Identity* (PS-I) model aims at establishing a link between permanent identities like cultural identity and flexible identities like political identity to merge geographic clusters. Finally, the *Social Identity Look-Ahead Simulation* (SILAS) model is dedicated to permanent and flexible identities like ethnical, political and religious memberships. The authors attempt to determine how internal conflicts between identities may be solved by modeling common enemy dynamics.

Thus, all these models have the advantage to highlight the dynamic character of social identification. However, they present some lacks, because they fail at investigating carefully enough a crucial characteristic of social identities, namely their salience. In particular, none of these models proposes a faithful representation of the notion of situational fit.

PURPOSE AND PRESENTATION OF PSYMDEV

In the context of asymmetric conflicts where the Armed Forces must intervene, they need to make the local population an ally. For that purpose, they spread out psychological messages to generate the desired thoughts, feelings and behaviors in members of the population selected by social criteria like age or educational level (the *info-targets*). PSYMDEV aims at helping the military analyst construct messages that trigger in the info-targets specific feelings directed towards another social group (the *aim-group*, that may be identical to the info-target group) like anger or fear or absolute feelings like despair or boredom. The inputs of the system are, on the one hand, the info-targets and possibly the aim-group, that are characterized by their respective values for social criteria like age, gender, social status, etc., and, on the other hand, a feeling that the info-targets must feel. This feeling is directed towards the aim group if it is specified. The output of the system is a twofold-situation described by means of a *categorization-situation* and an *action-situation* that will induce the feeling in the info-targets. These situations are illustrated by means of images or a film or some auditive elements, thanks to adapted media generally used by the military like tracts or video clips, for example. Therefore, the twofold-situation gives birth to a psychological message. The process that will trigger the feeling comprises two steps.

First Step of the Process

The first step of the process consists in inducing a positive or a negative initial state of mind in the info-targets depending on the type of feeling to be triggered. The underlying psychological mechanism is based on the Social Categorization Approach. Indeed, as we showed in the section devoted to this approach, people tend to show favoritism towards the members of their ingroup(s) and to feel prejudice towards the members of their outgroup(s). So, if the feeling to trigger is positive, the idea is to shape the perception of the info-targets in order for the aim-group to be considered as an ingroup. For that purpose, making salient the common social identities that both group share or in the absence of such identities, highlighting a common goal or fate for both social groups, will induce the impression of belonging to a common ingroup and will then imply a positive bias from the info-targets towards the aim-group. This strategy stems from the Common Ingroup Identity Model. It is the purpose of the categorization-situation to highlight the salience of the relevant social identities. This salience must be expressed through their normative and their structural fit. Thus, the categorization-situation of a twofold situation must include representatives of the aim-group and possibly the info-target group having the typical appearance and behavior of their respective relevant identities (normative fit). The salience of the concerned identities is also increased by introducing in the categorization-situation a contrasting background (structural fit). If the feeling to trigger is negative, conversely, the perception of the info-targets must be shaped in order for them to consider the aim-group

as an outgroup and to feel prejudice towards it. For this purpose, the categorization-situation will outline for both groups the social identities they don't share with each other. If such identities don't exist, the identities of both groups will be equally salient and no initial state of mind is induced prior to the triggering of the specific feeling that will be achieved by the action-situation directly.

Second Step of the Process

The second step of the process aims at effectively triggering the concerned specific feeling. The underlying idea comes from the Appraisal Theory of Emotions and the Intergroup Emotion Theory: through the evoking, by means of the action-situation, of a situation that the info-targets will assess as members of their social group, they will feel a feeling corresponding to the result of this appraisal. Indeed, the Appraisal Theory clearly mentions that a situation needs not be lived, but may only be imagined or remembered for individuals to assess it and feel the corresponding feeling. Actually, the Intergroup Emotion Theory is used in the reverse order. The feeling to be triggered, that is given as an input to the system, determines a set of values for the assessment criteria and the system designs the action-situation so that its characteristics match the values of the assessment criteria. For example, if the assessment of the criterion « the action matches the values of the social group of the info-targets » is positive, then the action performed in the action-situation must respect those values. The action-situation designed that way should then trigger the proper feeling in the info-targets.

In this paper, we will focus only on the first step that we will described in detail in the following sections.

Overall Structure of the System PSYMDEV

Figure 1 presents the overall structure of the system PSYMDEV.

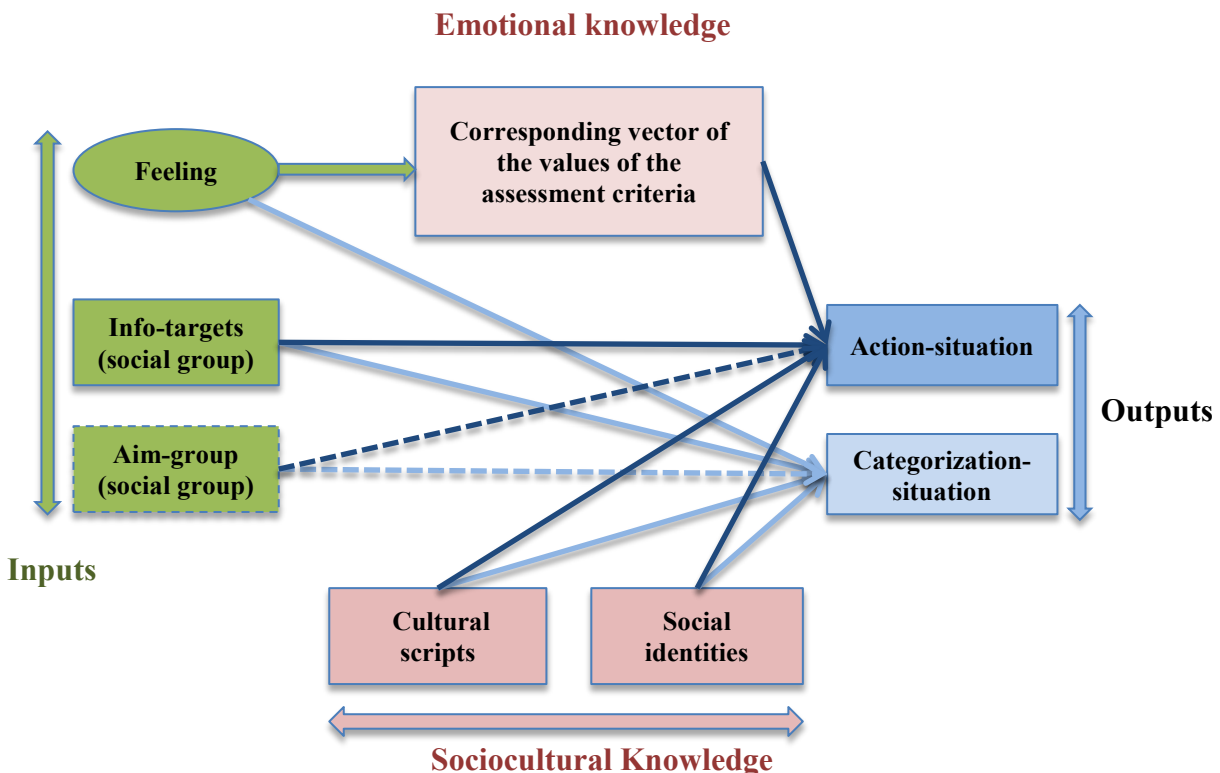


Figure 1. Overall structure of the system PSYMDEV

DESCRIPTION OF PSYMDEV COMPONENTS

Emotional Knowledge

We previously mentioned that, according to the Appraisal Theory, situations are assessed along 4 dimensions: Relevancy, Implications, Coping and Normative Significance. In the context of the Intergroup Emotion Theory, there is no need to assess the relevancy of a situation, because our purpose is to conceive a situation that is relevant to the info-targets. So we focus on the assessment criteria that belong to the three other families. Like the other components of the system, feelings are described using a frame-based representation (Minsky 1975). Given the assessment criteria shown in table 1, a feeling is an instance of **frame Feeling** defined by: slot **Conduciveness/Obstructiveness**, slot **Causal agency**, slot **Responsibility**, slot **Control**, slot **Power**, slot **Adaptability**, slot **Normative significance**. Slot Conduciveness/Obstructiveness can take the values “obstruction”, “facilitation” or “open”. Slot Causal agency can take the values “ingroup”, “outgroup”, “nature”, “open”. Slot Responsibility can take the values “intentional”, “chance”, “open”. Slot Control, Power, Adaptability and Normative significance can take the values “very high”, “high”, “medium”, “low”, “very low”.

Social Groups

Social groups are defined along 12 different dimensions or *social criteria*: age (c_1), gender (c_2), professional status (c_3), ethnic group (c_4), religion (c_5), political opinion (c_6), location (c_7), social status (c_8), educational level (c_9), marital status (c_{10}), parental status (c_{11}) and language (c_{12}). Each social criterion has a set of possible values: $V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, V_{10}, V_{11}$ and V_{12} .

A given social group may have an undefined value for one or more criteria. Let G be a social group,

$G=(v_1, v_2, \dots, v_{11}, v_{12})$, means that criterion c_i has an undefined value for social group G .

We define the set of social criteria: $C=\{c_1, c_2, \dots, c_{11}, c_{12}\}$. For each social criterion, c_i , V_i is structured hierarchically so that, let G_{v_1} and G_{v_2} be the sets of the members of the social group defined respectively by value v_1 and value v_2 for criterion c_i if $v_1 \leq v_2$ (v_2 is an ancestor of v_1), then $G_{v_1} \subset G_{v_2}$.

A social group is described as an instance of frame **Social-group** defined by the following slots : **Age, Gender, Professional-status, Ethnic-group, Religion, Political-Opinion, Location, Social-status, Educational-level, Marital-status, Parental-status, Language**. The possible values of these slots depend on the culture and the country that constitute the framework within which PSYMDEV is used.

Sociocultural Knowledge

Social Identities

Each value of one of the 12 social criteria mentioned previously is defined as an instance of Social-Identity frame that represents the concept of social identity and is described by means of the following slots:

- **Label (L)**: value of the social criterion (e.g. “old” associated with criterion Age or “Muslim” associated with criterion Religion),
- **Criterion (C)**: name of the associated criterion (e.g. Age or Religion),
- **Appearance (A)**: describes clothes and accessories worn by an individual having this social identity,
- **Values (V)**: (e.g. honesty, friendship, etc.),
- **Norms (N)**: describes typical behaviors (i.e. holding books, writing on a blackboard, etc.), ways of speaking relative to the identity. Not only characteristics of an individual that require vision are taken into account. Audible characteristics are also mentioned. Indeed, depending on the medium that will be used by the military analyst to convey their message, audible and visual features may be used to illustrate a situation. For instance, if, via loudspeakers, the analyst wants to describe a situation involving an old man, it's his voice that will help the listeners to categorize him as such.
- **Goals (G)**: Maslow's Pyramid (Maslow, 1943) represents a hierarchy of goals, in the framework of the Theory of Human Motivation, that remained popular for a long time. However, most scientists nowadays reject this model, because it has been studied only for western populations. In other models of society, it is not relevant, that's why we chose to elaborate a hierarchy of goals that depends on sociocultural characteristics. For each specific identity, we have defined a hierarchy that determines which goals are important to this identity and what are their respective priorities.

We use Schank's classification of goals: Satisfaction Goals (SG), Enjoyment Goals (EG), Achievement Goals (AG), Preservation Goals (PG), Crisis Goals (CG), Instrumental Goals (IG) and Delta-Goals (DG), (Schank, 1977). Instrumental Goals and Delta-Goals will appear in the framework of the representation of situations, because they actually are pragmatic sub-goals used for reaching the other goals of higher level.

A social identity can also be defined by the conjunction of the values of several criteria taken together (e.g. a married Muslim male). In most cases, the frame that represents this compound identity is not mentioned explicitly, because the respective values of slots Appearance, Values, Norms and Goals of the different identities add up to form the compound identity. However, it may happen that the compound identity has to be defined explicitly:

- If the previous slots have contentious values between the different identities;
- If the compound identity has not quite the same semantics as the conjunction of the original identities (e.g. a Jewish mother has their own characteristics that neither a Jewish, nor a mother has) or there are additional values for the previous slots that are not coming from the individual identities.

Social identities relative to each social criterion are also organized within two hierarchies :

- a hierarchy of specialization: for most criteria like Political opinion or Religion, for example, their values, that is the corresponding social identities, can be ordered hierarchically. For instance, social identities relative to the criterion Religion can be ordered as « monotheistic religion » and « polytheistic religion », then « monotheistic religion » can be specialized into « christianism », « judaism », « islam », etc.
- a hierarchy that relies on power: minority social groups as regards to a given criterion are situated lower in such a hierarchy than majority social groups. For example, it is the case for small communities speaking a particular language within a country whose official language is different.

Goals

The goals associated with a social identity are organized within a hierarchy including all types of goals. At the top of such a hierarchy are situated high-level goals like Satisfaction Goals or Achievement Goals whereas at the bottom of the hierarchy are situated pragmatic goals like Instrumental Goals and Delta-Goals. These hierarchies may be multiple-inheritance hierarchies, that is, an Instrumental Goal can be a sub-goal of different higher-level goals.

Situations

In PSYMDEV, we have defined a base of cultural scripts. A script describes a temporal sequence of scenes or situations (Schank and Abelson, 1977) relative to a global event like « going to the restaurant », « visiting friends » or « voting for the elections », etc. This kind of scripts is indeed cultural in the sense that the different scenes composing them may vary across countries and cultures. Cultural scripts are classified in different categories depending on the context within which they take place. There are 5 contexts, the *family context*, the *professional context*, the *educational context*, the *political context* and the *religious context* that respectively highlight the following sets of social criteria, marital status, parental status, gender, ethnic group and age, professional status and social status, educational level, political opinion and religion. In other words, each context allows to make salient the corresponding social identities.

We model two kinds of action-situations: those that depict situations that must trigger in the info-targets feelings that are not directed towards other people (absolute feelings) like boredom or happiness and situations that must trigger feelings felt towards another social group (directed feelings). In this paper, we will focus on the latter. Among them, we distinguish three types of situations:

- SAGIT situations, where the aim-group is directly responsible for the info-targets' feeling, because the latter are the target of an aim-group's action. Both groups must then be actors of such a situation.
- SAG situations, where the aim-group is acting in such a way that the info-targets experience the feeling, but the action is not made directly against them. Only the aim-group is then acting in this kind of situation. The cause of the feeling is then at least partially due to the fact that the action goes against the info-targets' values, goals or norms.
- COM situations, where both groups are collaborating on a common task with an equal status and a common goal.

We saw previously that the categorization-situation aims at making salient relevant social identities in the aim-group and the info-targets. So such a situation depicts representatives of those groups dressed and acting in accordance with their social identities to be made salient. To design a categorization-situation, the social identities to be made salient are first determined, then the context that corresponds to these identities is chosen as well as scripts defined within this context. We won't detail in this paper how a possible action-situation is determined from such a script, but as we will see, it takes place in a location and a period of time mentioned in the script. So does the associated categorization-situation.

A twofold-situation is represented by a frame that includes a *categorization-situation* and an *action-situation* and the common components of these situations, in particular the period of time and the location where both situations are taking place as well as the info-targets and the aim-group.

Frame Twofold-situation is then described by the following slots: slot *Salient-Aim-Group-Identities* and slot *Salient-Info-targets-Identities* which mention the social identities to be made salient respectively in the Aim-group and in the Info-targets, slot *Type-of-feeling* that specifies the type of feeling to be triggered by the twofold-situation (absolute or directed feeling), slot *Type-of-situation* which mentions the fact that the action-situation is a SAGIT, SAG or COM situation, slot *Context* whose values are Family, Professional, etc., slot *Categorization-situation* and slot *Action-situation* which mention the corresponding instances of frames Categorization-situation and Action-situation (see below), slot *Info-targets* and slot *Aim-group* whose values are social groups, slot *Location* (whose value is « downtown », for example), slot *Period* (whose value is « election-time », for example).

Frame Categorization-situation is described by: slot *Info-targets' behavior* and slot *Aim-group's behavior* whose values are typical behaviors of the social identity(ies) to be made salient for each social group (values of slot *Norms* for these social identities), slot *Info-target's appearance* and *Aim-group's appearance* (values of slot *Appearance* for these social identities).

Frame Action-situation is described by slot *Action-agent*: the agent of the action (the Aim-group), slot *Action-target*, this slot may be empty if the Info-targets are not involved in the situation, slot *Attendees*: possibly a social group whose role is merely to provide a contrasting background, slot *Action*: an action typical of the agent's identity(ies) in the specific context of the situation and slot *Goal* which mentions the goal pursued in the framework of the situation. The values of the latter are pragmatic goals that specifically concern the described Action-situation. These goals correspond to Instrumental and Delta Goals and must be achieved to reach the other more general goals mentioned in the representation of social identities seen previously. The specialization hierarchies specific to social identities including all the types of goals they pursue (that we saw in the section describing the components of the system) are modeled aside from the frames representing the social identities and the Action-situations.

We can notice that Instrumental Goals mentioned in the description of an Action-situation are objective and concrete, but the linking to higher-level goals depends on the goal hierarchies proper to each social identity. In other words, the interpretation of an Action-situation in terms of high-level goals is specific to each observer as a member of a social group characterized by several social identities.

We won't detail the description of actions, that is not relevant during PSYMDEV's first step. Roughly, there is a correspondence between their characteristics and the values of the assessment criteria that define the feeling to be triggered.

CONCEPTION OF CATEGORIZATION-SITUATIONS

Which Social Identities to Be Made Salient in a Categorization-situation?

Let G_{it} and G_{ag} be respectively the info-targets and the aim-group: $G_{it} = (v_1, v_2, \dots, v_{12})$, $G_{ag} = (v'_1, v'_2, \dots, v'_{12})$. Let define a social identity $si_{ij} = (c_i, v_{ij})$ and the predicate $Salient(s, si)$ that means "situation s makes salient social identity si ". Let Sit be the Categorization-situation to be generated.

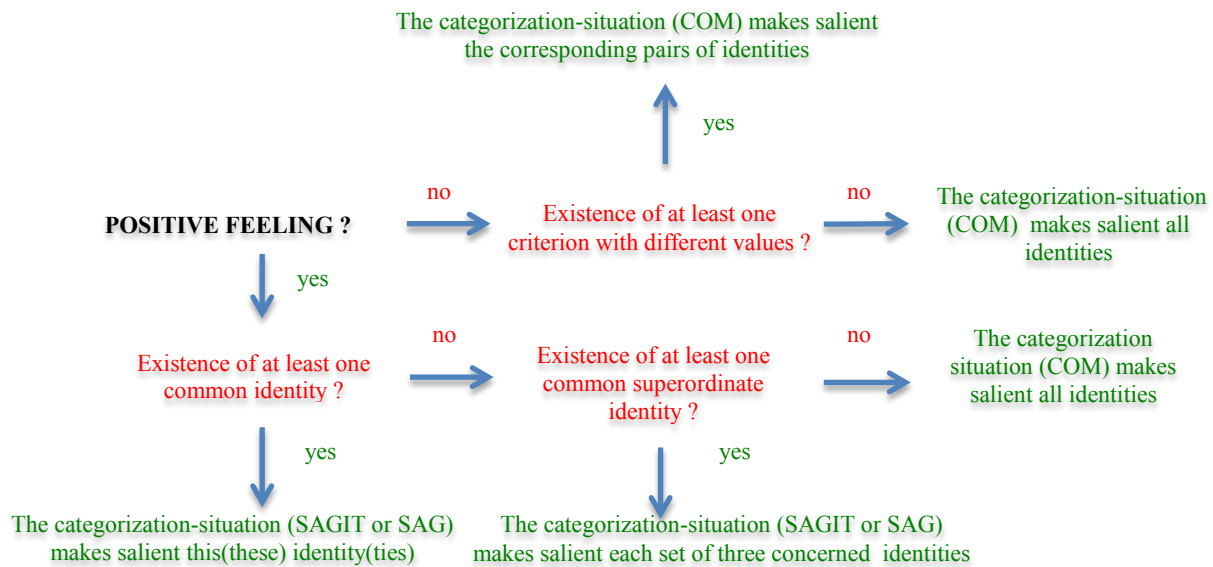


Figure 2. Determination of the salient social identities in a categorization-situation

How to Express the Salience of Social Identities?

In a previous section, we showed that the salience of an identity in a given situation is determined by its accessibility and its situational fit. An identity must be salient for a social group globally, so we cannot take into account the specific characteristics of an individual that make an identity more or less salient in their eyes, nor can we change the natural accessibility of an identity. So, we only have to put the stress on both components of the situational fit of a social identity in a given situation to make it salient: the *structural fit* and the *normative fit*.

In every case, the representative(s) of the aim-group and the possible representative(s) of the info-target group must have the external appearance and the behaviors corresponding to the typical values of their respective salient social identity(ies), to maximize the *normative fit*.

As for maximizing the structural fit, the idea is to emphasize the relevant social identity(ies) thanks to a contrasting background. Whatever the kind of feeling to be triggered, positive or negative, two cases may occur:

- the stress must be put on common identities between the identity(ies) of the info-target group and the aim-group (SAGIT situation) or on the identities of the aim-group (SAG situation). We then merely have to add in the categorization-situation of the twofold-situation, a social group in slot Attendees which has one or several identities consisting of a criterion whose value is identical for both groups, with a value that differs from the common value, to create a contrasting background.

- the stress must be put on identities that differ in the info-target group and the aim-group, then there is no need to create a contrasting background in a SAGIT or a COM situation, as the contrast is brought to each group by the presence of the other one in the categorization-situation. If it is a SAG situation, a social group with identities contrasting with those of the aim-group must be added in slot Attendees.

AN EXAMPLE ILLUSTRATING A FEW CASES

The context is Egypt a few months after the arrest of ex-president Muhammad Morsi. ~~The info~~ ^{target info} is the social group SGit consisting of married parent males, aged over 25, Egyptian, Muslim, pro-Morsi, living in Egypt, and speaking Arabic. The aim-group is the social group SGag consisting of married parent males, aged over 25, military, Egyptian, Muslim, Anti-Morsi, living in Egypt and speaking Arabic.

SGit=(over-25, male, 0, Egyptian, Islam, Pro-Morsi, Egypt, 0, 0, married, parent, Arabic)

SGag=(over-25, male, Military, Egyptian, Islam, Anti- Morsi, Egypt, 0, 0, married, parent, Arabic).

Different social identities are defined, for instance (we use previous abbreviations).

Appearance, values, norms, goals

Pro-Morsi (L: Pro-Morsi, C: Political Identity, A: civilian clothes, V: Justice, Freedom of expression, Democracy, N: to shout slogans, to carry Morsi's portraits, G: (AG, to restore Morsi's duties), (IG, to demonstrate)))

Muslim (L: Muslim, C: Religion, V: Family, Fraternity, N: to pray as a body, to give money or food to poorer than oneself, G: (AG, to achieve one's Muslim duties), (IG, to pray at the mosque, to attend the Friday preachment, to fulfill the charity for the poorest (zakat)))

Egyptian (L: Egyptian, C: Ethnic Group, V: Sharing, Honor, Respect, N: to offer bed and board, to offer a present when invited)

Egyptian male (L: (Egyptian, Male), C: (Ethnic Group, Gender), N: to support financially one's family, to make the important decisions, G: (AG, to marry, to have a family, to have a good job, to be respected, to have good relationships with others), (PG, to maintain a good harmony within one's family, to take care of one's family, to perpetuate traditions))

Military (L: Military Forces (Anti-Morsi), C: Professional status, A: soldier-clothes, V: Order, N: to watch, to supervise the civilians, G: to maintain order)

In the categorization-situation, all the representatives of the identities that must be salient have to conform to their respective typical appearance and norms and must be highlighted by a contrasting background.

• If the feeling is positive:

The common identities: over-25, male, Egyptian, Islam, Egypt, married, parent and speaking Arabic must be made salient.

- either a situation within the Family Context: it can be a SAGIT situation: in this case, it pictures an individual from the Aim-Group, a soldier, over 25, with his wife and children, speaking Arabic, who is invited for dinner in the house of an individual from the Info-Target group, who is over 25, with his wife and children, who speaks Arabic, and the soldier brings a trinket (norm: to offer a present when invited). It can be a SAG situation: in this case, a soldier is present, putting on his uniform to go working and talking to his wife and children (norm: men work to support financially their family).

- or a situation in the Religious Context: SAGIT or SAG situation: a group of soldiers (Aim-group) are patrolling near a mosque (norm: to watch, to supervise). The representatives of slot *Attendees* value are civilians.

• If the feeling is negative:

The identities that differ between both groups: no specified professional status vs. military, pro-Morsi vs. anti-Morsi must be made salient. The chosen context is then the Political Context: SAGIT or SAG situation: demonstrators are walking in the streets of an Egyptian city (norm: to shout slogans, to carry Morsi's portraits), military are present.

CONCLUSION

In this paper, we have described the functioning and the overall structure of the system PSYMDEV that aims at helping a military analyst to conceive psychological messages. Given a specific feeling to be triggered in the info-targets towards an aim-group, PSYMDEV generates a twofold-situation that must induce this feeling. We have then focused on the first step of the system. Our work has been influenced by some existing computational tools that deal with culture, like the Upper Ontology of Culture (UOC, (Blanchard, Mizoguchi, and Lajoie 2011)), the Target Audience Simulation Kit (TASK, (Taylor et al. 2010)) and other works from the Soar Technology Laboratory (Taylor et al. 2007). The UOC project focuses on the conceptualization of three culture-related domains: models of cognitive domain, affective domain and context. UOC provides guidelines for cultural systems to ensure their interoperability. The TASK system goal is the same as ours: developing effective messages, but in a marketing context. Both tools, UOC and TASK, use the Appraisal Theory for modeling the emotion process. Our system takes its originality from the fact that the Appraisal Theory is interpreted in an intergroup context thanks to the Intergroup Emotion Theory. Moreover, the system uses this theory in the reverse order and comprises a previous step that aims at inducing a positive or negative state of mind in the info-targets depending on the feeling to be triggered.

REFERENCES

- Allport, G. W. (1954), « *The nature of prejudice* ». Adisson-Wesley. pp. 28–47.
- Arnold, M. B. (1954), « *A theory of human and animal learning* », in: *The human person: An approach to an integral theory of personality*, Arnold M.B. (Ed.). pp. 331- 370. New York: Ronald Press.
- Arnold, M. B. (1960), « *Emotion and personality, Volume 1: Psychological aspects* ». Columbia University Press.
- Blanchard, E. G., Mizoguchi, R., Lajoie, S. P. (2011), « *Structuring the cultural domain with an upper ontology of culture* », in: *The Handbook of Research on Culturally-Aware Information Technology: Perspectives and Models*, Blanchard Emmanuel.G. and Allard Danièle (Eds.). IGI Global.
- Crisp, R., Hewstone, M. (2007), « *Multiple Social Categorization: Processes, Models and Applications* ». Taylor & Francis.
- Dovidio, J. F., Gaertner, S. L., Hodson, G., Riek, B. M., Johnson, K. M., Houlette, M. (2006), « *Recategorization and crossed categorization: The implications of group salience and representations for reducing bias* ». Psychology Press.
- Garcia-Prieto, P., Scherer, K. R. (2006), « *Connecting social identity theory and cognitive appraisal theory of emotions* », in: *Social identities: Motivational, emotional, cultural influences*, Brown R., Capozza D. (Eds.). Psychology Press, Hove, UK. pp. 189-208.
- Kopecky, J., Bos, N., Greenberg, A. (2010), « *Social identity modeling: past work and relevant issues for socio-cultural modeling* », in: *Proceedings of the 19th Conference on Behavior Representation in Modeling and Simulation*, Charleston, USA. pp. 203–210.
- Lazarus, R.S., Averill, J.R., Opton E.M. (1970), « *Towards a cognitive theory of emotion* », in: *Feelings and emotion*, Arnold Magda B. (Ed.). New York, NY: Academic Press.
- Lazarus, R. S. (1991), « *Progress on a cognitive-motivational-relational theory of emotion* », *American psychologist*, Volume 46 N°8. pp. 819-834.
- Mackie, D. M., Devos, T., Smith, E. R. (2000), « *Intergroup emotions: explaining offensive action tendencies in an intergroup context* », *Journal of personality and social psychology*, Volume 79 N°4.
- Minsky, M. (1975), « *A framework for representing knowledge* », in: *The Psychology of Computer Vision*, Winston Patrick (Ed.). McGraw-Hill, New York. pp. 211–277.
- Sander, D., Scherer, K. (2009), « *Traité de psychologie des émotions* ». Hachette Editions.
- Schank, R. C., Abelson, R. P. (1977), « *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures* ». Psychology Press.
- Scherer, K. R., Schorr, A. E., Johnstone. (2001), « *Appraisal processes in emotion: Theory, methods, research* ». Oxford University Press.
- Tajfel, H., Turner, J. C. (2004), « *The social identity theory of intergroup behavior* ». Psychology Press.
- Taylor, G., Quist, M., Furtwangler, S., Knudsen, K. (2007), « *Toward a hybrid cultural cognitive architecture* », in : *Proceedings of CogSci Workshop on Culture and Cognition*.
- Taylor, G., Knudsen, K., Marinier III, R., Quist, M., Furtwargler, S. (2010), « *Target audience simulation kit: Modeling culture and persuasion* », in: *Advances in Cross Cultural Decision Making*, Schmorrow Dylan, Nicholson Denise (Eds.). CRC Press. Taylor & Francis Group. pp.270–279.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., Wetherell, M. S. (1987), « *Rediscovering the social group: A self-categorization theory* ». Basil Blackwell.

ACKNOWLEDGEMENTS

The system PSYMDEV is being developed thanks to a grant funded by DGA (Direction Générale de l'Armement) and Thales Communication and Security, Gennevilliers, France (CIFRE-Defense n° 007/2012/DGA).

Impact of Network Connectivity and Agent Commitment on Spread of Opinions In Social Networks

David Galehouse ^a, Tommy Nguyen ^a, Sameet Sreenivasan ^a, Omar Lizardo ^b, G. Korniss ^a
and Boleslaw K. Szymanski ^a

^a Center for Network Science and Technology
Rensselaer Polytechnic Institute
Troy, NY 12180-3590, USA

^b Department of Sociology
University of Notre Dame
Notre Dame, IN 46556, USA

ABSTRACT

The spread of opinions in social networks underlies other important socio-economical processes such as the spread of innovations, the acceptance of new technologies, or the speed of modernization of infrastructure and industries. Here, we present our results on spread of opinion and the fundamental role played in this spread by two characteristics of the agents, average interconnectivity and level of commitment to the current opinion. Those in turn are dependent on many cultural characteristics of the agent's society, such as trust in others, openness to new views, willingness to discuss certain opinions outside the narrow group of family and close friends, and so on. We also report on our initial investigation how the first factor, average connectivity of social network nodes, varies across cultures based on a study of a real social network. In our *Binary Adoption Model (BAM) with committed agents*, the spreading of opinions involves two stages. Without sufficient support, the committed minority can advance its cause only by increasing its ranks to the so-called *tipping point fraction*. After such a fraction is achieved, the opinion spreads rapidly across the society. We discuss how the tipping point fraction is impacted by the two social-based factors mentioned above.

Keywords: Spread of Opinions, Committed Agents, Tipping-points

INTRODUCTION

Spread of opinions in social networks (Castellano et al., 2009, Durlauf, 1999, Galam, 2008) underlies other important socio-economical processes such as spread of innovation, acceptance of new technologies, or even the speed of modernization of infrastructure and industries. In this paper, we present our results on spread of opinion and the fundamental role played in this spread by two factors characteristic of community or society in which such spread is going on. One is the average interconnectivity of the opinion holders, while the other is the level of commitment of opinion holders to their opinion. Despite their simplicity, both factors are dependent on many cultural characteristics of the society of which the opinion holders are members, such as trust in others, openness to new views, willingness to discuss certain opinions outside the narrow group of family and close friends, and so on. We will also report on our initial investigation how the first factor, average connectivity of social network agents,

varies across cultures based on a study of a real social network.

Previous studies on binary opinion dynamics (Xie et al., 2011, Xie et al., 2012, Zhang et al., 2012, Szymanski et al., 2012) have demonstrated how the presence of a critical fraction of committed individuals (Lu et al., 2009), i.e., individuals who steadfastly hold onto one of the two opinions, is sufficient to cause the entire network to quickly converge to the same opinion. Here we study an extension of this binary agreement model to the case where the commitment of an initial set of committed nodes can be eradicated as a result of repeated exposures to the opposite opinion. The rules governing the opinion dynamics are identical to those utilized in (Xie et al., 2011) and based on the Naming Game (NG) (Baronchelli et al., 2006, Dall'Asta et al., 2006), and we refer the reader to those articles for a detailed description of the rules. The spread of opinions from the holders of minority opinion of a society to its majority under the binary agreement model involves two stages. Without sufficient support, the committed minority can advance its cause only by increasing its ranks to the so-called tipping point fraction, denoted as p_c . After such a fraction is achieved, the opinion rapidly spreads across the society.

Among a number of potentially relevant details, two important factors strongly impact the actual value of the tipping point fraction in a social network with two committed minorities competing for majority, ranging from 2% to 16.25%. One factor is the average interconnectivity of the network agents. The higher is the average degree of the agents in the network, the higher must be percentage of the minority committed to the opinion to enable its spread. The second factor is the level of commitment of minority members to their opinion. This is measured by the number of consecutive social interactions with holders of the opposing opinion that a member requires to switch the opinion. An uncommitted, or regular, member exposed to the single contrary opinion held by a friend, initially gets into the mixed state in which it will accept the opinion proposed by its next interlocutor, thus switching to the opposing opinion if this interlocutor holds the same. In contrast, an agent committed to an opinion at the commitment level $k > 1$ enters the mixed state only after k subsequent interactions with friends holding the opposing view. Naturally, the higher the level of commitment of minority members, the lower is the critical size of the minority group needed for fast spread of the opinion. We also analyze real social networks of freshmen at a US university and students in US high-schools in order to determine how the two social-network-based factors discussed above depend on different traits of the agents such as personality, religion, social class, gender, and race.

The *Binary Agreement Model with committed agents* used in this paper allows only for two opinions, A , and B , that agents can choose from and therefore it is a special case of more general Naming Game models (Dall'Asta et al., 2006) where a single asymmetry is introduced through the random inclusion of a minority fraction of committed agents whose opinion are fixed for all times to be A , say. The key observable is the expected time to consensus of the A opinion and its dependence on the committed fraction and the average agent interconnectivity. In (Zhang et al., 2011, Xie et al., 2012), we show that, for a complete graph, when the committed fraction grows beyond a critical value $p_c \approx 0.0979$, there is a super-exponential decrease in the time taken for the entire network to adopt the A opinion. Specifically, using a straightforward mean field approach, coarse-grained stochastic analysis, and direct simulations of the NG, we show that for $p < p_c$, the mean consensus time T_c is large, $T_c \sim e^N$, while for $p > p_c$, it is small, $T_c \sim \ln N$.

In the presence of committed agents of opinion A , the only absorbing state in the associated random walk Markov chain model (Zhang et al., 2011) is the consensus state of opinion A , while the near-consensus state where all susceptible agents have the B opinion becomes a reflecting state. Similarly, the averaged or mean field system of two coupled nonlinear differential equations (Xie et al., 2012), undergoes a saddle-node bifurcation when $p = p_c$, in which the saddle point (symmetric in phase plane in the case with no committed agents) merges with a node to form a new equilibrium point of saddle node type (Strogatz, 1994). The $T_c \sim \ln N$ time scale comes from the slow dynamics along the center manifold between the saddle node and the consensus of A opinion, and is the same order as the symmetric case where there are no committed agents. In contrast, for $p < p_c$, the $T_c \sim e^N$ time scale is due to the additional numerous time steps spent along that portion of the center manifold which is between the stable fixed point and the saddle-point where the latter corresponds to a state with a larger fraction of agents of opinion A than the former.

Simulations on a range of sparse random networks with 100 to 10,000 nodes have shown, after extensive and costly numerical experiments, that the above tipping point effect of the NG with a minority fraction of committed agents is a very robust phenomenon with respect to the underlying network topology. Of particular significance is the numerical and empirical observation (Xie et al., 2012) that as one lowers the average degree of the underlying random network, the tipping fraction p_c decreases.

In (Zhang et al., 2012), we analytically establish the numerical discovery using a refined mean field approach

(Strogatz, 1994) and report on precise changes in NG dynamics with respect to the average degree $\langle k \rangle$ of an uncorrelated underlying network which are beyond the reach of the straightforward mean field model discussed in (Zhang et. al., 2011, Xie et. al., 2012). Specifically, as shown in Figure 1, the critical tipping fraction in the binary agreement model decreases to a minimum of 5 percent when the average degree $\langle k \rangle = 4$ from a maximum of nearly 10% for complete graphs. This shows that the new mean field model is in better agreement with the numerical results reported above in (Xie et. al., 2012) and provides a much improved approximation to NG dynamics on large random networks in comparison to the straight forward mean field model in (Xie et. al., 2012).

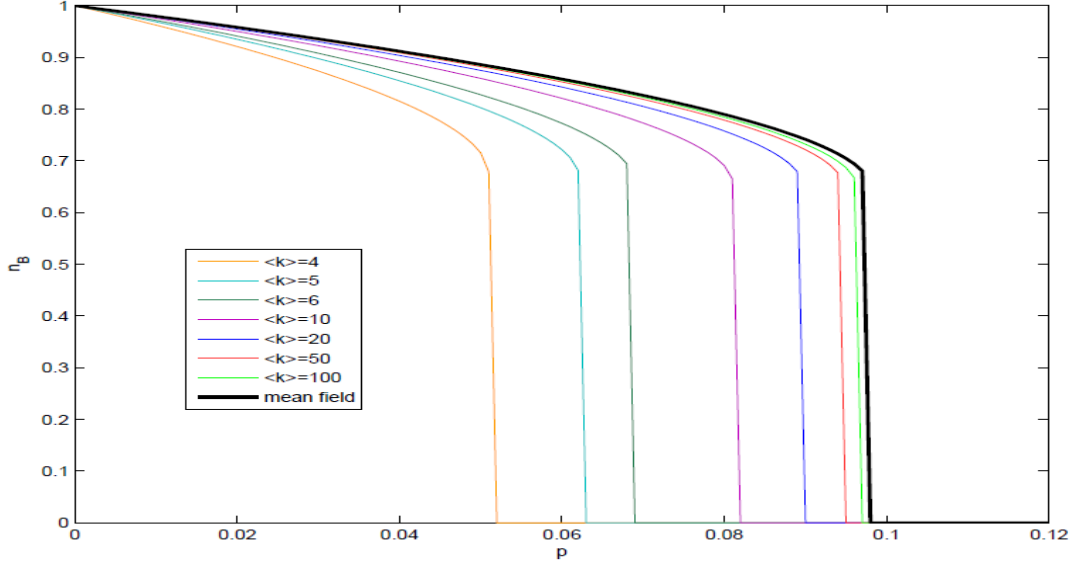


Figure 1. Dependence of the critical committed fraction p_c on agent's connectivity $\langle k \rangle$.

COMMITTED INDIVIDUALS WITH WANING COMMITMENT AND TIPPING POINTS IN OPINION DYNAMICS

The two-word NG rules (corresponding to the binary opinion dynamics) discussed in the paper can be summarized as follows. At each simulation time step, a randomly chosen speaker voices a random opinion from his list to a randomly chosen neighbor (designated as the listener). If the listener has the spoken opinion in his list, both speaker and listener retain only that opinion, or else the listener adds the spoken opinion to his list.

The important distinction in the study presented here is that a committed individual is assumed to become uncommitted after w consecutive pairwise interactions with individuals voicing the opposing opinion. The parameter w thus represents the strength of commitment. Interestingly, as a result of this distinction, any existing stable (attractive) fixed point of the underlying dynamical system is necessarily an absorbing fixed point corresponding to consensus in one of the two opinions. In contrast, when commitment is indefinite ($w = \infty$, equivalent to model of (Xie et al., 2011), when the committed fraction is below the critical value, one of the stable fixed points corresponds to a mixed state where densities of all opinions are non-zero. We designate the two opinions A and B , and denote their respective densities by n_A and n_B respectively, and the initial fraction of committed nodes in opinion A is denoted by p_A . To quantify the degree of dominance of one of the opinions over the other, we utilize the order parameter $m = (n_B - n_A)/(1 - p_A)$. Thus, the case of $m = 1$ represents the absorbing fixed point where all individuals have arrived at a consensus on opinion B , while the case of $m = -1$ corresponds to the absorbing fixed point which is equivalent to consensus on opinion A . Shown in Figure 2 are 20 representative evolutions of the order parameter for different initial committed fractions p_A , and with commitment strength $w = 10$, as obtained from simulations. Each simulation begins with all uncommitted nodes holding opinion B ; all committed nodes hold opinion A . For $p_A = 0.10$, all committed individuals eventually lose their commitment, and the network comes to consensus on opinion B [Fig. 1(a)]. However, as p_A increases, we observe increasing numbers of simulation runs where the initial committed fraction is able to propel the system to consensus on opinion A [see Figures 2(b), (c), and (d)]. This is indicative of the existence, in the asymptotic limit, of a critical fraction $p_c(w)$ above which the system reaches the all- A consensus

state with probability one, and below which it becomes exponentially unlikely for the system to reach the all- A consensus state, despite the existence of an initial committed fraction proselytizing this opinion. In other words, above $p_c(w)$, the all- A consensus is the only attractive fixed point remaining in the system, while below it, both the all- A consensus and all- B consensus fixed points are attracting, but the bias towards opinion B in the initial population ensures that in the asymptotic limit of network size, the system eventually converges on opinion B .

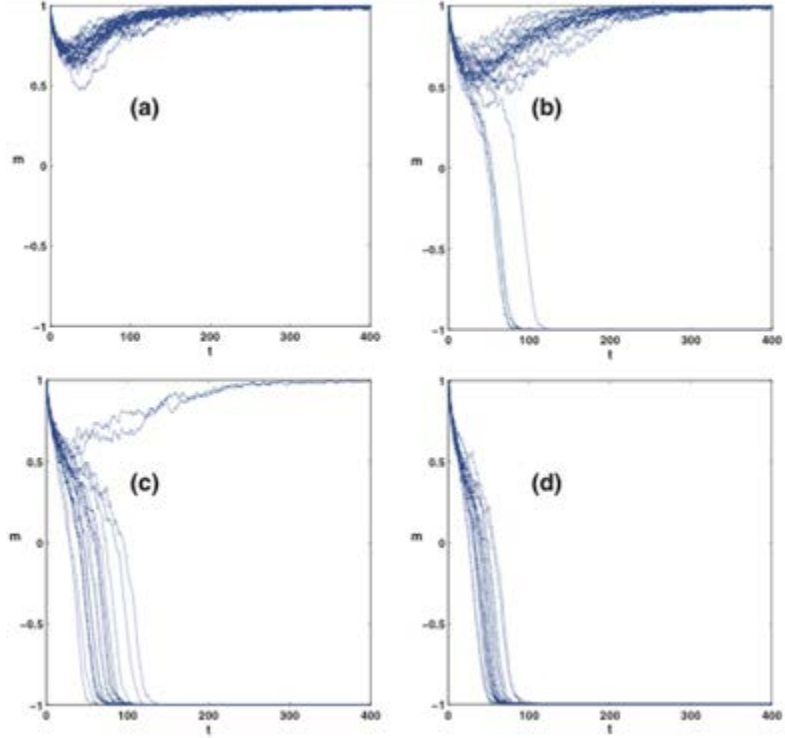


Figure 2. Trajectories of the order parameter m for different values of the committed fraction p_A with commitment strength $w = 10$. The four plots correspond to values of committed fraction (a) $p_A = 0.10$, (b) $p_A = 0.11$, (c) $p_A = 0.12$, and (d) $p_A = 0.13$. The social network is a Barabási-Albert network (Barabási and Albert, 1999) with $N = 1000$ nodes and average degree $\langle k \rangle = 10$.

We also study how the critical value $p_c(w)$, at which the transition occurs depends on the commitment strength w . Figure 3 shows the variation in $p_c(w)$, estimated from simulations, as a function of w . The results suggest an asymptotic scaling form $p_c(w) \approx p_c(\infty) + a/(b+w)^d$, where $d \approx -1.73$.

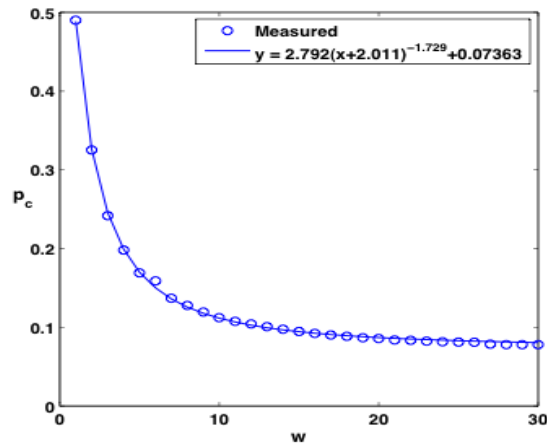


Figure 3. The critical committed fraction $p_c(w)$, as a function of commitment strength w for Barabási-Albert networks of size $N = 1000$ with average degree $\langle k \rangle = 10$.

Next, we consider the case where there are initially two committed groups of individuals within the network, one proselytizing opinion A and the other proselytizing opinion B , with both groups having a commitment strength w .

Previous studies (Xie et al., 2012, Szymanski et al., 2012, Verma et al., 2012, Verma et al., 2014) on such competition between committed groups with infinite commitment ($w = \infty$) have demonstrated the existence of a beak-shaped phase diagram in parameter space (p_A, p_B), where the boundaries of the beak are lines indicating saddle-node bifurcations – points at which the system undergoes a transition where one of the two stable fixed points disappears. Thus, inside the beak the system has two stable fixed points with transitions between them being exponentially rare, whereas outside the beak the system has a single fixed point favoring the opinion with the larger committed fraction. Importantly, for the case of infinite commitment, the stable points both inside and outside the beak represent mixed states with non-zero densities of both opinions.

When the commitment parameter is finite, we get a qualitatively similar picture as shown in Figure 4(a). However, in this case the region within the beak has two absorbing stable fixed points corresponding to the all- A and all- B

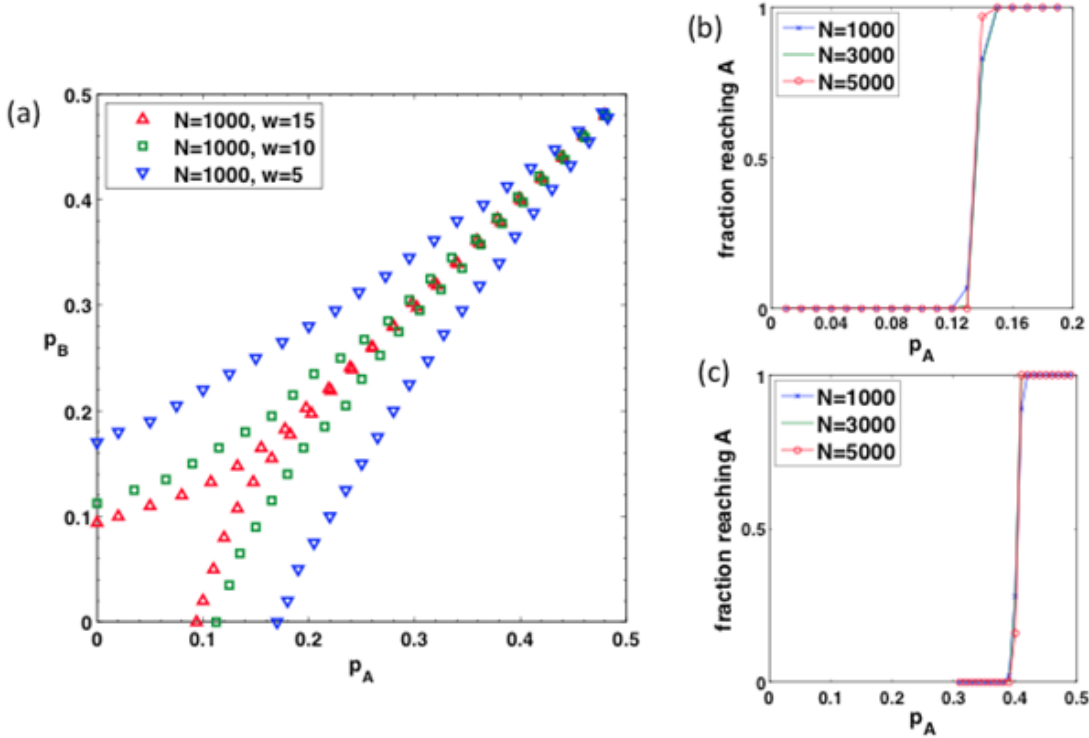


Figure 4. (a) Phase diagram of the system in the presence of competing committed groups holding opinion A and B with initial sizes p_A and p_B , respectively for Barabási-Albert networks with average degree $\langle k \rangle = 10$ (see text for details). (b) The fraction of runs reaching the all- A consensus state along a trajectory $p_A = 0.2p_B$ in parameter space. (c) The fraction of runs reaching the all- A consensus state along a trajectory $p_A = 0.8p_B$ in parameter space.

consensus states. Outside the beak, for example in the region below the diagonal, only one absorbing point (the all- A consensus state) is attracting, and the system deterministically ends up there. A point on the boundary of the beak can be detected in the following manner. Suppose that initially the system consists of all uncommitted nodes holding opinion B . Then for values of (p_A, p_B) that place the system within the beak, the system eventually ends up in the all- B consensus state as a result of the initial bias, coupled with the fact that the all- B consensus state is an attracting fixed point. However, if we systematically follow a path in parameter space that takes the system from being inside the beak to outside the beak (and beneath the diagonal), then we expect that once the beak boundary is crossed, the system will end up in the all- A consensus state, since it is the only attractive fixed point that remains. The location of this phase transition provides an estimate of a point on the beak boundary. The entire phase diagram, as shown in Figure 4, can be constructed by systematically exploring paths $p_A = c \cdot p_B$ in parameter space, for distinct values of c ranging from 0 to 1, and estimating for each path, the point [i.e., (p_A, p_B) values] at which the transition occurs. Two such transitions encountered along paths with $c = 0.2$ and $c = 0.8$ are shown in Figure 4(b) and 4(c) respectively.

A notable feature in the phase diagram, Figure 4(a), is that as the commitment strength w increases, the region where both absorbing points are attractive, i.e. the region within the beak becomes narrower. As $w \rightarrow \infty$, the phase diagram converges to that obtained in (Xie et al., 2012).

EVIDENCE OF IMPACT OF AGENT SOCIO-DEMOGRAPHIC CHARACTERISTICS ON ACTIVITY AND ASSORTATIVE MIXING

In this section we first summarize the findings concerning socio-demographic differences in both *activity* (Simmel and Hughes, 1949) and *assortative mixing* (Newman, 2003). The analysis is based on a social network of freshmen at an US university. The network contains 186 vertices (agents) and 647 undirected edges formed between August and December of 2011. The network is constructed by placing an edge between two vertices if we observed at least one two-way communication (text, or voice call) between the two agents in the given time period. We restrict our analysis to within study dyad because we have complete socio-demographic and attitudinal information on this subset of agents.

Sociability effects refer to the higher likelihood of members of certain categories to form more connections to other persons regardless of the socio-demographic traits of those persons. Studying assortative mixing in the context of modelling opinion spread is important for the following reasons. It is well known that opinions, beliefs, and other cultural commitments tend to be different for members of different socio-demographic groups. This means that when persons interact with those who are similar to them, their current beliefs and values are likely to be reinforced. Opinion change in real social networks is more likely to happen for those individuals who are more likely to venture outside of their own group for sociable interactions (Aral, 2009, Robins, 2007, Granovetter 1973). A social system characterized by high levels of assortative mixing is also one that is likely to exhibit opinion polarization rather than homogeneity.

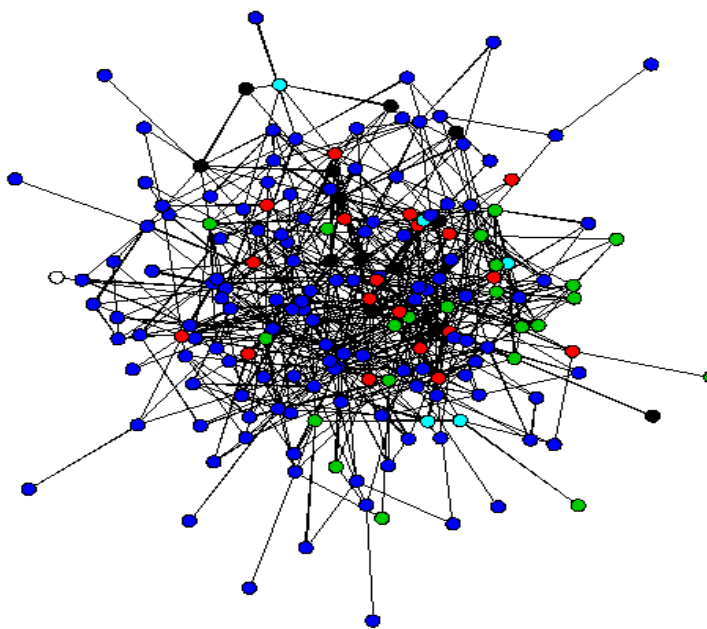


Figure 5. NetSense within-study network. Vertex color indicates race, Blue = White, Red = African American, Green = Latino, Black = Asian, Teal = Other.

We identify both sociability and assortative mixing effects using *exponential random graph models* [ERGMs] fitted to the network (Hunter, 2008). These models allow us to ascertain statistical dependences on the network while adjusting for lower order network properties, such as the overall tendency to form edges (density) and the tendency for agents who share neighbors to also be connected to one another. The purpose of ERGMs is to provide a summary description of the local relationship formation mechanisms that generate the global structure of the

network. For instance, the NetSense network (depicted in Figure 5), may be considered similar to the response or outcome variable in a regression or classification model. Here the predictors of whether we observed an edge (or not) are such things as “the propensity of individuals of the same gender/race/religion/class background to form connections” or “the propensity of individuals who share friends to be connected to one another.” In Figure 5 for instance, it is clear that individuals who share the same racial/ethnic status tend to cluster together. The ERGM can help us quantify the strength of this assortative mixing effect, while holding constant other confounding relationship formation mechanisms.

Interpreting ERGM parameters

The exponential random graph model specifies the conditional log-odds of an edge existing between two nodes vs. not existing, conditional on values of covariates and a network configuration:

$$\frac{\Pr(X_{ij} = 1|Y = y, X_{ij} = x_{ij})}{\Pr(X_{ij} = 0|Y = y, X_{ij} = x_{ij})} = e^{\theta^T \delta_{ij}(x) + \theta_a(y_i + y_j) + \theta_b y_i y_j}$$

Where refer to θ_a as the (additive) activity parameter, and θ_b is the (multiplicative) assortative mixing parameter; y_i, y_j are the values of each vertex on a given socio-demographic trait of interest (in this example gender, which is a categorical trait with two values: male, female). The other term is for the configuration (endogenous) part of the model. In our case, this involves a parameter for the overall number of edges in the graph (density) and a parameter that counts the number of shared neighbors for each node in the graph (transitivity). For the purpose of interpreting the activity and mixing parameter, we can ignore that part of the model because when taking odds ratios it drops out (assuming that the configuration model is the same for all edges).

What we can compute from this model is the conditional odds ratio. The reference category for these odds is $y_i = y_j = 0$, i.e. the (0,0) cell in the $n \times m$ table cross classifying edges by whether they have or do not have the trait $Y = y$, which in our case is a 2×2 table cross-classifying edges by the gender mix of the vertices attached to them. Thus, for the example used here Y is gender and $Y = 1$ indicates that the node is female.

The conditional odds ratio for an edge having one node with $Y = 1$ versus having both vertices with $Y = 0$ (e.g., odds ratio of male-female tie versus male-male tie) is e^{θ_a} , the activity parameter. Its value defines the odds that, conditional on one node being male, the other will be female. If the activity parameter is 0, then the odds ratio is 1, making the odds of male-female tie the same as the odds of male-male tie. This means the counts of male-male and male-female ties should be the same. If the activity parameter is greater than 0, then male-female ties are more prevalent than male-male ties, if less than 0, male-male ties are more prevalent than male-female ties. The activity parameter reflects difference in the degree of male versus female, because the only way male-female ties can be more prevalent than male-male ties (assuming random matching, i.e. no homophily) is if females have more ties than males.

If there is no homophily (i.e., random matching), then the odds of a female-female tie versus a male-male tie is $e^{2\theta_a}$. That is we expect more female-female ties solely because there are females having more edges (i.e. are more likely to be at the end of a random edge), so that the odds of a female being at both ends of an edge is the exponential of 2 times the activity parameter. If there is homophily (positive θ_b parameter) then the odds of a female-female tie versus a male-male tie is $e^{2\theta_a + \theta_b}$. The $2\theta_a$ captures the greater odds of a female-female tie due to the higher activity of females, while the θ_b term adds in increased odds of female-female ties resulting from assortative mixing. Note that the odds of the female-female tie versus a male-male tie is now $e^{\theta_a + \theta_b}$. Relative to the male-female edge, the female-female edge gets a boost over the male-female edge because females have higher activity and because of activity differences female-female ties are more prevalent.

For instance, from our estimates (Table 1, Model 4) we can see that $\theta_a = 0.14$ and $\theta_b = 0.46$. We coded female as 1 so the reference like above is male-male ties. The odds of a male-female versus a male-male ties is $e^{0.14} - 1 = 15\%$, i.e., the former are 15% more likely than the later, reflecting the higher activity of women. Without the homophily effect this would imply that female-female ties would be about 32% more prevalent than male-male ties (e.g. $e^{2*0.14} - 1$). But taking into account the assortative mixing effect and the odds of a female-female vs a male-male tie is $e^{2*0.14 + 0.46} = 2.09$, implying that female-female ties are more than twice as likely (209% more likely) than male-male ties. Finally, contrasting female-female versus male-female ties, we see that female-female ties are $e^{(0.14 + .46)} = 1.82$ times more likely than male-female ties.

In our estimates (Table 1) female-female ties are more prevalent, than male-female ties, which are more prevalent

then male-male ties. This is indicative of strong female assortative effect coupled with females having high degree, which implies a tendency for males to not associate with other males because they randomly associate with other people who happen to be more likely to be females because of their higher activity.

Additional Parameters

Density - All of the ERGMs models that we estimate hold two structural parameters constant. The first one is the *density* parameter. This is equivalent to the single Bernoulli probability of observing an edge in the Erdos-Renyi model. A model including only this parameter would be in fact equivalent to the E-R model for random networks. This parameter holds constant the propensity to form edges in the network independently of vertex level activity and assortativity effects.

Transitivity - There are various ways to model transitivity (the tendency of friends of friends to be friends) in the ERGM framework. Here we opt for specification described in (Hunter, 2008) which accounts for transitivity in the network by fitting the geometrically weighted edgewise shared partners (GWESP) statistic. This parameter is essentially equivalent to a penalized count of the number of shared neighbors for each edge in the network and therefore essentially is the number of edges that serve as a common base for i distinct triangles.

Table 1: Gender-Based Activity and Assortative Mixing

	Model 1	Model 2	Model 3	Model 4
Density	-3.24 (0.04)	-3.76 (0.08)	-4.15 (0.08)	-4.51 (0.10)
Activity (Girls)		0.25 (0.05)		0.14 (0.04)
Mixing		0.51 (0.08)		0.46 (0.07)
Transitivity			0.99 (0.07)	0.97 (0.07)

Race

Table 2 shows the result of fitting an ERGM model with separate activity parameters for each racial and ethnic group and a common assortative mixing parameter for all race groups. We find that, in terms of activity, minority groups are more active than Whites in their first semester on campus. Recall that the activity coefficients compare the odds of observing an edge featuring a non-white respondent versus the odds of observing an edge containing two white respondents. Across all racial groups (as indicated by the positive activity effects) we find that mixed race edges are more prevalent than the all White edges.

Table 2: Race-Based Activity and Assortative Socializing

	Model 1	Model 2	Model 3	Model 4
Density	-3.24 (0.04)	-4.53 (0.10)	-4.15 (0.08)	-5.13 (0.10)
Activity (Blacks)		0.99 (0.10)		0.75 (0.08)
Activity (Latinos)		0.84 (0.08)		0.64 (0.07)
Activity (Asians)		0.79 (0.08)		0.59 (0.06)
Activity (Other)		0.75 (0.19)		0.64 (0.13)
Mixing (All Races)		1.33 (0.10)		1.12 (0.08)
Transitivity			0.99 (0.07)	0.95 (0.07)

In addition, we find that same race dyads are statistically more prevalent than mixed-race dyads ($\theta_b = 1.33$), even

after adjusting for the propensity of friends of friends to also be connected to one another ($\theta = 0.95$). Thus, even though Whites are less active than non-whites, when they do form ties, they are disproportionately likely to be directed at other Whites. Non-whites on the other hand, have similar same-race preferences, but due to their higher activity, end up forming more mixed-race connections than we would expect by chance.

Religion

We find that (Table 3, Model 3) students that claim no religion are less socially active than students who report being Catholic. We find (Table 3, Model 2) that edges are more likely to form between agents that share the same religious status (e.g. Catholic, Protestant, or No Religion); $\theta_b = 0.45$. However, the strength of the assortative mixing effect is not the same across the different religion categories. Instead, the assortative mixing effect is very strong among agents with no religion ($\theta_b = 1.62$), it is fairly strong among Protestants ($\theta_b = 0.80$) and is non-existent among Catholics ($\theta_b = -0.23$).

Table 3: Religion-Based Activity and Assortative Socializing

	Model 1	Model 2	Model 3
Density	-4.20	-4.55	-3.91
	(0.08)	(0.11)	(0.17)
Activity (Other)	0.11	0.33	-0.17
	(0.06)	(0.07)	(0.15)
Activity (None)	0.08	0.27	-0.46
	(0.06)	(0.06)	(0.15)
Mixing (All Religions)		0.45	
		(0.09)	
Mixing (Catholic)			-0.23
			(0.16)
Mixing (Other)			0.80
			(0.32)
Mixing (None)			1.62
			(0.22)
Transitivity	0.99	0.98	0.97
	(0.07)	(0.07)	(0.07)

Class Background

We find that (Table 4, Model 3) there are no activity differences based on class background, once we adjust for the statistical propensity to find class-matched edges over class mismatched ones. In our data, we coded class background as follows: first generation college students were coded as *working class*, students with only one parent who had completed college were coded as *middle class*, and students with both parents having completed college were coded as *upper middle class*.

Table 3: Class-Based Activity and Assortative Socializing

	Model 1	Model 2	Model 3	Model 4	Model 5
Density	-3.24	-3.71	-3.56	-4.15	-4.17
	(0.04)	(0.14)	(0.21)	(0.08)	(0.08)
Activity (Working Class)		0.50	0.21		
		(0.10)	(0.12)		
Activity (Upper Middle Class)		0.11	0.13		
		(0.09)	(0.19)		
Mixing (All Classes)		0.31			
		(0.10)			
Mixing (Working Class)			1.06		0.90
			(0.27)		(0.14)
Mixing (Middle Class)			-0.44		
			(0.38)		
Mixing (Upper Middle Class)			0.11		
			(0.22)		
Transitivity				0.99	0.98
				(0.07)	(0.07)

Looking at the assortative mixing effects, we find that the statistical propensity to match class is restricted to individuals who share a common background as working class ($\theta_b = 1.06$). Neither middle class nor upper class students display any tendency to form ties disproportionately with members of the same class group. Note that the working class affinity effects persist even after adjusting for the tendency towards transitivity.

Also of interest is time evolution of network activities for difference classes of agents, classified according to their personality, gender and race. The resulting diagrams, shown in Figure 6, provide also range of differences between the respective classes of agents as well as trends in time. The strong personality, race, and gender effects on active network size of students after joining the university evolve differently. As shown in Figure 6, race effects persist through the observation period but gender effects decline.

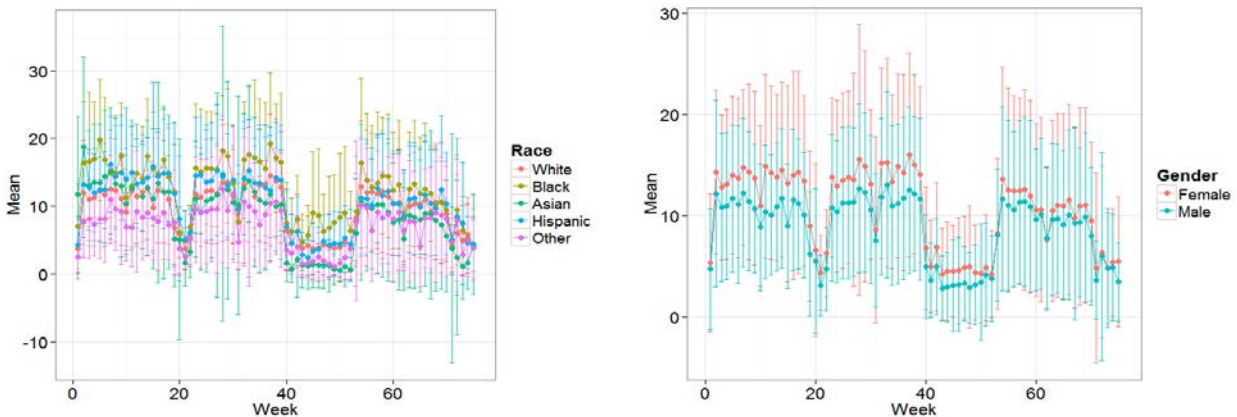


Figure 6. Race, and gender effects on active network size.

High-School Social Networks

We also used a set of social networks (high-school friendship networks), which were provided by the National Longitudinal Study of Adolescent Health (Add Health)¹. The high-school friendship networks investigated here were constructed from a paper-and-pencil questionnaire in the AddHealth project (Moody, 2001). Here, nodes represent students and edges represent mutual relations or friendships. Two students are considered to be friends and have an edge between them when one nominates the other as her/his friend and both of them participated in some activities, e.g., talked over the phone, spent the weekend together, etc., in the last seven days. We selected ten largest high school networks consisting of 20,020 students in total from the available data set.

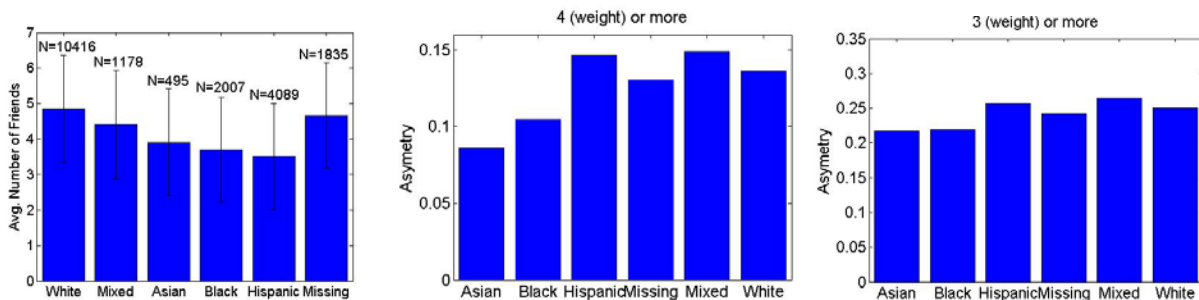


Figure 7. Race effects on active network size and asymmetric friendship in high-school networks. The effects are strong for large asymmetry (4 edge weight units) but small from small imbalance (3 or less units).

¹ This research uses the network-structure data sets from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. For data files contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524, (addhealth@unc.edu, <http://www.cpc.unc.edu/projects/addhealth/>).

Figure 7 presents the results. We observed (see left panel of Figure 7) similar race effects on active network size for high-school networks as reported above for the US university student network. In addition, we used strength of interactions between friends that were self-reported by each student at the level from 1 to 6 to denote their weighted friendship edges. We define friendship as asymmetric with difference x when a student assigns the strength of the friendship to his friend by x levels higher than that friend assigns to this student. The center panel in Figure 7 shows that race has strong effect on asymmetric friendships with difference 4 but this effect disappears when the level of asymmetry decreases (the right panel of Figure 7 shows asymmetry with difference 3). Since asymmetric friendship shows commitment to a friend against opinions of others, we believe that this is a good indicator of willingness of a person to commit to an opinion, and the results presented indicate that such willingness could also vary across the cultures.

CONCLUSIONS

In this paper, we first presented a theoretical model of spread of opinion in social networks. In this model, we demonstrated that in the presence of committed agents, the committed minority of society members can rapidly spread their opinion to the entire society if the number of such committed minority members is large enough. There is a critical fraction of the committed minority which enables such a rapid spread and below it, the spread is extremely unlikely. We then identified properties of network agents that impact the value of the tipping point fraction. First, we showed that the tipping point fraction decreases as the average connectivity of an agent decreases from the highest level (fully connected graph) to lower values, when each agent is connected to only a few other nodes. Then, we investigated how changing commitment of the agents to their opinion influences the tipping point fraction. We used the binary agreement model tipping fraction as the baseline for comparison, since the commitment in this case is infinite, meaning that no finite number of interactions with agents holding the opposite opinion will convince a committed agent to abandon the opinion to which it is committed. We then showed that as a result of waning commitment, the critical fraction is increased over the baseline value, and that this increment grows as the number of interactions required to erode commitment decreases. These theoretical results raise the question of whether the differences in the interconnectivity and commitment are culturally or socio-demographically dependent. To answer this question, we first studied a real-time network of freshmen in a US university and observed and quantified the evolution of students' activity network and their level of assortative mixing over their first semester at university. We observed the large differences both in activity and assortative mixing between groups classified by personal, gender and race criteria. Then we analyzed high-school social networks for average number of friends and the willingness to maintain asymmetric friendship. We found that both of these metrics are race dependent. The later one increases as the imbalance of friendship increases. We believe that maintaining strongly imbalance friendship against opinions of others is indicative of willingness to commit to an opinion. Thus, the immediate conclusion is that the different groups of students analyzed in our paper will likely respond differently to new opinions and ideas, and they might be differently receptive to innovation and new products.

In the future works, we plan to extend our study of real social networks in two directions. First, we want to study the level of asymmetric interactions that demonstrate commitment to weak ties against the objective evidence. Next, we want to extend the geographical scope of the network to include students from different countries and cultures including those not nested within the US. Finally, we plan to extend the study of these networks by tracking spread of opinions among them.

ACKNOWLEDGEMENT

This work was supported in part by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053, by the Office of Naval Research Grant Number N00014-09-1-0607, and by the Army Research Office Grant Number W911NF-12-1-0546. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

REFERENCES

- Aral, S., Muchnik, L., & Sundararajan, A. (2009), "*Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks*", Proceedings of the National Academy of Sciences, **106**(51), 21544-21549.
- Barabási, A.-L., & Albert, R. (1999), "*Emergence of scaling in random networks*", Science **286**, 509–512.
- Baronchelli, A., Loreto, V., & Steels, L. (2006), "*Sharp transition towards shared vocabularies in multi-agent systems*", J. Stat. Mech. P06014.
- Castellano, C., Fortunato, S. & Loreto, V. (2009). "*Statistical physics of social dynamics*". Rev. Mod. Phys. **81**, 591-646.
- Dall'Asta, L., Baronchelli, A., Barrat, A., & Loreto, V. (2006), "*Nonequilibrium dynamics of language games on complex networks*", Phys. Rev. E **74**, 036105.
- Durlauf, S.N. (1999). "*How can statistical mechanics contribute to social science?*", Proc. Natl. Acad. Sci. USA **96**, 10582-10584.
- Galam, S. (2008), "*Sociophysics: A Review of Galam Models*". Int. J. Mod. Phys. C **19**, 409-440.
- Granovetter, M. (1973), "*The strength of weak ties*", American journal of sociology, **78**(6), 1360-1380.
- Hachen, D., & Lizardo, O. (2013), "*The effects of individual traits on the formation of social networks among first-year college students*", International School and Conference on Network Science (NetSci), Copenhagen, Denmark.
- Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008), "*ergm: A package to fit, simulate and diagnose exponential-family models for networks*", Journal of Statistical Software, **24**(3), nihpa54860.
- Lu, Q., Korniss, G., & Szymanski, B.K., (2009), "*The Naming Game in social networks: community formation and consensus engineering*", J. Econ. Interact. Coord. **4**, 221.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001), "*Birds of a feather: Homophily in social networks*", Annual review of sociology, 415-444.
- Moody J (2001), "*Race, school integration and friendship segregation in America*", American Journal of Sociology **107**:679–716
- Newman, M. E. (2003), "*Mixing patterns in networks*", Physical Review E, **67**(2), 026126.
- Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007), "*An introduction to exponential random graph p^* models for social networks*", Social networks, **29**(2), 173-191.
- Simmel, G., & Hughes, E. C. (1949), "*The sociology of sociability*", American Journal of Sociology, 254-261.
- Strogatz, S.G. (1994), Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering. Da Capo Press.
- Szymanski, B.K., Sreenivasan, S., Xie, J., Zhang, W., Emenheiser, J., Kirby, M., Lim, C., & Korniss, G. (2012), "*On influence of committed minorities on social consensus*", Proceedings of the NATO RTO HFM-201 Specialists Meeting, Tallinn, Estonia, April 16-18, 2012, pp. 32-1 – 32-15.
- Verma, G., Swami, A., & Chan, K., (2012), "*Effect of zealotry in the naming game model of opinion dynamics*", MILCOM 2012, pp.1-6, IEEE Computer Science Press.
- Verma, G., Swami, A., & Chan, K. (2014), "*The impact of competing zealots on opinion dynamics*", Physica A **395**, 310-331.
- Xie, J., Sreenivasan, S., Korniss, G., Zhang, W., Lim, C., & Szymanski B.K. (2011), "*Social consensus through the influence of committed minorities*", Phys. Rev. E **84**, 011130.
- Xie, J., Emenheiser, J., Kirby, M., Sreenivasan, S., Szymanski, B.K., & Korniss, G. (2012), "*Evolution of opinions in social networks in the presence of competing committed groups*", PLoS ONE **7**(3):e33215.
- Zhang, W., Lim, C., Sreenivasan, S., Xie, J., Szymanski, B.K., & Korniss, G. (2011), "*Social influencing and associated random walk models: Asymptotic consensus times on the complete graph*", Chaos **21**(2):025115.
- Zhang, W., Lim, C., & Szymanski, B.K. (2012), "*Analytic treatment of tipping points for social consensus in large random networks*", Phys. Rev. E **86**, 061134.

Behavioral Influence Assessment (BIA): A Multi-Scale System to Assess Dynamic Behaviors Within Groups and Societies Across Time

Michael L. Bernard, George A. Backus, and Asmeret B. Bier

*Cyber Engineering Research Institute
Sandia National Laboratories
Albuquerque, NM 87123-1327 USA*

William C. Busch and Regan W. Damron

*Strategic Foresight Branch
U.S. European Command
Stuttgart, Germany*

ABSTRACT

Sandia National Laboratories, in cooperation with the United States European Command's (EUCOM) Strategic Foresight (SF) branch, has developed an initial capability to better understand and anticipate likely responses to events by groups within countries under EUCOM's area of responsibility. The Behavioral Influence Assessment (BIA) system is a theory-based analytical capability that is intended to enable analysts to better assess the influence of events on groups interacting within a country or region. These events can include changes in policy, man-made or natural disasters, war, or other changes in environmental and economic conditions. To help achieve this, BIA models the dynamic social/political/economic actions and counter-actions between groups in response to events over time. This paper outlines the rationale and general results produced by this effort. This includes a discussion of: 1) underlying psychological, social, and economic theories that are synthesized within its structure; 2) inclusion of data and expert opinion into the modeling structure; 3) methods used to computationally instantiate theories, data and opinion; 4) types of assessments that are generated; and 5) implications of these assessments in comparison to current events.

Keywords: Cognitive Modeling, Social Modeling, Systems Modeling, Country Assessments

INTRODUCTION

A common problem associated with the effort to better understand dynamic behaviors of different groups within various countries is the sheer difficulty in gathering appropriate subject matter expertise across relevant domains of interest. Even with a sizable collection of subject matter experts (SME), one's cognitive ability to fully comprehend the dynamic nature of populations, particularly over time and considering feedback effects, can be limited. That is, humans' ability to contemplate higher-dimension interaction effects within and between groups is more easily

restricted to a small number of behaviors and counter-behaviors across time. In addition, the contribution of behavioral, social, and economic theories regarding how groups and countries, as a whole, make decisions is often not fully considered. Yet, an understanding of dynamic human behaviors is typically considered important when assessing country behaviors and, thus, should be addressed. A common question is how can this be accomplished in a more systematic manner? Recently, the phenomena underlying the societal dynamics that drives stability and instability in countries has become understandable enough to pose testable hypotheses amenable to simulation (Acemoglu & Robinson, 2009). Ultimately, we believe this can produce the ability to better hypothesize and assess impacts associated with various policies and actions.

The main focus of this work is to present an approach taken by Sandia National Laboratories to assess the dynamics and key psychosocial and economic processes underlying how people make decisions and express behaviors over time. Included in these simulations are behaviors that affect the decision making of others, creating complex feedback loops within and between individuals and groups. Each simulated behavior is a function of individual psychosocial characteristics (described below) along with environmental and group dynamic factors. The goal of this work is to minimize the likelihood of decisions that lead to undesirable consequences by providing a more systematic analysis of group perceptions, beliefs, intentions, and behaviors. Its focus is on likely dynamic repercussions of actions, not on point predictions.

The basis for this computational framework, known as the Behavioral Influence Assessment (BIA) tool, is a synthesis of data-supported psychosocial theories of human behavior (Bernard & Backus, 2009; Bier et al., 2011). The legitimacy of this synthesis is supported by an independent assessment of theory-based analytical studies of historical socioeconomic data. In other words, we have integrated the set of elements from psychosocial theory that are consistent with economic theory, experimental data, and historical data pertaining to human behavior. The result is a unified framework that connects the multiple scales of human behavior (from individual to societal interactions) to the external (geopolitical, physical, and socioeconomic) world. The simulation framework is thus a model of human behavior determined by local perceptions of world conditions, contained in a feedback process that links behaviors, conditions, and events and shows how they unfold over time. Our analysis emphasizes these response and counter-response progressions, whose recognition can prevent counter-productive behaviors. □

BIA uses a hybrid cognitive-system dynamics computational approach that integrates psychological, social, socioeconomic, and system dynamics theory relevant to a region. This includes characterization of people within a society, interactions between governmental and nongovernmental groups, and external variables such as global changes in the geopolitical and economic climate. This is designed to capture outcome distributions used to investigate attitudinal and behavioral reactions to policies and actions within countries. The theories used are limited to ones that can be: 1) integrated into a representation of behavior; 2) translated into a set of computational equations; and 3) instantiated, tested, and verified using accessible data. The structure is populated using psychosocial, economic, and geopolitical information from subject matter expert guidance, reports, opinion polls, social media, and economic data. While the information that populates the BIA structure is specific to particular countries, its structure is general enough to be applied to any geographical region.

Computationally, the BIA structure consists of a modeling framework, model simulators, and an analysis engine. The current structure allows for various cross-modeling domains (i.e., different countries and groups), information sharing, and visualization. For example, Figure 1 shows a simplified conceptual representation of a hypothetical BIA structure that involves the modeling of two interacting groups and several leaders. Exogenous inputs to the model (e.g., global economic factors and general population-support) influence the dynamic interactions within and between the entities. Each simulated behavior is a function of psychological characteristics along with environmental and group dynamic factors. This enables the assessment of group behavior as it reacts to the perceptions of others and world conditions.

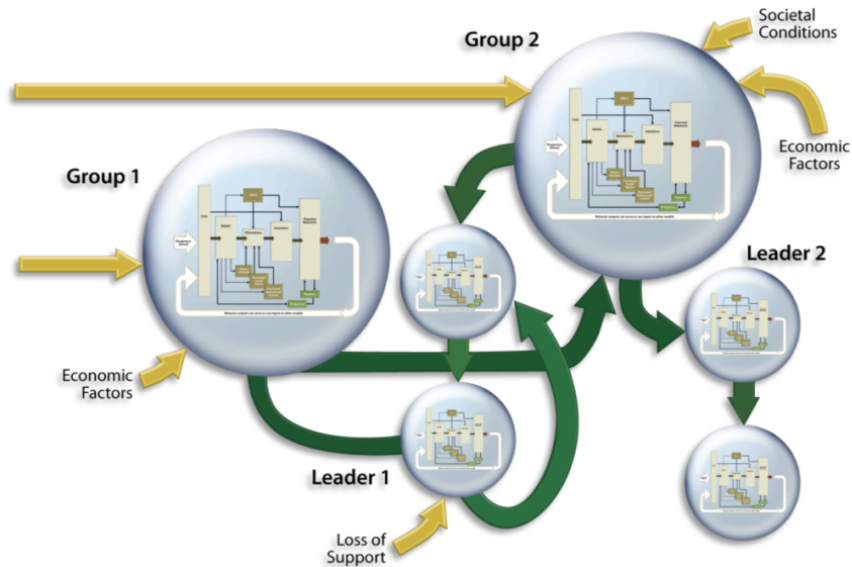


Figure 1. Conceptual diagram of the full systems view of BIA

RESEARCH DESIGN AND METHODOLOGY

The underlying psychosocial theories

In order to make sense of and predict our environments, humans attempt to find patterns in stimuli. If relevant, the stimuli can be perceived as cues that can stimulate particular beliefs. However, because of differences in culture and experiences among individuals, the same stimuli may be interpreted differently and stimulate different beliefs. Beliefs may stimulate other cognitive processes such as emotional reactions (which we more broadly characterize as positive and negative affect), attitudes, expectations associated with perceived social norms, and perceived behavioral control over potential behaviors associated with that belief or series of beliefs. These things may help stimulate a motivation to perform some specific type of behavior. If the motivation is high enough, it can stimulate an intention or set of intentions to perform some type of behavior. The intention to perform a specific behavior is typically a function of what is actionable. Thus, upon assessing the environment, intentions that are not attainable will lose strength while intentions that are attainable will gain strength. Moreover, the valence associated with affect (low to high positive, low to high negative) will mediate the selection of behavior (Bernard & Smith, 2006). The actual behavior that is realized is a function of the intent, associated affect, and the perceived environment indicating that behavior is indeed actionable. Additional factors that affect the likelihood of a behavior being realized include how often and how recently that behavior has been previously acted upon. That is, previous behaviors are a good predictor of future behaviors (Ouellette & Wood, 1998). This process is exemplified in the conceptual diagram of the BIA psychosocial (decision) model, shown in Figure 2.

Formation of beliefs

In the BIA decision model, a belief is considered to be an estimate of some attribute or state in an environment that may affect an existing attitude or give rise to a new one. When a belief rises to full consciousness it will be compared to a belief “template” that is stored in long-term memory. These templates store semantic perceptions of self and environment as beliefs and serve to categorize/classify and structure one’s belief of one’s world. In this sense, semantic memory is made up of categories (class of objects that belong together) and concepts (mental representations of a category) for later retrieval.

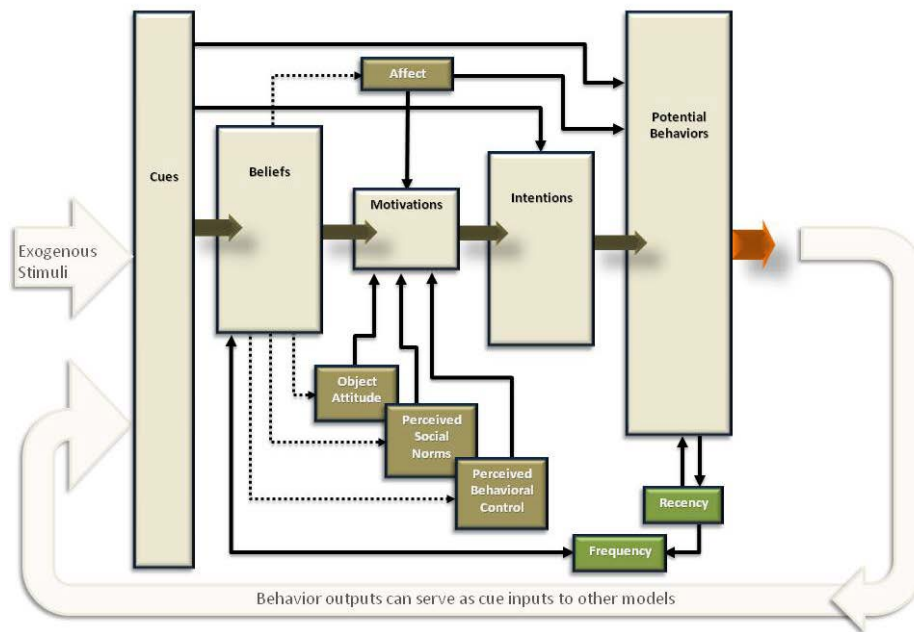


Figure 2. Conceptual diagram of the psychosocial (decision) model

To make a decision about category membership within a belief, an item is compared against some stored cue representation of the category (Markman, 1999). If the cue similarity is strong enough, then it is concluded that the item is part of that belief category. If the differences are great enough between belief categories, then new categories can be spawned to reflect this diversity. Each cue membership and its associated belief category are specific to an individual or group. That is, one individual may have a different set of potential beliefs than another individual. Moreover, similar beliefs may contain different cue memberships, with different percentages of cue evidence associated with each belief. Groups that share similar cultures and/or religious backgrounds may have a common set of general beliefs and related set of cue memberships that may differ from other, less similar, groups. It stands to reason that the further apart one group is from another, the less similar the cue memberships and beliefs will be.

In our simulations, the instantiation of this concept involves the representation of environmental cues and relevant knowledge in a manner that accommodates pattern recognition. Patterns of cues may be associated with a known or potential situation or state. If the activation of cues indicates a match between an on-going situation and a cognitive representation associated with a particular belief, a belief is generated that might ultimately direct behavior. In this simulation, each cue provides some degree of evidence (a continuum from 0 to 1) associated with one or more beliefs. Conceptually, the system represents the fundamental processes associated with both individual and group/organizational (aggregate) decision-making. The notion that cues, in many circumstances, can trigger a belief without the need for extensive deliberation has been proposed by Klein and colleagues (1993) in their model of recognition-primed decision making (RPD). With respect to attitudes, beliefs are thought to be associations or linkages that people establish between attributes of attitude objects (Fishbein & Ajzen, 1975). An attitude can be thought of as a general and relatively enduring evaluative response to an attitude object, where an attitude object can be a person, a group, an issue or a concept (Visser & Clark, 2003). This evaluative response generally has some degree of favor or disfavor, approach or avoidance, or attraction or aversion toward that object (Ajzen, 1991; 2005). This is expressed in differences in affective valence or direction in that they can be “bifurcated into positive and negative evaluations” (Eagly & Chaiken, 1993, p. 4). A configuration of attitudes can form an ideology, which are attitudes clustered around some societal theme, such as communism verses capitalism (Converse, 1964).

Modeling behaviors

A general theoretical model supporting the notion that attitudes and intent play a large role in predicting behavior is the theory of planned behavior (TPB). The TPB postulates a process in which behaviors are influenced by (a) current attitudes towards a specific behavior, the (b) subjective norms associated with acting out that behavior, and (c) the perception that carrying out this behavior is within the person’s control. The combination of these factors

forms a behavioral “intention” state, which then can serve to drive that person’s actual behavior (Ajzen, 1991; 2005; Madden, Ellen, & Ajzen, 1992). It is asserted in the TPB that an individual’s intentions capture the factors that influence some type of behavior, which is indicative of one’s effort to perform that behavior. While the TPB is most certainly one of the most used and cited behavioral models within social psychology (Cooke & Sherran, 2004), it has been critiqued for being too restrictive in scope (e.g., Perugini & Bagozzi, 2001). Its parsimony is typically considered one of its strengths, enabling it to be robust across a wide variety of behaviors and cultures. However, this can come at a cost of reduced predictive power. In principle, the model can, and was intended to be, open to additional predictor variables, such as emotion (Ajzen, 2005). Indeed, researchers including Ajzen have developed extensions to the TPB so as to increase its predictive strength for various types of behaviors (for example, see Beck & Ajzen, 1991).

An extension to the TPB in both depth and breadth was developed by Perugini and Bagozzi (2001). Their conceptual model, termed the model of goal-directed behavior (MGB), is intended to explain a larger percentage of variance associated with behavior than the TPB. The MGB asserts predictor variables of attitude, positive and negative affect, subjective norms, and perceived behavioral control, which drive desires. Desires, in turn, drive intentions, which drive behaviors. Frequency of past behavior mediates desires, intentions, and behaviors. In the MGB, Perugini and Bagozzi chose the more specific term “desire” instead of motivation. BIA uses the more hierarchically expansive term, “motivation” to represent a broader range of drives. As with the TPB, perceived behavioral control also mediates behavior. In addition, recency of past behavior serves to mediate current behaviors. Perugini and Bagozzi argue that individuals take into account both their attitudes and affect regarding potential achievement or failure with respect to a sought after goal. That is, an attitude is an “evaluative response towards an object or act that, once learned, is triggered automatically” (p. 82). The processes underlying affect, on the other hand, are “more dynamic and entail self-regulation in response to feedback” (p. 82). As discussed above, adding these predictor variables should, in many circumstances, explain a higher percentage of the variance associated with behavior. In addition, including frequency and recency of behavior as predictor variables, the MGB further broadens the TPB, providing greater predictability. Indeed, in comparing the MGB to the TPB across two common behaviors revealed that the MGB does explain a larger percentage of accounted variance (approximately 25%) than the TPB (Perugini & Bagozzi, 2001).

Processes underlying economic behaviors

In parallel with the described psychosocial theories, a set of behavioral economic theories—also extensively evaluated with experimental and historical data—have been incorporated into the general BIA framework. The theories perfectly mesh with the discussed psychosocial theories. This is not surprising, since economics can be described simply as people making choices. In fact, there is abundant time-series data on economic decisions across culture, which overlaps with psychosocial views of those same decisions.

The physical and economic behavioral implications are readily simulated using basic aspects of conventional simulation methods such as system dynamics, engineering, and economics. Societal and economic behaviors can be thought of as a consequence of behavioral decisions and, thus, decisions can be thought of as the process of making choices. Accordingly, all behaviors are the consequence of choices that are made. This notion is outlined by McFadden (1984), who pioneered the use of (psychologically framed) qualitative choice theory (QCT). QCT quantitatively determines the importance people place on information, tastes, beliefs, and preferences when making decisions. The robust parameterization of QCT is often based on data readily obtainable in the field. Other techniques can further determine the correct functional representation of the QCT utility formulation for the problem at hand (Keeney & Raiffa, 1976). A key part of the decision process is the filtering of information and the extent to which experience biases the decision process. At a group level, the probabilistic nature leads to a mean-value response because random variation in one direction by a single individual is balanced by the reverse variation of another individual. The enduring aspects of the population (society) dominate the group behaviors. The identification of the transient and stable components of the decision process use co-integration (also Granger Causality) methods pioneered by Granger (1969). These same methods also ascertain the filtering and delayed-response processes associated with information perception and behavior. These methods and others are summarized in Backus and Glass (2006). These techniques can integrate disparate perspectives and information, qualitative as well as quantitative, into analysis and decision support systems. The methods are compatible with orthodox macroeconomic assumptions and used for all matter of choices (including those associated with security). The actions taken to repeat or approximate an individual’s reference perceptions are known as purposive actions. Purposive actions exist in a complex and constantly changing environment (e.g., ecological forces, such as temperature and human forces, such as social structure). For this reason, humans maintain multiple reference perceptions for various aspects of their lives and continually compare these reference perceptions to perceptual cues

from the environment. (For instance, the way we want something to taste is a reference perception and the amount of salt we add is a purposive action to alter the taste to match our reference perception.) Any purposive action that is taken will be compared to the respective environmental cues and to the individual's reference perception of that event. Purposive actions continue as long as there are undesirable discrepancies between the reference perception and the environment. The psychosocial model described is embedded within a system dynamics model to make (economic) decisions for the individuals and groups of individuals that the model simulates in the form of "cognitive entities." Each of these entities has a separate cognitive model. For example, there may be a cognitive model of a type of individual (leader), a group supporting that leader, and a group opposing that leader. In this case there would be three cognitive entities that will be assessed.

Computational foundations of the BIA framework

Within BIA, modeled entities structure and process information in the manner illustrated in Figure 2, which is mathematically represented in Figure 3. Figure 3 shows an example mapping between the conceptual, psychosocial model structures (left) and the mathematical implementation of that model structure (right). (A larger visualization of the mathematical representation is shown in Figure 4). Here, stimuli are the physical realization of world conditions and human actions. When an individual places these stimuli in context, they become cues that inform or effect behaviors. As discussed above, the grouping of cues forms a pattern. (For example, the observation of asphalt, cars, sidewalks, and buildings act as cues, giving you the belief that you are on a city street). Beliefs typically take on importance when they are incongruous with or different from expectations. Expectations are often the memory of the status quo or the anticipation of future conditions. Cognitive resources are employed to produce a learned attitude toward a condition (the condition being a perceived notion or incongruity) or our learned ability to respond to a condition. Our cognitive resources and beliefs of a situation (via notions and incongruities) act together to help us evaluate the choices we have to respond to those conditions. The result represents our intentions. The execution of those intentions further depends on the level of the incongruity and our attitudes toward that behavior. Once a behavior is initiated, it takes time before it becomes an action affecting the external world (including other individuals). Depending on the proximity or our social network, the realized consequence of our actions becomes the cues to some individuals but not to others. The feedback logic of one entity's behavior becoming another entity's cues, possibly through the intermediation of external physical processes, explicitly captures the social network considerations that are often seen in more abstract, agent-based modeling.

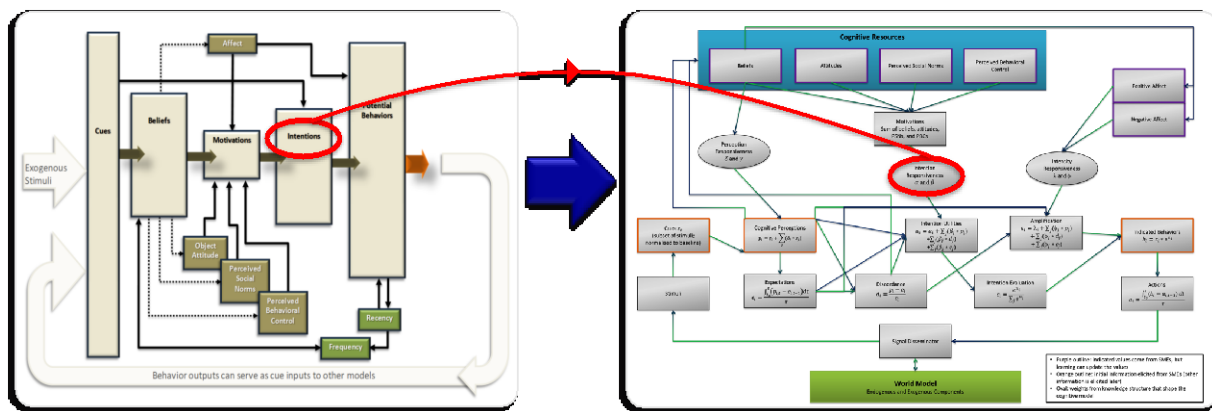


Figure 3. Conceptual diagram of the psychosocial model

Each block in the math diagram contains equations that are fully elucidated in Backus et al., (2010). Note that each block can process large flows of data. There are typically a large number of stimuli generating a large number of notions, leading to a large number of potential choices and behaviors, across a number of individuals. Some of the differences to note between Figure 2 and Figure 4 are, for example, the decomposition of "intentions" in Figure 2 into several subcomponents shown in Figures 4, such as intention utilities and intention evaluation.

It takes time to cognitively recognize a set of cues. Cues can also produce emotive notions that characteristically occur faster than cognitive notions and use minimal information. The emotive notions can set the “mood” for processing the cognitive information, often adding a risk aversion element to the choice invoked by the cognitive information (Forgas, 1995; Lerner & Keltner, 2000). Research shows that emotive and non-motive components are both part of the normal processing that leads to behavior (Martin, Ward, Achee, & Wyer, 1993; Zajonc, 1984). The model explicitly recognizes and uses both these categories of information flow. In addition, Figure 4 contains what are noted as “Tiering Loops.” Specific notions (such as you realizing there is a fire in your house), can dramatically amplify your realization of stimuli/cues, such as the location of doors and other occupants of the house. Similarly, making one decision may affect your selection of a related decision. The same is true for executing behaviors. Attitudes affect the importance you may place on information. Attitudes are explicitly calculated in the model and are based on cognitive resources (experiences, abilities, and beliefs). Learning is noted as conditioning in the model and is an effort to reduce an incongruity by developing the ability to accommodate or effectively respond in the presence of a notion. Attitudes, emotive content and cognitive information all act to determine the utility of a choice. These utilities come together to shape the probability of making a specific choice. Limitations in mental processing and physical response mean the individual must prioritize notions and behaviors when either becomes potentially excessive (Dolan, 2002; Gigerenzer & Goldstein, 1996). For example, changing the radio station when you hear a song you dislike is quickly neglected when you see the car ahead of you hit another car. Moreover, Figure 4 depicts the psychological components that interact, feedback, and combine to produce behavior. People are constantly exposed to a large number of stimuli. They attempt to find patterns in these stimuli to help predict their environments. Only a small fraction of stimuli can be processed and recognized as relevant cues for prediction. A specific pattern of these cues can produce a belief and general notion regarding the current environment. In the model, relevant cues include political, social, physical and economic conditions. The inflation rate, for example, is an economic cue that may lead to a notion about the health of the economy.

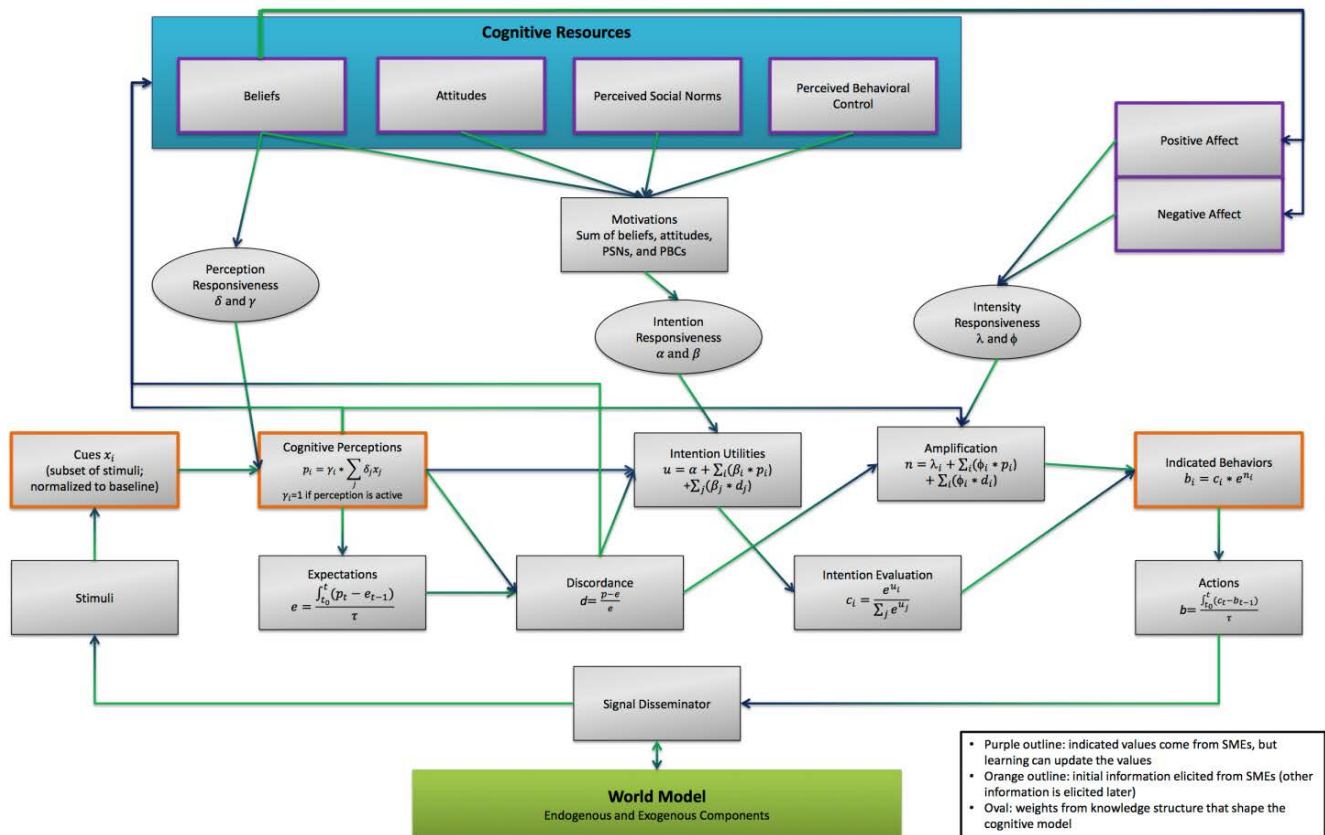


Figure 4. The full representation of the BIA system

The process of representing behavior

The BIA modeling process begins at a systems level. At this level, we seek to understand the key influencers from organizations and conditions to represent the overall dynamics within and between entities. Sub-system structures may also be included to represent physical processes within the broader system, such as a supply chain. The general scope of the structure is determined from the overarching question(s) that are posed via an analyst consumer. At this stage, the process of generating possible sub-questions begins. Vetting the sub-questions with an analyst consumer helps to further refine this structure. The specific expression pertaining to each influencer and what choices or behaviors those influences can invoke has to be determined through the use of SME guidance and available data. SMEs can hypothesize beliefs and more abstract “notions” that are not reflected in the data. Analytical methods can allow an estimate of how those hypothesized behaviors could occur based on knowledge of an individual’s behaviors in other circumstances.

Representing potential behavioral responses and counter-responses is first achieved through causal-loop diagramming, which causally relates all the interactions embodied in the theories (see Bernard & Bier in this issue of the proceedings). The causal loop diagram is then mapped to a stock-and-flow diagram that explicitly details the flow of information and physical quantities through the system. A key feature is the designation of stocks that represent the accumulation of information, experience, monetary, or physical quantities. These stocks are called “state variables” and they largely characterize the nature of the system and its responses. The change or difference in the value of stocks over increments of time is the “differential” part of the differential-equation approach to computational modeling. The exact mathematical expression of the theory is anchored in the accumulation of flow into and out of the stocks. The mathematical expression of the flows comes from a causal interpretation of the theory into the language of mathematics. Only those theories that have a measurable meaning, supportable, at least in principle, by historical or experimental data, are included in the model. The data determines the parameters that control the progression of the simulated values through time. Rigorous statistical techniques determine the appropriate parameters and the uncertainty associated with their use. This uncertainty can later define the confidence in the results of an intervention analysis. This general process is shown in Figure 5.

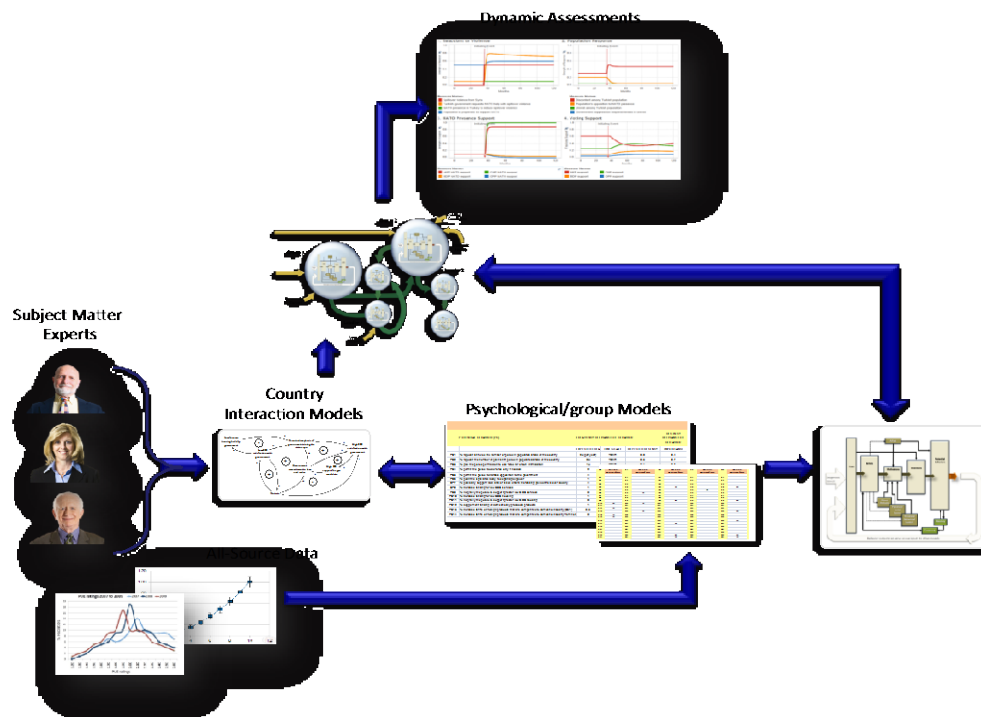


Figure 5. A simplified example of the BIA modeling process

Uncertainty analysis techniques can determine the potential for behaviors to affect the policy selection of external security interventions. (The assessment of these intercessions is the actual purpose of the model.) However, the model cannot generate potential behaviors that are beyond the imagination of SMEs and are not reflected in available data. These unknown unknowns are a limitation in all realms of physical and social science. Nonetheless, the use of decision theory, SME guidance, and data should produce the best representation possible, despite uncertainties. To model the consequence of behavioral influences, it is necessary to not only model the initial behaviors of affected individuals and groups, but to also determine how interactions with other individuals and the physical world, over time, can alter the outcome. The feedback processes among individuals and the physical world unfold dynamically, which could cause the outcome of an intervention to initially go in the desired direction, but in the long-term lead to counter-responses that that are contrary to the desired goal. That is, the delay between behaviors and impacts can cause secondary dynamics that make it extremely difficult to know whether the ups and downs of behavioral responses and counter-responses will ultimately lead to the desired outcome.

As the model is developed, domain information is used to add specificity to the structure. It is determined which entities (individuals or groups) should be modeled at a more detailed level. The domain information consists of quantitative data and SME guidance. This information is recorded via the BIA's "Knowledge Structure" datasheets. The BIA Knowledge Structure is consistent with the psychosocial theories of decision-making and is organized in a manner that characterizes the decision processes of specific individuals or group of individuals. Knowledge structures capture information such as beliefs, motivations, norms, attitudes, general affect, intentions, and previous and current behaviors of specific groups, organizations, and/or individuals. For example, formulas that are circled in Figure 6 are populated with information, via the Knowledge Structure, also shown in Figure 6. As these structures are developed, increasingly more detailed domain information is used to populate the models and to help ensure that the systems-level and detail-level structures are consistent. That is, this information can be used to further strengthen the overall systems structure. The quantitative data consist of survey polls, economic output reports and projections, demographics, and the like that provide useful information pertaining to beliefs, attitudes, behaviors, and trends.

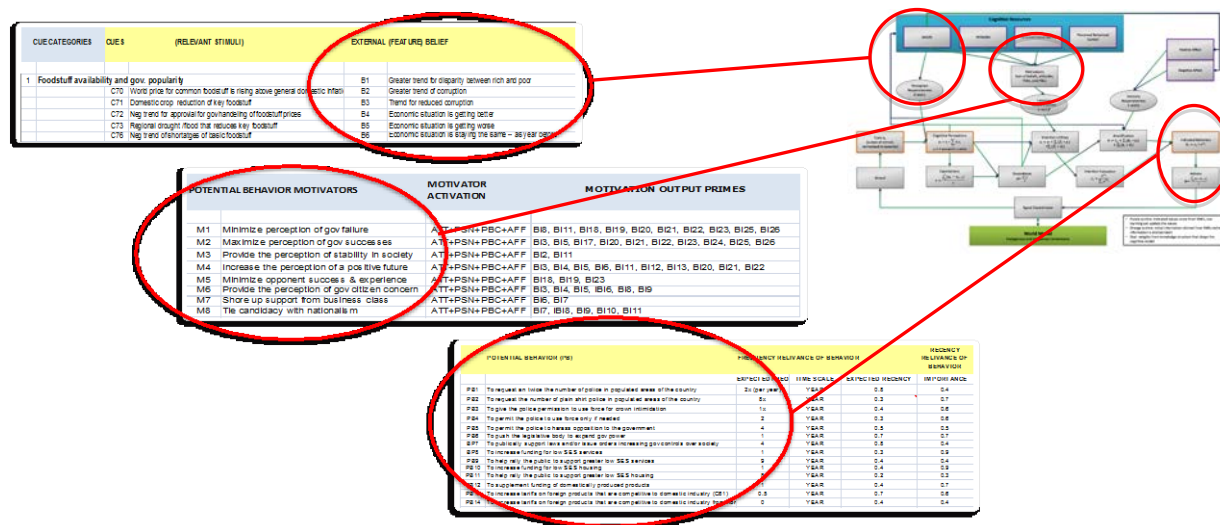


Figure 6. Example of a Knowledge Structure showing belief, motivation, and potential behavior information

This approach to modeling is made possible by assuming a fixed (but potentially very large) set of potential behaviors embodied in a representation of the individual or group. The representation contains the preferences and personality characteristics pertinent to the relevant decision-making. While the magnitude of interactions may change, the model does not produce new paths of cognition. All potential interactions are determined via initial parameterization of the model. Over time (at most a couple of years and often on the order of weeks), the simulation will be less predictive in that the modeled individuals or groups will change their behaviors outside the domain of their historical experience and habits. This will require updating the parameters within the models.

BIA ASSESSMENTS

An example of actual BIA assessment output is depicted in Figure 7. The assessments shown below are designed to reveal likely responses (and counter-responses) to potential actions or events as well as the geopolitical and economic processes behind the responses for specific groups of interest. The BIA architecture underlying these assessments consists of a modeling framework, simulators, and an analysis and visualization engine. The current BIA modeling architecture provides for the storage of BIA models, the execution of those models by compatible simulators, and the analysis of run results by various analysis engines. The integration with simulators and analysis engines leverages a plug-in architecture to convert data repository records into a compatible format for various commercial visualization tools. Other arbitrary types of records that could represent source references or other supplemental information can be tied to models to provide deep traceability from a run or an analysis. The current BIA architecture also includes a flexible database engine within the modeling framework to support organization of models, runs, and references. We have developed a progressively more complex BIA modeling architecture to enable more efficient cross-modeling domain, simulation and assessments, information sharing, knowledge structure development, and visualization. For each model run, and change in a model run, BIA provides a dynamic hypothesis and analysis. The dynamic hypothesis shows the core drivers of the system dynamics and the model output presents the actual model runs. The duration of a modeled scenario ranges from as short as one week to as long as two decades. Figure 7 shows an example of the assessment interface built for EUCOM.

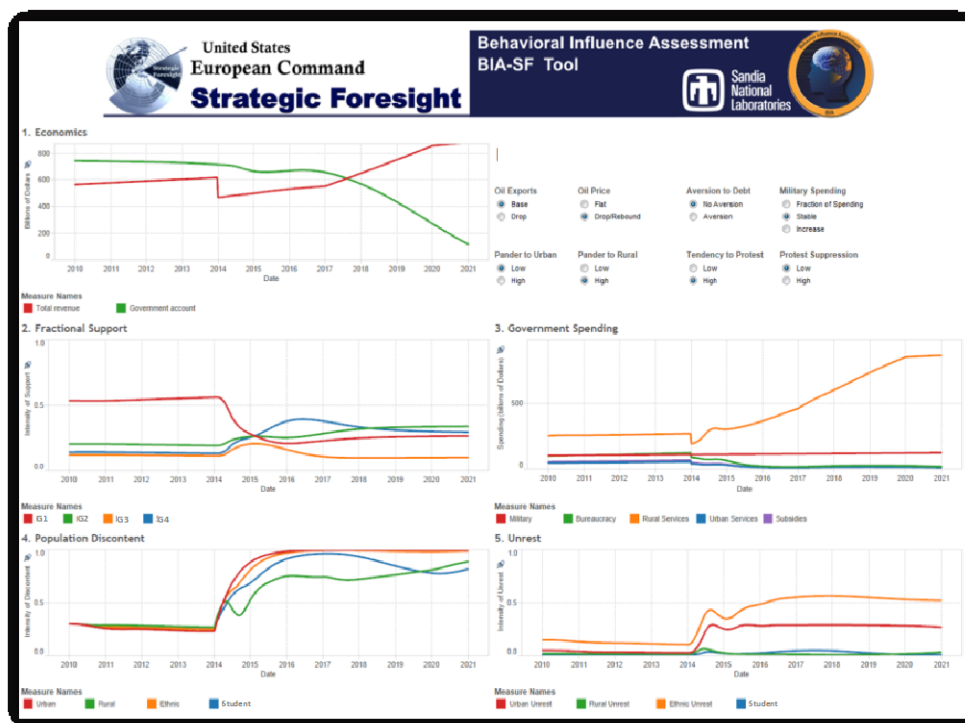


Figure 7. An example of BIA assessment output

Working with EUCOM's Strategic Foresight branch, which is focused on using advanced information technology to inform command decision-making, strategy development, and planning, BIA assessed such things as: (1) political and social reactions to various governmental actions in response to perceived internal and/or external influences or threats; (2) the long-term effect that changes in economic and/or social conditions have on the amount of support given to various groups; (3) the long-term social/political stability-effects related to various funding priorities within countries; and (4) potential interactions between countries due to changing geopolitical conditions. Strategic

Foresight is currently providing insight into strategic risks and opportunities within EUCOM's operational environment through the analysis of socio-cultural futures in the mid- to long-term (greater than 24 months). The BIA assessment interface shown in Figure 7 analyzes individual/group perceptions, beliefs, intentions, and behaviors associated with potential societal actions (over time). These actions are in response to the behaviors of government and non-governmental entities, as well as exogenous (outside) variables. The exogenous variables are often hypothetical conditions that could occur, such as changes in global economic conditions, changes in global oil/gas markets, cross-border violence, migration into or existing a country, conflict, and the like. As these variables change, so does the model's assessment of potential political, social, and economic reactions of groups and societies.

Model confidence

Within BIA, historical data and SME information become the raw data used to calibrate and parameterize the models. Uncertainty in the data is explicitly determined through the statistical process of uncertainty quantification to develop model parameters. This process provides confidence intervals on the results of the model analyses that test interventions. By simultaneously performing uncertainty quantification for model parameters and potential interventions, BIA can determine the portfolio of intersessions that have the highest (quantified) probability of success despite uncertainty. It can also quantify the risk associated with the intervention not performing as anticipated. Additionally, BIA can perform sensitivity analyses to determine what minimal additional information is needed to maximally reduce uncertainty and further assure the proposed interventions produce the desired outcome throughout the time horizon of interest. Moreover, because the model is causal, decision-makers can reach-back into detailed results of the simulation to independently evaluate the nuanced processes that caused the predicted outcomes and find leverage points that would be maximally effective at altering outcomes. Furthermore, the same process can determine early warning fingerprints whose measurement today or during the initial implementation of an intervention can verify or exclude the possibility of critical conditions and outcomes.

CONCLUSIONS

The intent of this document is to provide a brief discussion of the BIA framework. Results suggest that BIA can provide a data-driven analytical capability, consistent with psychosocial and economic theory, usable for assessing national security policies and operations. We are still developing the secondary components needed for greater usability. Most of the sub-components have been thoroughly tested in previous studies. Still, the full integration of all the parts into a comprehensive framework is a research and development effort. The framework is currently being tested with detailed data sets and realistic scenarios. Albeit limited, our experience with the use of the system to-date indicates that the approach we are taking is sound and can produce the expected capabilities. Initial assessment of BIA suggests that it is consistent with the general processes underlying human behavior, inclusive of cultural, biological, and institutional constraints, and conditions. The BIA framework is based on first principles that can encompass an unlimited number of entities with any number of alternative decisions, and with any level of interrelationship complexity. The theories used were limited to ones that 1) were mutually self consistent, 2) would integrate into a complete representation of behavior from stimuli through to action, 3) would translate to a unique set of computational equations, and 4) could be instantiated, tested, and verified using accessible data. This structure and discipline allowed for 1) use of readily available data on individual or regions to calibrate the model, 2) use of subject matter expert data to sparsely augment data as needed, 3) testing of hypotheses surrounding alternative interventions and behavioral responses, 4) quantification of the uncertainty (risk) that an intervention will produce the desired results, and 5) time-dependent consequential counter-responses to follow from an intervention. Most importantly, the BIA framework is designed to naturally capture the implications of new (even unique) information flows that may be considered in information operations or other interventions.

REFERENCES

- Acemoglu, D. & Robinson, J. (2009). *Economic origins of dictatorship and democracy*. Cambridge University Press.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211.

- Ajzen, I. (2005). *Attitudes, personality, and behavior*. (2nd Ed.). New York: McGraw-Hill.
- Backus, G., Bernard, M., Verzi, S., Asmeret, B., Glickman, M. (2010). *Foundations to the unified psycho-cognitive engine*. Sandia National Laboratories technical report SAND Report 2010.
- Backus, G. & Glass, R. (2006). *An agent-based model component to a framework for the analysis of terrorist-group dynamics*. Sandia National Laboratories technical report SAND2006-0860.
- Bernard, M. L. Backus, G. (2009). Model the interaction between leaders and society during conflict situations. *Annual Meeting of the System Dynamics Conference*, Albuquerque, NM.
- Bernard, M., L. & Smith, B. (2006). *The effects of emotional states and traits on risky decision-making*. Sandia National Laboratories technical report SAND2006-7642.
- Beck, L., & Ajzen, I. (1991). Predicting dishonest actions using the theory of planned behavior. *Journal of Research in Personality*, 25, 285-301.
- Bier, A., Bernard, M. L., Backus, G., Glickman, M., & Verzi S. (2011). Using a hybrid cognitive-system dynamics model to anticipate the influence of events and actions on human behaviors. *29th International Conference of the System Dynamics Society*, Washington, D.C.
- Converse, P. E. (1964). The nature of belief system in mass publics. In D.E. Apter (Ed.), *Ideology and discontent* (pp. 206-261). New York: Free Press.
- Cooke, R., & Sheeran, P. (2004). Moderation of cognition–intention and cognition–behavior relations: A meta-analysis of properties of variables from the theory of planned behavior. *British Journal of Social Psychology*, 43, 159-186.
- Dolan, R. J. (2002) Emotion, cognition, and behavior, *Science*, 298, (8) 1191-1194.
- Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.
- Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM), *Psychological Bulletin*, 117 (1) 39-66.
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424-438.
- Gigerenzer, G. & Goldstein, G. D. (1996). Reasoning the fast and frugal way: Models of bounded rationality, *Psychological Review*, 103 (4) 650-669.
- Keeney, R. L. & Raiffa, H. (1976). *Decisions with multiple objectives*. New York NY: John Wiley & Sons.
- Klein, G. A., Orasanu, J., Calderwood, R., & Zsombok, C. E. (Eds.). (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex Publishing Corporation.
- Lerner, J. S. and Keltner, D. (2000) Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition & Emotion*, 14 (4) 473-493
- Madden, T. J., Ellen, P. S. & Ajzen, I. (1992). A comparison of the theory of planned behavior and the theory of reasoned action. *Personality and Social Psychology Bulletin*, 18, 3-9.
- Markman A. B. (1999). *Knowledge representation*. Mahwah, NJ: Erlbaum.
- Martin, L. L., D. W., Ward, J. W., & Achee, Wyer, R. S. (1993). Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology*, 64, 317-326
- McFadden, D. (1984). *Econometric analysis of qualitative response models*. Handbook of Econometrics, Volume II. Chapter 24. Elsevier Science Publishers BV.
- Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life : The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, 124, 54-74.
- Perugini, M., Bagozzi, R. P. (2001). The role of desires and anticipated emotions in goal-directed behaviors: Broadening and deepening the theory of planned behavior. *British Journal of Social Psychology*, 40, 79-98
- Visser, P. S., & Clark, L. (2003). *Attitudes*. In A. Kuper & J. Kuper (Eds.), *Social Science Encyclopedia*. London: Routledge.

ACKNOWLEDGEMENTS

This research was possible in part by funding from the United States Department of Defense. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly-owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. Approved for unlimited release: SAND 2014-1539C.

Applying Factor Analysis to Population Surveys in Afghanistan to Facilitate Improved Decision Making

*Capt. Joseph K. Maddux
United States Marine Corps
Operations Analysis Division
Marine Corps Combat Development Command
Quantico, Va. 22134, USA*

ABSTRACT

As the U.S. Military's involvement in Afghanistan has changed from a traditional kinetic fight to counter-insurgency to nation building, so too has our need to understand the population which we seek to assist. Multiple agencies have taken on the task of surveying Afghans at the national, provincial and district levels. Conducting these surveys carries great risk due to the presence of Taliban fighters and terrorist safe havens strewn throughout the country. While conducting the surveys has its share of obstacles, properly interpreting the results carry its own unique set of challenges. We use three surveys that have been mainstays as sources of information for the U.S. Military and other nations to understand popular perceptions. Using Factor Analysis on these three surveys collected over the last 3-5 years allows us to analyze all questions and responses in the surveys by identifying the most relevant data analytically vice a subjective analyst or commander deciding what is most important. We do so by using a methodology that establishes quality, analytically derived indicators (groups of survey questions) by using these three survey instruments collected throughout Afghanistan. We generate indicators from survey data using Factor Analysis then assess them by using nonparametric statistics techniques to detect spatial and temporal changes throughout the country. The final result is a full analytics suite that provides perspective to an analyst regarding indicator change across the country over time. Further, we improve the commander's ability to allocate scarce resources to particular districts and provinces which need the most attention.

Keywords: Factor Analysis, Survey, Non-Parametric Statistics, Mann-Whitney, Military, Afghanistan, Spatial and Temporal change, Spreadsheet, Hypothesis Testing, Analytics

INTRODUCTION

The United States Military has been conducting Operation Enduring Freedom in Afghanistan at various commitment levels since 2001. While the initial U.S. interests were to eliminate Osama Bin Laden and his Al-Qaeda forces, efforts shifted to more of a counter-insurgency (COIN) engagement. In order to victor in a counter-insurgency environment, it is typically important to gain the trust and confidence of the local populace by ensuring that they are being treated fairly. On one hand, it is fairly easy to assess our efforts to repel and defeat the enemy. We use intelligence reports to understand where the enemy is and what their plans are and we conduct operations to destroy their resources, deny them access to population centers, crush their will to fight, or dismantle their leadership network to cause disarray in their efforts. The usual device for assessing our actions is to gauge our performance against a set of Measures of Effectiveness (MOE) (United States, 2011). These MOEs are set in place before operations are conducted and our progression towards them is gauged during and after the operation. Meeting all MOEs is not necessarily required to call a mission a success, but having a plan in place to judge where we were against where we are allows us to understand whether our actions were successful. Understanding how much, or even if, we are changing the populations attitudes is much harder. Does the population support the fact that we are in *their* country, sometimes fighting *their* husbands, brothers and sons, sometimes destroying *their* infrastructure, or interrupting *their* way of life for what *we* promise to be a better, more fulfilling lifestyle? This type of question is at the forefront within the higher levels of the International Security Assistance Force (ISAF).

One way that we have tried to understand the people is by using survey data. Surveys are a proven, useful way to gauge public perception. They are used every day for all types of reasons from making predictions about an upcoming election to deciding whether a company should introduce a new product to market. Surveys are conducted in Afghanistan for mainly one reason. That reason is to answer the question: Are we winning the hearts and minds of the Afghans? Typically a survey will ask a series of questions that are demographic in nature, i.e. how old are you, are you employed, where do you live and so on. The meat of the questions, though, lies in the perception questions; i.e. how has your district government performed over the last six months or do you feel that the removal of the Taliban is good for the country? In some surveys, hundreds of these perception questions are asked all in an effort to answer one real question. We analyze these hundreds of questions asked to thousands of respondents to gauge our progression towards the MOE of winning hearts and minds of Afghan locals. Unfortunately when analysts do this type of work what they tend to do is pore over the list of questions and select some subset of questions that they feel will provide the best gauge of the MOE. This method of selection is a form of bias that the analyst is putting on their own analysis whether intended or not. Using a western mindset to decide what is important to Afghans is not a wise choice if the chooser does not have an intense understanding of the culture and history of Afghanistan which most analysts do not possess. Conducting these surveys is extremely expensive both in terms of money and man-hours and sometimes quite dangerous, especially in a place like Afghanistan. Why then would we spend so much time and money asking thousands of people hundreds of questions then only analyze one or two of them as a basis for our decisions?

In this research I intend to show that using Factor Analysis on these perception questions is a better choice. By using Factor Analysis we can use *all* of the perception questions in the survey and let the survey tell us which ones are important. Factor Analysis will group the questions of a survey in to categories known as Factors. We find that these Factors are groups of similarly-themed questions. Each question in the group has a correlation coefficient depicting how well a particular question correlates to the group. These values can also be thought of as weights of each question within the Factor. The questions serve as an observed variable (i.e. a question that was asked directly) and the Factors serve as unobserved variables (i.e. topics that were never brought up during the interview). The interesting thing about Factor Analysis is that we can now make judgments about how a topic has changed over time and where it has changed over time (through using the demographic data) without ever having to ask about the topic directly. In this research we use these Factors as indicators of progression toward an MOE. We generate indicators (factors) from survey data using Factor Analysis and assess them by using nonparametric statistics techniques to detect spatial and temporal changes throughout the country. Then, we use the concept of Exploratory Data Analysis (EDA) by displaying spatial and temporal changes that have occurred as part of an analytics suite to build a better knowledge base of popular perceptions. While EDA will not only give us a better understanding of *if*, and *where*, positive changes are happening, it will also give us the opportunity to further explore *why* positive changes are or are not happening by digging deeper into a particular area of interest to improve the situation.

BACKGROUND OF AFGHANISTAN SURVEY INSTRUMENTS

History

Once the focus of Operation Enduring Freedom shifted from a traditional, kinetic fight to counter-insurgency so did the emphasis on improving infrastructure, enhancing youth education, increased employment, the provision of healthcare for the populace along with better governance and development. Efforts were introduced to gauge the perceptions of the local population through surveys. Each of the surveys used in this research carry a slightly different purpose and scope. The surveys started as early as 2008 and have all continued in varying degrees of granularity and recurrence (every 2-4 months) since then. Afghanistan's administrative borders follow a model of Country, Province, District, and City similar to the way the United States follows a Country, State, County, City model. Afghanistan's Ministry of Interior officially recognizes 34 Provinces and 398 districts (CIA, 2014). The level of granularity to make judgments on a particular geographic level depends largely on the survey's sample size in that geographic level. Due to proprietary and classification reasons, these surveys will be referred to as Surveys A, B and C throughout this report. Each of these surveys is described in more detail below.

Survey A

The first iteration (each iteration is known as a Wave) of Survey A was conducted in the summer of 2008. This survey is conducted on a roughly quarterly basis (i.e. every 2-3 months) at a nationwide level. It can be parsed to the Provincial level or higher due to its sample size in each province. As with the other surveys in this research, even though each province and district has representation, not all areas of a district are reachable due to various reasons, chiefly security of the area. Each wave of this survey has approximately 12,000 respondents from all 34

provinces and all 398 districts. The pros of Survey A include the national scope and its longevity; 22 Waves of this survey have been conducted to date. The con of the survey is simply the degree to which the questions in the survey have changed over time. This is to be expected sometimes since the ideas and topics researchers are trying to understand will change just as the level of stability in the country has changed.

Key areas of discussion within Survey A include (in no particular order):

- Security and safety of roads and community
- The performance of the Government of the Islamic Republic of Afghanistan (GIROA), the Provincial Governor (PG) and the District Governor (DG)
- The corruption of various government officials
- Family's economic state, their satisfaction with healthcare and education, and their perceived future outlook of the same
- The Taliban's ability and propensity to rejoin the Afghanistan society
- Satisfaction with infrastructure such as roads, electricity and water
- Propensity for crimes and violence in the community
- The ability of GIROA and the Afghan National Army (ANA) to succeed without ISAF

Survey B

The first Wave of Survey B was conducted in the spring of 2010 and has been conducted on a roughly quarterly basis at a nationwide level. It can be parsed to the provincial level or higher due to its sample size in each province. Similar to Survey A not all areas are represented in Survey B for similar reasons; representation exists from only 26 provinces and 137 districts. 14 Waves of Survey B have been conducted with approximately 7,000 respondents per Wave; the most recent Wave taking place in October 2013. This survey has had a more consistent set of questions over its duration with only minimal change as compared to Survey A.

Key areas of discussion within Survey B include (in no particular order):

- The performance of the PG, ANA, and Afghan National Police (ANP)
- Satisfaction with infrastructure such as roads, electricity and water
- Abuse of power and corruption amongst various government officials
- Family's current quality of life and its future outlook

Survey C

Survey C is the survey that is most familiar to the author and is quite different than the previous two surveys in many ways and similar in others. Survey C is only focused on one province of Afghanistan, Helmand Province. This granularity of the survey is down to the district level since the sample size of each district is large enough to allow for statistically significant judgments on of each district. Survey C was initially conducted in October 2010 and has been conducted on a roughly quarterly (2-4 months) basis ever since. Each wave of this survey has approximately 4,000 respondents that are fairly evenly distributed amongst nine of Helmand's 12 districts.

Key areas of discussion within Survey C include (in no particular order):

- The performance of the DG in governance, infrastructure and employment opportunities, etc.
- General satisfaction with infrastructure, education and health services
- Performance of ANA and ANP
- Ability of various levels of leadership to solve disputes fairly, address corruption and provide security to their districts
- Expectations of future infrastructure, education and health services
- Fairness of district, provincial and national level judicial system

METHODOLOGY

Question Selection and Recoding

In order to conduct Factor Analysis we must have a complete data set to work with where each respondent has answered each question. A single missing answer to a question will not allow the mathematics of Factor Analysis to work. The first step in doing this is to determine which questions are the mainstays of the survey we are working with; we'll refer to these questions as Enduring Questions. In order to compare waves of the survey over time each

question must be asked each time. This is almost never the case because situations and interests eventually change over time. Determining the Enduring Questions is typically a labor intensive, time consuming process when starting out; one question not asked in any wave of the survey means that question in all other waves must be dismissed. Unfortunately by doing this we will be losing some of the data but we still have a high number of questions in each of our three surveys to work with; Survey A has 67 Enduring Questions, Survey B has 50 Enduring Questions and Survey C has 91 Enduring Questions.

Once we have our Enduring Questions we must carefully recode the answers numerically. Each of our Enduring Questions requires a three or five-point Likert Scale response. For example a five-point Likert Scale response might look something like “Strongly Agree”, “Somewhat Agree”, “Neither Agree nor Disagree”, “Somewhat Disagree” or “Strongly Disagree” while a three-point Likert Scale response might be “Yes”, “Neither” or “No”. Note that each of these examples has a neutral point and an equal number of positive and negative possibilities as options. For this research we use a numeric recoding scheme of 2 for most positive, 1 for the slightly positive, 0 for neutral, -1 for slightly negative and -2 for most negative for the five-point Likert Scale and a similar 2, 0 or -2 structure for three-point Likert Scale responses. Also, note that I used the words positive and negative instead of agree and disagree. This is due to the nature of question. For example, a question might be asked “Do you support the idea of the Taliban returning to power in Afghanistan?” From ISAFs perspective, a response of “Strongly Agree” would be negative. Conversely a response of “Strongly Agree” to the question “Do you support ISAFs role in reconstruction of your area?” would be seen as a positive to ISAF personnel.

Imputation of Missing Data

Next we must deal with those respondents who simply did not answer the question for any number of reasons. Initially these answers are not given a numeric value and are simply recoded as “NA”. But, as stated earlier, these NAs will not allow the mathematics of Factor Analysis to work. One approach to resolving NAs is to perform casewise deletion. For our analysis of survey data, this means that if a respondent has an NA for just one of the Enduring Questions then all of his/her data would be removed from the data set as if the interview never happened. If we used this casewise deletion approach we would be doing a great disservice to ourselves by losing a plethora of information. In order to save this data we can impute the missing values. Imputing the missing values means that we substitute an NA with some other value of -2, -1, 0, 1, or 2. We have to be careful here not to impute “too much” missing data. Unfortunately, there is no good metric that indicates that too much data is missing to use imputation (Stuart, 2010). For this research, we conclude that if a question has less than 20% of the responses missing we allow imputation of the missing values. Because of the measures already discussed in the previous section (i.e. the careful selection of Enduring Questions), none of our Enduring Questions violate this criterion. One way to approach imputation would be to randomly choose the value; this approach is a poor choice because we would simply be guessing. Another approach is to take the mean of the non-missing responses and substitute that value as the response. We could even use the median of smaller subsets of demographic data such as the district the respondent is from or their religion to try to find a more likely response from a person of similar background. But instead we will use a process known as Hot Deck Imputation (HDI). HDI finds a value for the missing value by comparing demographic data of the respondent whose value is missing to demographic data of respondents who answered the particular question.

The initial step of HDI is to make a subset of the data into a “donor class;” for this analysis the donor class for all Waves comes from the district column of the demographic data. Within each district subset, the data is split into two further subsets of “donor” and “receiver.” The receiver subset consists of data where responses are missing for a particular question and the donor subset contains data where all responses have recoded, numeric values. We choose the donor that is most like the receiver based on the receivers’ religion and tribal affiliation. As an example of imputation, say there is only a single donor who has matching demographic data to the receiver. In this case we would use the donor’s recoded response as the value for the receiver’s response. If there are multiple donors whose demographic data is an exact match to the receiver then one of the donors is picked at random and their recoded response is used as a replacement value for the receiver. If there is no exact matching donor then we look for donors that are closest to matching the receivers’ demographics. From this list the same rule applies amongst ties here as it did in the exact matching cases. The HDI process is done for all waves of data as necessary in order to produce complete numerically recoded sets of data. We are now able to use Factor Analysis on our survey data.

Factor Analysis

At its core Factor Analysis serves as a data reduction technique (Garrett-Mayer, 2006). Its use with survey data is ideal because we are typically dealing with a large number of variables (i.e. the questions asked in the survey) and we would like to be able to pare this data set down to some smaller number of variables without losing any of the

surveys intrinsic value. Factor Analysis not only does this for us but it also provides us with a correlation coefficient (r), also known as a Factor Loading, that indicates how well each question correlates to the derived factor; $-1 \leq r \leq 1$. These correlation coefficients serve as explanations of variance of each question within the factor and can be used later as weights. Factor Analysis helps to eliminate correlation between Factors while increasing correlation within Factors (Garrett-Mayer, 2010). In short, Factor Analysis provides us with groups of questions that are similar in nature and allows us to work with a smaller set of correlated variables (questions). In the end we have our observed variables, the questions and a new set of unobserved variables, the factors. While we could take the time to look over the questions and group these questions together manually, we would essentially be doing the same thing warned against in the introduction; using a western mindset to decide what is best for Afghans. Additionally, we would be, unintentionally or not, giving each question the same weight of 1. By using Factor Analysis we are able to analytically derive a set of correlated questions, called factors or indicators, and use these indicators as our basis for measuring if and how situations have changed across a province or district.

For this research we use the latest wave of each survey to conduct Factor Analysis which allows us to establish a baseline for each survey to make comparisons from. As a result of using Factor Analysis on the three surveys in this research we reduce our data sets as follows: Survey A is reduced from 67 observed variables to 13 unobserved variables, Survey B from 50 observed to seven unobserved, and Survey C from 94 to 22. For our purposes, we keep only those questions that have a Factor Loading of 0.4 or higher and -0.4 or lower and set the loadings of all other questions to 0 (Example in Table 4). Each factor consists of at least two and up to as many as 12 observed variables.

Results of Factor Analysis

Survey A

The 13 factors produced by Factor Analysis on Survey A can be viewed in Table 7 of the Appendix. Generally speaking the topics covered by the factors are the satisfaction and performance of various government entities in the areas of security, governance, provision of infrastructure, employment opportunities, and education for children. Two factors in particular are worthy of discussion here and they both deal with corruption. Corruption is very much common place in Afghanistan. For example, anecdotal stories suggest that if a District Police Chief does not keep at least some portion of his policemen's salary for himself then something must be wrong. I point this out because this type of context is important to understanding corruption in Afghanistan. One of the factors in particular has three questions in it and individually they ask how well GIRoA, the Provincial Governor and the District Governor reduce corruption in their respective governments. The other factor has four questions that more generally cover corruption by asking how serious of a problem corruption is, if it is a serious problem in the government and how those who conduct illicit activity are seen in the community. For U.S. citizens, any level of corruption would call for outrage by the public. In Afghanistan, it is the norm. Later in this research we will discuss various metrics to gauge how this factor has changed over time. We must keep in mind that what we think of as an acceptable level of corruption is much different than what our survey respondents think.

Survey B

The seven factors produced by Factor Analysis on Survey B can be viewed in Table 8 of the Appendix. A few of the factors produced for Survey B are similar to those in Survey A in that the questions in the factor deal with reduction of corruption and the satisfaction with governments provision of infrastructure, public services, and security. Two types of factors are worthy of discussion here for Survey B. The first is made up of six questions having to do with the improvement of the Afghanistan National Security Force (ANSF) in providing security and their ability to do so in the future without ISAFs assistance. ANSFs ability to maintain security gains post ISAFs departure is one of the key concerns of the entire operation. Regardless of their actual ability, which will not be truly known for years, we can now gauge the popular perception of this topic. The other factors have to do with abuses of power not only by different levels of leadership, but the Taliban as well. While abuse of power is similar to corruption, it is in fact different. Not only can we gain better understanding of the popular perception of this topic, we can do it at three different levels of leadership. These factors can help us understand these abuses at the formal level of leadership (Provincial and District government) the informal level (the Tribal Elders and the Maliks) and the enemy's leadership (the Taliban).

Survey C

The 22 factors produced by Factor Analysis on Survey C can be viewed in Table 9 of the Appendix. The first and most obvious observation of the results is there are quite a few more factors here than in the other surveys. The second notable observation is that this survey allows forces or other responsible parties to gain a much deeper understanding of Helmand Province in particular. While the general nature of the questions making up each factor is not much different than those in the other surveys, their level of detail and specificity to a particular group is quite

valuable. For example, several separate factors use questions that specifically discuss the Afghan National Police, Afghan National Army, ISAF, the Taliban, GIROA in general, members of Afghanistan's Parliament, Provincial Government, District Governments, and so on. Other factor topics which are different than our other two surveys include local militias, fair use of taxes, and acceptance of western societal norms. The Factor Analysis results for this survey allow an analyst to gain a much more specific perspective of this province by covering various areas of concern. This type of analysis could not truly be conducted on our other surveys due to their broad range of topics.

Deriving Indicators from Factors

The Factors that we derive are simply labeled initially as Factor 1, Factor 2, etc. Since our desired end state is look at progression towards a particular MOE we will give each factor a more descriptive name. For example, if we look at a selection of Factor Analysis results from Surveys A (Table 1) and C (Table 2) we can see the thought process on naming the indicators and how there are sometimes similarities between the different surveys.

Question	Factor Loading
How is the security situation in your mantaqa (village)?	0.4778
How safe do you feel using the roads in your district?	0.6055
Is security in your mantaqa better, the same, or worse than it was 6 months ago?	0.4079
How safe are the children in your village when they go to school and study in school?	0.5034
How safe do you feel traveling outside of your mantaqa during the day?	0.7209

Table 1. Factor 4 as derived by Factor Analysis on Survey A.

Question	Factor Loading
Do you agree or disagree that the District Government has improved security along roads in your area over the last three months?	0.4727
Do you agree or disagree that the District Government has improved security along roads in major urban areas over the last three months?	0.4149
Do you agree or disagree that the District Government has improved security in your area over the last three months?	0.4124

Table 2. Factor 5 as derived by Factor Analysis on Survey C.

The results displayed in Tables 1 and 2 from Surveys A and C, respectively, carry a similar theme. Namely, each question asks something about security in a particular area. Due to the fact that the security theme, along with safety, persists throughout the questions in Table 1, we give this factor a more descriptive name, "Satisfaction with Security and Safety in Local Area", and refer to it as an indicator from here on. Similarly, in Table 2, the security theme appears again, this time with more of a focus on roads. In this case we name the indicator "Satisfaction with Security along Roads." While a different author may come up with a slightly different name for the indicator, the theme of the indicator would have to remain the same. The exact name of the indicator is not important; however, what is important is that we now have an analytically derived indicator that we can measure against. Each question asked carries a particular weight, the Factor Loading, which will be used in the next step to make these comparisons.

Perhaps the most important takeaway from this is that no one person chose this grouping of questions. In essence, the group of respondents as a whole decided that the topics of these questions were important enough that, individually, each respondent tended to answer each of the set of questions in a similar manner. That last sentence often causes confusion. It's not that everyone in the survey answered, or even tended to answer, all questions of the indicator the same way. It's that each person, individually, tended to answer each of the indicators' questions similarly. A notional example of how one case of Factor Analysis might work on a dummy data set is helpful to understand this concept better. Table 3 shows a dummy data set with six questions and six respondents. Each answer is recoded similarly to the structure we are using in this research. At first glance, an obvious pattern may not be noticeable. But a closer, more studious inspection reveals something about two groups of the six questions. Each respondent tended to answer questions 1, 4, 5 and 6 the same way. In fact, in this extreme example they each answered them exactly the same way. Similarly, they each answered questions 2 and 3 the same way. While

between respondents the answers to questions 2 and 3 are sometimes different and on each end of the integer spectrum of -2 to 2, the respondent that answered Q2 with a particular value always answered Q3 with that same value. Even if the reader was able to quickly notice the pattern immediately within this dummy data set, it would be impossible to expect anyone to notice these patterns on data sets of 67 questions and upwards of 12,000 respondents, like Survey A, or 93 questions and 4000 respondents, as is the case in Survey C.

Respondent	Questions					
	Q1	Q2	Q3	Q4	Q5	Q6
A	1	-2	-2	1	1	1
B	2	0	0	2	2	2
C	-2	1	1	-2	-2	-2
D	1	-2	-2	1	1	1
E	0	2	2	0	0	0
F	-1	2	2	-1	-1	-1

Table 3. Example of Matrix used for Factor Analysis

Question	Factor 1	Factor 2
Q1	0.53	0
Q2	0	0.68
Q3	0	0.72
Q4	0.47	0
Q5	0.49	0
Q6	0.62	0

Table 4. Example of Factor Analysis output

All of the indicators (44 in total amongst the three surveys) are named in the same manner as described above. A full list of each surveys indicator, its questions, and the factor loadings can be found in Appendix A.

Calculating Respondent Indicator Scores

After recoding all of our survey responses to numeric values and conducting Factor Analysis on the last wave of each of our surveys, we want to calculate an indicator value for each respondent in each survey. We'll refer to this value as the Score. To do these calculations we will be using two matrices. The first matrix, **A**, is the recoded survey responses. It is an $n \times m$ matrix where n = the number of respondents in the survey wave and m = the number of Enduring Questions in the survey. The second matrix, **B**, is the survey's Factor Analysis output. This matrix is an $m \times p$ matrix where m is, as in **A**, the number of Enduring Questions in the survey and p is the number of factors that were generated from Factor Analysis. To calculate each Score, we use simple matrix multiplication on our two matrices as described in Equation 1. The output of Equation 1 for each survey wave will be an $n \times p$ matrix where n and p are as described previously.

$$AB_{ij} = \sum_{k=1}^m A_{i,k} B_{k,j}$$

Equation 1. Matrix Multiplication

The results of Equation 1 on Tables 3 and 4 can be seen in Table 5. We can see in Table 5 that each respondent in our survey now has a Score for each indicator. A layman's example of the math follows: for respondent A in Table 3 we would take each of his recoded answers and multiply them by the corresponding Factor Loading and sum those products. So we find the sum of $(1 \times .53) + (-2 \times 0) + (-2 \times 0) + (1 \times .47) + (1 \times .49) + (1 \times .62)$ equals 2.11. This means respondent A's Score for Factor 1 for our dummy survey is 2.11. We use this process for each respondent in each wave of each of our actual surveys to calculate all factor Scores for all survey respondents. Throughout this process we have saved the respondents' home district or province to represent changes geographically which will be discussed later.

Respondent	Score	
	Factor 1	Factor 2
A	2.11	-2.8
B	4.22	0
C	-4.22	1.4
D	2.11	-2.8
E	0	2.8
F	-2.11	2.8

Table 5. Respondent Scores

Measuring Change of Indicators Spatially and Temporally

For this research, we want to know if change has occurred for a particular indicator from one wave of a survey to the next for each surveyed province or district of Afghanistan. Surveys A and B will show provincial level change while Survey C will show district level change. In order to do this we'll use the Mann-Whitney test to compare the Scores of one wave of a survey to the Scores of another wave of the same survey and determine if the Scores tend to be equal between the two survey waves or larger in one wave than the other. The Mann-Whitney test is the nonparametric equivalent to the two-sample t-test that is used to compare the means of two data sets and assumes the means are normally distributed. One key difference between the Mann-Whitney test and the two-sample t-test is that the t-test is used to determine if the means of two sets of data are equivalent while the Mann-Whitney test is

used to determine if one set of data tends to have larger values (Scores) than the other regardless of the mean of each data set. As such, the Mann-Whitney test is robust to the influences of outliers.

To conduct the Mann-Whitney test we will use geographic subsets of each survey wave. Surveys A and B will be parsed in to subsets of each province while Survey C will be parsed in to subsets by district. For simplicity sake, we'll refer to provinces or districts as regions. Our hypotheses to be tested are formally stated below where X is the set of Scores from a particular region in one wave of a survey and Y is the set of Scores from the same region in some later wave of the same survey:

$$\begin{aligned} H_0: P(X > Y) &= P(X < Y) \\ H_1: P(X > Y) &\neq P(X < Y) \end{aligned}$$

A less formal statement of the hypotheses is:

H_0 : The probability that an X observation exceeds a Y observation is 0.5

H_1 : The probability that an X observation exceeds a Y observation is not 0.5

To conduct the test we let n equal the number of respondents in the wave we are measuring from, we'll call it the "From Wave," and let m equal the number of respondents in the wave we are measuring to, the "To Wave" and $N = n + m$ (these variables will be used later in this section). We then create a table with N rows and 3 columns. Column 1 is the indicator Score, Column two ("Source") is used to indicate which sample the Score comes from (i.e., "From Wave" or "To Wave"), and the third column indicates the rank of the Score. We then sort the table by Score from lowest to highest. In the case of ties amongst Scores, of which there are many due to the nature of our data, we average the ranks of each of the ties and each tie score receives the same ranking. Table 6 shows the output of these calculations for a hypothetical example. We then calculate a test statistic, T, which will determine any statistical significance of difference between the two waves.

Score	Source	Rank
2.2	From Wave	1.5
2.2	From Wave	1.5
2.3	From Wave	4
2.3	From Wave	4
2.2	From Wave	4
2.4	To Wave	7
2.4	To Wave	7
2.4	To Wave	7
2.5	To Wave	9.5
2.5	To Wave	9.5
3.8	To Wave	11
5.8	From Wave	12

The calculation of T is quite simple as shown in Equation 2. It is simply a summation of the ranks of each Score in our From Wave.

$$T = \sum_{i=1}^n \text{Rank}(\text{From Wave}_i)$$

Equation 2. Test Statistic for Mann-Whitney

Unfortunately, as mentioned before, there are a high number of tied ranks. Even in the small example above ten of the 12 Scores are tied with at least one other Score. Due to this high number of ties in our survey datas' Scores we must "subtract the mean number of observations in each data set from T and divide by the standard deviation" (Conover, 1999) to get a new test statistic known as T_1 (Equation 3). In this equation, $\sum_{i=1}^N R_i^2$ refers to the sum of squares of all N ranks or average ranks actually used in both samples" (Conover, 1999).

$$T_1 = \frac{T - n \frac{N+1}{2}}{\sqrt{\frac{nm}{N(N-1)} \sum_{i=1}^N R_i^2 - \frac{nm(N+1)^2}{4(N-1)}}$$

Equation 3. Test Statistic for Mann-Whitney with Ties Amongst Ranks

Once the T_1 Test Statistic is derived we use this to find a p-value. This p-value will tell us, at a given confidence level, whether to accept or reject the null hypothesis; for this research we use a 95 % confidence level ($\alpha = 0.05$) for each hypothesis test. If the p-value tells us to accept the null hypothesis then we know that there has not been a statistically significant change of the indicator from one wave to the next. However, if the p-value tells us to reject the null hypothesis then we only know that there *is* a statistically significant difference between the two waves but we don't know whether the change is higher or lower (i.e. better or worse). To fix this we create a function in R which, instead of requiring the output of a p-value, requires the output to be a -1 if the change is negative, a 1 if the change is positive, or a 0 if there is no change. In order to do this we use the same T_1 test statistic that we calculated above and return a -1 if T_1 is in the upper tail, a 1 if it is in the lower tail, or a 0 otherwise.

As an example of this, we calculate Equations 2 and 3 using the data in Table 6. Here $T = 27$ and $T_1 = -1.956$. This T_1 value would tell us to reject H_0 with $\alpha = 0.05$. Rejecting H_0 is equivalent to saying there is a statistically significant difference between the two data sets and the value of T_1 tells us that, in this case, the Scores in the To Wave tend to be larger than those in the From Wave. In fact, we would further state that there has been a positive change for this indicator. We use this particular hypothetical example to point out the value of using this

methodology vice relying on comparing the means of two sets of numbers. The mean of our From Wave Scores is 2.85 while the mean of our To Wave Scores is 2.67. This illustrates the virtue of using the Mann-Whitney test to assess trends. While the mean of the population's Scores would falsely lead us to believe the population's Scores trended down from the From Wave to the To Wave, the Mann-Whitney test concludes the To Wave Scores tended to be larger than the From Wave Scores in our example. It is clear by inspection that even with the presence of an outlier from the From Wave, the To Wave Scores in Table 5 tended to be higher than those of From Wave.

DISPLAYING SPATIO-TEMPORAL INDICATOR CHANGE BY SURVEY

The statistical significance of change is an important first step in understanding what is happening in our area of interest. Relaying that information to decision makers is perhaps the most challenging aspect. To accomplish this, we will use one of the most common, readily available pieces of software; Microsoft Excel®. Using Excel, we have created an analytics suite, dubbed MICH (Measure of Indicator Change), and designed it to be user friendly and only in need of minor explanation to a first time user. We'll use the MICH analytics to show three separate metrics of each indicator within the spreadsheet application; the change of the indicator across waves as derived from our Mann-Whitney test, the mean of the indicator in each wave by region, and the proportion of support for each indicator over the duration of the survey by region. Due to constraints we will only show examples from Survey C.

MICH uses the indicator Scores and the 1, 0, and -1 codes indicating positive, neutral, or negative change as the basis of its analytics suite. A form in Excel allows the user three options: First, which wave they would like to measure from; our From Wave, Second, the wave they would like to measure to; our To Wave, and lastly, an option to choose one of the indicators. In Figure 1 we see the default options for Survey C where clicking the "View Map" radio button would show the user the three metrics for the "Satisfaction with Security Along Roads" indicator. To choose different waves or a different indicator, the user selects from the dropdown box. Upon clicking View Map the user is presented with a display of our three metrics.

Figure 1. Wave and Indicator Choice Form

On the map displayed in Figure 2 we see an area of Afghanistan that was commanded and controlled by ISAF's Regional Command Southwest (RC (SW)). The area actually covered two provinces even though Survey C only polled respondents in one, Helmand. Also we see each district of Helmand Province. Three of the Helmand districts are "greyed out", this indicates that our surveys did not poll any respondents from these districts and therefore they receive a rating of "Not Measured."

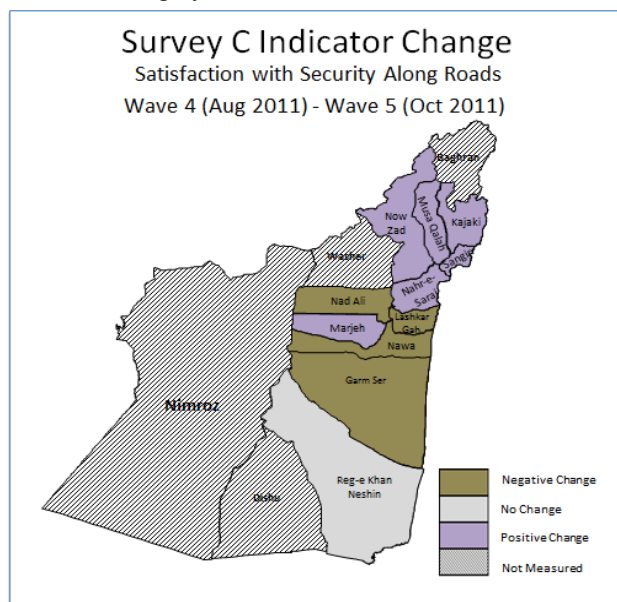
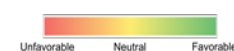


Figure 2. Map of Statistically Significant Change

Next we see six districts with a purple coloring. Purple indicates a positive change for the indicator. We see that the Reg-e Khan Neshin District is grey; this indicates that there was no statistically significant change here. Lastly we see four districts in brown; this indicates that there was a negative change for the indicator in each of these four districts. What we have shown here is simply whether there has been a change or not, and if so in which direction, but this map alone does not necessarily afford a clear idea of how the general sentiment is to this indicators topic. In other words, we know that our northern districts are more satisfied with the security along the roads in Wave 5 than they were in Wave 4, but we don't know where, on any type of scale, the actual sentiment lies. To help answer that question we look at Figure 3 which is simultaneously displayed on the screen to the user.

Figure 3 displays a color coded chart that ranges from red to green for each of our measured districts. A red cell indicates the sentiment of the indicator (during the

time period of the surveys wave) to be most unfavorable, green indicates most favorable and yellow indicates neutrality. Rather than simply using the three colors, we use a gradient scale to give us all possible color indications from red to yellow to green. The color of each cell represents the mean Score of the indicator for each district in the province. The extremes of these Scores are set by multiplying -2 by each factor loading for the indicator to get the most unfavorable Score possible and multiply the same factor loadings by 2 to get the most positive Score possible; each indicator will have a different value for the most favorable and most unfavorable Score because each factor has different questions and different loadings. Using the color scheme shown in Figure 3 with the statistical significance of change information from Figure 2 allows us to have a better understanding of the full picture. For example, we know from Figure 2 there was a negative change for the “Satisfaction with Security Along Roads” indicator in the Lashkar Gah District. Figure 3, however, brings new information and broadens our understanding a little more. We can see that in Lashkar Gah for both Waves 4 and 5 the general sentiment for this indicator is positive. When we consider both pieces of information, the takeaway is that while there has indeed been a decline from Wave 4 to Wave 5 for Lashkar Gah’s Satisfaction with Security along Roads, it is still quite favorable.



District	Wave 4	Wave 5
Garmser		
Kajaki		
Lashkar Gah		
Marjah		
Musa Qalah		
Nad Ali		
Nahr-e-Saraj		
Now Zad		
Nawah		
Reg-e Khan Neshin		
Sangin		

Figure 3. State of Indicator

The third set of metrics displayed to the screen are charts for each district and the province showing the proportion of support for an indicator over the duration of the survey, as shown in Figures 4 and 5, regardless of the Waves selected by the user. These charts give a slightly different perspective of the indicator than shown in Figure 3. The values displayed by the line charts are again derived from the set of respondent Scores by district or province similar to Figure 3. This time, however, we are showing them as a percentage where the higher values indicate more support for the indicator; additionally we start to get a broader trend perspective here as well. As an example, if the range of unfavorable to favorable were 0 to 10, respectively, and the average Score for a district were 7.5, then we would say there proportion of support for the indicator was 75%. Additionally we see error bars displayed on the charts which represent a 95% Confidence Level of the proportion. These error bars are directly affected by the number of respondents for the survey. Therefore, the error bars are much narrower in the provincial level charts than they are in the district level charts.

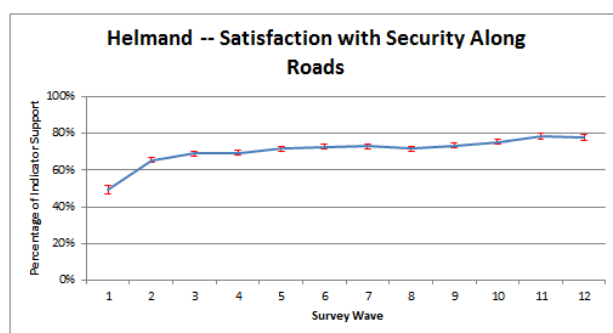


Figure 4. Proportion of Support for Helmand Province

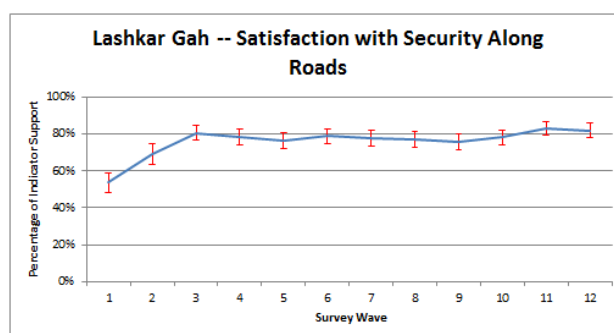


Figure 5. Proportion of Support for Helmand Province

As our third and final piece of the analytics suite shows in Figures 4 and 5, we can see that satisfaction with road security has been quite high in Lashkar Gah, and across Helmand Province in general, since about Wave 3 of the survey. We can now state the following from our survey data: According to Survey C the sentiment regarding the Lashkar Gah’s satisfaction with security along roads declined from Wave 4 to Wave 5, the average Lashkar Gah resident remained generally positive about road security, and that sentiment has continued to be quite steady for the duration of Survey C.

While we only have space to show a small number of indicator metrics in this paper, understand that the MICH is designed to be able to do display these metrics for all surveys in the research for any two time periods for every indicator; this equates to thousands of comparisons to be measured at the district and provincial levels of Afghanistan with only a few clicks. The metrics are relatively easy to understand and the availability of Excel to run the analytics suite is nearly universal.

CONCLUSIONS

One thing the MICH cannot do is determine the why. We can help leaders use survey data to understand popular perceptions of topics and provide them a means to use that information to formulate better decisions, but we cannot use MICH to tell us why something changed for the better or worse. To do that, one must dive even deeper into the topic of interest. In the end, the MICH serves an analytically derived exploratory data analysis tool designed to inform decisions and to incite further areas of study.

This research has demonstrated the integration of proven statistical techniques (Factor Analysis and the Mann-Whitney test) on large, complex data (surveys) into an exploratory data analysis tool that provides great potential for an organization to help understand and track a population's perception over time. The MICH analytics suite takes the information one step further by providing quick and easy to understand metrics of the analytically derived and measured indicators. The MICH is scalable and modifiable to provide an end user a means to quickly inform leadership in a wide range of organizations when developing and analyzing long-term plans involving populations.

REFERENCES

- CIA World Factbook, (2014), *Afghanistan*. The CIA World Factbook Website: <https://www.cia.gov/library/publications/the-world-factbook/geos/af.html>
- Conover, W. (1999). Some methods based on ranks. In W. Conover, *Practical Nonparametric Statistics* (pp. 269–286). New York: John Wiley & Sons.
- Garrett-Mayer, E. (2006). *Statistics in Psychosocial Research*. Johns Hopkins University Website: Retrieved Feb 17, 2014: <http://ocw.jhsph.edu/courses/statisticspsychosocialresearch/pdfs/lecture8.pdf>
- Stuart, E. A. (2010, June). *Academy Health: Advancing research, policy and practice*. Retrieved Feb 19, 2014, from Academy Health: <http://www.academyhealth.org/files/2010/sunday/StuartE.pdf>
- United States Department of Defense. (2011). Use of effects. In *Commander's handbook for assessment planning and execution* (p. 1-8). Suffolk, VA.

Appendix

Satisfaction with Security Along Roads	Satisfaction with Personal Quality of Life and Economic Situation
How is the security situation in your mantaqa?	Has your family's economic situation gotten better, worse, or stayed the same in the past 12 months?
How safe do you feel using the roads in your district?	How satisfied are you with the current quality of your life?
Is security in your mantaqa better, the same, or worse than it was 6 months ago?	Please tell me, in the future, do you think your life conditions will improve, worsen or stay the same?
How safe are the children in your village when they go to school and study in school?	Concerns of Corruption and Illegal Activity
How safe do you feel traveling outside of your mantaqa during the day?	Do you think it is possible for Afghan farmers to provide for their families without growing poppies?
Professionalism, Performance and Future Outlook of ANA and ANP to Provide Security	Do you believe that corruption in the Government affects your daily life?
Have you heard of or seen the police in your mantaqa doing anything improper?	Please tell me if you agree or disagree that corruption is a serious problem in the government?
Has the performance of police in your district worsened or improved in the last 6 months?	How are those who run and profit from the drugs trade seen by others in your community?
How capable are the police of protecting your mantaqa?	Overall Satisfaction with Provincial Governor to Provide Public Services (i.e. Security, Economy, and Reconstruction)
Do you believe local government will be capable of maintaining security after ISAF leaves your area?	How well does the Provincial Governor do his/her job overall?
Have you heard of or seen the National Army doing anything improper?	How well does the Provincial Governor do his/her job of securing the province?
In the last 12 months has anything prevented school aged boys and girls from going to school?	How well does the Provincial Governor do his/her job of improving the economy?
Satisfaction with Government Services -- Infrastructure	How well does the Provincial Governor do his/her job of development and reconstruction of the province?
How satisfied are you with the provision of roads in your area?	How well does the Provincial Governor do his/her job of providing essential public services?
How satisfied are you with the provision of water in your area?	Overall Satisfaction with District Governor to Provide Public Services (i.e. Security, Economy, and Reconstruction)
How satisfied are you with the provision of healthcare in your area?	How well does the District Governor do his/her job of development and reconstruction of the district?
How satisfied are you with the provision of electricity in your area?	How well does the District Governor do his/her job of improving the economy?
How satisfied are you with the provision of education for school-aged children in your area?	Propensity for Acts of Violence and Threats in Community
Reduction of Corruption in Government	How often have criminals committed crimes inside your community in the past three months (e.g. robbery, extortion)?
How well does the Government of Afghanistan do its job of reducing corruption in the government?	How often have Anti-Government Elements come to your community and threatened the people or committed acts of violence in the past three months?
How well does the Provincial Governor do his/her job of reducing corruption in their administration?	Overall Satisfaction with GIROA to Provide Public Services (i.e. Security, Economy, and Reconstruction)
How well does the District Governor do his/her job of reducing corruption in their administration?	How well does the Government of Afghanistan do its job of improving the economy?
Provision of Security by Government	How well does the Government of Afghanistan do its job of development and reconstruction of Afghanistan?
How well does the Government of Afghanistan do its job overall?	Propensity for ANA and ANP Presence in Community
How well does the Government of Afghanistan do its job of securing the country?	How often do you see the police in your mantaqa?
How well does the Provincial Governor do his/her job overall?	How often do you see the National Army (ANA) in your mantaqa?
How well does the Provincial Governor do his/her job of securing the province?	Propensity for Children to Attend School
How well does the District Governor do his/her job overall?	Are 6-14 years old boys in your household going to school or madrasa?
How well does the District Governor do his/her job of securing the district?	Are 6-14 years old girls in your household going to school or madrasa?

Table 7. Indicators Derived from Factor Analysis on Survey A.

Improvement of Security and Ability to Maintain Security by ANSF in the Future	
In the past 6 months, do you think the Taliban has grown stronger, weaker or remained the same?	
Do you think the ANSF will be able to defeat the Taliban in the next few years?	
How is the security in your mantaqa?	
Is the security in your mantaqa better, the same or worse than 6 months ago?	
Has the performance of police in your district worsened or improved in the last 6 months?	
Do you believe local government will be able to maintain security after ISAF leaves your area?	
Satisfaction with Government Services -- Infrastructure	
How satisfied are you with the provision of the roads in your area?	
How satisfied are you with the provision of water in your area?	
How satisfied are you with the provision of healthcare in your area?	
How satisfied are you with the provision of electricity in your area?	
How satisfied are you with the provision of education for school-aged children in your area?	
How satisfied are you with the provision of jobs/employment in your area?	
How satisfied are you with the provision of security in your area?	
Reduction of Corruption in Local and National Government	
How is corruption compared to 6 months ago within the District Government?	
How is corruption compared to 6 months ago within the Provincial Government?	
How is corruption compared to 6 months ago within the National Government?	

Overall Satisfaction with Provincial Governor to Provide Public Services (Security, Economy, and Reconstruction)	
How well does the Provincial Governor (PG) do his/her job overall?	
How well does the PG do his/her job of securing the Province?	
How well does the PG do his/her job improving the economy?	
How well does the PG do his/her job of reducing corruption in their administration?	
How well does the PG do his/her job of providing essential public services?	
Abuse of Power by Provincial and District Leadership	
Do you think the Provincial Governor misuses his/her power?	
Do you think the District Governor misuses his/her power?	
Do you think the Provincial Police Chief misuses his/her power?	
Do you think the following District Police Chief misuse his/her power?	
Abuse of Power by Local and Informal Leadership	
Do you think the Village/Tribal Elders misuse their power?	
Do you think the Malik misuses his/her power?	
Abuse of Power by Taliban or Other Opposing Government Elements	
Do you think the Taliban misuse their power?	
Do you think the following Mukhalfeen-e dawlat their power?	

Table 8. Indicators Derived from Factor Analysis on Survey B.

Satisfaction with Provincial Government and Court System	
Confidence in ability to resolve dispute in fair manner - Prov Governor	
Confidence in ability to resolve dispute in fair manner - District Huquq	
Confidence in ability to resolve dispute in fair manner - Provincial Huquq	
Confidence in ability to resolve dispute in fair manner - Dist Prosecutor	
Confidence in ability to resolve dispute in fair manner - Prov Prosecutor	
Confidence in ability to resolve dispute in fair manner - LKG Court	
Confidence in ability to address corruption - Provincial Governor	
Confidence in ability to address corruption - District Huquq	
Confidence in ability to address corruption - Provincial Huquq	
Confidence in ability to address corruption - District Prosecutor	
Confidence in ability to address corruption - Provincial Prosecutor	
Confidence in ability to address corruption - LKG court	
Confidence in the Legitimacy of the ANP	
Satisfaction/Dis - Role in keeping district secure - Police (ANP)	
Agree/Dis - National Police (ANP) is honest and fair with the local people.	
Agree/Dis - National Police (ANP) is unprofessional and poorly trained.	
Agree/Dis - National Police (ANP) help improve the security in my area	
Agree/Dis - National Police (ANP) efficient arresting to bring to justice	
Agree/Dis - National Police (ANP) in my region act interest of local people	
Agree/Dis - National Police (ANP) are not sanctioned when they misbehave	
Confidence in ability to resolve dispute in fair manner - Police	
Confidence in ability to address corruption - Police	
Satisfaction/Dis - Role in keeping district secure - Dist gov / authority	
Agree/Dis - (generally speaking) GiRoA are doing a good job.	
Satisfaction with Employment Opportunities, Bazaars, and Justice Systems	
Agree/Dis (3 mnths Dist) DG has improved employment opportunities	
Agree/Dis (3 mnths Dist) DG has increased effectiveness - justice (State)	
Satisfaction/Dis - services in area - Government Justice Services	
Expectations (next 6 mnths, district) - The central bazaar	
Expectations (next 6 mnths, district) - Availability economic oport / jobs	
Expectations (next 6 mnths, district) - Effectiveness state justice system	
Confidence in the Legitimacy of ISAF	
Satisfaction/Dis - Role in keeping district secure - ISAF / Foreign Forces	
Confidence in ability to resolve dispute in fair manner - Foreign forces	
Confidence in ability to address corruption - Foreign forces	
Satisfaction with Security Along Roads	
Agree/Dis (3 mnths Dist) DG has improved security - along roads in my area	
Agree/Dis (3 mnths Dist) DG has improved security - roads major urban areas	
Agree/Dis (3 mnths Dist) DG has improved security in my area	
Confidence in the Legitimacy of Local Militias	
Satisfaction/Dis - Role in keeping district secure - Local warlords militia	
Confidence in ability to resolve dispute in fair manner - Local militia	
Confidence in ability to address corruption - Local militia	
Satisfaction with Utilities (Electricity, Drinking Water, and Irrigation Systems)	
Agree/Dis (3 mnths Dist) DG has improved electricity supply	
Agree/Dis (3 mnths Dist) DG has improved irrigation services	
Agree/Dis (3 mnths Dist) DG has improved household/drinking water	
Expectations (next 6 mnths, district) - The supply of electricity	
Expectations (next 6 mnths, district) - Availability of irrigation water	
Expectations (next 6 mnths, district) - Availability of safe drinking water	

Expectation of District Improvements Over The Next Six Months	
Expectations (next 6 mnths, district) - Availability of education services	
Expectations (next 6 mnths, district) - Ability to move freely and safely	
Expectations (next 6 mnths, district) - The security in my area	
Expectations (next 6 mnths, district) - Availability of irrigation water	
Expectations (next 6 mnths, district) - Access to information and news	
Confidence in the Legitimacy of the ANA	
Confidence in ability to resolve dispute in fair manner - Army	
Confidence in ability to address corruption - Army	
Confidence in the Legitimacy of the Taliban	
Satisfaction/Dis - Role in keeping district secure - Taliban	
Confidence in ability to resolve dispute in fair manner - Taliban	
Confidence in ability to address corruption - Taliban	
Satisfaction with Availability and Access to Government Healthcare Services	
Agree/Dis (3 mnths Dist) DG has improved healthcare services	
Satisfaction/Dis - services in area - Government Health Services	
Expectations (next 6 mnths, district) - Availability of health services	
Satisfaction with Availability and Access to Government Education Services	
Agree/Dis (3 mnths Dist) DG has improved education services	
Satisfaction/Dis - services in area - Government Education services	
Expectations (next 6 mnths, district) - Availability of education services	
Confidence in the Legitimacy of Members of Parliament	
Confidence in ability to resolve dispute in fair manner - Member of Parliament	
Confidence in ability to address corruption - Members of Parliament	
Acceptance of Western Societal Norms	
Acceptable/Not - Listening to music	
Acceptable/Not - Watching television	
Acceptable/Not - Talking about issues related to money	
Confidence in the Legitimacy of the District Governor	
Confidence in ability to resolve dispute in fair manner - District Governor	
Confidence in ability to address corruption - District Governor	
Satisfaction with the Performance of GiRoA	
Agree/Dis - GiRoA listen/acts on behalf of people rather than own interests	
Agree/Dis - (generally speaking) GiRoA are doing a good job.	
Satisfaction with Security in General	
Rate - Change in security situation in District (compared to 6 mnths ago)	
Rate - (in general) security situation in area	
Satisfaction with District Governments Use of Taxes to Improve Services and Access to News	
Agree/Dis (3 mnths Dist) DG has provided better access to information news	
Agree/Dis (3 mnths Dist) DG has used taxes to improve public services	
Confidence in the Legitimacy of the District Court	
Confidence in ability to resolve dispute in fair manner - District Court	
Confidence in ability to address corruption - District Court	
Confidence in the Legitimacy of the Provincial Governor	
Confidence in ability to resolve dispute in fair manner - Prov Governor	
Confidence in ability to address corruption - Provincial Governor	
Confidence in the Legitimacy of the District Huquq	
Confidence in ability to resolve dispute in fair manner - District Huquq	
Confidence in ability to address corruption - District Huquq	
Confidence in the Legitimacy of the District Prosecutor	
Confidence in ability to resolve dispute in fair manner - Dist Prosecutor	
Confidence in ability to address corruption - District Prosecutor	

Table 9. Indicators Derived from Factor Analysis on Survey C.

Section 5

***Building Models for Decision Making in
Complex Cross-Cultural Domains***

Using Geographic Information Systems Analysis for Mapping Adverse Events in an Active War Theater

Erman Cakit ^a, Waldemar Karwowski ^b

*^aDepartment of Industrial Engineering,
Aksaray University, 68100, Aksaray, Turkey*

*^bDepartment of Industrial Engineering and Management Systems,
University of Central Florida, Orlando, FL 32816-2993, USA*

ABSTRACT

This study is a review of adverse events throughout the war in Afghanistan by representing the mapping of these events, where we considered three types of adverse events in terms of number of people killed, wounded and hijacked, and their total number in the active war theater of Afghanistan over the period 2004-2010. The country was divided into seven regions for pattern analysis, where each region has different numbers of provinces, districts, and number of records. A point-density analysis was conducted to detect those areas where a high density of data point locations was concentrated. Based on the results obtained, it was concluded that the frequency of adverse events has increased from 2004 through 2010. The south-western region had the highest mean by district values than other regions and the whole of Afghanistan for all variables. On the other side, the north-western region had the lowest mean by district values than other regions and the whole of Afghanistan for all variables. When we compared the variables against each other, the number of people hijacked had the lowest values in total and average by district than the other variables.

Keywords: Adverse Events, Geographic Information Systems, Spatial Analysis, Point Density Analysis

INTRODUCTION

Afghanistan lies in the Central Asia and divided into 34 provinces and these provinces are subdivided into 400 districts (Figure 1). It has borders with Pakistan, Iran, Turkmenistan, Uzbekistan, Tajikistan and China. Afghanistan has 647,500 square kilometers and it is somewhat smaller than Texas, US. The population is approximately 30 million. Based on the United Nations (UN) Human Development Index, that index is calculated according to the health, education, and economic life of people, Afghanistan has been ranked 175th out of 185 members states of the UN (The 2013 Human Development Report). The Afghanistan geography does not land itself to trade, military, and operations. Therefore, this situation makes it difficult to secure the population and to improve their economic situation.

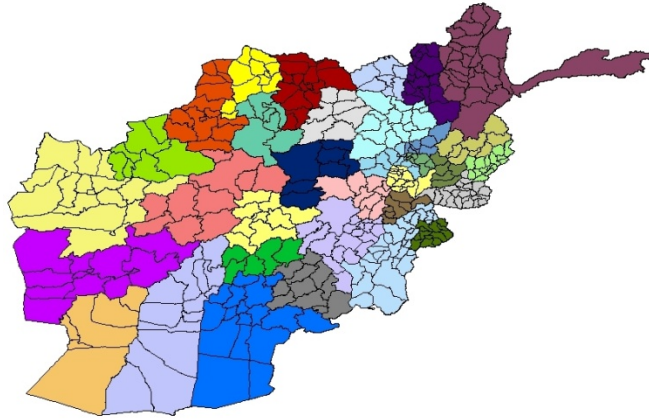


Figure 1: A map of the districts which are color grouped by province

Since terrorist attacks are not random in space and time, patterns do exist. It is possible to discover representative patterns in adverse activity or behavior over time and space by analyzing the geospatial intelligence on reported incidents, as stated by the Director of National Intelligence, Open Source Center (Federation of American Scientists, 2009). The study of spatial and temporal patterns in terrorist attacks is gaining increasing interest in the academic literature. LaFree et al. (2011) examined spatial and temporal patterns in all terrorist attacks attributed to the ETA (a Spanish separatist group) between 1970 and 2007, exploring how the approaches of terrorist groups might be related to their geospatial attack patterns over time. Berrebi and Lakdawalla (2007) considered how terrorists sought targets and focused on the spatial and temporal determinants of terrorism in Israel between 1949 and 2004. Based on the analysis, they found that space and time are necessary to describe the patterns of terrorism in Israel. Siebeneck et al. (2009) used historical data of terrorist incidents in Iraq from 2004 to 2006 to develop a series of analyses to describe terrorist activity spaces and counter-terrorist actions. This study applied clustering analysis, spatial and temporal statistics, and GIS methodologies to extract knowledge from patterns. Brown et al. (2004) focused on a specific type of event (suicide bombings) to explore difficulties in understanding and preventing terrorist attacks. They proposed a fusion model, combining spatial likelihood modeling of environmental characteristics and logistic regression modeling of demographic features, and concluded that the fusion model performs better than other methods such as kernel density estimation methods. Johnson and Braitwaite (2009) studied space-time clusters of improvised explosive device (IED) attacks and non-IED attacks in Iraq from January to June of 2005. Webb and Cutter (2009) described trends in terrorist incidents with respect to space and time in the United States spanning the years 1970 through 2004. In their paper, the authors highlighted several aspects of terrorist incidents by applying spatial statistics. In this study, we apply a geographic information systems (GIS) approach to understanding patterns of adverse events in Afghanistan.

METHODOLOGY

The dataset

This study uses data describing adverse events in Afghanistan provided by human social culture behavior (HSCB) program management. The adverse event data included information regarding the date of the event, the number of people killed, wounded, or hijacked, as well as the city, location in latitude and longitude coordinates and a simple event summary. Afghanistan was divided into seven regions for pattern analysis, where each region has different numbers of provinces, districts, and number of records (Table 1 and Figure 2). The dataset used in this study was acquired monthly at the district level, where it was limited to incidents occurring between 2004 and 2010. The dataset was sorted based on the order of year, month, and province info. ESRI's ArcMap® 10.1 software was used for mapping the dataset.



Figure 2: Regions of Afghanistan

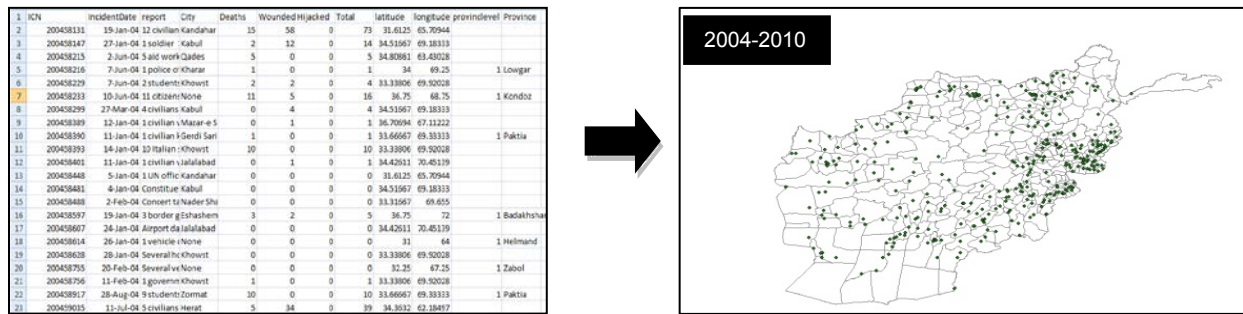


Figure 3: Mapping latitudes and longitudes into GIS format. The points illustrate the locations of adverse events in Afghanistan

Table 1: Province and district info for each region

Region	Number of Province	Number of District	Total number of records
Central	6	55	4620
Eastern	4	50	4200
North Eastern	4	67	5628
South Eastern	4	62	5208
Western	5	51	4284
North Western	5	55	4620
South Western	6	60	5040
Afghanistan (Total)	34	400	33600

Point density analysis

Mapping the number of adverse events by region and population provides location-based information to elucidate and understand the hidden patterns behind the adverse events. The density function spreads out a measured quantity of an input dataset throughout a landscape to produce a continuous surface (ESRI, 2013). Density maps are useful for illustrating where point locations are concentrated. In ESRI's ArcMap®, there are three density mapping tools including kernel density, line density, and point density. For this study, a point density analysis was conducted for the purpose of calculating a magnitude per unit area from point features that fall within a neighborhood around each cell. In other words, points that fall within the area are summed and then divided by the corresponding area size to calculate density value for each cell.

RESULTS AND DISCUSSION

The “WITSGEO” dataset was selected for the purpose of mapping adverse events and variables in this file are represented in Figure 3. As shown in Figure 3, we considered three types of variables in terms of number of people killed, wounded, and hijacked that took place during the time period 2004 through 2010. As a first step, all of these were variables mapped into GIS format and the output of the shape file is shown in Figure 3. The output provides information about the distribution of all adverse events throughout Afghanistan. Table 2 provides information for total number of people killed, wounded, hijacked and total number of adverse events and mean number of events by district in each region took place between 2004 and 2010. Based on the information in Table 2, the number of people killed, wounded, and hijacked and the total number of adverse events were the lowest in 2004 and the highest in 2010. Between 2004 and 2010, there was an increasing trend, in general. The corresponding mean by district values varied between 0 and 30. The south-western region had the highest mean by district values than other regions and the whole of Afghanistan for all variables. On the other side, the north-western region had the lowest mean by district values than other regions and the whole of Afghanistan for all variables. When we compared the variables against each other, the number of people hijacked had the lowest values in total and average by district than the other variables.

Table 2: Yearly total and average number of events by district in each region between 2004 and 2010

Region	Year	Number of people killed		Number of people wounded		Number of people hijacked		Total number of adverse events	
		Total	Mean by district	Total	Mean by district	Total	Mean by district	Total	Mean by district
Central	2004	24	0.43	56	1.02	3	0.05	17	0.31
	2005	27	0.49	62	1.13	1	0.02	49	0.89
	2006	92	1.67	296	5.38	5	0.09	112	2.04
	2007	268	4.87	391	7.11	30	0.55	151	2.75
	2008	213	3.87	489	8.89	65	1.18	186	3.38
	2009	295	5.36	1053	19.15	84	1.53	226	4.11
	2010	270	4.91	583	10.60	72	1.31	366	6.65
Eastern	2004	16	0.32	65	1.30	15	0.30	28	0.56
	2005	56	1.12	75	1.50	10	0.20	65	1.30
	2006	76	1.52	160	3.20	9	0.18	93	1.86
	2007	133	2.66	260	5.20	30	0.60	115	2.30
	2008	120	2.40	482	9.64	48	0.96	86	1.72
	2009	260	5.20	476	9.52	63	1.26	232	4.64
	2010	170	3.40	585	11.70	172	3.44	349	6.98
North Eastern	2004	20	0.30	14	0.21	0	0	7	0.10
	2005	6	0.09	9	0.13	0	0	9	0.13
	2006	18	0.27	62	0.93	0	0	32	0.48
	2007	137	2.04	227	3.39	11	0.16	46	0.69
	2008	41	0.61	61	0.91	12	0.18	51	0.76
	2009	134	2.00	185	2.76	49	0.73	202	3.01
	2010	307	4.58	479	7.15	110	1.64	321	4.79
South Eastern	2004	66	1.06	43	0.69	0	0	30	0.48
	2005	76	1.23	84	1.35	3	0.05	77	1.24
	2006	221	3.56	530	8.55	79	1.27	265	4.27
	2007	422	6.81	491	7.92	117	1.89	342	5.52
	2008	486	7.84	604	9.74	141	2.27	365	5.89
	2009	566	9.13	854	13.77	149	2.40	511	8.24
	2010	603	9.73	859	13.85	150	2.42	648	10.45

Western	2004	25	0.49	58	1.14	2	0.04	14	0.27
	2005	28	0.55	35	0.69	8	0.16	24	0.47
	2006	130	2.55	124	2.43	12	0.24	83	1.63
	2007	194	3.80	187	3.67	16	0.31	97	1.90
	2008	132	2.59	202	3.96	248	4.86	115	2.25
	2009	287	5.63	313	6.14	133	2.61	288	5.65
	2010	329	6.45	338	6.63	192	3.76	421	8.25
North Western	2004	0	0	2	0.04	0	0	3	0.05
	2005	17	0.31	53	0.96	0	0	17	0.31
	2006	30	0.55	32	0.58	0	0	34	0.62
	2007	25	0.45	59	1.07	1	0.02	34	0.62
	2008	51	0.93	37	0.67	1	0.02	44	0.80
	2009	92	1.67	66	1.20	35	0.64	115	2.09
	2010	143	2.60	253	4.60	141	2.56	284	5.16
South Western	2004	138	2.30	111	1.85	5	0.08	51	0.85
	2005	472	7.87	432	7.20	91	1.52	250	4.17
	2006	684	11.40	903	15.05	66	1.10	342	5.70
	2007	766	12.77	781	13.02	85	1.42	333	5.55
	2008	954	15.90	1016	16.93	76	1.27	374	6.23
	2009	1144	19.07	1240	20.67	104	1.73	551	9.18
	2010	1380	23.00	1750	29.17	113	1.88	916	15.27
Afghanistan	2004	289	0.72	349	0.87	25	0.06	150	0.38
	2005	682	1.71	750	1.88	113	0.28	491	1.23
	2006	1251	3.13	2107	5.27	171	0.43	961	2.40
	2007	1945	4.86	2396	5.99	290	0.73	1118	2.80
	2008	1997	4.99	2891	7.23	591	1.48	1221	3.05
	2009	2778	6.95	4187	10.47	617	1.54	2125	5.31
	2010	3202	8.01	4847	12.12	951	2.38	3305	8.26

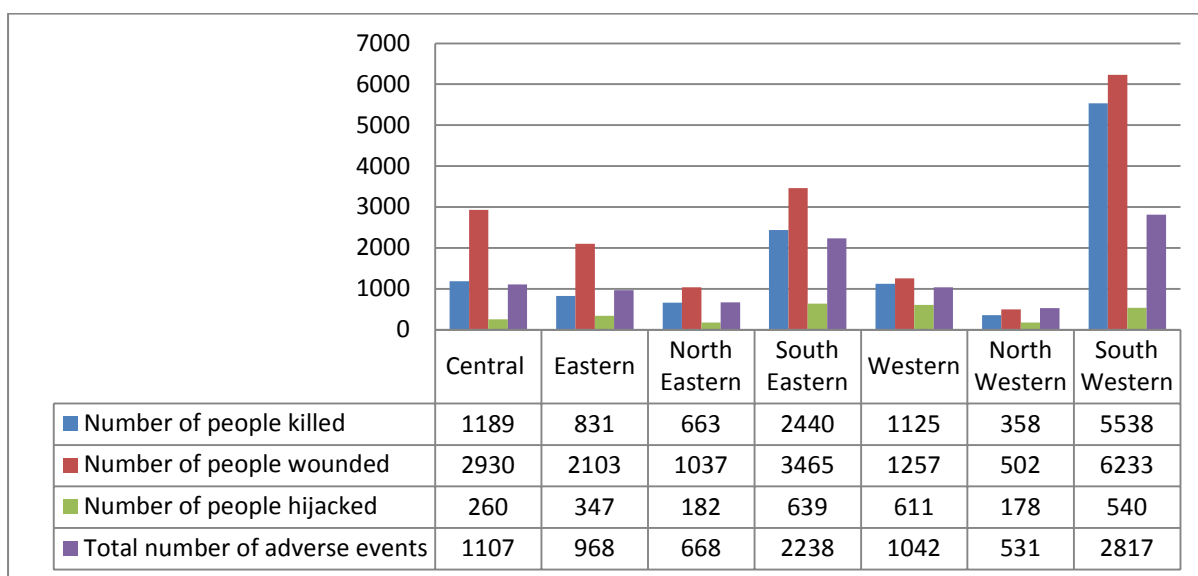


Figure 4: Regions with total number of each variable (2004-2010)

Figure 4 provides information about regional differences for number of people killed, wounded, hijacked and total number of adverse events occurred during the time period 2004 through 2010. Based on Figure 4, south western region has experienced the highest number of people killed, number of people wounded, and total number of adverse events, with a total of 5583, 6233, and 2817 respectively. South eastern region had the second highest number of people killed, number of people wounded, and total number of adverse events, with a total of 2440, 3465, and 2238 respectively. For number of people hijacked, south eastern region had the highest event levels, with a total of 639 and western region had the second highest event levels, with a total of 611.

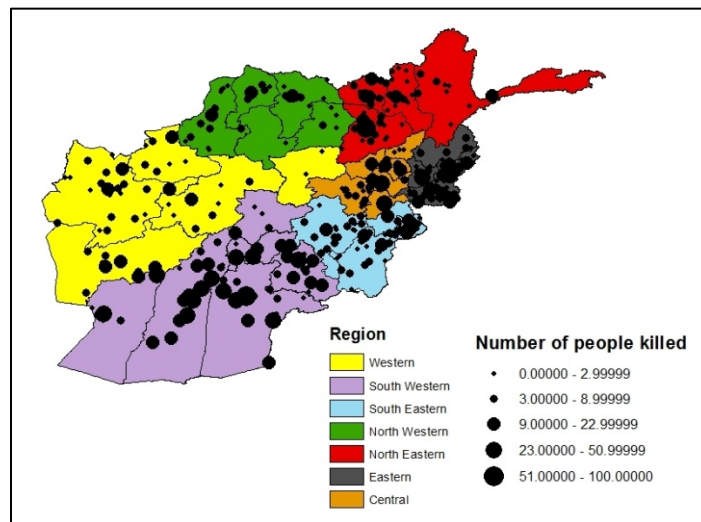


Figure 5: Number of people killed by region (2004-2010)

Visualizing the number of people killed, wounded, and hijacked by region reveals useful information about patterns. During the time period 2004 through 2010, most of the adverse events are concentrated in the Pakistan border along the south western, south eastern, and eastern regions (Figure 5, 6 and 7).

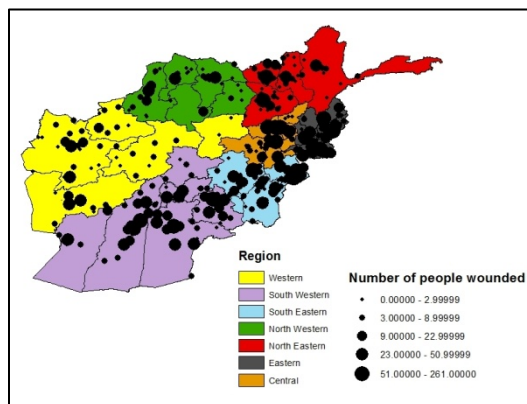


Figure 6: Number of people wounded by region (2004-2010)

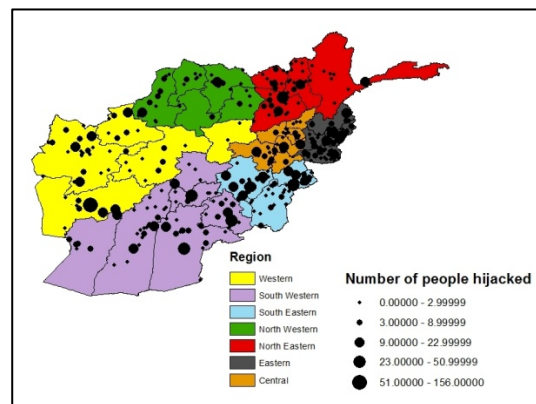


Figure 7: Number of people hijacked by region (2004-2010)

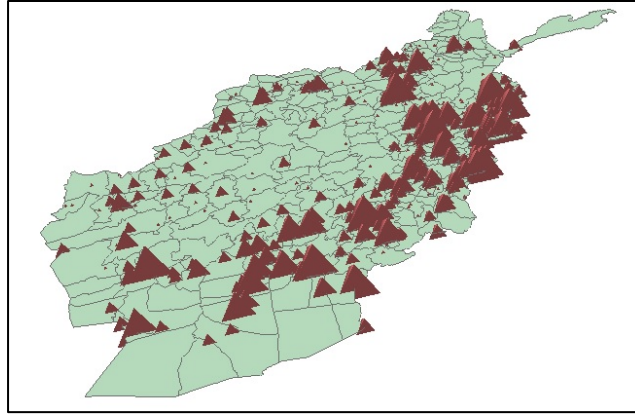


Figure 8: Three dimensional view of total number of people killed, wounded, and hijacked density (2004-2010)

Figure 8 provides information about three dimensional intensity view of total number of people killed, wounded, and hijacked. Three dimensional view is useful to compare the number of events happened in different locations. Based on Figure 8, total number of people killed, wounded, and hijacked was cumulated mostly in south eastern of Afghanistan.

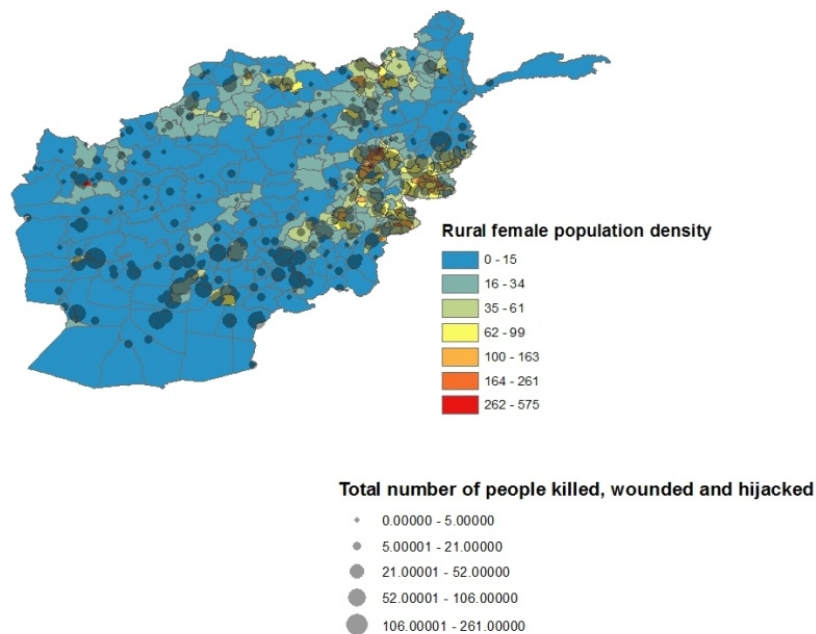


Figure 9: Mapping rural female population density with total number of people killed, wounded, and hijacked

The rural female, rural male, urban female, and urban male population density was mapped with the total number of people killed, wounded, and hijacked and corresponding information is provided in Figures 9 through 12. The eastern, the northeastern, and the central regions that have a large number of people killed, wounded, and hijacked also have high rural female and male population densities. On the other hand, in addition to the northeast region, there are two high urban male and female population density areas (red colors) shown in the central and southwestern regions that have a large number of people killed, wounded, and hijacked. These areas contain major urban population centers located in Kabul and Kandahar province. These centers are two of the most convenient locations in Afghanistan. As expected, the regions had a high number of adverse events where the population density was high.

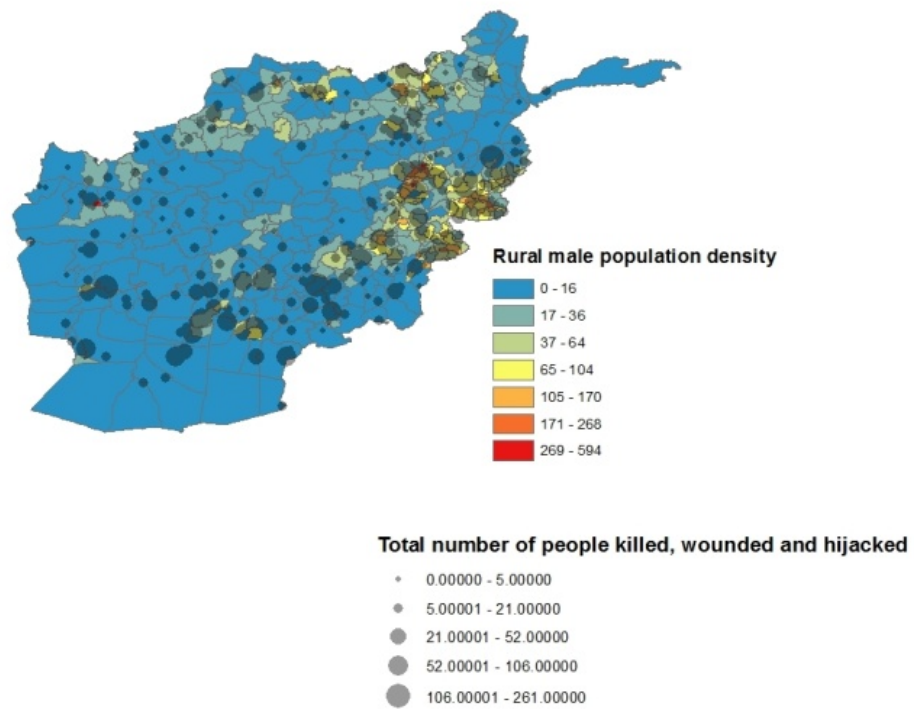


Figure 10: Mapping rural male population density with total number of people killed, wounded, and hijacked

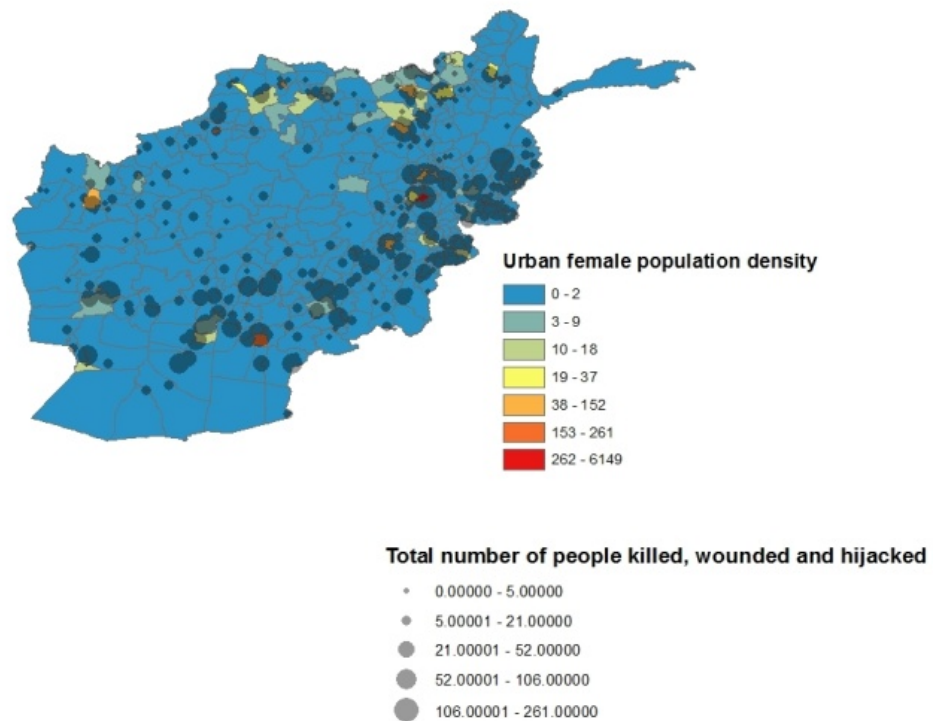


Figure 11: Mapping urban female population density with total number of people killed, wounded, and hijacked

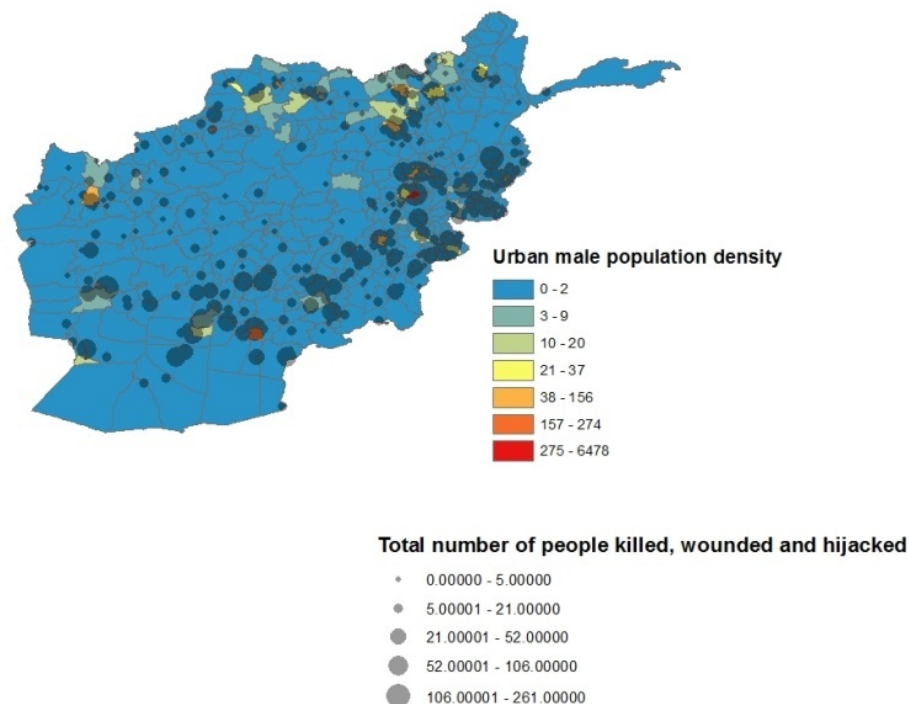


Figure 12: Mapping urban male population density with total number of people killed, wounded, and hijacked

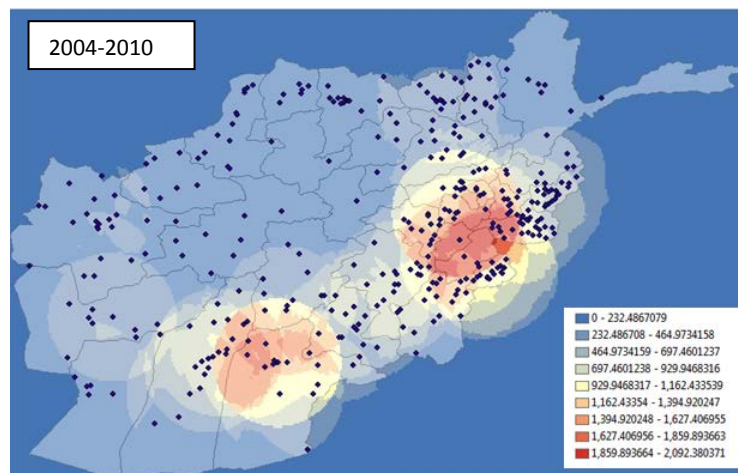


Figure 13: A point density analysis of total number of people killed, wounded, and hijacked from 2004 through 2010

Using ESRI's ArcMap® 10.1 software, the variable called "Total" (the sum of the "number of people hijacked", "number of people wounded", and "number of people killed" variables) in the "WITSGEO" event file (from 2004 through 2010) was added to the data frame. After adverse events were mapped in the ArcGIS environment, a point-density analysis was conducted to detect high adverse event density areas.

The blue color represents low density and the red color shows a high-density area. A legend is provided in the right corner of each map. These numbers show the density value of the total number of people killed, wounded, and hijacked. All data during the period between 2004 and 2010 are represented and individual years that show differences between years can be observed visually (Figure 14).

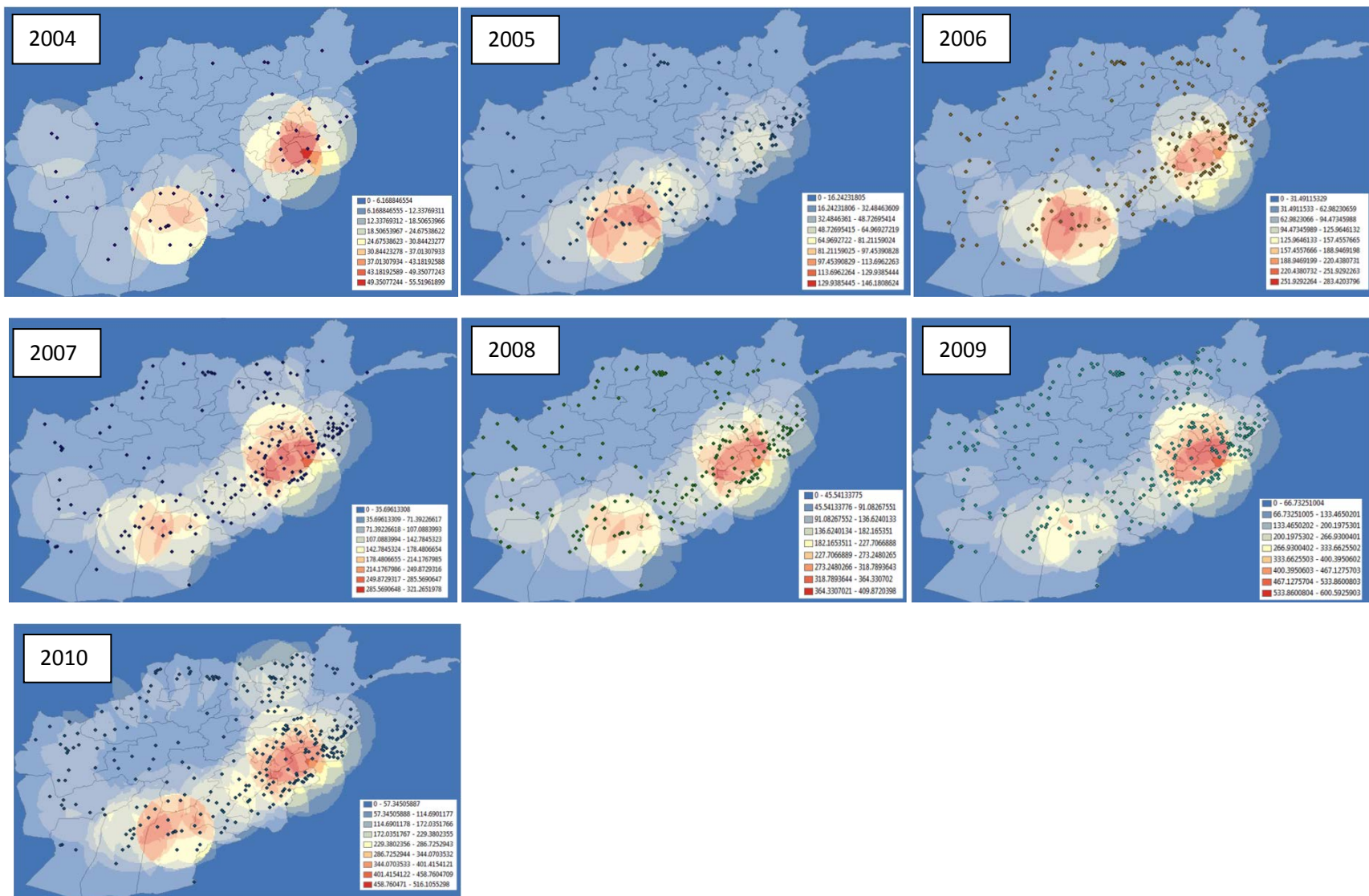


Figure 14: A point density analysis of total number of people killed, wounded, and hijacked for each year (2004-2010)

CONCLUSION

This study represented the adverse event data in terms of number of people killed, wounded, and hijacked in the active war theater of Afghanistan over the period 2004-2010. The visualization of each map helps the reader for understanding implicit trends and provides new information about the adverse events represented by a dataset with over 30,000 records where the information about adverse event patterns is represented in regional level and yearly. Based on the results obtained, it was concluded that the frequency of adverse events has increased from 2004 through 2010. The number of people killed, wounded, hijacked and the total number of adverse events were at their lowest point in 2004 and at their highest point in 2010. When the regions were compared to each other, the south-western region had the highest mean by district values than other regions and the whole of Afghanistan for all variables. On the other side, the north-western region had the lowest mean by district values than other regions and the whole of Afghanistan for all variables. When we compared the variables against each other, the number of people hijacked had the lowest values in total and average by district than the other variables. As expected, the regions had a high number of adverse events where the population density was high. Understanding these patterns may provide useful information about strategies and tactics to be employed for similar scenarios in an active war theater in the future.

ACKNOWLEDGMENTS

This study was supported in part by Grant No. 1052339, Complex Systems Engineering for Rapid Computational Socio-Cultural Network Analysis, from the Office of Naval Research. The authors acknowledge the helpful guidance of ONR program management, and the contributions of the technical team.

REFERENCES

- Berrebi., C. &Lakdawalla., D. (2007). How Does Terrorism Risk Vary Across Space And Time? An Analysis Based On The Israeli Experience. *Defence and Peace Economics*, Vol. 18(2), April, pp. 113–131.
- Brown, D., Dalton, J., & Hoyle, H. (2004). Spatial Forecast Methods for Terrorist Events in Urban Environments. *Intelligence and Security Informatics*, 426-435.
- ESRI, 2013. /<http://www.esri.com> (accessed 08.01.13)
- Federation of American Scientists.(2009). 'A GEOINT analysis of terrorism in Afghanistan.' *Secrecy News*. Retrieved May 3, 2012 from the World Wide Web: <http://www.fas.org/blog/secrecy/2009/05/05>.
- Johnson., S. D. & Braithwaite., E. (2009). 'Spatio-temporal modeling of insurgency in Iraq.' In Freilich, J. D. and Newman, G. R. (eds), *Reducing Terrorism Through Situational Crime Prevention: Crime Prevention Studies*, Vol. 25. Monsey, NY: Criminal Justice Press, pp. 9–32.
- LaFree.,G, Dugan.,L, Xie.,M. and Piyusha Singh, P. (2012) Spatial and Temporal Patterns of Terrorist Attacks by ETA 1970 to 2007. *Journal Of Quantitative Criminology* Volume 28, Number 1, 7-29.
- Siebeneck., L.K, Medina., R.M, Yamada.,I, &Hepner., G.F. (2009). Spatial and Temporal Analyses of Terrorist Incidents in Iraq, 2004-2006. *Studies in Conflict & Terrorism*, 32:591-610.
- The 2013 Human Development Report – "The Rise of the South: Human Progress in a Diverse World"". HDRO (Human Development Report Office) United Nations Development Programme. pp. 144–147. Retrieved May 20 2013.
- Webb, J., &Cutter, S.L. (2009). The geography of U.S. terrorist incidents, 1970-2004. *Terrorism & Political Violence*, 21: 428-449.

Developing An Agent-based Architecture to Model Population Displacement

John A. Sokolowski and Catherine M. Banks

*Virginia Modeling, Analysis and Simulation Center
Old Dominion University
Suffolk, VA 23435, USA*

ABSTRACT

United Nations High Commissioner for Refugees data reports the dislocation of millions from their native or accustomed environments. Statistical models authenticate the numbers exiting and ascertain proximate reasons for decisions to leave. These models employ static data and thus provide static outputs. This is problematic because when a decision to exit is concluded factors driving that decision can change or additional variables can present. A more efficient model is needed to envisage a broader range of outputs proffering why, when, and where migration will occur. Additionally, possessing the means to assess *what-if* scenarios allows for anticipating a range of expected/unexpected outcomes. This paper presents a unique architecture to state (environment) and population (agent) representation for agent-based modeling (ABM) as a means to analyze the decision-making process of individuals threatened with population displacement. This architecture facilitates agent-based modeling that can represent both fluid conditions in the environment and fluctuations in the decision-making process by people under duress. The predominant population displacement modeling application has been statistical, exclusive of dynamic inputs. The conclusions and recommendations within the literature validate implementing a new architecture for this research, which can model root, proximate, and triggering variables associated with this multi-layered, human-factors laden phenomenon.

Keywords: Population Displacement, Early Warning Model, United Nations Human Rights Council (UNHRC), Agent-Based Modeling (ABM), ABM Environment Matrix, ABM Agent Matrix

INTRODUCTION

The global environment is experiencing a grand-scale complex dilemma in the form of the dislocation of millions of individuals from their native or accustomed environments. The data from the United Nations High Commissioner for Refugees speaks to the severity of the problem (Year of Crisis, 2011):

* 12 million	stateless individuals
* 7.1 million	people in a protracted situation
* 4.3 million	displaced individuals due to conflict or persecution
* 3.5 million	internally displaced individuals
* 2.7 million	refugees spread throughout 70 states
* 876,000	individuals applied for asylum or refugee status
* 800,000	individuals deemed refugees

The various above designations are assigned individuals premised on *why one leaves* and *where one goes* (e.g., refugee, or internally displaced person). For purposes of this research we have taken an inclusive approach to these designations by assigning the term **population displacement** to this phenomenon.

The literature readily supports that individuals within a threatened population consider all factors known to them and then choose the best option based on assessment of the particular circumstances, perceived risks, costs, and benefits. Statistical models are developed along these lines to authenticate the numbers of those exiting and to ascertain the proximate reasons for that decision. This type of model holds to static data and thus provides a static output. This is problematic because even when a decision to exit is concluded, factors driving that decision can change during the course of the journey, or additional variables can come into the situation. A more complex modeling capability is required for these scenarios to envisage broader potential outputs that include why, when, and where migration will occur. Additionally, possessing the means to assess *what-if* scenarios allows for anticipating a broader range and/or unexpected outcomes. As such, this paper presents a unique methodology to environment and agent development for engaging agent-based modeling to analyze the decision-making process of individuals in the threatening situation of population displacement.

The paper flows as follows: **Part 2** reviews the current literature on both assessing and modeling population displacement; **Part 3** presents two models developed by prominent social scientists in the field of refugee movement. These models serve as a baseline for crafting a more detailed matrix that can be used for representing “environment” for an agent-based model; **Part 4** previews the two matrices developed by the authors as a methodology for environment representation and agent characterization to be used for an agent-based model; and **Part 5** summarizes the approach, proffers suggested uses for the methodology; and tenders concluding comments.

LITERATURE REVIEW ON MODELING POPULATION DISPLACEMENT

There is an earnest discussion on how to assess population displacement within the sphere of interested scholars from across the social science disciplines (political science, sociology, legal studies, anthropology, economics, *etc.*). To shape the course of our approach to population displacement analysis, we did a cursory review of the theoretical approaches to this research: the criticisms, challenges, and options put before intersecting scholars in this field. We reviewed a number of papers to get better idea of how the topic was being researched. This review was necessary to better understand the crisis-like problem of population displacement, how the phenomena is being researched, how it is being assessed (modeled), and where do we as modelers from the engineering world fit in. In sum, we found that the research approaches are driven by discipline and the intent of the research and that the modeling was overwhelmingly statistically based.

Existent Theoretical Approaches Evaluating Population Displacement

A lesson in the exclusivity of population designations, *i.e.*, refugee or internally displaced, came from *Forced Migration Studies: Can We Just Agree to Date*, in which Hathaway questions the soundness of integrating refugee studies into the widely accepted forced migration studies (Hathaway, 2007). His rationale: this approach to assessing the problem might result in failure to consider the specificity of the refugee’s circumstances as the result of becoming better *focused on the phenomena instead of the personal predicament*. This finding alerted us to ensuring that we represented population, entities, and environment with equal significance. Another deficiency in the empirical research is the notion of *constraining population representation* by disaggregating, standardizing, and categorizing groups (Bakewell 2008).

This finding shaped our research approach as one that should present a comprehensive characterization of displaced populations, causal relationships, representations of the environment, or state of the state as we address it, and the human factors that are indisputably difficult to capture but integral to understanding events and outcomes. In *Refugee Studies and the International Refugee Regime*, Scallettaris (2007) notes that a wider migration approach has often proved more appropriate and profitable in studying refugee related situations.

Our review of this aspect of the empirical data led us to conclude that these differing theoretical approaches adapted either a *people perspective* or *phenomena perspective* and as such they are unable to present a comprehensive depiction that conveys the totality of an imminent or crisis situation. Repeatedly, however, this literature rec-

ommends the need to blur the lines or allow for overlap of mixed populations and mixed environments (situations on the ground) as these entities evolve over time.

Existent Modeling Approaches to Assessing Population Displacement

The modeling approach for assessing population displacement consists primarily of statistical applications. We found a single game theory treatment and a simplified modeling and simulation technique, both of which we viewed as creative. Almost in its entirety the problem of population displacement has been analyzed from a statistical standpoint. Of the 10 relevant papers we investigated, 8 make clear this is the approach many social scientists employ. Overall, these models presented a static look at population displacement primarily the result of model constraints associated with statistical applications. Still, we aimed to evaluate the literature with the simple goal of investigating what was done, what are the strengths and deficiencies of these models, what can we adopt from these efforts.

Prominent among these research efforts is the statistical modeling of Moore and Shellman, whose numerous iterations focused on expanding the both model and the hypothesis. Moore and Shellman (2004) have determined the leading cause of forced migration as violent behavior from the state and as such they focused on *push factor emanating from the state*. They propose that people monitor the violent behavior of the government as well as dissidents and assess the threat it poses on them. Their statistical model goes beyond the modeling the structural characteristics of the state (Schmeidel 1997, 1998, 2000) which, in their estimation fails to connect the human factor and renders a binary assessment in the form of inputs – violence, and outputs – population displacement. The purpose of this model was to examine the exiting decision-making process of a threatened population.

Based on probability developed along the lines of a lottery (each person is eventually going to be a victim), the Moore and Shellman model characterized the state or habitual residence as experiencing four sources of threat: government forces, dissident forces, the interaction of these two forces, and foreign forces. The dependent variable was the decision to leave. The model was limited to forced migrants produced by a state with the flow measure calibrated from 0 to 1 million. There were two government variables for threat: genocide / politicide (using data from the Harff 2003 study) and human rights violations (with data from Gibney and Dalton 1996). Two models were executed for the periods 1952-1995 and 1976-1995. The models included data from 175 states with dependent variable being forced migrant flow specific to people responding to a single-information set such as behavior of state or dissidents. Thus, people's decisions are linked via a common set of information accessible to all; their decisions are not independent. Descriptive statistics included these additional variables: genocide, dissident violence, civil war, international war, government terror, democracy, transition, gross national product, and forced migration. There was a data matrix of 7000 observations, but many of the countries were missing data or variables in the 2 time periods. With the outputs from their model of co-variates of forced migration, Moore and Shellman unabashedly note state that their model results "tell us precious little about the specific impact of co-variates in any given forced migration event."

Another Moore and Shellman model investigated *push and pull factors* of states (Moore and Shellman 2006). This model is designed to represent an individual's decision hinging on two things: victimization (to go or not) and socio-politico-economic opportunities (where to go). Characterization of the state held to the push factors such as sources of threat and institutions that provide freedom and rule of law. The pull factors of neighboring state included its political institutions, wage opportunities, cultural opportunities. The results from the statistical runs indicated as dissident and state violence increases, migrants likely to remain as IDPs than cross borders because dissidents have a history of providing pockets of safety. Thus, it is the state that is likely to target the civilian population. Moore and Shellman findings were quite stark: refugee production vice IDP production is a function of state targeted terror and civil wars produce more IDPs while genocides produce more refugees. This Moore and Shellman model makes clear the need to represent within the model both the state and region at-large. Although on the right track in appreciating the need to represent the decision-making process of the population, the model does not provide satisfactory outputs that capture trigger points forcing a decision and a move. Addressing this deficiency would result in a useful predictive model that could be used to gauge how long a population will remain and where they will go once they have decided to leave.

The research surrounding push/pull factors is significant because it goes straight to the question of why people migrate. A study of 129 states (spanning 1964-1989) engaged a thesis that the state creates a threatening environment, either actively or passively (Davenport 2003). Davenport *et.al* compared other studies that attempt this thesis only to find some suffered from selection bias because their models assessed only states that produced refugees or reviewed the data at a global rather than national level (Hakorirta 1986, Apodaca 1998, Gibney 1996). These studies fail to represent a major sector of the population, internally displaced peoples, and they address only

bi-variate relationships leaving the models wanting in adequate consideration of push / pull factors. The Davenport model was successful in developing a model beyond bi-variate relationships. This enabled an assessment of push/pull elements by using statistical analyses using fixed effects least squares on a pooled cross sectional time series data set as mentioned above. The findings supported a somewhat common-sense conclusion that threats to personal integrity are of primary importance in the decision to abandon a locale. Moreover, measures of state threats to personal integrity, dissident threats to personal integrity, and joint- state-dissident threats have a statistically significant effect on migrant production. Using net migration as the dependent variable the model included a number of independent, widely recognized variables such as nature of threat, structure of polity, interaction of state and dissidents.

Perhaps the most profound finding in the Davenport model, and worthy of representation in model development, is that states moving toward democracy tend to have a greater number of forced migrants. Notably, this counter-intuitive conclusion was drawn when less weighty variables were considered: environmental cues implying threat; economic security; and social networks. This unanticipated, significant result speaks volumes of the need to develop comprehensive models that can find correct values for all factors affecting the decision-making process.

At the time of this writing we found only one game theoretical model (Azam 2002) that analyzes motives of violence against civilians during internal wars. The model addresses two theses: soldiers terrorize because they need to loot as a means to augment resources; and, looting is simply a function of conflict. From the state's perspective, displacement of large factions of the population reduces fighting efficiency of the enemy as they cannot hide or obtain supplies. Both hypotheses are investigated in game theory format with presenting the government and rebels simultaneously deciding on the level of forces engaged in violence against civilians (Stage 1) before these opposing forces decide the level of forces they will engage in fighting in each other (Stage 2).

Azam's analysis and conclusions, drawn from excellent empirical research, bring much to the subject and modeling approach of population displacement. Introducing more counter-intuitive findings, the game theoretical model allows for deriving clearly the different implications to the presence of looting in civil conflicts. As a function of conspicuous atrocity, the Azam model suggests that violence against civilians is a military tactic not just a by-product of war. The outputs from this model suggest numerous factors that merit representation that have not been evaluated in any other models; most significant perhaps is the notion that there will be more refugees if a state government gets more resources and this is the result of the type of conspicuous atrocity – war against civilians.

Lastly, Bruzzone (2011) presents a conceptual model of a use-case to investigate the impact of migrant seasonal workers on communities. He characterizes the two populations, local and migrant, in an effort to measure the effects of conflict and disorder in the community and the response of local law enforcement. The model is conceptual in design and as such is not an executable model with simulation runs and outputs. What the model does provide is a perspective on situational population displacement due to the economic pull of another state. Bruzzone conveys a comprehensive representation of the effects of population movement typically deemed acceptable, as with migratory workers. He suggests that under difficult economic periods more workers come, bringing higher competition for work, overall and escalating social unrest takes place, and crime increases. These factors merit attention within the more traditional view of population displacement as a result of conflict and/or persecution as they highlight what unwelcoming conditions could exist in locales serving as destinations for displaced populations.

Probably most prolific in the field of assessing and measuring factors that affect population movement is the work of Susanne Schmeidel. Her research (1997) assesses the role of generalized structural factors in the formation of forced migration. Her models support the predominating thesis that political violence is the most important cause of migration; more specifically the level and type of violence determine likelihood and size of refugee displacement. A unique addition to the qualitative analysis on the subject of population displacement is Schmeidel's adaptation of the *Early Warning of Refugee Flows* model by Clark (1989) in which Schmeidel emphasizes the significance of *intervening factors* or *intervening conditions*, such as refugees meeting with various unexpected obstructions to their flight or humanitarian facilities being held hostage. Schmeidel expands the Clark model by delineating root causes, proximate causes, and intervening causes all of which are necessary in the development of a theoretical model designed to explain an entire situation or behavior with the idea that the assessment and model would eventually be able to predict behavior. But the mapping of qualitative data to a quantitative analysis poses some deficiencies as these statistical studies cannot conclusively determine the role of significant variables. Nor is statistical modeling able to clarify the degree of impact of certain variables. Simply, statistical modeling cannot represent intervening factors in a meaningful way so as to yield insightful predictive power.

Our literature review indicated that the predominant modeling application to assessing population displacement has been a static approach excluding dynamic inputs. The conclusions and recommendations presented in the literature validate our decision to implement the methodology selected for this research which can model causal,

correlative, and corresponding factors associated with the multi-layered, human-factors laden phenomena of population displacement.

Population displacement is a complex network of relationships that must be modeled as more than just a linear link between agents or entities. As such qualitative analysis, intangibles and subjective data affecting human behavior, must be represented. Not to be excluded is the incorporation of what is most commonly used, quantitative data, to allow for an examination of population displacement through numerical representation. The mixed-methods data mining from the social science contribution to this model will provide a summarization of large bodies of subjective data as well as a generalization based on quantitative projections (objective data) as a means of representing the totality of the issues surrounding population displacement (Sokolowski and Banks, 2009).

TWO SIGNIFICANT SOCIAL SCIENCE MODELING APPROACHES

Probably most prolific in the field of assessing and measuring factors that affect population movement is the work of Susanne Schmeidl. As a social scientist, her research assesses the role of generalized structural factors in the formation of forced migration (Schmeidl, 1997). Her models support the predominating thesis that political violence is the most important cause of migration; more specifically, she posits that the level and type of violence determines likelihood and size of refugee displacement. A unique addition to the qualitative analysis on the subject of population displacement is Schmeidl's adaptation of the *Early Warning of Refugee Flows* model by Lance Clark (see Figure 1) in which Schmeidl emphasizes the significance of intervening factors or intervening conditions, such as refugees meeting with various unexpected obstructions to their flight or humanitarian facilities being held hostage (Clark, 1989).

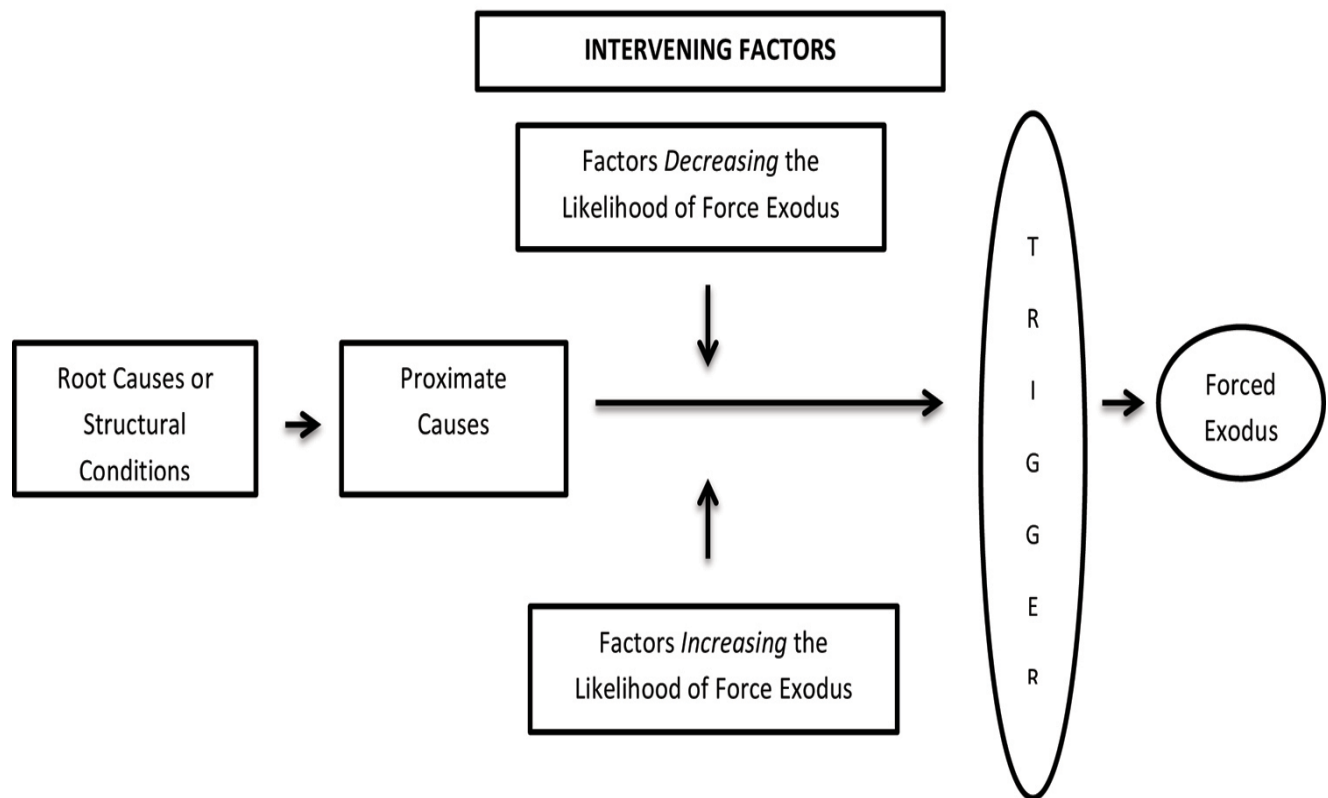


Figure 1. Early Warning Model of Forced Migration - Clark Model

Schmeidl expands the Clark model by delineating root causes, proximate causes, and intervening causes to population displacement all of which are necessary in the development of a theoretical (conceptual) model designed to explain an entire situation or behavior with the idea that the assessment and model would eventually be able to predict behavior (see Figure 2).

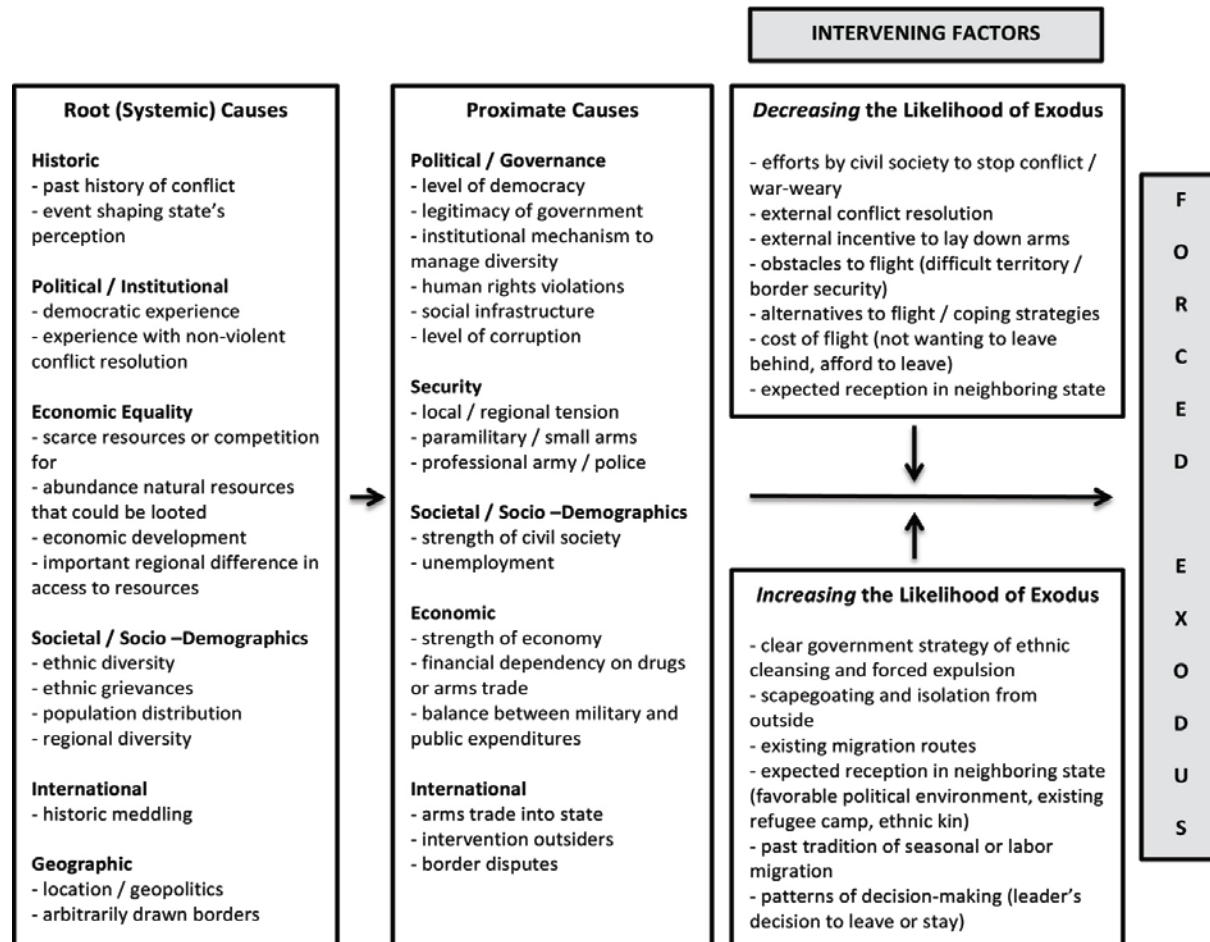


Figure 2. Indicators for the Early Warning Model of Forced Migration - Schmeidl Model

The challenge with the Schmeidl model is the mapping of qualitative data to quantitative values that could be used for simulations. Again the deficiency of this study is that it cannot conclusively determine the role or significance of specific variables. Nor is Schmeidl's statistical modeling able to clarify the degree of impact of certain variables.

We have concluded from the literature review and the assessment of the social science models (statistical and theoretical modeling respectively) cannot represent intervening factors in a meaningful way so as to yield insightful anticipatory or prognostic analysis. The theoretical models do, however, allow for a comprehensive characterization of the data as well as a means to represent triggering events which can serve as a template for assimilating population displacement data.

METHODOLOGY FOR REPRESENTING THE STATE (ENVIRONMENT) AND CHARACTERIZING THE PEOPLE (AGENT)

The authors of this research gleaned much from the literature review, especially the Clark and Schmeidl models, in constructing a methodology to represent the environment and to characterize the agent(s) for an agent-based model. We conclude this is the most effective modeling paradigm for assessing and analyzing the data that all of the modelers discussed in the Literature Review deemed significant to understanding the phenomena of population displacement.

Agent-based Modeling

To effectively understand the influence of one's environment in a distressed population one must begin by studying, then representing the factors that control the environment both from the threatened population's perspective and from the perspective of those causing the threat. Specific to the Syrian case study is the existence of two competing forces, the rebels and the government, that are each vying for control of the state and arguably the hearts and minds of the people. The relative importance of each of these groups and the methods they use directly dictate how the threatened population behaves.

Modeling and simulation, specifically the agent-based paradigm, has become a recognized tool for exploring real world systems, especially those that cannot be readily manipulated for experimentation purposes, such as human and social systems (National Science Foundation, 2013). The challenge is to develop a computational representation of these systems in an efficient and accurate manner.

Some definitions are in order to clarify the above steps: a model is a representation of an event or system at some level of abstraction; a simulation is the operation of the model over a period of time. The concept of abstraction is important because it is unlikely that the model will be an exact replica of the system it is intended to represent. The modeler must decide what level of abstraction is sufficient to meet the purpose of building the model in the first place.

Agent-based modeling, on the other hand is a bottom up approach to representing a system usually at the micro level. They are dynamic models that imitate the actions and interactions among the units of analysis. Agent-based modeling focuses on these units of analysis, or Agents, and the sequence of actions and interactions of the Agents over a period of time. Agents may represent people, organizations, countries—any type of social actor. These actors may act in parallel, may be heterogeneous, and may learn from their actions. Each Agent responds to the prior action of one or more of the other Agents or the environment in the model (or system). This, in turn, produces an extended and often emergent sequence of behaviors which can be analyzed for different things. The action of the Agent can be regarded as a variable; inaction of the Agent is also a variable and it can be considered hostile, neutral, or information seeking. This modeling paradigm was chosen for this case study because we needed to observe the individual changes in the population regarding their treatment by the government and the core group of rebels.

To conduct agent-based modeling one begins by defining the basic behavior of an Agent. This behavior may be captured in many ways. The most often used method is through a series of simple rules that the Agent must follow. These rules help describe the fundamental goals that the Agent is trying to achieve. Probably the best known example of an agent-based model is Craig Reynolds' Boids which simulates the flocking behavior, emergence, of birds (Reynolds, 1987). This idea of emergence is central to agent-based models. Because of the Agent's autonomous, goal-seeking behavior, complex behaviors emerge that are not pre-programmed. This methodology is an effective way to simulate complex social behaviors through the application of relatively simple rules that each Agent follows.

Thus, one can see that agent-based modeling is intrinsically social in that the actions and characteristics of the Agents are influenced by the actions and characteristics of the other Agents in the social system. An agent-based model consists of autonomous software entities that interact with the other entities in the simulation to achieve their own set of goals. Agent-based models are a viable means to explore complex adaptive systems such as the insurgency in the Delta (Axelrod, 1997).

For our ongoing study of the Syrian Civil War, the environment will be developed based upon the Environment Matrix shown in Figure 3. The agents are drawn from groups of people and entities listed in the ABM Agent Matrix in Table 1 below. The simulation will present these agents while moving about their region, encountering other agents or entities, and interacting with the agents and entities in some manner depending on their state of being.

Crafting Matrices for Environment and Agents

We begin by crafting an ABM Environment Matrix that mirror-images the specific factors outlined by the United Nations Human Rights Council (UNHRC, 1996). These variables are divided among Prompting Departure, Intervening Factors, and Triggering Events. Here is a sampling from the UNHRC:

Prompting Departure

- ethnic and racial tensions
- social tensions
- religious tensions
- human rights abuses
- political instability

Intervening Factors

- alternatives to international flight
- international relief in place of origin
- international protection force in place of origin obstacles to flight

Triggering Events

- new types of people affected
- spread of problems in region
- increase intensity of situation
- changes in viability of flight

To be sure, population displacement is a complex network of relationships that must be modeled as more than just a linear link between agents or entities. As such qualitative analysis, intangibles and subjective data affecting human behavior, must be represented. Not to be excluded is the incorporation of what is most commonly used by the vast majority of modelers of population displacement, quantitative data, to allow for an examination of population displacement through numerical representation.

The authors have a history of engaging multi-disciplinary, mixed-methods research such as drawn from the social science contributions to provide a narration and summarization of large bodies of subjective data as well as a generalization based on quantitative projections (objective data) as a means of representing a holistic view of the facts and concerns surrounding population displacement. With this in hand, the authors have developed an ABM Environment Matrix that can be individually mapped per case study to ensure adequate and correct representation of the model environment reflecting the specific factors delineated by the UNHRC. We also included an OUTCOMES section into which the outputs of the simulation will be categorized (see Figure 3).

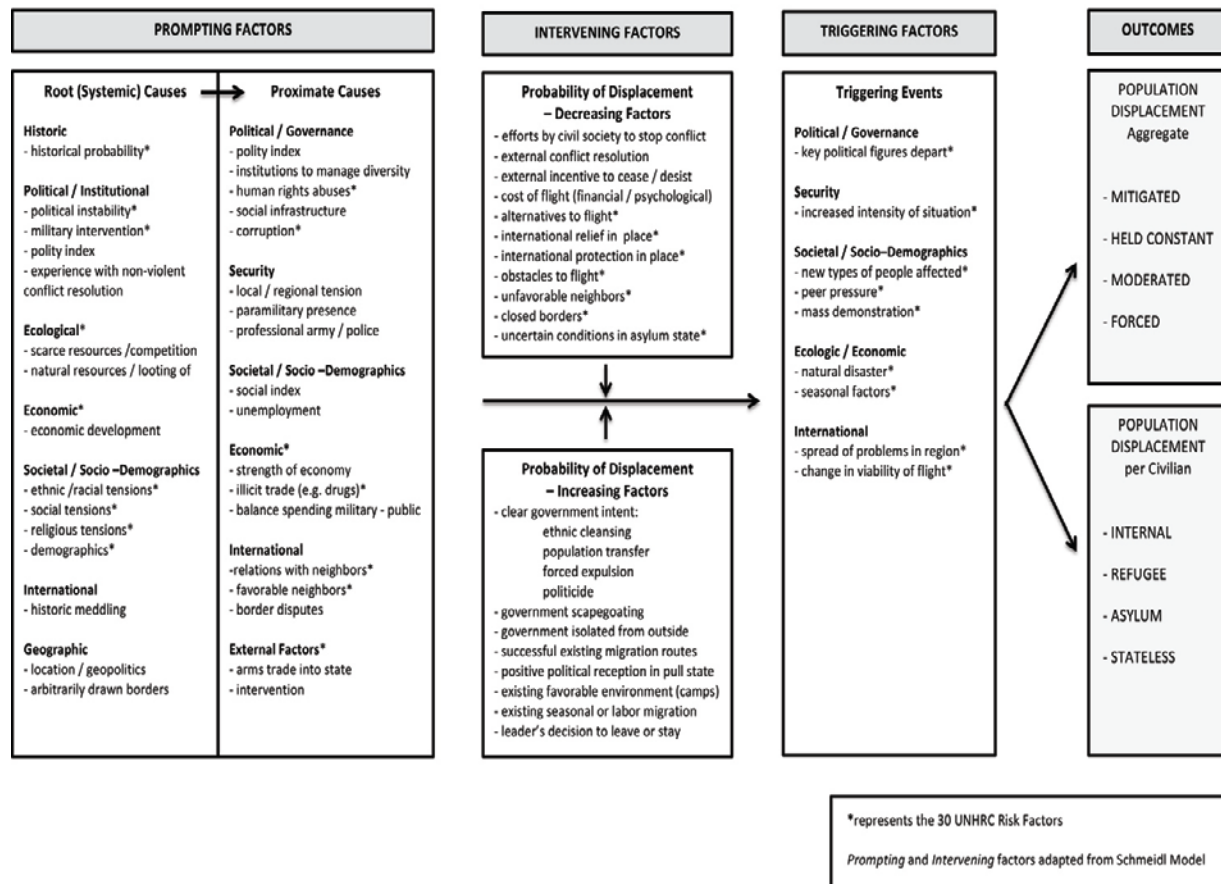


Figure 3. ABM Environment Matrix - Sokolowski-Banks Model

Both the Clark and Schmeidl models provided little in the way of developing the agent portion of our methodology. The inability to characterize the key component to analyzing population displacement – the person aka agent – is a critical shortfall in statistical modeling we reviewed in the literature. Integral to assessing population displacement is the need to monitor motivation for migration as reflected in the agent. Our aim at this juncture is to craft a detailed agent characterization matrix to observe how an agent performs to achieve a goal such as survival (civilian: decision to stay or move) or to accomplish a task (regime: effort to retain political control).

For purposes of this paper (and as an introduction to a future use of this methodology to analyze the Syrian civil war), the agents populating the Agent Matrix will reflect the current actors and segments of Syrian society – participants – involved with or affected by the current conflict (Office UNHCR, 2013; US Department of State, 2013; US Agency on International Development, 2013). The matrix in Table 1 assigns location (state internal or external) and the actors' intent or goal (sustain power, over-throw government, or survive the conflict).

Table 1. ABM Agent Matrix of the Sokolowski-Banks Model

PARTICIPANTS	LOCATION	INTENT / GOAL	FACTORS
	INTERNAL – SYRIA		
National Defense Army (Assad Regime: Shia and Alawite)	Military	Political-Military	(to be drawn from the Environment Matrix)
Shabiha (pro-government gangs)			
Free Syrian Army	Rebels		
EXTERNAL – SYRIA			
Iran	respective states	Ideological-Political	
Russia (antagonists – pro-Assad)			
Nusra militia – Salafists linked to al-Qaida and Taliban	Rebel Supporters	Ideological-Political	
Al-Farouq Brigades linked to Muslim Brotherhood (both protagonists – anti-Assad)			
INTERNAL – SYRIA			
Syrian Civilians			
1) 75% Sunni	1) majority, anti-Assad	1) survival, gain political control	
2) 15% Alawites	2) minority, pro Assad	2) retain political control	
3) 10% Christian	3) minority, pro Assad	3) survival	

This matrix cites participants at-large and allows for additional segments of the population and/or a more granular look at these populations. For example, under the Location variable a population can be assigned both an internal and external representation; or, under the Intent / Goal variable a population might include Religion as a motivator (intent) and goal (*e.g.*, overthrow a secular government).

CONCLUSIONS

Agent-based modeling is an important tool for investigating human and social phenomena. This type of modeling has a number of advantages over traditional statistical modeling used to investigate the phenomenon of population displacement. First, subjective data that can be represented in the model plays a role: these computer agents can mirror-image real people given the correct inputs from the research. Second, these models can closely represent how an agent (human) interacts with its surroundings and the other agents (persons) in it. Third, agent-based models can dynamically coordinate communication and activity (Sokolowski and Banks, 2009). The authors acknowledge the strengths to this type of analysis and as such have developed the first step in agent-based modeling of population displacement to reflect an integration of previous modeling efforts. This methodology deliberately incorporates criteria set out by the UNHRC in representing specific factors to population displacement respecting the UNHRC as the authority on the subject.

The purpose of this effort was to develop a comprehensive representation of population displacement using ABM as the means to characterize accurately individuals, entities, and environment via the integration of qualitative (fuzzy or soft) data as well as quantitative values. Our methodology provides both a different approach and a different perspective to researching and understanding population displacement. As such, when populated and executed using our methodology, an agent-based model can proffer insight on how to prevent, hold constant, or moderate escalating effects of threats to populations whose rights to citizenship, and all that it encompasses, are in jeopardy.

Our next effort will be to develop an agent-based model that fully represents events in Syria 2011-2012. We intend to execute and engage the model as a predictive or anticipatory tool to provide computationally sound conclusions as to where the crisis is heading over the stipulated time-steps, observing the effects of degenerating conditions and trigger points that could lead to external involvement (foreign military intervention) or collapse of the state.

REFERENCES

- A Year of Crisis: Global Trends 2011.* (2012), UNHRC. Geneva: Media Relations UNHRC. UNHRC Website: www.unhcr.org/statistics/populationdatabase
- Apodaca, C. (1998), "*Human Rights Abuses: Precursor to Refugee Flight?*", Journal of Refugee Studies 11(1): 80-93.
- Azam, J.P., Hoefler, A. (2002), "*Violence Against Civilians in Civil Wars: Looting or Terror?*", Journal of Peace Research, 39: 461. doi: 10.1177/0022343302039004006
- Bruzzone, A.G. (2011), "*Modelling and Stabilisation, Normalisation and Reconstruction Activities and their Impact over Internally Displaced Persons, Refugees, and Immigrants*", NATO-CAX presentation. Conference Website: <https://transnet.act.nato.int/WISE/COE/Individual/MS/Events/Archive/6thNATOCAX/Presentation/Container5/ModellingS/file/WFS/Modelling%20Stabilisation%20Normalisation%20%26%20Reconstruction%20Activities%20over%20IDPs.pdf>
- Clark, L. (1989), "*Early Warning of Refugee Flows*", Washington, DC: Refugee Policy Group.
- Davenport, C., Moore, W.H., Poe, S.C. (2003), "Sometimes You Just Have to Leave: Domestic Threats and Forced Migration, 1964-1989", International Interactions, 29:27-55. doi: 10.1080/03050620390182966
- Gibney, M., Dalton, M. (1996), "*The Political Scale*", Policy Studies and Developing Nations 4:73-84.
- Hakorirta, H. (1986), "*Third World Conflicts and Refugeeism: Dimensions, Dynamics, and Trends of the World Refugee Problem*", Helsinki: Finnish Society of Sciences and Letters.
- Harff, B. (2003), "*No Lessons Learned from the Holocaust? Assessing Risks of Genocides and Politicides Mass Murder since 1955*", American Political Science Review 41:1079-103.
- Moore, W.H., Shellman, S.M. (2004), "*Fear of Persecution: Forced Migration 1952-1995*", Journal of Conflict Resolution Vol. 40, No. 5. doi: 10.1177/0022002704267767
- Moore, W.H., Shellman, S.M. (2006), "*Refugee or Internally Displaced Person: To Where Should We Flee?*", Comparative Political Studies Vol. 39, No. 59. doi: 10.1177/0010414005276457
- National Science Foundation. *Simulation-Based Engineering Science: Revolutionizing Engineering Science Through Simulation*, Report of the National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering Science. May 2006. Website: http://www.nsf.gov/pubs/reports/sbes_final_report.pdf
- Office of the United Nations High Commissioner for Refugees Website: <http://www.unhcr.org>
- Reynolds, C.W. (1987), "*Flocks, Herds, and Schools: A Distributed Behavior Model*", Computer Graphics; 21(4):25 – 34.
- Schmeidel, S. (1997), "*Exploring the Causes of Forced Migration: A pooled time-series analysis, 1971-1990*", Social Science Quarterly 78 (2): 284-308.
- Schmeidel, S., Jenkins, J.C. (1998), "*Early Warning of Humanitarian Disasters: Problems in Building an Early Warning System*", International Migration Review 32(2): 471-486.
- Schmeidel, S. (2000). "*The Quest for Accuracy in the Estimation of Forced Migration*", in Stephen C. Lubkemann, Larry Minear, and Thomas G. Weiss (Eds.), Humanitarian Action: Social Science Connections. (Occasional Paper No.3, pp. 127-157). Providence, RI: Thomas Watson Jr. Institute for International Studies.
- Sokolowski, J.A., Banks, C.M. (2009), *Modeling and Simulation for Analyzing Global Events.* Hoboken: John Wiley & Sons, Inc.
- UNHCR (1996), "*Contingency planning guidelines: A practical guide for field staff. Division of Operation Support, Geneva.*" Website: <http://www.the-ecentre.net/resources/e-library/doc/2-CNTPLN>
- US Department of State Website: <http://www.state.gov/>
- US Agency for International Development Website: <http://www.usaid.gov/>

Validating a Hybrid Cognitive-System Dynamics Model of Team Interaction

Asmeret B. Bier and Michael L. Bernard

*Cyber Engineering Research Institute
Sandia National Laboratories
Albuquerque, NM 87185-1327, USA*

ABSTRACT

Computational models of human behavior can lead to important insights regarding how people interact with each other and with their environments; but validation of these models is difficult and data is generally hard to collect. Better validation strategies could help to make these models more strongly justified and thus more useful. This paper describes an effort to study cooperation between teams in cyber security training exercises by building a model that captures the interactions between them. Real-world exercises provide a useful source of validation data and can serve to help calibrate the model. In this study we simulated two cyber scenarios where the primary difference was the intensity of cyber attacks experienced by two organizations. The model simulated the potential outcomes and decision-making processes involved in cooperative cyber security agreements designed to reduce redundant work. Insights learned from the model are intended to improve future versions of the cyber exercise. We also used validation data and insights from the exercise model to create and justify decision-making strategies in a model of a real-world counterpart to the situation exercises: information sharing programs for cyber defense.

Keywords: Cyber Security Training Exercise, Cooperation, Behavioral Influence Assessment, System Dynamics, Cognitive Model

INTRODUCTION

Computational models of human behavior can lead to important insights regarding how people interact with each other and with their environments, but validation of these models is often limited. By finding ways to more effectively validate cognitive and behavioral models, the results could become more strongly justified and thus more useful. We studied cooperation between teams in cyber security training exercises by building a model that captures the interactions between them. Real-world exercises provided a useful source of validation data and served to help calibrate the model. Insights learned from the model are intended to help to improve the future versions of the exercise.

Tracer FIRE (Forensic Incident Response Exercise) is a cyber security training program developed by Sandia National Laboratories and Los Alamos National Laboratory. It combines traditional classroom and hands-on training with a competitive game forum. Participants work in teams to solve a series of challenges based on real-world incidents. Teams are sometimes chosen by the participants themselves (which often results in teams segregated by organizational affiliation) and sometimes chosen by organizers based on the expertise of the participants. One planned goal of Tracer FIRE is to improve learning by promoting cooperation during the exercises. The organizers hope this might improve participants' desire to cooperate in real security situations (which can increase effectiveness and reduce effort required to combat threats) and help to build relationships between participants during the Tracer

FIRE exercises. Sandia National Laboratories built a model of team interaction during the Tracer FIRE exercises to provide insight into the drivers and effects of cooperation in this learning environment. The model utilizes the Behavioral Influence Assessment (BIA) framework, a hybrid cognitive-system dynamics structure for simulating systems that involve human behavior and decision making. The theoretical framework of the BIA is based on psychological, social, and economic theories that have been incorporated into a single structure that is both self-consistent and dynamic. Cognitive models are implemented using system dynamics and embedded into an encompassing system dynamics model, which simulates interactions between people, groups, and physical, economic, or other system components.

We collected data from Tracer FIRE exercises to motivate, calibrate, and validate a model of team interaction dynamics. Data included specific decision-making strategies of subject matter experts based on the BIA structure (including cues, perceptions, motivations, intentions, and potential behaviors), environmental data (such as distance between teams and noise level), personality data, and game data (scores, etc.). This information was incorporated into the computational model and used to conduct sensitivity analysis of the information/data and uncertainty quantification of the model output.

This project and the Tracer FIRE BIA model provided insight into how teams made decisions about cooperation and how cooperation might affect performance and learning during the exercise. The model and assessment were used to assess Tracer FIRE challenge designs that would promote cooperation and learning. Simulation exercises proved to be a highly useful source of validation data for this project, and future work is planned to assess how validation data collected from simulations might be used to inform models of real-world cyber security work. Data collected for this model, as well as insights gained from the model, were used to inform a subsequent model of a real-world counterpart of the Tracer FIRE exercises: an information sharing program for cyber defense.

COOPERATION IN CYBER DEFENSE AND THE DATA PROBLEM

Cyber attacks pose a major threat to modern organizations. These attacks can have nefarious aims and serious consequences, including disruption of operations, espionage, identity theft, and attacks on critical infrastructure. Organizations must put substantial resources into protecting themselves and their customers, clients, and others against cyber attacks. However, even with a substantial investment in cyber defense resources, the risk of harm from a cyber attack is significant for many organizations. The effectiveness of cyber defense can likely be enhanced if programs are implemented that allow organizations that face similar cyber threats to share information and resources. The threats faced by different organizations may be similar or even identical (figure 1). Thus, much of the work done by cyber defenders at these organizations may actually be redundant (Hui et al., 2010). By sharing personnel and information regarding effective defense strategies pertaining to cyber attacks, organizations may better protect themselves against cyber threats while maintaining or even reducing the resources dedicated to cyber security.

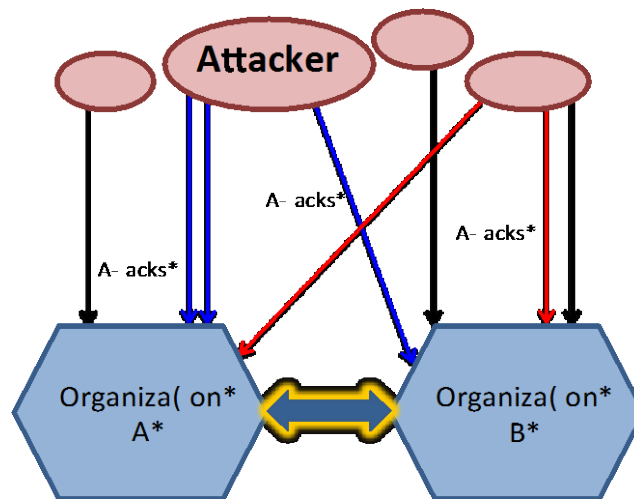


Figure 1: Cooperation can guard against attacks with similar traits or sources

Despite these potential benefits, cooperative cyber defense strategies are not common. Cyber defense teams must balance the potential benefits of cooperation against motivations not to cooperate. For example, if its vulnerabilities are made publicly known, an organization might become more susceptible to cyber attacks and might face damage to its reputation. Trust in cooperating organizations is therefore necessary for successful cooperative cyber security programs. Since organizations that are likely to cooperate with each other are those that face similar threats, they might also be in similar industries and have competitive relationships. Competition for customers, clients, or funding may raise concerns about motive and competitive advantage, making organizations less likely to trust each other. Finally, group inertia is a significant factor to overcome and individual habits may be even more difficult to change than organizational strategy.

This project was designed to begin elucidating how human decision making regarding participation affects an information-sharing program for cyber defense. We created a model that simulates information sharing between multiple organizations. The model focuses on how staff and management in these organizations make decisions about whether and how much to participate in the program. In each organization, decision makers weight the risks and benefits of participation, and their desires to contribute determine the organizations' participation rates. As we built this model, we realized there was a severe lack of information concerning how people might make decisions in these programs. Only a few functional information-sharing programs exist, and the people involved are often wary to share insights on how they do their jobs. Furthermore, it is difficult to collect information on decision making in general (since the mechanisms behind the decisions are often hidden and often not fully understood, even by the decision maker).

To study ways to improve cooperation we used of a training program for cyber defenders called Tracer FIRE. We worked with Tracer FIRE organizers to collect data that could be used in a model of the exercises as well as to inform the information-sharing model. We believe that using controlled exercises as data collection test beds for real-world counterpart systems has immense potential. The exercises were controllable, so that data on behavior determination in different situations could be collected. They also provided a relatively large number of participants who were willing to serve as participants, and who were also willing to discuss their decision making strategies with us at length.

THE TRACER FIRE EXERCISES AND COOPERATION

Sandia National Laboratories and Los Alamos National Laboratory, realizing the increasing threat from cyber attacks, created a training program called Tracer FIRE (Forensic and Incident Response Exercise) to increase the effectiveness of cyber security incident response teams (CSIRTs). Tracer FIRE combines traditional classroom and hands-on training with a competitive game forum. In the classroom portion, students cover incident response topics and are given hands-on training with tools commonly used by CSIRT personnel. In the game portion of the exercise, students form teams and use these tools to solve a series of challenges based on real-world incidents. The challenges cover a variety of cyber defense topics, and the number of points awarded is based on the difficulty of the challenge. The size of the teams varies from 4-10 players. An effort is made to ensure that each team has a balanced skill set and that all teams have roughly the same skill level. Tracer FIRE has been used to train almost 1,000 incident responders from the Department of Energy and other U.S. government agencies, critical infrastructure teams, and academia. In fact, the most recent Tracer FIRE event was held online and had hundreds of participants from over 10 countries around the world. Tracer FIRE also presents an opportunity for human-focused research on cyber security and training. The exercise offers a controlled environment with a variety of challenges and an opportunity for data collection that does not often exist in traditional security environments. A variety of research projects have used Tracer FIRE to study individual and group characteristics in relation to the effectiveness of cyber defense and training.

Tracer FIRE has begun to explore incorporating challenges that encourage cooperation between players. By cooperating with other organizations (sharing information about cyber attacks, effective defense strategies, and personnel with specific expertise), cyber defenders might increase the resources and information available for solving a particular cyber problem and thus better protect their organizations. Researchers have begun to explore the possibility of organizational cooperation in cyber defense (Hui et al., 2010; Sandhu et al., 2010; Luna-Reyes, 2006; Ring & Van de Ven 1994; Oliver, 1990; Luna-Reyes et al., 2008), and the Tracer FIRE team is exploring methods for enhancing cooperation both during and after the exercise. The current design of Tracer FIRE encourages cooperation within teams (points are rewarded by team) and does not prohibit cooperation between teams. Some teams do cooperate with each other to solve challenges, but the point structure, combined with a tendency toward a culture of individualistic work in cyber security (Gates & Whalen 2004), does not always encourage high levels of cooperation.

We collected a large amount of data from the Tracer FIRE exercises to inform, populate, and validate the Tracer FIRE model described below. Our ultimate goal was to inform the information sharing model. We conducted extensive interviews with three Tracer FIRE participants who also work as cyber security professionals. We also conducted shorter, more informal interviews with many other Tracer FIRE participants. We observed multiple rounds of Tracer FIRE and collected data on levels of interaction within and between groups, including personality survey data, and data on environmental condition—such as how close groups were located to each other, the degree of ambient noise, lighting in the room, and the location of shared information. This enabled us to examine two cyber attack (consistent versus uneven attack) scenarios.

THE TRACER FIRE BEHAVIORAL INFLUENCE ASSESSMENT (TF-BIA) MODEL

In order to study the dynamics of cooperation in Tracer FIRE, the Tracer FIRE Behavioral Influence Assessment (TF-BIA) model was created. The model was populated based on interviews with subject matter experts, who were past participants in the Tracer FIRE program and cyber security professionals. It was calibrated using data collected during Tracer FIRE exercises. The model is based on the BIA framework, which was designed to model decision making using well-established psychological, social, and economic theories, all within a system dynamics structure.

Behavioral Influence Assessment (BIA)

Behavioral Influence Assessment (BIA) is a system dynamics-based modeling framework for simulating systems that involve human behavior and decision making. The theoretical framework of the BIA is based on well-established psychological, social, and economic theories that have been incorporated into a single structure (figure 2) that is both self-consistent and dynamic. BIA uses a hybrid, cognitive-system dynamics architecture. Cognitive models are implemented using system dynamics and embedded into an encompassing system dynamics model, which simulates interactions between people, groups, and physical, economic, or other system components.

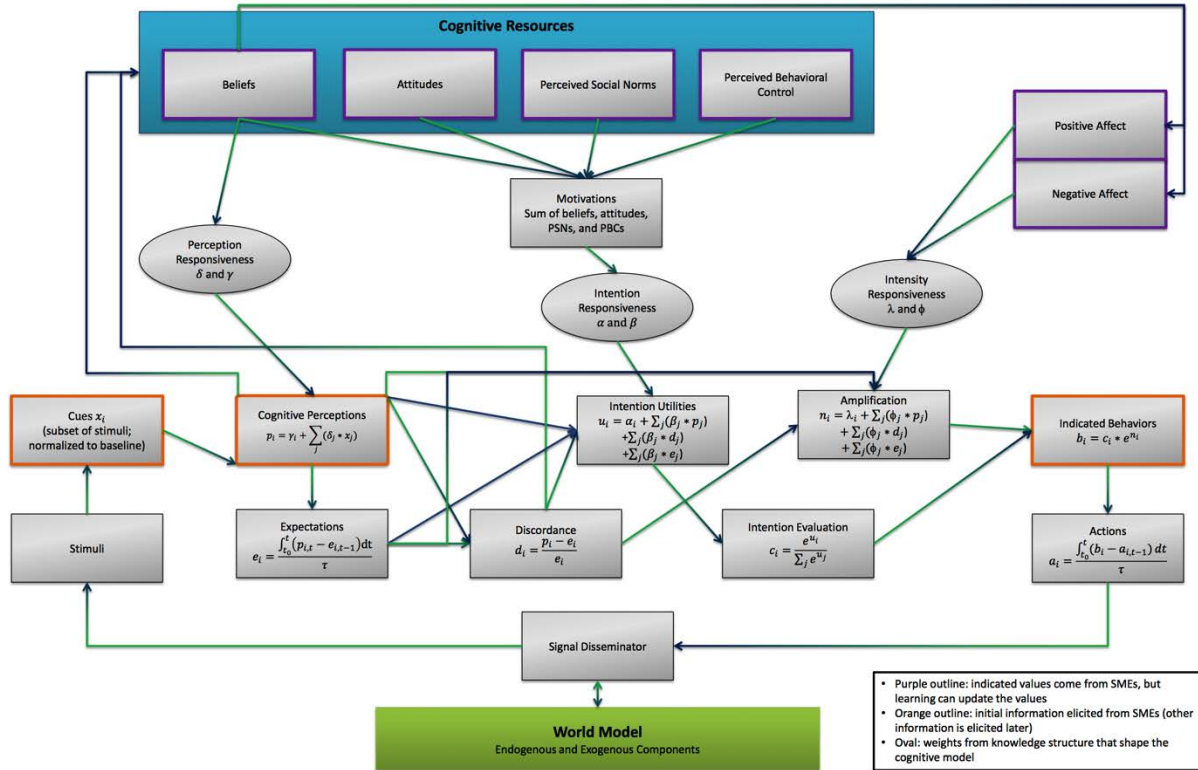


Figure 2: Computational structure of the BIA framework

The cognitive portion of the BIA begins with individuals or groups being exposed to cues (stimuli relevant to the decision-maker). These cues are processed to create cognitive perceptions, the decision-maker's assessment of the world or situation. Over time, cognitive perceptions become expectations, which are compared to cognitive perceptions to determine discordance with the current situation. Discordance and cognitive perception affect beliefs, a category of cognitive processes that includes the components of the theory of planned behavior (attitudes, social norms, perceived behavioral control) (Ajzen, 1991) and affect. Intentions are calculated using utility functions. A multinomial logit function (McFadden, 1982) compares intentions to determine realized behaviors, and over time those behaviors become physical realized actions. One of these cognitive models is populated for each individual or group being included in the system. These cognitive models are connected to each other and to a world model sector using system dynamics. The world model sector includes all of the non-cognitive components of the system of interest, including physical systems, economics, etc. Outputs from the world model and the cognitive models act as inputs, or stimuli, for the cognitive model in subsequent time steps. For a broader discussion of BIA see Bernard, Backus, & Bier and Bernard & Bier in this issue of the proceedings).

Tracer FIRE BIA (TF-BIA)

The Tracer FIRE BIA (TF-BIA) model uses the BIA framework to simulate behaviors of participants in Tracer FIRE. The model simulates six teams, each with the same basic cognitive structure (cognitive parameters can vary between teams). Each team determines the amount of effort it spends working individually versus working cooperatively with other teams. Considering the difficulty of the remaining challenges, individual and cooperative progress is calculated. Cooperative progress also takes into account the amount of work required to cooperate with other teams and shared knowledge available through cooperation. Shared knowledge available depends on the amount of knowledge that each team has and the effort that each team puts toward cooperation.

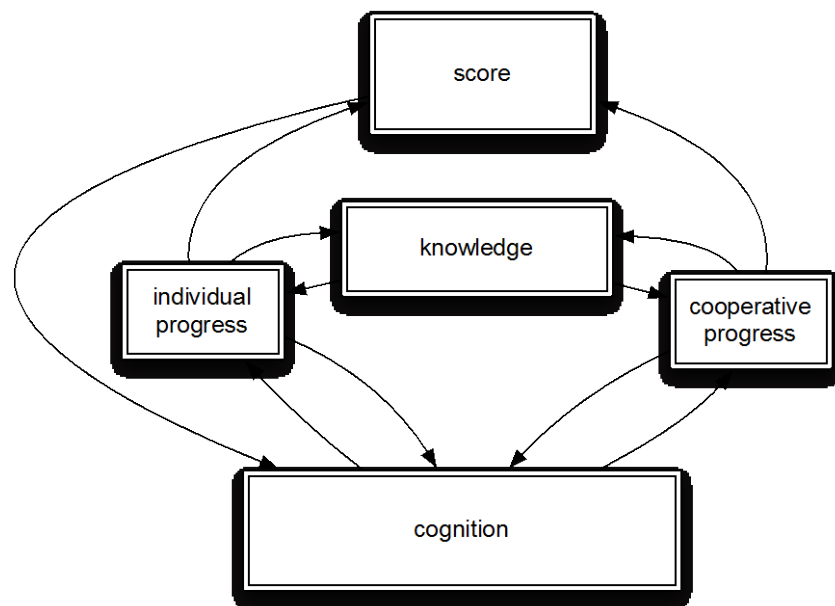


Figure 3: Model structure overview

Individual and cooperative progress for each team is combined to determine the increase in the overall score. As teams solve more challenges, remaining challenges become more difficult. Increase in score and challenge difficulty are used as indicators to determine learning for each team. As knowledge increases, teams become more efficient at solving problems and have more to contribute to cooperative efforts if they choose to do so.

Both behavioral and non-behavioral portions of the model feed into the cognitive models as cues. Interviews with subject matter experts (SMEs) were held to determine how decisions are made during Tracer FIRE. The SMEs were previous participants in the exercise and also work as cyber security professionals. These interviews were used to determine the structure of the decision process (which cues and perceptions are considered, how cues determine perceptions, etc.) and to understand the relative importance of each input for model parameterization. The cues and cognitive perceptions that feed into each potential behavior are shown in table 1.

Table 1: Cues, cognitive perceptions, and potential behaviors

Cues	Potential behaviors →	Work Individually			Work Cooperatively	
	Cognitive Perceptions →	Competition	Benefit of indiv. work	Time pressure	Benefit of cooperation	Frustration
	Effect on behavior →	+	+	+	+	+
	Score difference from nearest competitor	-				
	Team rank	+				
	Recent individual progress		+			
	Recent cooperative progress				+	
	Recent total progress					-
	Difficulty of remaining tasks					+
	Time remaining in game			-		

Each team determines how much effort it puts into individual versus cooperative work. Teams tend to increase individual work when they feel time pressure, competition (based on team rank and having competitors close in score), or when individual work has increased the team's score in the recent past. They tend to work cooperatively when they are frustrated (due to lack of progress or high task difficulty), or when cooperation has recently produced benefits. These factors are compared to determine the effort that goes toward each type of work (individual and cooperative), which then affects score and knowledge, as described above.

Select model results

A key goal of Tracer FIRE participants is to win the game (by generating a higher score than any other team), but the primary goal of Tracer FIRE is to increase participants' knowledge about cyber security incident handling. Cooperation allows teams to learn from others, but requires effort and may give competitors an advantage. Teams must decide how much effort to put into cooperation versus individual work, and this decision affects both learning and scores.

There are four adjustable inputs in the TF-BIA model. The first two, initial knowledge (for each team) and baseline cooperation (for each team) are characteristics of the teams but can be altered by the Tracer FIRE designers. In the simulations discussed here, we assume that all teams have the same initial knowledge and baseline cooperation unless otherwise indicated. The other two variables of interest can be directly manipulated by the white cell (the people running Tracer FIRE). The white cell can modify the difficulty of the challenges, which is represented in the model by a maximum task difficulty variable. It can also make it easier or more difficult for teams to cooperate with each other. This might involve changes to communication infrastructure (instant messaging, shared message boards, etc.), locating players in the same room, challenges that encourage cooperation between teams, verbal encouragement to cooperate from the white cell, or other strategies.

The base case simulation is shown in figure 4. In the base case, each team begins with 25% of the knowledge necessary to complete all of the Tracer FIRE challenges. Work required to cooperate is 25% (in other words, only 75% of the effort put into cooperation actually goes toward progress in the exercises). Challenge difficulty is .75 (of a maximum of 1), and each team begins the exercises with a baseline 25% of effort going toward cooperation. The teams end up with about 78% of the maximum score and about 52% of the total knowledge that can be gained from the exercises, doubling their knowledge over the course of the exercise. Cooperative effort starts out at 25% (the baseline), but declines after the beginning of the exercise. Since all the teams have similar, relatively low levels of initial knowledge, not much can be gained from cooperation and teams put more focus into individual work. Competition remains stable in this scenario because the teams' scores are equal. Near the middle of the time horizon, learning and frustration encourage more cooperation. All teams are gaining knowledge, so the potential benefit of cooperation is increasing. The remaining challenges are getting harder (teams tend to solve the easiest challenges first), so frustration is also increasing. At the end of the exercises, time pressure causes teams to focus more on individual work.

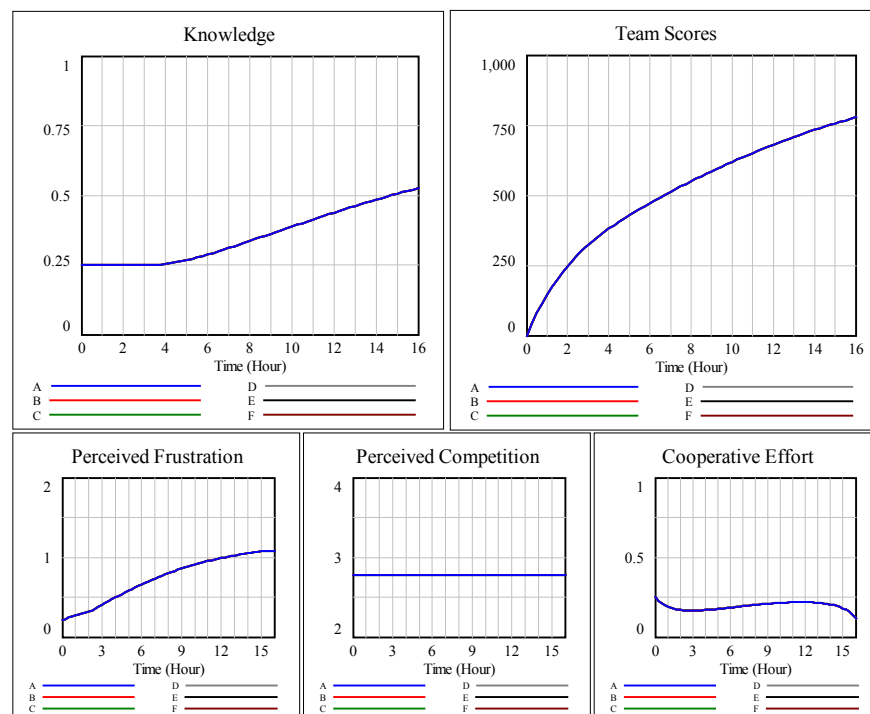


Figure 4: Base case simulation (initial knowledge = 0.25, baseline cooperation = 0.25)

Figures 5 and 6 show scenarios where teams have a higher baseline rate of cooperation (50%) than in the base case (25%). This could represent a situation where teams or participants were chosen specifically for characteristics (personality traits, familiarity with other players, etc.) that encourage cooperation. It could also represent an exercise where teams are encouraged to cooperate before the game starts, or where challenges are designed to encourage cooperation between teams. Both scenarios show that learning increases from the base case. The final knowledge variable for each team nears 66% when baseline cooperation increases to 50% (figure 5), and if barriers to cooperation are removed to make work required to cooperate 5% (rather than 25%), knowledge reaches 70% (figure 6).

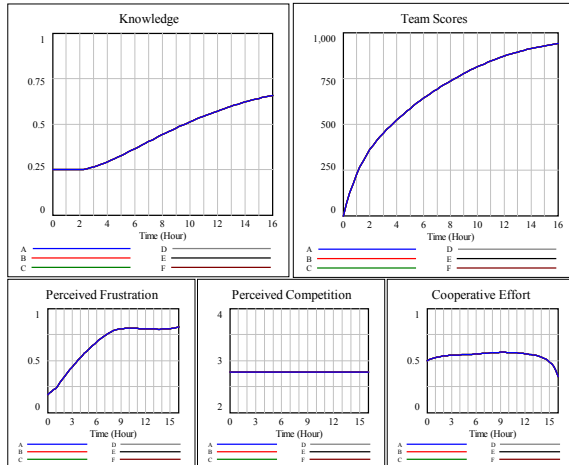


Figure 5: Baseline cooperation = 50%; work required to cooperate = 25%

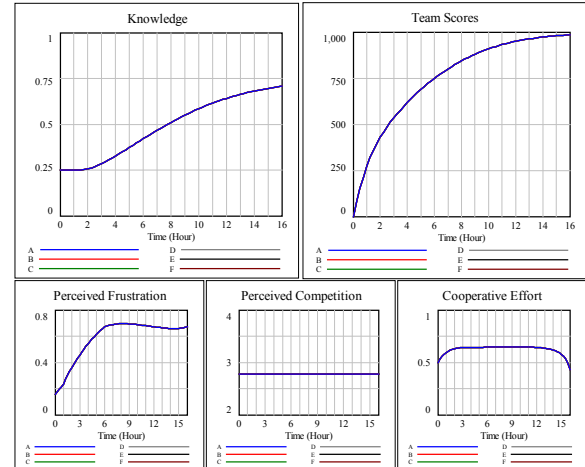


Figure 6: Baseline cooperation = 50%; work required to cooperate = 5%

Learning can be further improved by increasing the difficulty of tasks, as in the scenario shown in figure 7. This scenario is the same as the one shown in figure 5, except that the task difficulty is at its maximum. Participants learn more with higher task difficulty in this scenario, but frustration is also higher. This could cause participants to reduce overall effort levels or to dislike the Tracer FIRE program, discouraging their colleagues from participating in the future. While this model does not consider distraction or future participation in the program, it is a consideration for exercise design and implementation.

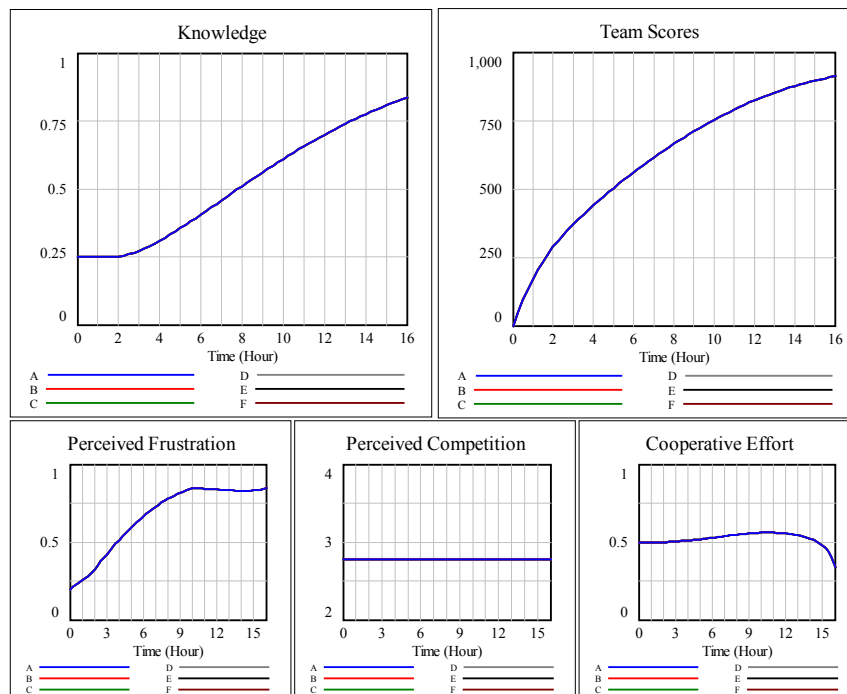


Figure 7: Baseline cooperation = 50%, task difficulty = 1

It is also likely that different teams will have different baseline cooperation levels. Figures 8 shows a scenario in which five teams have baseline cooperation of 25% and one team has a higher level of baseline cooperation (50%). Learning and score both increase a small amount for the team that cooperates more than the others. Figure 9 shows a scenario in which three of the six teams have the higher (50%) baseline level of cooperation. Because more teams are more willing to cooperate, the pool of shared knowledge increases and these teams see an even higher increase in score and knowledge than the others. These scenarios assume that work required to cooperate is the same as in the base case. As barriers to cooperation increase, benefits of cooperation will decrease, at some point (around 50% work required for cooperation in this scenario) creating a negative incentive to cooperate.

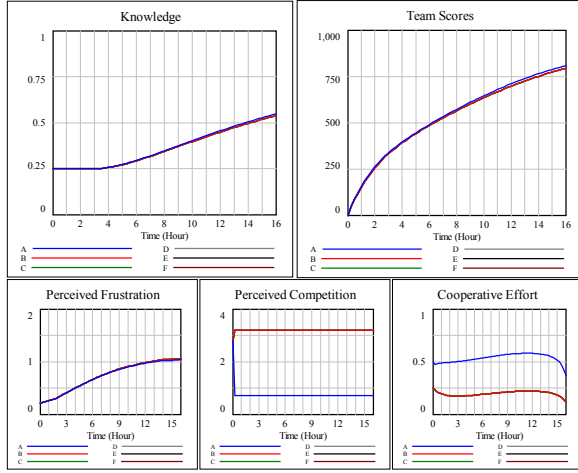


Figure 8: One team with baseline cooperation = 50%

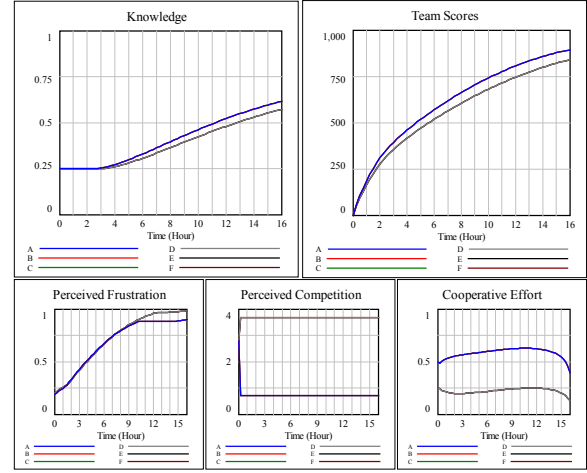


Figure 9: Half of teams with baseline cooperation = 50%

The goal of Tracer FIRE is to increase the participants' knowledge about cyber security incident response. Sensitivity analysis was conducted to indicate which of the four adjustable inputs to this model were most important in determining the teams' average knowledge at the end of the simulation. Partial correlation coefficients are shown in table 2. All of the inputs have high correlation with the knowledge output with high confidence. The maximum task difficulty has the highest (negative) correlation, but the others are also important.

Table 2: Partial correlation coefficients for average knowledge at end of simulation

Variable	Partial correlation coefficient	p-value
Maximum task difficulty	-0.93516	7.8392e-90
Work required to cooperate	-0.92539	4.2709e-84
Average initial knowledge	0.81709	1.5894e-48
Average baseline cooperation	0.75821	4.5148e-38

CONCLUSION: DATA COLLECTION AND VALIDATION

Tracer FIRE presented a great opportunity for human-focused research on cyber security and training. The exercise offers a controlled environment with a variety of challenges and an opportunity for data collection that does not often exist in traditional security environments. A variety of research projects have used Tracer FIRE to study individual and group characteristics in relation to effectiveness of cyber defense and training. These results suggest various strategies that the white cell might try to improve learning during Tracer FIRE. They might make challenges more difficult, remove barriers to cooperation, increase the initial knowledge of participants, or increase participants' baseline levels of cooperation by altering teams based on personality types of participants, composition of teams, familiarity of players with each other, structure of the game, or other strategies. The BIA framework proved useful for modeling cooperative behavior in the Tracer FIRE exercises. It is even more useful in providing a cognitive structure that can be applied both to the exercises and to the real-world information sharing program model. Because the framework includes an explicit cognitive model, we can use the model to understand intermediate phases in participants' decision-making process, such as cognitive perceptions, affect, and motivations. This might be more useful for understanding problems like learning than the decision rule method most common in system dynamics models. The BIA framework shows promise for modeling human behavior, especially in situations where details of cognition may be important.

Using a simulation exercise to study decision making strategies and collect data for models of real-world counterparts proved highly useful for this project. We believe that this strategy can be applied in many other situations. The exercises were alterable, so that we could put decision makers into scenarios that would provide us with the most useful data possible. The large number of participants and relatively casual nature of the situation made it easy for the researchers to discuss decision making strategies with participants. Finally, the large potential for data collection during the exercises (including information on actions taken by participants, scores, and environmental data) provided much more information than would have been readily available directly relating to the information sharing programs we were ultimately interested in.

REFERENCES

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211.
- Bernard, M. L., & Bier, A. B. (2014). Analytical capability to better understand and anticipate extremist shifts within populations in failing states. In *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*.
- Bernard, M. L., Backus, G. A., & Bier, A. B. (2014). Behavioral Influence Assessment (BIA): A multi-scale system to assess dynamic behaviors within groups and societies across time. In *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*.
- Gates, C., Whalen, T. (2004). Profiling the defenders. *Proceedings of the 2004 workshop on new security paradigms (NSPW '04)*. ACM, New York, NY, USA, 107-114.
- Hui, P., Bruce, J., Fink, G., Gregory, M., Best, D., McGrath, L., Endert, A. (2010). Towards efficient collaboration in cyber security. *International Symposium on Collaborative Technologies and Systems (CTS)*, 489-498.
- Luna-Reyes, L. F. (2006). Trust and collaboration in interagency information technology projects. *Proceedings of 2006 International Conference of the System Dynamics Society*, Nijmegen, The Netherlands.
- Luna-Reyes, L. F., Black, L. J., Cresswell, A. M., Pardo, T. A. (2008). Knowledge sharing and trust in collaborative requirements analysis. *System Dynamics Review*, 24(3), 265-297. doi:10.1002/sdr.404
- McFadden, D. (1982). *Qualitative response models*. In *Advances in Econometrics*, Ed. Werner Hildenbrand, Cambridge University Press, New York.
- Oliver, C. (1990). Determinants of interorganizational relationships: Integration and future directions. *Academy of Management Review*, 241-265.
- Ring, P.S., Van de Ven, A.H. (1994). Developmental processes of cooperative interorganizational relationships. *Academy of Management Review*, 90-118.

Sandhu, R., Krishnan, R., White, G. B. (2010). Towards secure information sharing models for community cyber security. *6th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom)*. (pp. 1–6).

ACKNOWLEDGEMENTS

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. Approved for unlimited release: SAND: 2014-1679C

Managing Safety-Related Compliance of Machines in Global Market

Juha Vasara and Jouni Kivistö-Rahnasto

*Center for Safety Management and Engineering
Department of Industrial Management
Tampere University of Technology
Tampere P.O. Box 541 FI-33101 Finland*

ABSTRACT

In global trade a machine manufacturer must localize their products for different customers and market areas and they need to manage large variety of product safety requirements, conformity declarations and product liability issues. The aim of this study is to determine 1) which kind of problems there are involved in managing product safety-related requirements of machines intended for use at work internationally and 2) how globally operating companies designing and manufacturing machines have managed this issue in global market. The study is based on literature review and interviews of representatives of two large internationally operating European companies manufacturing machines intended for use at work. The companies' representatives experienced that the European integration has clarified the product safety requirements, but the actual practices may still vary between different member countries within the EU. The compliance with European product safety requirements were seen as a good basis for re-engineering the machines to the global market. The typical strategies to localize the products were 1) to meet the requirements locally in the front line, and 2) identify and take into account the local requirements and needs during the initial design and manufacturing of the machine.

Keywords: Product Safety, Requirements, Compliance Management, Global Market

INTRODUCTION

In order to succeed in global trade a machine manufacturer must be able to localize and modulate the products for different customers and market areas. The manufacturers must have tools to recognize the individual local needs and requirements but they need to be able to design the products for wider market as well. Within the European Union (EU), the Directive 2006/42/EC on machinery and the related harmonized standards define clear requirements and guidelines for product safety considerations in the design process of most machines. However, when the market area is wider than the EU the manufacturers may confront difficulties in gathering and managing all the existing local information. In addition, they may have difficulties in following forthcoming requirements. The globally operating machine manufacturer needs to be aware of the differing product safety requirements, conformity declarations and product liability issues within their market areas. They must be aware of and understand (as well as to comply with) the local or regional legislation and jurisdiction, the local operating conditions, duty types, the customers' fields of operation as well as to be able to apply the valid standards and specifications. In addition, the manufacturers must find suitable methods to take these issues into account in practice. These are e.g. the concepts of compliance management, mass customization and supply chain management.

The design of safe and ergonomic products can be an integrated part of engineering and design (e.g. Rausand and Utne; 2009; Hale, et al., 2007; Karwowski, 2005). Similarly, concepts of supply chain management as well as mass customization of products and services are quite widely publicized (e.g. Fogliatto et al., 2012; Marucheck et al.,

2011; Mentzer et al., 2001). However, despite the fact that safety design, compliance management and the supply chain management are well established disciplines, the product safety management of machines intended for use at work is not widely discussed from the compliance management and the supply chain management point of views. The aims of this study are to determine 1) which kind of problems there are involved in managing product safety-related requirements of machines intended for use at work internationally and 2) how globally operating companies designing and manufacturing machines have managed this issue in global market. In this study the management of safety-related requirements of machines is explored in two different market areas, the EU region and Australia. The study is based on literature review and interviews of representatives of two large internationally operating European companies manufacturing machines intended for use at work.

METHODOLOGY

Information on the present situation of the problems encountered in managing product safety-related requirements globally and the management practices was gathered by interviews. The interviews were conducted during 2012 and 2013 in two large companies manufacturing machines intended for use at work. Both of these companies are operating globally but they have notable business in Finland as well. The EU region and Australia are essential market areas for both of these companies. The interviewees represented product safety team, design and product line management. Representatives of the product safety teams were interviewed as individuals. Representatives of the design and product line management were interviewed in groups of 2 to 5 persons. Altogether 11 interview events with 25 interviewees were conducted; 4 representatives of the product safety teams, 16 representatives of design and 5 representatives of the product line management.

The interviews were semi-structured, including pre-prepared questions which were open-ended. The topics were covered in a nearly similar sequence in each interview. However, the group interviews were more like discussions including mutual discussion between the interviewees. The framework of the interviews was based on tentative discussions with the companies and literature review related to compliance management, differing requirements, regulatory strategies and the authorities' role. Questions for the interviews were divided into different categories including follow up of the requirements, determination/detection of the requirements, management of the requirements (compliance), liability issues and authorities' role in different market areas. The respondents also had the opportunity to discuss other issues that had not been covered in the interviews. The results of this study are qualitative.

The literature review was complemented with topics tangential to themes of the interviews. These topics consist of e.g. supply chain management, mass customization, modularization and postponement. Relevant publications were searched from the databases of electronic journals, other electronic publications and books.

THEORY

Globalization can be both challenging and an opportunity for a company. A company must decide e.g. which geographical markets they should offer their products, should they have similar product for customers throughout the world or do they need a variety of products with different specifications, and where they will develop and manufacture their products. Companies have to understand customer requirements and comply with regulations together with other requirements in many countries. (Stark, 2011; Sadiq and Governatori, 2010; Drahos and Braithwaite, 2001) The companies must also be aware of the regulatory agencies' (or authorities') strategies and styles to enforce the requirements in different market areas and what the atmosphere to comply is (Tallberg, 2002; Sutinen and Kuperan, 1999).

The basis of compliance can be either affirmative or negative (May, 2004). From the authorities perspective the implementation and enforcement can be distinguished by different approaches: cooperative/ accommodative approach and coercive/sanctioning approach. (Bluff, 2011; National Research Centre for Occupational Health and Safety Regulation, 2002; Ayres and Braithwaite, 1992; Braithwaite, 1985) The authorities may also apply different styles and approaches of interaction with the companies depending, for example on the field of business, size of the company, history of implementation and the actual legislation. (Bluff, 2011) All in all, the global market makes it more difficult to assure and enforce the product safety (Marucheck et al., 2011). Underlying issues with compliance

management are also the increasing of so called regulatory burden or overload and complex formulation/incoherence of the requirements. In general, there is a structural tendency for regulation to increase over time. (Hale et al., 2011)

Compliance management can be defined as ensuring that business processes, operations, and practices are in accordance with a set of prescribed and/or agreed requirements. It should not be a distinct activity but as a part of the business practice. (Sadiq and Governatori, 2010) Compliance with the requirements indicates e.g. that a company conforms the stated and applicable external requirements concerning it and its products and/or services (e.g. Carroll and McGregor-Lowndes, 2002). The basis of compliance is the comprehensive enough acquisition and management of the required information. However it may be asked what comprehensive enough is. The process of compliance (with the requirements) can be divided into e.g.: Identifying and discovering requirements, Interpreting requirements, Identifying changes (impact analysis), Decision of compliance, Specifying method of compliance, Communication, Implementation and applying, and Evaluation and monitoring (e.g. El Kharbili et al., 2008; Henson and Heasman, 1998). Compliance management should rather have a preventive focus, aiming to achieve compliance by design (Sadiq and Governatori, 2010; Lu et al., 2008). As a part of the compliance process the companies can try to influence the upcoming requirements (regulation, standards etc.) beforehand, try to effect on the requirements in force or their enforcement. In practice this can be carried out by the companies or the associations representing the companies. (Tala, 2001; Henson and Heasman, 1998) It can be presented that the companies have different strategic choices in responding to new or previously unknown regulations: opportunism, full compliance, partial compliance, noncompliance or influencing regulator/enforcer (Henson and Heasman, 1998). The choice may be based on e.g. the possible consequence of noncompliance (Bluff, 2011; Tala, 2001). Noncompliance can have both short-term and long-term consequences for a company, positive or negative (El Kharbili et al., 2008).

Regulation has both direct and indirect effects on product design. The companies must pay attention to existing and pending regulations that have direct and indirect effects on the designed products. (Baram, 2007) The principal instrument for regulating machinery safety in Europe is the Directive 2006/42/EC on machinery. Member nations may set additional requirements consistent with this Directive, but the freedom of movement of products must be assured (Baram, 2007). In Australia the Commonwealth government and each state and territory government regulates occupational safety and health in its own jurisdiction. There is variation in responsibilities and the duties to ensure that machine is designed and manufactured so that risks to safety or health do not exist. However, the harmonization has been initiated in the beginning of 2012. ("Occupational safety and health country profile: Australia", 2014; National Research Centre for Occupational Health and Safety Regulation, 2002; "New work health and safety (WHS) laws", 2014) There has also been an ongoing process of comparing the Australian way to regulate machinery safety to the EU and convergence to the EU's regulations. The Australian regulations and codes of practice rely on general duty requirements, performance standards, process requirements and documentation requirements. The performance standards do not describe companies exactly how they should achieve compliance, instead they define the obligations of the companies by the goals they must achieve, or the problems they must solve. (Bluff and Johnstone, 2004) There are mutual features between the Australian Occupational Health and Safety (OHS) regulatory regime for plant and the regime created by the EU's Machinery Directive. Both of these regimes have a risk management approach, require provision of information, utilize technical standards and require companies' self-assessment of machinery. In addition, they provide a third party verification for specific types of machinery. However, there are some differences as well; e.g. different compound of performance-outcome, systematic process and specification provisions used in the regimes. The Australian regime is mainly process-based but the EU's regime focuses more on achieving performance-outcomes. (National Research Centre for Occupational Health and Safety Regulation, 2002; Bluff, 2004)

In addition to the compliance management and the related information flows, companies must be able to react to changes in customer requirements and implement the design, manufacturing and distribution of the products and services. In their literature review Kara et al. (2002) introduce three basic forms of flexibility that companies need: External flexibility, Inter flexibility and Intra flexibility, which are then further divided into more detailed subgroups. The development flexibility is seen as a part of the companies' technology and inter flexibility and it describes the economic costs of product modification caused by the response to external or internal changes. Similarly, the design change and modularity flexibility is a part of the technology flexibility, but it refers more on the time rather than the costs that a company needs to implement the changes to a specific part. (Kara et al., 2002)

A commonly used framework for modelling and managing the flows of products, services, information and the financing is the supply chain management (Mentzer et al., 2001). One of the key concepts of the supply chain management is the decoupling point, the point of supply chain where the customer specific and varying requirements are added to more general standard requirements and decoupled to product structures or to a specific

customer order (Brun and Zorzini, 2009; Gosling and Naim, 2009). The management of the variation in customer needs is also extensively discussed in the mass customization literature (Fogliatto et al., 2012). The typical strategies of mass customization to tackle the problems caused by the variation are postponement and modularization (Brun and Zorzini, 2009). In the case of the postponement, the decoupling point of the customer needs and the final product customization are typically located and implemented at the later parts of the supply chain as an addition to a more pre-designed or even pre-manufactured standard product. In the case of modularization, on the other hand, the variation of the customer needs is managed by the pre-defined modules and their combinations on the basis of the customer preferences and selections.

RESULTS AND DISCUSSION

The interviewees' understanding of the problems involved in managing product safety requirements of machines and of practices the globally operating companies apply when managing these requirements are presented in this chapter. The views are discussed in general but paying particular attention to perspective of the EU region and Australia.

The problems and practices can be divided into six categories: Flow of information, Competence, Requirements, Standards, Interpretation of the Requirements and Standards, and Actions of the Organization. The essential problems remarked are presented in Table 1 and the essential practices correspondingly in Table 2. These are presented from two companies' (A and B) perspective, the results are quite consistent in both of these companies.

Problems experienced

Flow of information

Flow of information was conceived as a problem in several situations. In many cases the information of the requisite requirements is coincidental, i.e. the requirements are not systematically searched and discovered. The information may also flow inadequately both inside the different units of the company as well as between the units. For example, the information may not be exchanged between product lines and projects and therefore; similar information is searched concurrently and repeatedly. One reason may be the lack of systematic documentation. It is also possible that the company's local unit, front line, in other market area, e.g. Australia, modifies the product to meet the local requirements (see Brun and Zorzini, 2009). However, the content of these modifications are not always informed to the manufacturing unit.

Competence

Several interviewees perceived that sales and marketing, as well as the so called front line had insufficient understanding and competence in product safety. They wished that sales and marketing should consult more designers and product line when negotiating and informing potential customers. They should both receive all the requisite information related to the requirements and not promise in excess to the customer. On the other hand, the interviewees highlighted the essential role and expertise of the front line when ensuring that the local requirements are managed. However, the front line usually represents more sales and marketing without competence related to safety issues.

Requirements

Requirements and attaining compliance were related to both the problems in information management and the technological difficulties to fulfill requirements in some countries. The problem may be the lack of an adequate and available system to process the requirements. The interviewees called feasible requirements' management system, comprehensive list of safety-related requirements and the proper comparison of the similar requirements between different market areas which the products are intended.

Table 1: The essential problems in managing product safety requirements of machines

	Company A	Company B
Problems		
<i>Flow of information</i>		
Lack of documentation	x	x
Flow of information, inside of a unit	x	x
Flow of information, between the units	x	
Coincidental information	x	x
Local modifications not informed	x	
<i>Competence</i>		
Lacking competence of the sales and marketing	x	x
Lacking competence in the front line	x	
<i>Requirements</i>		
No requirements' management system	x	
No lists list of the safety-related requirements	x	
Lacking comparison of the requirements between different market areas	x	
Requirements of Australia	x	x
Requirements of the UK		x
Requirements of Sweden	x	x
<i>Standards</i>		
The apprehension and interpretation	x	x
Lack of type C standards	x	x
<i>Interpretation of the requirements and standards</i>		
No guidelines for interpretation	x	
No code of practice for design		x
Lacking global alignments for safety		x
<i>Actions of the organization</i>		
Non-uniform practices in projects	x	
Unclear responsibilities	x	
Operations in the front line		x

Table 2: The applied practices in managing the requirements

	Company A	Company B
Practices		
<i>Flow of information</i>		
Requirements' management system under construction	x	
Product data management system		x
Documentation of the stages in design	x	
Risk assessment	x	x
System to reach all their customers		x
<i>Competence</i>		
International networking meetings	x	
Participation to the drafting	x	x
<i>Requirements</i>		
The standard product platform based on European markets requirements	x	x
All the main market areas simultaneously taken into account in standard products		x
Products are based on standards		x
Certain common and minimum requirements and the specific local regulatory requirements defined	x	
Divergent requirements in contracts or orders	x	x
Local information from the front line	x	x
Sales companies' and dealers' assistance	x	x
<i>Standards</i>		
Participation to the drafting	x	x
Comparison between different market areas	x	
<i>Interpretation of the requirements and standards</i>		
Utilization of external bodies	x	x
Benchmarking	x	x
<i>Actions of the organization</i>		
Multiprofessional product safety team	x	
The product safety team follows the requirements	x	x
The product safety team determines requirements in general	x	x
Project-based determination	x	
The product safety team compiles a safety plan for a project	x	
The product lines gather the requirements	x	
Global product council		x
The product safety team as a link for the design	x	
Customers help in tracing and verifying the local requirements		x
Products are modified to meet the local requirements by the front line	x	x
The front line asks permission for modification from designing and manufacturing unit		x

Even though the EU region seems to have harmonious legislative requirements related to safety of machinery the country-specific practices, requirements and their enforcement, may vary (see Tallberg, 2002; Sutinen and Kuperan, 1999). Especially the United Kingdom and Sweden were stated to be difficult. In Australia, the other market area studied, the requirements are occasionally stricter than in other market areas. The area is not uniform, the requirements diverge territorially. Lists of requirements mentioned in sales contracts are also extensive and contract techniques must be well mastered. In addition, a personal liability related to safety falls upon the designers.

Standards

Standards are differentiated as a separate part of the requirements, due to their nature of being instructions to fulfill the requirements. The apprehension of standards and their interpretation were perceived to be difficult for the companies. In addition, the information is not necessarily up-to-date. The lack of more detailed type C standards (machine safety standards) was seen either negative or positive. Type C standards clarify the design but they may also be experienced to complicate yielding added value of safety solutions.

Interpretation of the requirements and standards

The problems related to interpretation were related to the lack of guidance and alignments. According to many interviewees clear and generally available guidelines for interpreting requirements would be appreciated. Due to this similar requirements can be interpreted repeatedly and in a different way. Code of practice for design and global alignments for safety could also unify the operations.

Actions of the organization

From actions of the organization point of view the perceived problems were around non-uniform practices in projects, unclear responsibilities and the operations in the front line. The global requirements' management system and code of practice for design could facilitate controlling these problems. The responsibilities of complying the requisite requirements during different stages of projects life cycle must be defined as well as the parent company must be aware of the actions in the front line.

Practices in use

Flow of information

The flow of information is facilitated with information management systems and documentation. The information must be stored, available and understandable for several parties. The requirements' management system, product data management system and the documentation of stages in design are examples of applied practices. A company may also have a system to reach all their customers. The system of this nature is needed e.g. to gather feedback or inform customers for detected hazards or deficiencies.

Competence

The practices brought out for ensuring or improving competence were mostly indirect. The practices around competence were internal international networking meetings, participating the drafting of the requirements, benchmarking from other companies' products and learning from accidents. Maintenance has a significant role in adding competence. They are a direct contact to customers, as well as customers working with competitors machines.

Requirements

The practices related to the requirements and attaining compliance was around how to detect the requirements, which requirements are incorporated into global products (see Sadiq and Governatori, 2010) and possibility to effect on drafting stage (see Tala, 2001; Henson and Heasman, 1998). The product safety team is following the requirements in a general level and determining specifically for a project or product. The product line may also gather the requirements for the project or product, by way of the front line, at first. A company may have differing strategies for decoupling the additional local requirements to their standard products. The standard product platform can be based on single (for example European) market requirements or the company may take all the main market areas simultaneously into account (see Stark, 2011). The requirements for customized products are partially following the basis of standard products. All the products must meet certain common and minimum requirements and the specific local regulatory requirements as well as the customer specific requirements are decoupled to these

(see Fogliatto et al., 2012). Both mass customization strategies, modularization and postponing, are applicable (see Brun and Zorzini, 2009; Gosling and Naim, 2009). If the customization strategy is based on the postponing, the front line may modify the product to meet the local requirements and they may be required to ask permission from designing and manufacturing unit.

Standards

The direct practices related to standards were mentioned to be participating the drafting process of standards, their application as a part of the design and comparison between similar standards in different market areas (see Baram, 2007). The participation covers both national and international drafting committees. Besides offering possibility for proactive influencing the committees offer an area for inter-company benchmarking as well.

Interpretation of the requirements and standards

The companies hardly expressed specific practices for interpretation. They may utilize research institutes, consultants and inspection bodies. In Australia the authority was described to be also proactively interested. They can be asked for advice as well as they are educating companies' representatives. The versatile product safety team and international networking are promoting the interpretation as well.

Actions of the organization

The actions of the organization were the mostly mentioned practices for managing the requirements. The importance of clear distribution of responsibilities between the design, product line, local front line and product safety team is the most important of the actions. Product safety team is a support but their role may be varying in different projects and with different product lines. The determination can be carried out under a different body. The boundaries of local modifications, informing and permission of these are required to deal with the front line as well. The products may be modified to meet the local requirements by the front line (see Brun and Zorzini, 2009). However, it should not be executed uninformed. Another strategy to meet the local requirements is to by the help of the local front line and customers design and manufacture ready at once.

CONCLUSIONS

The companies' problems in managing product safety requirements of machines globally and the practices when managing these requirements can be divided into six categories: Flow of information, Competence, Requirements, Standards, Interpretation of the Requirements and Standards, and Actions of the Organization. According to the interviewed companies' representatives most of the perceived problems were related to the flow of information and requirements. The essential problems were not distinctly reflected to the current management practices. Most of the practices in use were instead around the actions of the organization.

In this study a particular attention was paid to two market areas: the EU region and Australia. The EU region's general requirements seem to be well under control and the European integration has clarified the product safety requirements as well. However, the varying country-specific practices, requirements and their enforcement may pose difficulties. In the Australian market the requirements are occasionally stricter than in other market areas. In addition, the area is not uniform. The requirements diverge territorially, though the harmonization process is ongoing.

The required information from another market area is typically gathered with the help of the company's local unit, the front line, and the customers. It is also possible that the designing and manufacturing unit does not determine all the local requirements, i.e. the front line may modify the product to meet the local requirements. A company may have differing strategies for decoupling the additional local requirements to their standard products. The standard product platform can be based on a single market's requirements or a company may take all the main market areas into account simultaneously. If the customization strategy is based on postponing, the problem may be that the front line does not always inform the content of these modifications to the designing and manufacturing unit. Related to the flow of information a notable problem is also that information may not be exchanged between product lines and projects and therefore; similar information of the requirements is searched concurrently and repeatedly. More specific, a generally available and accessible documentation is required.

REFERENCES

- Ayres, I., Braithwaite, J. (1992), *“Responsive regulation: Transcending the deregulation debate”*, Oxford University Press.
- Baram, M. (2007), *“Liability and its influence on designing for product and process safety”*, Safety Science Volume 45. pp. 11-30.
- Bluff, L. (2004), *“Regulatory Strategies for the Safe Design of Plant”*, Working Paper 24. 26 p. Retrieved February 26, 2014, from <https://digitalcollections.anu.edu.au/bitstream/1885/39959/3/WorkingPaper24.pdf>
- Bluff, L. (2011), *“Something to Think About – Motivations, Attitudes, Perceptions and Skills in Work Health and Safety. A Review of the Literature on Socio-Psychological Factors and Their Influence on Organisations’ and Individuals’ Responses to Regulation”*. Report prepared for Safe Work Australia. 71 p. Retrieved February 26, 2014, from http://www.safeworkaustralia.gov.au/sites/SWA/about/Publications/Documents/604/Something_to_think_about.pdf
- Bluff, L., Johnstone, R. (2004), *“The Relationship Between ‘Reasonably Practicable’ and Risk Management Regulation”*, Working Paper 27. 45 p. Retrieved February 26, 2014, from https://digitalcollections.anu.edu.au/bitstream/1885/42687/2/Pract_RiskManWPfinal.pdf
- Braithwaite, J. (1985), *“To punish or persuade: Enforcement of coal mine safety”*, State University of New York Press.
- Brun, A., Zorzini, M. (2009), *“Evaluation of product customization strategies through modularization and postponement”*, International Journal of Production Economics Volume 120. pp. 205-220.
- Carroll, P., McGregor-Lowndes, M. (2002), *“Managing Regulatory Compliance”*, Paper presented at the Current Issues In Regulation: Enforcement and Compliance Conference. Australian institute of criminology.
- Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast). Official Journal of the European Union 9.6.2006 L 157/24.
- Drahoš, P., Braithwaite, J. (2001), *“The globalisation of regulation”*, The Journal of Philosophy Volume 9 No. 1. pp. 103-128.
- El Kharbili, M., Stein, S., Markovic, I., Pulvermüller, E. (2008), *“Towards a Framework for Semantic Business Process Compliance Management”*. Proceedings of GRCIS 2008.
- Fogliatto, F.S., da Silveira, G.J.C., Borenstein, D. (2012), *“The mass customization decade: An updated review of the literature”*, International Journal of Production Economics Volume 138. pp. 14-25.
- Gosling, J., Naim, M. (2009), *“Engineer-to-order supply chain management: A literature review and research agenda”*, International Journal of Production Economics Volume 122. pp 741-754.
- Hale, A., Borys, D., Adams, M. (2011), *“Regulatory Overload: A Behavioral Analysis of Regulatory Compliance”*, WORKING PAPER No 11-47. Mercatus Center, George Mason University. 31 p.
- Hale, A., Kirwan, B., Kjellén, U. (2007), *“Safe by design: where are we now?”*, Safety Science Volume 45. pp. 305-327.
- Henson, S., Heasman, M. (1998), *“Food safety regulation and the firm: understanding the compliance process”*, Food Policy Volume 23 No. 1. pp. 9-23.
- Kara, S., Kayis, B., O’Kane, S. (2002), *“The Role of Human Factors in Flexibility Management: A Survey”*, Human Factors and Ergonomics in Manufacturing Volume 12, No. 1. pp. 75-119.
- Karwowski, W. (2005), *“Ergonomics and human factors: the paradigms for science, engineering, design, technology and management of human-compatible systems”*, Ergonomics Volume 48, No. 5. pp. 436-463.
- Lu, R., Sadiq, S., Governatori, G. (2008), *“Compliance Aware Business Process Design”*, in: Business Process Management Workshops, Hofstede, A., Benatallah, B., Paik, H-Y (Ed.).
- Maruchek, A., Greis, N., Mena, C., Cai, L. (2011), *“Product safety and security in the global supply chain: Issues, challenges and research opportunities”*, Journal of Operations Management Volume 29. pp. 707-720.
- May, P.J. (2004), *“Compliance Motivations: Affirmative and Negative Bases”*, Law & Society Review Volume 38 No. 1. pp. 41-68.
- Mentzer, J.T., DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.D., Zacharia, Z.G. (2001), *“Defining supply chain management”*, Journal of Business Logistics Volume 22. pp. 1-25.
- National Research Centre for Occupational Health and Safety Regulation (2002). *“Towards a Regulatory Regime for Safe Design. A review of regulatory approaches and enforcement strategies”*, Retrieved February 26, 2014, from http://www.safeworkaustralia.gov.au/sites/SWA/about/Publications/Documents/288/TowardsARegulatoryRegimeForSafeDesign_2000_PDF.pdf
- New work health and safety (WHS) laws (2014), Retrieved February 26, 2014, from <http://www.business.gov.au/BusinessTopics/Occupationalhealthandsafety/Pages/OHandSActsRegulationsandCodesofPractice.aspx>
- Occupational safety and health country profile: Australia (2014), Retrieved February 26, 2014, from <http://www.ilo.org/safework/countries/asia/australia/lang--en/index.htm>
- Rausand, M., Utne, I.B. (2009), *“Product safety – Principles and practices in a life cycle perspective”*, Safety Science Volume 47. pp. 939-947.
- Sadiq, S., Governatori, G. (2010), *“Managing Regulatory Compliance in Business Processes”*, in: Handbook on Business Process Management 2 Strategic Alignment, Governance, People and Culture, vom Brocke, J., Rosemann, M. (Ed.). 616 p.
- Stark, J. (2011), *“Product Lifecycle Management. 21st Century Paradigm for Product Realisation”*, Decision engineering. 562 p.
- Sutinen, J.G., Kuperan, K. (1999), *“A socio-economic theory of regulatory compliance”*, International Journal of Social Economics Volume 26 No. 1. pp. 174-193.
- Tala, J. (2001), *“Lakien vaikutukset. Lakiuudistuksen tavoitteet ja niiden toteutuminen lainsäädäntöteoreettisessa tarkastelussa”*, Oikeuspoliittisen tutkimuslaitoksen julkaisuja 177. Hakapaino Oy. 390 p.
- Tallberg, J. (2002), *“Paths to Compliance: Enforcement, Management, and the European Union”*, International Organization Volume 56. pp. 609-643.

Cross-Cultural Differences in Preference for Relationally Framed Decision Alternatives

Lelyn D. Saner^a, Andrew Mathis^b, Sergey Blok^b, Sharon Glazer^c, Ivica Pavisić^c, and Susannah Paletz^a

^a*Center for Advanced Study of Language,
University of Maryland, College Park, MD 20742, USA*

^b*Booz Allen Hamilton, Washington, DC, 20005, USA*

^c*University of Baltimore, Baltimore, MD 21201, USA*

ABSTRACT

Much research on cross-cultural psychology has focused on either culture-level dimensions or individual values as starting points for explaining the influence of culture on individual reasoning and decision-making. Culture is a complex concept, however, determined by the beliefs and behaviors of both individuals and of social systems. To understand how culture predicts the behavior of an individual in a situation requires lower-level descriptors of how individuals and groups interact in different contexts. We investigated the application of Relational Models Theory (Fiske, 1992) as a way both to describe social situations and to distinguish cultures by which relational models their members consider to be the most appropriate in different situations. We presented decision scenarios to participants from different cultural backgrounds through a survey and asked them to rate the appropriateness of several responses to each scenario that were oriented toward different relational models. We observed significant interactions between cultural background, scenario, and the ratings given to options associated with each relational model. We concluded that relational models might provide a valuable tool for understanding cultural differences in individual decision-making, but that the context of the situation itself also has a significant impact on the options people consider to be most appropriate for resolving situations.

Keywords: Relational Model, Decision Scenario, Situation, Domain

INTRODUCTION

One of the challenges in the study of cross-cultural decision-making is to identify factors that directly connect the cultural values to decisions made by individuals in different situations. Culture is manifested in the values, norms, behaviors, and beliefs that are implicitly shared among members of a social system and passed through the generations (Rohner, 1984). Social systems vary in size and formality, including countries, nations, regions, organizations, ethnic groups, genders, families, and teams. Individuals who belong to a culture are shaped by it and contribute to its continuous reshaping. (Casnir, 1999). The goal of cultural and cross-cultural research is to understand not only culture's influence on *what* people think, but also its influence on *how* they think, through identification of culture features that impact human cognitive processes (Oyserman & Lee, 2007).

Much research has been done at several levels of analysis. Various studies have examined high-level cultural values, such as *horizontal vs. vertical individualism and collectivism* (HVIC; Triandis & Gelfand, 1998), general *human values*, that is, principles that motivate cognition, affect, and behaviors (e.g., benevolence, power, security; Schwartz, 1992), and *self-construals*, which are how people think about themselves in relation to others (e.g., independent vs. interdependent; Markus & Kitayama, 1991). Individualism and collectivism, for example, are culture-level dimensions that describe a culture's values regarding how people should identify themselves and behave in relation to society. Individualistic societies encourage their members to see themselves as unique and

independent from others. Collectivistic societies, in contrast, encourage people to see themselves as interrelated with one another and to prioritize the well-being of the group.

Although the individualism-collectivism distinction should not be used to compare individuals, it is often the starting point for cultural studies (e.g., Hofstede, 1984; Kim & Markus, 1999; Markus & Kitayama, 1991; Triandis, 1989). People in individualistic cultures are more likely to consider themselves in independent terms and people in collectivistic cultures are more likely to consider themselves in interdependent terms (i.e., independent vs. interdependent construals of self). But although the actions of individuals reinforce cultural values and help individuals internalize cultural values as their own, a person's reasoning processes do mediate the relationships between cultural values and actions. Triandis (1996) noted that culture is too complex to be adequately characterized by one cultural dimension. Just as a person can be described by many personality traits, so are cultures distinguished by many cultural qualities (e.g., Hofstede, 1984; Schwartz, 1992). As such, using general characterizations of culture to explain specific choices in particular contexts may ignore the more immediate influences of the context on actions.

In this study, we examined the constructs of Relational Models Theory (RMT; Fiske, 1992, 2004) as potentially more descriptive and discriminating of cultural differences in how individuals respond to group decision situations. RMT addresses relationship structures by focusing on the interplay between a focal person and others. The relational models characterize an individual's preferences or expectations surrounding the behaviors of others in one's social sphere, expressing four ways in which people engage in interpersonal interactions. *Equality Matching* (EM) refers to a relationship based on tit-for-tat reciprocity, matched contributions, and equally divided distributions. *Market Pricing* (MP) refers to proportional reciprocity and relationships based on equitable (not equal) distribution. The focal consideration is the utility of others in one's social sphere. *Authority Ranking* (AR) refers to status differentials and hierarchical structures. And finally, *Communal Sharing* (CS) refers to in-group equality, whereby members of in-groups provide and take resources as needed. Each of the four models helps people organize, coordinate, and adapt to relational situations; they likewise contribute to the ways people "construct and construe social action" (Fiske, 2004, p. 21).

We investigate relational models as generalized preferences or tendencies across situations. This study examined whether relational models capture cross-cultural differences in which alternatives are considered to be most appropriate for resolving decision-making scenarios. Prior research on Relational Model Theory (RMT) has established that all four of the models are utilized across different cultures (Fiske, 1992). However, cultures are likely to differ, not only by which models are preferred in general, but also by which models should characterize each type of situation. Despite compelling evidence of the universal features of the models, there have been relatively few empirical studies examining cultural differences in the way the models are represented and used across different situations, actors, and domains (e.g., morality, identity).

Koerner (2006) used the *Modes of Relationship Questionnaire* (MORQ; Haslam & Fiske, 1999) to contrast North American and Singaporean students regarding both acquaintance and friendship relationships. North Americans were found to give higher ratings to Communal Sharing (CS) oriented responses than Singaporeans did, whereas Singaporeans gave higher ratings to Authority Ranking (AR) oriented strategies than did North Americans. In another study, Woodhull and Louis (2009) constructed a novel scenario-based measure, with four choices that corresponded to each of the four relational models. For example, in one of the scenarios a business division is assigned a complex job to complete and participants had to choose among four relational model-related options for dividing the tasks of the larger project. These scenarios were constructed based on seven contextual "domains" of interaction (reciprocal exchange, distribution, work, moral judgment and ideology, decision-making, social influence, and identity). Woodhull and Louis contrasted US/Canadian and Mexican corporate managers, all of whom worked for US/Canadian multinational companies. They reported that Mexican managers consistently ranked Communal Sharing-oriented options higher than US/Canadian managers did, across all tested domains, whereas preferences for the other relational models differed between domains.

These studies provide evidence that relational model endorsement in any given domain can vary by cultural context. Applications of relationship structures that may be dominant in one society will not necessarily be dominant in another society. Furthermore, even within a given society, it is important to consider the domain that contextualizes a given relationship when interpreting patterns of relational model priority. However, additional research is needed to understand the extent to which relational model endorsement differs generally across cultures. In this study, we compared participants of European American background to three other groups living in the USA with different cultural backgrounds: Afghan, Chinese, and Iranian. We chose these three groups for several reasons. First, they represent three kinds of "collectivistic" cultures, even though they may differ in relationship structure emphasis (Dien, Blok, & Glazer, 2011; Hamedani, Purvis, Glazer, & Dien, 2012). Second, there is some research on both

Euro-American and Chinese-American cultures that makes them good points of reference for understanding how culturally different Afghan and Iranian culture are from them and each other (e.g., Hong, Morris, Chiu, Benet-Martinez, 2000). Finally, because relational models are a different aspect of culture than has been studied in the past, this study also extends current research in a new direction, and a comparison of multiple groups with respect to these important cultural dimensions allows for a rich interpretation of the findings (Ross, 2003).

METHOD

Participants

There were 213 participants in this study, including 71 participants of Euro-American background, 57 of Chinese background, 24 of Afghan background, and 61 of Iranian background. It is important to note that, although the cultural background of participants was rigorously verified, participants were neither necessarily native to the countries of their cultural heritage nor living in those countries at the time they participated. Therefore, as a matter of precision, we refer to participants as being of a cultural “background” rather than of a “culture.” Age and sex demographics for the final sample are reported in Table 1.

Table 1: Sample size, sex, and age by culture group

Culture Group	n	Female	Male	Age		
				M	Range	SD
Euro-American	71	42 (59.2%)	29 (40.9%)	31.2	18-64	12.13
Afghan	24	9 (37.5%)	15 (62.5%)	26.0	18-47	6.13
Chinese	57	39 (68.4%)	18 (31.6%)	27.4	18-60	10.29
Iranian	61	33 (54.1%)	28 (45.9%)	29.5	18-64	12.58

In the demographics questionnaire on the online version of the survey, participants were asked to name their birthplace, the birthplace of their primary caretakers, and whether they spoke a language other than English at home. The distribution of participants on these measures is shown in Table 2.

Table 2: Home language and country of birth by cultural background group

Culture Group	Spoke a language other than English at home		Born in Focal Country	
	n	%	n	%
Euro-American	2	2.8	71	100.0
Afghan	22	91.7	16	66.7
Chinese	55	96.5	39	57.0
Iranian	57	95.0*	34	61.0

* Out of total including 1 missing value.

Procedure

This study was conducted via survey that was administered to participants in either electronic format online or in person using a paper-and-pencil version. Participants were recruited through a variety of channels, and their cultural

affiliations were authenticated according to a predefined decision tree in order to reach a diverse population for each cultural group.

Participants for the online version were recruited through online mailing lists and relevant organizations (e.g., University of Maryland message boards, culture-specific professional and social organizations, and foreign language instructors), flyers and handouts posted in libraries, cafes, and restaurants throughout the District of Columbia (DC) metropolitan area, print advertisements placed in DC-area publications, web advertisements, social networking on the part of both the researchers and the contacted participants, and survey service organizations, including Amazon Mechanical Turk (AMT) and Affordable Samples. For the paper-and-pencil version of the survey, project staff set up a booth outside of a main campus center on the University of Maryland campus and administered the survey to Afghan and Iranian students. Project staff also visited DC area Afghan restaurants and recruited Afghan participants to complete the survey while at the restaurant.

In order to be eligible to participate in the study, all potential participants had to pass a pre-screening process before receiving an invitation to take the survey. Eligible participants were assigned to one of two versions of the survey that differed only in the order of question item presentation. Online survey participants were sent an e-mail containing an invitation link to the online survey powered by the LimeSurvey software (version 1.91+), hosted on the University of Maryland's web server. Paper surveys were generally completed in a single session. However, because of some logistic challenges associated with gaining the participation of people in some groups, a few participants were allowed to complete the survey in two sittings, no more than one day apart.

Measures

The entire survey included a demographics questionnaire and a series of scenarios measuring relational model endorsements. The demographic questionnaire replicated items on the pre-screening questionnaire, and also asked about religious observance, marital status, education level, and the location from which the participant took the survey (for the online version).

Using Woodhull and Louis's (2009) measurement approach as a foundation, we expanded their scenarios to cover a larger range of social interaction situations beyond just the business context. A set of 35 scenarios was generated using all possible combinations of seven contextual domains (i.e., collective decision-making, distribution of resources, reciprocal exchanges, identity, influence, morals, and work or labor organization) applied in five different situations (i.e., a psychology class project, a technology convention, a suburban neighborhood, engineers building a bridge, and a student newspaper). These contextual domains were a subset of 15 domains originally presented by Fiske (1992) that were selected by Woodhull (2006). Table 3 presents a sample scenario and response options.

Table 3: Example Scenario

Situation	Sample Scenario	Sample Options
Students in a psychology class who are working on a research project	Students need to use a computer program to analyze their results. This program is only available on a single laptop.	The professor should appoint a student to be the leader. That person should have priority with the laptop computer. [AR]
		All team members should share the use of the laptop computer with their fellow team members. [CS]
		The data analysis time period should be divided into equal parts and students should take turns with the laptop. [EM]
		The students should set up a fund to pay for project expenses. Time with the laptop should be proportional to how much money a student is willing to contribute to the project fund. [MP]

Each scenario had four options to choose from for how to respond to the situation and each option was representative of a different relational model. Participants were asked to rate each response option using a Likert-type scale, with five options, ranging from "Strongly Disagree" (coded as 1) to "Neutral" (coded as 3) to "Strongly Agree" (coded as 5). The ratings reflected the extent to which respondents agreed that each model was appropriate for addressing the problem in the situation. Respondents each rated the options in all of the 35 scenarios and the presentation order of the situations was randomized across participants.

RESULTS

There were 254 surveys submitted by eligible participants, but some participants with missing data or invalid rankings (e.g., double ranked) were eliminated from each of the cultural background groups, leaving a final sample size of 213 participants.

The focal research question for this study was that of the degree to which there are cultural differences in the judged appropriateness of applying each relational model in different circumstances. We employed a Repeated-Measures ANOVA to analyze the rating data. In order to control for differing scale use in the rating data, the rating ANOVA was performed with participants' overall mean ratings across all items as a covariate. The estimated marginal mean ratings presented below assume an overall mean rating of 3.39, which is the average across all participants. The between-subject factors were cultural background group and sex. The within-subject factors were the ratings by relational model (4 models), the situation of each decision scenario (5), and the domain of each scenario (7).

Effects of situation and domain were used to determine whether the scenarios themselves affected rating outcomes, but of particular interest were the interaction effects between cultural group and response ratings, while controlling for sex and scenario. This analysis unpacks cultural differences in the relative importance of the different relational models. Although post-hoc comparisons among culture groups were examined, there were very few significant differences that did not involve participants of European-America background. As such, unless otherwise noted, all reported comparisons are between the participants of European-American background and those in each of the other cultural background groups.

Differences in Scenario Response Ratings by Relational Model

When controlling for culture, sex, and scenario, we observed a significant effect for relational model in the scenario response ratings, $F(2.33, 475.61) = 9.96$, $p < 0.001$, $\eta_p^2 = .047$. In other words, ratings of response options associated with each model were significantly different. Post-hoc comparisons indicated that ratings for MP-related responses were significantly lower than ratings for responses associated with the other three relational models ($p < 0.001$), none of which differed from each other.

Interactions of Culture and Relational Model

Significant variation in rated appropriateness of different relational models was observed in the interaction between cultural background group and response option relational model orientation, indicating that cultural groups differed generally in their relative appropriateness assessments of options associated with each relational model, $F(6.99, 475.61) = 4.91$, $p < .001$, $\eta_p^2 = 0.067$. Estimated marginal means for this interaction are shown in Figure 1.

Post-hoc comparisons between cultural background groups on each relational model indicated that participants of European-American background rated *authority ranking* oriented options significantly higher ($M = 3.82$) than participants of both Afghan background ($M = 3.54$; $p < 0.01$) and Iranian background ($M = 3.40$; $p < 0.01$). Furthermore, those of European-American background rated *equality matching* oriented options significantly higher ($M = 3.74$) than those in the Chinese background group ($M = 3.56$, $p < 0.05$). With respect to actions framed in terms of *market pricing*, however, participants in the European-American group had significantly lower ratings ($M = 2.28$) than those of Afghan ($M = 2.58$, $p < 0.05$), Chinese ($M = 2.62$, $p = 0.001$), and Iranian ($M = 2.59$, $p = 0.001$) background. Cultural background groups did not differ significantly on *communal sharing*, and all of these effects were observed while controlling for sex and scenario. Further results showed that cultural differences in option ratings themselves differed by sex as well, revealing a significant sex by cultural group interaction, $F(6.99, 475.61) = 2.21$, $p < 0.05$, $\eta_p^2 = .032$.

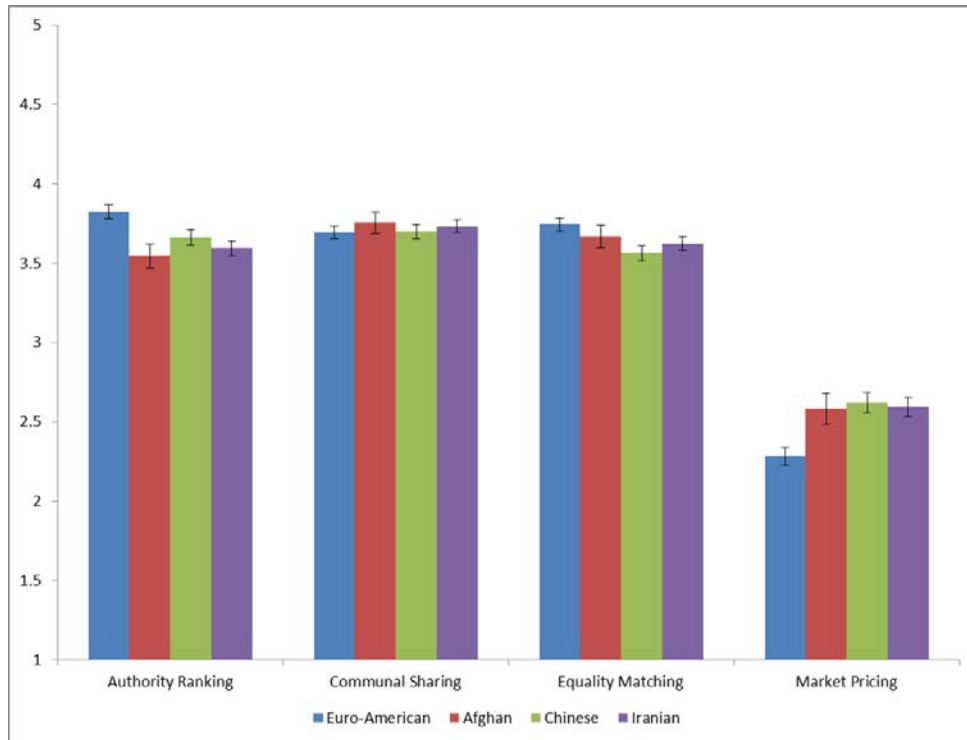


Figure 1. Estimated marginal mean ratings for interaction between Culture and Relational Model (with Standard Error bars)

Interactions of Situation, Domain, and Culture

In order to determine whether ratings were related to context, that is, to particular situations, domains, and scenarios, we first examined the interaction between scenario situation and scenario domain, and we found it to be significant, $F(19.86, 4052.01) = 2.41$, $p < 0.001$, $\eta_p^2 = 0.012$. Estimated marginal means for this interaction are shown in Figure 2, where each bar represents one of the 35 scenarios in terms of one situation in one domain.

As the graph shows, there was a great deal of variability between the scenarios with respect to the overall average ratings given to the options, regardless of the relational model orientation of the options and across all participants. A large number of post-hoc comparisons between the scenarios were significant, but because no one situation or domain clearly dominated all of the others, we do not describe them all in detail here.

Although Figure 2 does not address differences in ratings relative to cultural background, there was a significant three-way interaction between situation, domain, and cultural background group as well, $F(59.59, 4052.01) = 1.84$, $p < 0.001$, $\eta_p^2 = 0.026$. Post-hoc comparisons revealed additional significant pairwise differences between cultural background groups, even on ratings within scenarios. Both of these interactions illustrate the impact that context by itself has on how people make decisions about the appropriate way to respond to a situation. Individual scenarios led to more positively or negatively rated choice options by participants from different cultural backgrounds.

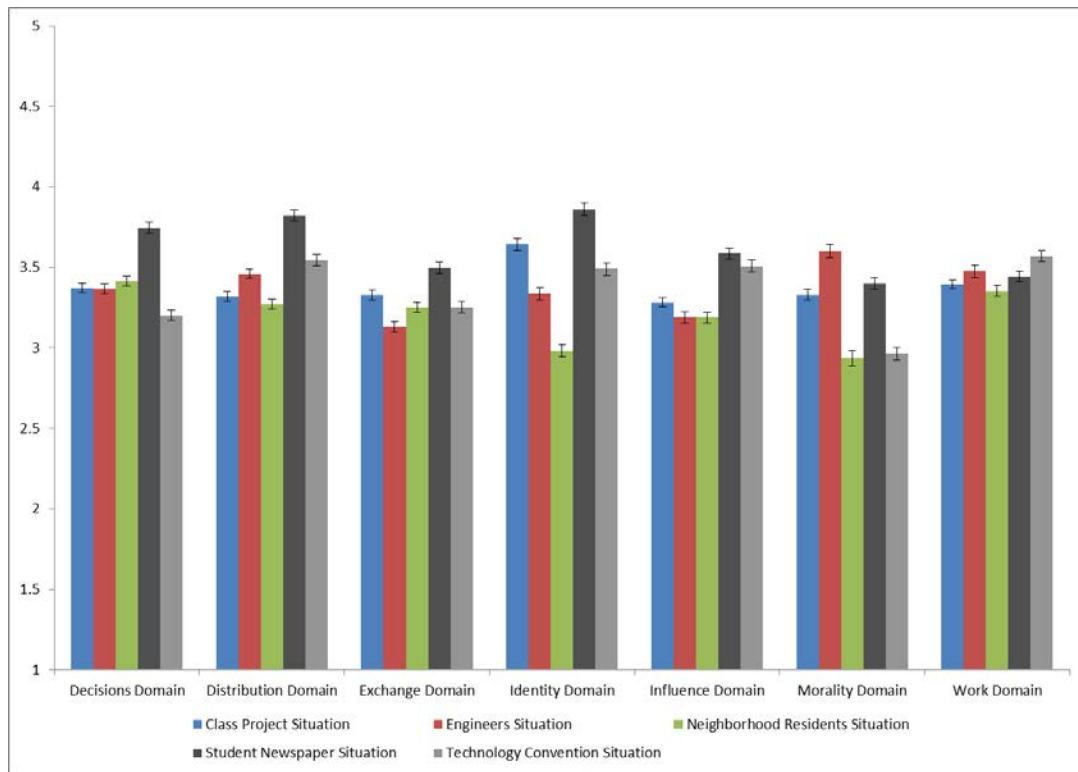


Figure 2. Estimated marginal mean ratings for interaction between Domain and Situation (with Standard Error bars)

Interactions of Relational Models with Situation, Domain, and Culture

The preceding findings demonstrated the importance of context on social decision-making, both within and across cultures. Taking relational models into consideration, there was a significant interaction between scenario situation, cultural background, and the relational model orientation of response options. Ratings of options associated with particular models, made by participants from different cultural backgrounds, varied significantly depending on the situation of the scenario being presented, $F(28.86, 1962.68) = 2.04, p < 0.001, \eta_p^2 = 0.029$. For example, although participants of European-American background rated *authority ranking* oriented actions more highly than did those of Iranian background generally, this pattern was more pronounced for scenarios involving the *engineers* situation than for those involving the *neighborhood residents* situation.

A similar result was found in the significant interaction of cultural background, scenario domain, and relational model orientation, $F(36.14, 2457.64) = 2.35, p < 0.001, \eta_p^2 = 0.033$. As an example, although those of European-American background generally rated *market pricing* oriented responses significantly lower than did those of Afghan background, the differences were more apparent in the scenarios involving the *work* domain than in those set in the *distribution* domain.

Finally, a significant four-way interaction between cultural background, domain, situation, and relational orientation of options, $F(116.20, 7901.69) = 1.71, p < 0.001, \eta_p^2 = 0.024$, demonstrated that context dependency of social decision-making extended to the level of each individual scenario. One of the largest differences observed ($M_{\text{diff}} = 0.84, SE = 0.15$) was with respect to the situation of *students in a psychology class* working on a research project (*work* domain – “This complex project has several different parts. How should the different tasks be assigned?”). Participants of Afghan background rated the *market pricing* oriented option significantly higher than did participants of European-American background (i.e., “The students should set up a fund to pay for project expenses. Students who contribute more to the fund should get to do less work and vice versa. Overall, the amount of work should be inversely proportional to the contribution to the fund.”).

All of these complex interactions reinforce that social decision-making is very specific to each case at hand and that cultural decision-making paradigms are more complex than the broad patterns distinguished solely on the basis of relational models. Context must be taken into consideration as well.

CONCLUSIONS

Implications of Findings

This study set out to answer the research question of whether relational models could capture cross-cultural differences in coordination and decision-making considerations related to a variety of scenarios. Overall, our results showed significant cultural differences in judged appropriateness of scenario response options associated with the *authority ranking*, *equality matching*, and *market pricing* relational models, but not with *communal sharing*.

European-American participants had significantly lower ratings for *market pricing* related options, relative to participants of Iranian, Chinese, and Afghan backgrounds; however, all participants within each of the cultural groups rated and ranked the *market pricing* related options significantly lower than options related to other relational models. It might be that the comparative evaluation of *market pricing* oriented approaches demonstrates a general indifference for the other relational models, but it could also be a function of a poorly developed set of options for the *market pricing* oriented choices. In terms of the first consideration, there is research indicating that self-enhancement values, which may be conceptualized as utility-based values, are universally least preferred compared to self-transcendence, conservatism, and openness to change values, though they are also generally preferred more in some cultures compared to others (Glazer, Daniel, & Short, 2004; Roccas & McCauley, 2004; Schwartz & Bardi, 2001). In terms of the latter consideration, the majority of the *market pricing* related options that were presented in the survey were framed in terms of monetary value and transactions. Numerous comments from participants indicated that they were upset at the idea of applying a dollars-and-cents framework to the social scenarios provided. Corroborating this interpretation is the fact that mean ratings were higher for *market pricing* related options in the *student newspaper* scenarios (paired t-test, $t(213) = 20.14, p < 0.001$), which were generally framed in terms of productivity, than the ratings of options with the same orientation in other situations, which were generally framed in terms of money. It may well be that respondents across the different cultural background groups would respond more positively to market pricing options that emphasized utility without money.

Further, cultural differences were found in ratings for *equality matching* and *authority ranking*, with participants of European-American background rating *authority ranking* oriented options higher than did those of Afghan and Iranian backgrounds, and rating *equality matching* oriented options higher than did those of Chinese and Iranian backgrounds. Together, these findings suggest that people of European-American background prefer people in authority roles to make decisions, whether the authority is based on status or on merit. The sex of participants was particularly influential on the cultural differences in ratings of *authority ranking* and *market pricing* options as well. Among female participants, those of European-American background only showed differences from those of Chinese and Iranian background, whereas males of European-American background only differed from males of Afghan background. These results suggest that patterns of contrast between cultures may be different when the sub-cultural factor of sex is considered. This is not unexpected as men and women, boys and girls, even within cultures, often respond differently to the same situation (Shiraev & Levy, 2004).

Study results also indicated that context (i.e., the combination of situation and decision domain) impacted the degree to which different relational model orientations were considered appropriate, both across and within cultures. Complex interactions of culture, scenario situation, and scenario domain created statistically distinct patterns of relational model ratings. This result is consistent with related findings in other cross-cultural decision-making and performance research (e.g., Earley, 1993; Yates, Lee, Sieck, Choi, & Price, 2002). Although the scenarios assessed were varied enough to point towards broad cultural differences, the prevalence of differences in ratings related to individual scenarios suggests that further research is required to obtain a clear understanding of the impact of various situations and contexts on the different relational decisions across cultures. To fully understand the application of relational models, variations need to be analyzed at the level of relevant situation and context for each culture. For example, studies of how decisions are made for *resource allocation* in specified cultures (e.g., a small village in rural Afghanistan), and the effects of violations of the decision schema on dyadic relationships, would permit comparisons across cultures within particular categories of situations; in this case, *resource allocation*. Admittedly, the situations presented in this study were only a sample of the full variety of possible contexts in real-world social life. The complexity inherent in determining when and where each relational model is applied means that generalizing about relational models on a culture-wide level should not be done lightly. Still, for the purposes of drawing out culture-general tendencies, the sampling of scenarios and domains used in this study provides important results that can be expanded on in further investigation of how culture might explain why some models were prioritized over others in different circumstances.

Even with nuanced context (situation-by-domain) effects, the culture-by-relational model interaction effects were observed across the contexts. The European-American group did have some usage patterns that were distinct from the other groups examined. These differences were striking considering the makeup of the culture groups in the sample. As scholars of culture know all too well, “culture” is not a clear-cut, monolithic property that can clearly be assigned to individuals or groups. The non-European-American culture groups in our sample mostly represented immigrants to the USA and not residents of their countries of origin. These participants also varied in the extent to which they had been acculturated, with variation in first vs. second generation, ages, and life experiences. Significant differences in ratings across the culture groups in our sample may represent only a small fraction of the possible distinctions between Americans and foreign nationals residing in their countries of origin. Even so, our analysis of the scenario response ratings found significant cultural differences, a testament to the powerful influence that “culture” might have across life experiences.

Challenges and Limitations

In this study, we focused on determining whether key variables exhibited expected interactions at a pan-situation and pan-domain level. For the sake of conciseness, however, a number of interesting, more detailed analyses had to be left for later investigation, such as a comparison of male and female ratings within cultures. Similarly, the analysis of how cultural background interacted with context verified that such interactions did occur among different combinations of variables, but further investigation is needed to complete more directed analyses of whether situation or domain was a stronger driver of the rating outcomes, as well as to compare effect sizes for specific combinations of relational model orientation and cultural background.

Future studies would also benefit from developing scenarios that are authentic and establishing an understanding of how people within each of the focal cultures would respond to them. A major challenge to the construction of the scenarios was to avoid biasing the participants towards any one particular relational model based on the content or framing of the option. This precluded the introduction of scenarios that may have served as stronger manipulations of interpersonal closeness. For example, a scenario that describes a family was initially considered but ruled out because even the use of the term ‘family’ is likely to prime *communal sharing* regardless of culture.

Another potential influence in this study was that, despite the measures taken to verify the authenticity of participants’ cultural backgrounds, subtle population differences between sample groups may have had some impact on the observed effects. Most of the significant effects were differences between participants of European-American background and those of other cultural backgrounds. To be included in the European-American background group, potential participants had to be citizens of the USA or children of citizens, but could be of any unidentified European ethnic heritage. It is possible that some unmeasured demographics among those of European-American background drove some of the significant contrasts. It is notable, based on Table 2, that this group had the lowest rate of multilingualism in the home and that all members of the sample group were born in the USA. Additionally, it was particularly challenging to recruit participants of Afghan background for this study. A replication using larger numbers of participants from different ethnic groups is necessary to increase the generalizability of this study. Regardless, obtaining and verifying “pure” samples of cultural perspective will always be a challenge in empirical work. We adopted a rich decision aid for testing the sample in this study and future work would provide opportunity for further honing of such tools.

Finally, even with pilot testing of realistic decision scenarios, it is always possible for spurious differences in framing of alternatives to lead to differences in ratings. For example, *market pricing* was least preferred of all relational models across all cultures, suggesting that this might be a universal indifference. However, it is also possible that the *market pricing* response options were inappropriate (i.e., numerous MP options were framed around money). It may take multiple deployments of a measurement instrument to achieve full validation. However, considering the robustness of the broad effects and the range of significant post-hoc comparisons spanning many condition combinations, there is strong basis for concluding that the scenarios used in this study were diagnostic of real differences in participant response.

REFERENCES

- Casrnir, F. L. (1999). *Foundations for the study of intercultural communication based on a third-culture building model*. International Journal of Intercultural Relations, 23(1), 91-116.
- Dien, J., Blok, S., & Glazer, S. (2011). *Cultural priming: Adopting the adversary’s mindset to improve analysis*. College Park,

- MD: University of Maryland Center for Advanced Study of Language.
- Earley, P. C. (1993). *East meets West meets MidEast: Further explorations of collectivistic and individualistic work groups*. Academy of Management Journal, 36, 319-348.
- Fiske, A. P. (1992). *The four elementary forms of sociality: Framework for a unified theory of social relations*. Psychological Review, 99(4), 689-723.
- Fiske, A. P. (2004). *Relational models theory 2.0*. In N. Haslam (Ed.), *Relational models theory: A contemporary overview* (pp. 3-25). Mahwah, NJ: Lawrence Erlbaum Associates.
- Glazer, S., Daniel, S. C., & Short, K. M. (2004). *A cross-cultural study of the relationship between organizational commitment and human values*. Human Relations, 57, 323-345.
- Hamedani, N. G., Purvis, T. M., Glazer, S., & Dien, J. (2012). *Ways of manifesting collectivism: An analysis of Iranian and African cultures*. College Park, MD: University of Maryland Center for Advanced Study of Language.
- Haslam, N., & Fiske, A. P. (1999). *Relational models theory: A confirmatory factor analysis*. Personal Relationships, 6(2), 241-250.
- Hofstede, G. H. (1984). *Culture's consequences: International differences in work-related values*. London: Sage.
- Hong, Y.-Y., Morris, M. W., Chiu, C.-Y., & Benet-Martinez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist*, 55, 709-720.
- Kim, H., & Markus, H. R. (1999). *Deviance or uniqueness, harmony or conformity? A cultural analysis*. Journal of Personality and Social Psychology, 77(4), 785-800.
- Koerner, A. F. (2006). *Models of relating-not relationship models: Cognitive representations of relating across interpersonal relationship domains*. Journal of Social and Personal Relationships, 23(4), 629-653.
- Markus, H. R., & Kitayama, S. (1991). *Culture and the self: Implications for cognition, emotion, and motivation*. Psychological Review, 98(2), 224-253.
- Oyserman, D., & Lee, S. W. S. (2007). *Priming "Culture"*. Handbook of cultural psychology.
- Roccas, S., & McCauley, C. (2004). *Values and emotions in the relational models*. Relational models theory: A contemporary overview, 263-285.
- Rohrer, R. P. (1984). Toward a conception of culture for cross-cultural psychology. *Journal of Cross-Cultural Psychology*, 15, 111-138.
- Ross, N. (2003). *Culture and Cognition: Implications for Theory and Method*, SAGE Publications.
- Schwartz, S. H. (1992). *Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries*. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1-65). New York, NY: Psychology Press.
- Schwartz, S. H., & Bardi, A. (2001). *Value hierarchies across cultures: Taking a similarities perspective*. Journal of Cross-Cultural Psychology, 32, 268-290.
- Shiraev, E. B., & Levy, D. A. (2004). *Cross-cultural psychology: Critical thinking and contemporary applications*. Boston: Allyn & Bacon.
- Triandis, H. C. (1989). *The self and social behavior in differing cultural contexts*. Psychological Review, 96, 506-520.
- Triandis, H. C. (1996). *The psychological measurement of cultural syndromes*. American Psychologist, 51(4), 407.
- Triandis, H. C., & Gelfand, M. (1998). *Converging measurement of horizontal and vertical individualism and collectivism*. Journal of Personality and Social Psychology, 74(1), 118-128.
- Woodhull, M. (2006). *Comparing relational models across cultures: Discovering how maquiladora and parent company managers construct their social relationships at work*. (Doctoral dissertation). Retrieved from ProQuest Information and Learning Company. (UMI Number 3233975)
- Woodhull, M. D., & Louis, D. J. (2009). *Why did they do THAT??? A cross-cultural study of daily decision-making by Mexican maquiladoras and US/Canadian managers*. Business Journal of Hispanic Research, 3(1), 77-93.
- Yates, J. F., Lee, J. W., Sieck, W. R., Choi, I., & Price, P. C. (2002). *Probability judgment across cultures*. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 271-291). New York: Cambridge University Press.

Reading Patterns of Life: Practical Tools from Ethnography

Tracy St. Benoit^a and Clarissa Graffeo^b

*^aUniversity of Central Florida
Orlando, FL*

*^bMESH Solutions, LLC
Orlando, FL*

ABSTRACT

Cross-cultural liaisons and advisors face numerous challenges understanding and adjusting to different cultures and working cross-culturally to accomplish capacity building goals. While pre-assignment culture-specific training is important, it cannot adequately prepare individuals for the full range of relevant information, requiring personnel to build and refine socio-cultural knowledge “on the ground.” Ethnography offers best practices for performing such cultural sensemaking in the field. Further complicating the issue, however, is the understanding that “culture” is not a unified whole, but rather a set of complex interacting systems; globalization, disconnections between politically and ethnically defined boundaries in host nations, and the interaction of cultural and practical or situational contexts mean that learning “culture” may not be sufficient. To provide sensemaking skills and interpretive tools that are actionable across cultural boundaries and accommodate social complexity we propose a Patterns of Life (PoL) framework that incorporates observable practices, social structures, interactions, and relationships not only among humans but also including the environment, objects, and non-human actors. We then outline our Ethnographically-informed Sensemaking Protocol (ESP), an iterative, reflective process drawing from ethnographic methods (e.g., participant observation and ethnographic interviewing protocols) to develop a holistic understanding of the human domain. We finally present an illustrative use case.

Keywords: Culture, Sensemaking, Social Complexity, Cross-Cultural Competency, Ethnography

INTRODUCTION

Ongoing trends toward globalization in a variety of industries, as well as increased U. S. government involvement in providing capacity-building and disaster relief assistance to partner nations has resulted in the deployment of individuals from a variety of fields and disciplines as subject matter experts, liaisons, and advisors overseas. These liaisons and advisors are selected on the basis of their technical or field expertise, but may have little or no previous cross-cultural experience, or knowledge of the particular cultural environment in which they are expected to operate. Additionally, in many cases do not know the language of the host-nation and must work through interpreters, thus complicating processes of adjustment and interaction, and placing additional emphasis on the need to interpret body language and other non-verbal forms of expression or communication (Zbylut, et al., 2008). Given the critical role of language and interpretation, it is essential for advisors to have as many tools as possible at his or her disposal in order to fully grasp the meaning of cross-cultural communications, meaning observational techniques and recommendations as well as key points for consideration when speaking with others. The short terms of some of these assignments create an additional need for skills to facilitate quick yet detailed assessment of local needs and

social systems.

In this paper, we leverage ethnographic expertise and provide sensemaking tools for reading the human domain that will be applicable across a range of complex environments in order to augment the capabilities of cross-cultural liaisons and advisors. We first outline the Patterns of Life (PoL) framework that extends beyond culture to incorporate a fuller sense of social complexity, defining particular environmental, behavioral, and contextual elements that facilitate accurately reading the human domain. We then present our Ethnographically-informed Sensemaking Protocol (ESP), a nested, iterative observational protocol drawing from inductive ethnographic and sociological methodologies. Finally, we provide a use case demonstrating the limitations of applying cursory or generalized cultural orientations and sensibilities to Key Leader Engagements, and illustrating how the application of PoL and ESP provide improved capability.

LIMITATIONS OF CROSS-CULTURAL TRAINING AND THE NEED FOR HUMAN DOMAIN SENSEMAKING

A long history of research and training development across multiple domains (e.g., management, health care, and military operations) has been applied to define the knowledge and skills needed for cross-cultural interaction. Cross-cultural training programs and recommendations frequently include region-specific or culture-specific training; meta-knowledge of particular cultures has indeed proved useful in successful cross-cultural interaction, reinforcing use of this training component (e.g., Byram, Nichols, and Stevens, 2001; Leung, Lee, and Chiu, 2013). Much of this instruction is necessarily brief, however, and may operate on or lead to broad generalizations; in some cases, it may even consist of inaccurate information or erroneously apply Western models to foreign cultural groups. Such incorrect meta-knowledge of cultures can result in negative performance during cross-cultural interaction (Leung et al., 2013). Moreover, these short-term, overly simplified cultural training programs may inadvertently *worsen* learner stereotypes and give personnel a false impression of cultural knowledge and competence (Gertsen, 1990; Abbe and Halpin, 2010). Even where the knowledge gained is correct and appropriate, however, trainees may still face difficulty understanding the practical significance of these cultural concepts and translating that generalized information into applied performance.

Efforts to engender transferable cross-cultural competencies and augment insufficient or unavailable culture-specific training led to the development of “good stranger” culture-general training and the identification of aptitudes and skills related to a concept of overarching cultural intelligence (see Bhawuk and Brislin, 2000; Earley and Ang, 2003). The emphasis on empathy, openness, and perspective-taking emphasized in much of this training is useful, but ultimately both culture-specific and culture-general training appear to leave gaps in relevant cross-cultural skills, and evidence of effectiveness has been inconsistent (e.g., Littrell et al., 2006; Puck, Kittler, and Wright, 2008; Abbe and Halpin, 2010). Furthermore, like culture-specific training, culture-general training may rely on broad generalizations or frameworks like Hofstede’s cultural dimensions, which have come under some criticism for their accuracy and validity (e.g. Erez and Earley, 1993; McSweeney, 2002) and may lead trainees to develop a conception of static national culture that does not adequately prepare them for interactions with real individuals.

Ultimately the concentrated nature of much cross-cultural training, whether region-specific or general, cannot fully prepare an advisor or liaison for the demands of dynamic, ongoing interactions in a complex foreign environment. While it may indeed provide a useful foundation to ease culture shock and facilitate adjustment, as well as a starting point, cross-cultural workers need to augment and revise their knowledge on the ground during the course of their job. Recent research on cross-cultural competency has begun to incorporate cultural learning and sensemaking as a necessary skill to fill this gap (e.g., Rasmussen, Sieck, Crandall, Simpkins, and Smith, 2011; Sieck, Smith, and Rasmussen, 2013; Osland and Bird, 2000; Sands, 2012; Queensland Health, 2010; Byram, et al., 2001). Currently, however, existing training programs and recommendations have yet to provide specific instruction on how exactly trainees should perform such sensemaking.

We propose that ethnographic methods present just such skills and processes for robust human domain sensemaking. Weick (1995) explains sensemaking as a process through which individuals develop plausible explanations based on their knowledge and perceptions, and that it is an enacted process (i.e., the sensemaker’s presence and activities affect the state of the environment and the progression of the situation). Ethnography, as “a scientific approach to discovering and investigating social and cultural patterns and meanings in communities, institutions, and other social

settings” depends on the researcher as instrument and sensemaker to observe and interpret the complex events and meanings in a cultural group (Schensul, Schensul, and LeCompte, 1999, p. 1). Like sensemaking, ethnography is framed and shaped by pre-existing information and experiences. Furthermore, a key methodology within ethnography is participant observation, in which the researcher becomes an active participant in the community and shares in the experiences, as well as being involved on a personal level, which echoes notions of enacted sensemaking (see Bernard, 2005; Murchison, 2010; Schensul et al., 1999; and DeWalt and DeWalt, 2011).

Some early research and commentary has begun to voice such an interest in ethnography for cross-cultural training (e.g., Sands, 2012; Knipper, 2013), but practical solutions and processes are still needed. Here, we provide what we hope is a useful first step toward developing this sort of training. Before we discuss the particulars of our observational protocol, however, we recommend a move away from the present emphasis on “culture” as a static, bounded object available for scrutiny or as a complete context for social attitudes and practices, and we present Patterns of Life (PoL) as an alternative.

FROM "CULTURE" TO PATTERNS OF LIFE

The Problem with Culture

The extant social science and cultural studies literature present a wide variety of definitions for culture. Within some fields of anthropology and sociology, culture is conceived as a primarily cognitive phenomenon, constituting “... an historically transmitted pattern of meanings embodied in symbols, a system of inherited conceptions expressed in symbolic forms by means of which men communicate, perpetuate, and develop their knowledge about and attitudes toward life” (Geertz, 1973, p. 89) or “the acquired knowledge that people use to interpret experience and generate social behavior” (Spradley, 1979, p. 5). In some cases, culture is positioned as metacognitive, providing the structure for how individuals think and perceive; in this sense, “culture does not provide a cognitive map, but rather a set of principles for map making and navigation” (Frake, 1977, p. 7). Cultural materialist frameworks and practice theory, on the other hand, turn more focus toward culture as rooted in, and responding to, material concerns of production and resource management, and emphasize the observation of visible behaviors and practices rather than reported values or cognitive processes (e.g., Harris, 1979; Reckwitz, 2002; Rouse, 2007; Schatzki, Cetina, and Von Savigny, 2000). Indeed, some scholars argue that the very effort of Westerners to define and lay out foreign “cultures” are, in and of themselves, problematic impositions of power onto others (e.g., Mitchell, 1995; Said, 1978). There is, in short, little agreement on the precise content or boundaries that constitute a culture.

Cross-cultural instructional researchers and training developers are also beginning to gain awareness of issues already circulating among anthropologists and other social scientists, namely that culture (whatever its definition) is not a discrete object, nor “a unified corpus of symbols and meanings that can be definitively interpreted” but rather is “contested, temporal, and emergent” (Clifford and Marcus, 1986, p. 19). Groups of people identified as cultural communities consist of a number of subcultural groups, each with their own values and practices. Even at the individual level, people may simultaneously belong to multiple cultural groups (Handwerker, 2001). Furthermore, individuals are affected by aspects outside of culture, and driven by competing cultural and extra-cultural motivations. Culturally based motivations may be trumped by situational or practical concerns, meaning that advisors or liaisons cannot solely rely on culture to evaluate others’ behavior or to make decisions (Osland and Bird, 2000). Using culture as a framework for interpretation and decision making overseas is further complicated by the fact that areas of operation at regional and national levels are defined by politically drawn—frequently by outsiders with little knowledge of or consideration for the residents’ ethnic and cultural groups—geographical boundaries, and therefore consist of a variety of different cultural groups or “ethnoscapes” which continue to shift, change, and mix due to modern patterns of globalization and international media distribution (Appadurai, 1990). Social systems are complex, open, and adaptive, introducing change over time and in response to various conditions and pressures, as well as in response to contact and interaction with other systems or component sub-systems (Byrne and Callaghan, 2014). Therefore, we argue that continuing to emphasize culture as a central factor may not facilitate the sort of holistic mental models that must be applied to appropriately understand and act effectively within the human domain as a complex system. An alternative is required that will be more effective in an applied environment. Reframing the area of concern from “culture” to Patterns of Life (PoL) and highlighting several areas of interest, as well as providing some terms to facilitate communication and shared mental model building, may assist advisors and liaisons in developing this more complex, holistic view.

Defining Patterns of Life

To mitigate these issues with the emphasis on culture our research has moved to a more generalized theory of Patterns of Life (PoL). The term “patterns of life” was defined by anthropologist Ward Goodenough as “the regularly recurring activities and material and social arrangements which give a community the semblance of a homeostatic system” (Simpson, Gerard, Goodenough, and Inkeles, 1961, p. 521). Though it provides a useful starting point, we find this definition insufficient, as it does not fully reflect the degree to which these aspects interrelate with and influence one another. Its focus on society as “homeostatic” also fails to adequately address intra-cultural variation and divergence, as well as the influence of practical, economic, or other extra-cultural factors. The nature of social systems as “far from equilibric, open and...complex adaptive” signifies that the human domain is *not* homeostatic, and therefore generalizations and broad interpretations at the system level are problematic (Byrne and Callaghan, 2014, p.26). Observable patterns do still emerge, however, within localized groups as a result of the aggregate behavior and interactions of individual actors within multiple overlapping and competing contexts; sensemaking and interpretation of the human domain must use these patterns. As Paul Cilliers notes, the behavior of complex systems such as societies “is not determined primarily by the properties of individual components of the system, but is the result of complex patterns of interaction” (1998, p. 91). These interactions are driven by *local* information and action, within *local* contexts, and the degree of social variation and interacting sub-systems mean that rich sensemaking and interpretation must occur at this local level.

While operating on this local level, however, care must be taken not to lose a holistic orientation that can address robust interactions among components of the system. Advances in complexity theory reveal that traditional deconstructive methods of analysis are inadequate for complex systems, including social systems. Within a biological context, Garfinkel (1987) discusses this problem of complex systems, that “in classical reductionism, the behavior of holistic entities must ultimately be explained by reference to the nature of their constituents ... Although it may be true in some sense that systems ‘are just’ collections of their elements, it does not follow that we can *explain* the system’s behaviour by reference to its parts, together with a theory of their connections” (p. 202-203). Assessments of nations and large social groups frequently subdivide the system into different categories which are then treated as independent factors; we argue instead that aspects of the human domain (for instance geography, culture, economy, and political structure) cannot be considered in isolation but must be assessed as interrelated. For complex systems, then, an approach that can attend to local conditions and contexts while simultaneously implementing holistic models of research and interpretation is needed.

To facilitate this holistic orientation we propose shifting from a primary emphasis on culture to Patterns of Life (PoL), which incorporates multiple elements of the human domain and its context including, but not limited to, human and nonhuman actors, common movements, actions, use and exchanges of goods, housing situations, recognized kinship structures and relations therein, and practical necessities due to resource or economic pressures. PoL robustly encompasses and frames the human domain; it links and unifies the culturally rooted and practically necessary or contextually relevant aspects of life, and also considers the significance of, and human interactions with, objects and nonhuman actors which are frequently excluded within common mental models based heavily on culture.

Due to the variety of potential practices, situational contexts, and cultural groups that liaisons or advisors might encounter, we define PoL through an assortment of high level components and contextual lenses that will facilitate sensemaking, rather than a rigidly defined checklist of items. Though this does not exhaustively encompass or describe all elements of interest to observers, and all elements may not be relevant (or equally weighted in relevance) in any given situation, PoL directs observers to examine several key factors: actors present in the situation, the functions served by those actors, the centers of gravity in the environment or situation, and various forms of time: natural time, mythological time, socio-cultural seasons, and psychological seasons (see Table 1 below for more information on these components).

Table 1: Major components of PoL

PoL Component	Definition
Actors	Human or non-human that has the potential for, and may exert, agency

Functions	The interpretation of the actions and/or significance of an individual actor/object within, and produced by, a particular context and that can change over time; may overlap with roles, but are generally more discrete and contextually limited; functions provide meaning
Centers of gravity	Locations or individuals that exert force on the patterns of life in a given context, e.g., drawing people to them (physically or in terms of interaction/attention) or pushing them away
"Natural" time	Passage of time as defined by natural phenomena, e.g. day/night cycles and the passage of seasons
Mythological time	Religious and sacred conceptions of time; for instance whether time is linear or circular, as well as narratives and understandings of sacred origins and eras, often manifesting in a "nostalgia for origins" or narratives of divinely relevant endings of time (Eliade, 1967, p. 44; see also Eliade, 1963/1998; and Eliade, 1954/1971)
Socio-cultural seasons	Temporally limited or defined patterns of life specifically related to the local culture(s), such as religious holidays, agricultural practices, or nomadic migration periods; the cadence of a culture
Psychological seasons	Emotional and cognitive patterns occurring in discrete time frames and phases, whether linked to natural time (e.g., Seasonal Affective Disorder), cultural practices and seasons, or situational timelines (e.g., common emotional responses to disasters)

Observed elements of PoL do not have discrete meanings that transfer exactly from situation to situation, but must always be evaluated and interpreted in the local context. Relevant contexts may be cultural or situational, i.e., particular practices may be motivated by cultural values and traditions, or motivation and significance may stem from other requirements and frameworks, such as business contexts or economic factors. Due to the complexity of the human domain, certain component elements of PoL may also function as a framing context; for instance, a given socio-cultural season may bear its own practices and signifiers, but may also alter the way practices or functions that span outside of that time frame should be interpreted while in it. See Figure 1 below for an illustration of these overlapping component and contextual elements.

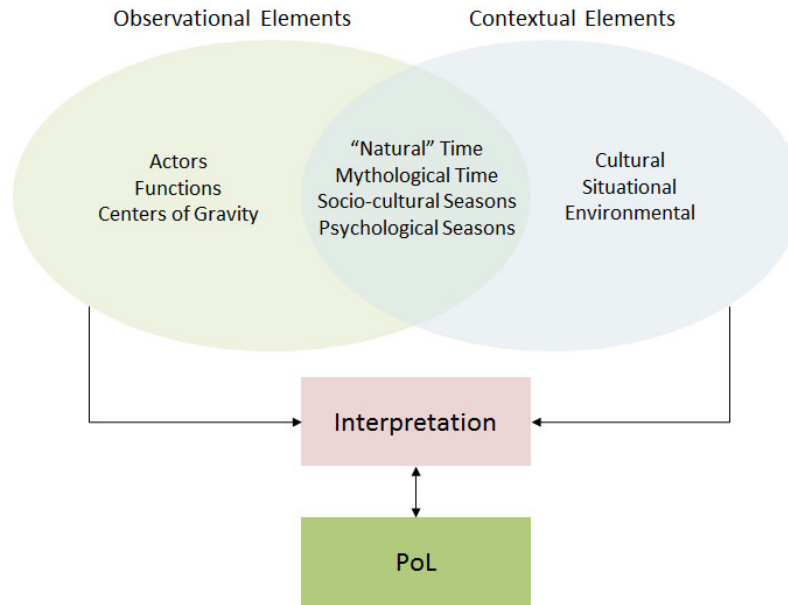


Figure 1. Components and basic structure of PoL

The observation and interpretation of PoL is conducted via an iterative, reflexive, and reflective sensemaking process which we outline in the following section.

OUTLINING AN ETHNOGRAPHICALLY-INFORMED SENSEMAKING PROTOCOL (ESP)

In order to construct an understanding and interpretation of the Patterns of Life within a given community or organization, individuals must engage in sensemaking. To provide actionable direction on how this sensemaking should occur, we define an Ethnographically-informed Sensemaking Protocol (ESP). As complex systems behavior emerges from, and is dependent on, local information and contexts, this process should be conducted at the local level; in other words, to avoid the erroneous application of generalizations and ensure sensitivity to particular situational and practical impacts, as well as to accommodate the unique combination of ethnoscares and cultural interactions within an environment, advisors and liaisons must augment broad cultural information and culture-general skills with a robust understanding of local PoL. It is these communities that aggregate up to develop large-scale emergent behaviors. Here again we emphasize the constant tension of complex systems between local and system-wide, between individual actors and larger Patterns of Life. Sensemakers can only operate locally, but they should maintain an awareness of the larger system and the framing contexts it and other systems may impose. Ethnography works within this structure, focusing on a deep understanding of the local environment but with an eye directed toward larger systems. As Schensul et al. (1999) explain, “all ethnography is local...[it] builds local theory-theories that explain events, beliefs, and behavior in the special site an ethnographer is studying” (p. 7); these understandings feed up into, and exist within, “a broader socioeconomic and political context” that is “essential in order to situate local experience and cultural observations” (p. 5).

Liaisons or advisors who need to perform cultural sensemaking on the ground must operate within limited time and space boundaries. Traditional models of ethnography, in which a researcher would live within a community for several years, are not feasible. Therefore, ESP has drawn from Rapid Assessment Processes and quick applied ethnographic methods that present a short-term, intensive, iterative process (see Beebe, 2001; Handwerker, 2001; Murchison, 2010). Operating at the local level and observing local actors (i.e., in a particular village or organization) with recourse to region or nation-wide information as a larger context facilitates this tight focus and quick implementation. We also focus on tools and best practices from rapid, applied ethnography such as embedded participant observation, because they enhances and refines the liaison/advisor’s deep knowledge and therefore

facilitate better recommendations and decision making for locally effective and sustainable capacity building.

ESP works on two levels: a long-term, high-level process and a per-encounter process. At the top level, individuals conduct an ongoing, iterative cycle of integration, interpretation, and mental model production/refinement that may continue over extended periods of time, for instance over the course of a liaison or advisor's entire term in the host nation. Sensemakers implement a cognitive cycle (see Figure 2 below) in which they identify relevant actors and/or agencies as sources for information; take their experiences and interpretations from a particular encounter, and share and collectively interpret the knowledge gained with partners and stakeholders as needed; update and refine their mental models based on new input (whether gained directly through the encounter or indirectly through others); identify knowledge gaps to ascertain where particular inquiry or further observation may be needed; and re-engage with the population of interest. Sensemakers should also regularly engage in validating and triangulating their knowledge by various means, for instance gathering additional information using ethnographic interview protocols (open-ended, semi-structured) to check their own observations and interpretations, or to verify data gained from other individuals and/or agencies.



Figure 2. The iterative top-level ESP cycle

We focus here on this overarching iterative process. ESP does feature a secondary level, per-encounter integrated process of observation and interpretation that facilitates establishing local baselines. For reasons of relevance to our case study and length limitations of the current paper, we will provide further detail on this second level in future publications.

CASE STUDY: INEFFECTIVE KEY LEADER ENGAGEMENTS

This case study, drawn from the primary author's work as an embedded social scientist with Regional Command East Afghanistan Provincial Reconstruction Team (PRT) efforts (see Rohde, 2007), illustrates the shortcomings of using the broad application of a fixed concept of culture when identifying stakeholders and conducting Key Leader Engagements. It also demonstrates the negative impact of ignoring the characteristics and contexts of the local community, as well as the need for a holistic reading of Patterns of Life within that local community context. The events of this case study occurred prior to our work on ESP, but involved use of a precursor to the current protocol

which incorporates many similar components and rests on the same fundamental theoretical and procedural basis. We have reframed the discussion here to fit with POL and ESP-relevant terminology, but have not altered the events, observations, or interpretations.

As part of the Coalition's strategic policy of winning Afghan "hearts and minds" (Department of the Army, 2009) the PRT had been negotiating terms with local tribal elders to provide a mobile medical clinic, in the hopes of gaining village support for local development and aid efforts. The village was part of a spike in insurgent activity, and the PRT hoped that providing essential services and winning local support would reduce the violence. PRT personnel received regional cultural training on identifying traditional tribal leaders based on cultural markers (e.g., clothing and age) with whom to negotiate and gain approval for aid and development, which they followed. When villagers directed the PRT personnel to key leaders, because these individuals fit the cultural profile provided, they assumed that these were in fact the appropriate, empowered stakeholders. These tribal elders arranged for the clinic to occur at an unfinished school construction site that they stated was central and easily accessible by foot, and that the PRT should prepare to treat men, women, and children. The primary author was assigned to accompany PRT personnel to the site in order to assess the effectiveness of the aid efforts in garnering local support and goodwill.

Prior to implementation of the clinic, the primary author performed an initial iteration of the ESP process and interviewed U.S. personnel who had been involved in the negotiations, including individuals from Civil Affairs and USAID. As part of the first step, identifying relevant actors and agencies, in addition to the local population such relevant individuals included U.S. personnel who had established a relationship with the villagers and had knowledge of village dynamics. These interviews demonstrated that U.S. personnel had used the fixed cultural models provided by pre-deployment training to identify key leaders with whom to negotiate, and that the village had been assessed merely at a broad level as an Afghan agricultural village. No attention had been directed to the local details of the economy and workforce. The primary author also conducted research through secondary sources about the area where the village was located, including requests for further information on the current socio-political situation in the hopes of pre-identifying centers of gravity that could be further observed during the event. Individuals and organizations, such as the Haqqani network, that had a potential ideological, economic, or political impact on the village were also identified in the hopes of understanding how they might affect the planned medical engagement. All of these component POL elements, whether environmental, political, economic, or security-based, had to be considered not only in isolation, but in how they interacted to develop a complex local system.

This pre-assessment, and the integrated observations and interpretations (high level ESP stage 2) that resulted, helped the primary author to prepare general research questions and key focus areas for detailed POL observation on the ground during the clinic event. Based upon this early synthesized mental model, the primary author identified a short list of guiding questions to help the PRT understand the results of the planned effort. The primary questions were: 1) whether the medical engagement was appropriate for the population; 2) was there a way to truly measure the results on the target population's sentiments toward the U.S.; 3) if so, how could a successful engagement be sustained, or what other efforts could be done with the local population to sustain goodwill; and 4) if the clinic did not address the needs of the village, then why not?

Upon arriving at the site for the mobile medical clinic it became clear that the negotiations for aid had not gone according to PRT estimations. The crowd waiting at the site consisted entirely of men age 30 and older, who were well-dressed and appeared to be in good health; none of the expected women and children were present. Two cars arrived shortly, and the crowd quickly separated to let them discharge the occupants at the entrance to the school. A young Afghan man, approximately 23-25 years old and wearing a spotless salwar kameez, jumped from the lead vehicle to open the door for a frail elderly gentleman in the second vehicle. Neither of these men had been present in the earlier negotiations; the PRT team had no knowledge of who these men were or their function. The elderly man leaned heavily on the younger man's arm and the pair slowly made their way to the front of the crowd. He was clearly a major center of gravity as the crowd, without any verbalized direction, approached the elderly man and in turn each embraced him and kissed his hand, then nodded to the younger man as they shuffled aside to let the next man offer greetings and respect. The procession of healthy, middle aged men and the ritualistic show of respect to the elderly Afghan troubled the observing PRT medical team. Despite the chief medical officer's request to one of the tribal elders to commence the clinic, he was instructed to wait until all those present had a chance to offer their respects; repeated requests to begin were rebuffed. With obvious impatience and a concern for security, the PRT medical team waited until all the men present had a chance to greet the elderly Afghan.

Once the last man paid his respects, the younger Afghan escorted the elderly man to the car, shut the door and stood

aside as the two vehicles sped away from the site. The PRT medical team without hesitation quickly instructed the crowd to line up to receive services. The young Afghan man then made an astonishing transformation, moving from his earlier function as the elderly man's quiet companion to the function of a very vocal man in charge, and the clear center of gravity. He firmly directed a cadre of men to organize the crowd and then went about ensuring all was in order. Two hours into the medical clinic it was readily apparent to PRT staff this was not a normal humanitarian project; the majority of men seen by the medical staff had no health complaints and instead asked for Viagra.

It was clear that the two unidentified men were of great importance from behavioral patterns; in addition, the tribal elders previously identified by PRT personnel were initially absent from the clinic, and upon arriving later showed deference to the two men. The position and function of these two men represented a knowledge gap in the previous research. Furthermore, the makeup of the crowd, their behavior, and the nature of the medical requests did not match expectations for a normal clinic operation. The mental model for the village and the aid situation was clearly incorrect, prompting the lead author to begin a second iteration of the ESP cycle, seeking new relevant actors to inform a revised mental model. Sources of information on these two men were needed, but had to be selected carefully to avoid misinformation or causing a conflict.

Observing PoL at the site, individuals needed to be identified that may have had a relationship with, and knowledge of, the two men but were accessible and less likely to provide misinformation or be resistant to speaking with the researcher. At the time of the event, children were perceived as neutral participants and could generally be trusted to provide honest information; moreover, unlike adult men or women, they were frequently easier to speak with because complicated procedures for introductions and security were not necessary. Though no children were present within the group of patients, there were some playing or observing near the site, all of whom were boys. In the area in which the village was located, it was normal to see children of both genders together in public, so the absence of girls was odd. It was discovered that females of all ages had been instructed to stay at home.

The primary author approached one of the children and asked casual questions about the village and his normal life (e.g. does he attend school and whether he is playing with school friends, and whether they are usually at this location during this time of day). Since sustained observation of general PoL within the village was not possible in the given time frame, this information had to be elicited from locals. The child mentioned that the women would normally be out in the fields working, but that his mother had been told to prepare dinner for unfamiliar guests. The primary author also asked about the two men, and the child responded that they were members of an important family. In deference to the children's security, more detailed questions were not pursued. Other children were asked similar general questions in order to triangulate the initial response and hopefully establish an understanding of normal village PoL. Responses from other children nearby matched those of the initial child. As an extension, to triangulate and validate the emerging information, questions were also asked of some of the accessible adult male attendees in order to develop a more holistic view and compare their knowledge to that of the PRT. In total, over 30 people were interviewed on site. These interviews continued to identify knowledge gaps not only for the researcher but for the overall PRT, and repeatedly triangulate and validate gathered information.

The researcher reported these initial findings to the commanding officer. Agreeing that something was amiss, the PRT commander requested an interview with the young Afghan man, who agreed. The primary author conducted the interview with the aid of an interpreter, and quickly discovered that the young man was the actual power broker in the village rather than the group of tribal elders. He and his uncle had used the tribal elders with whom the PRT negotiated as proxies to set up the clinic. The young man revealed that he was responsible for the construction of the site at which the medical clinic was being held. He and his family held some traditional cultural authority, but he had augmented this with additional economic authority in response to the circumstances in the village. The war and insurgency had resulted in the loss of many of the community's men, hampering the ability of the small agricultural community to maintain their output, and creating an additional financial and resource burden on the remaining men due to their duty to support the population of widows. Additional stress was placed on traditional agrarian production due to instability of the area as a result of the war, as well as Haqqani network activity. The upsurge in violence was a result of the local men taking jobs to place IEDs or otherwise support the Haqqani network's criminal activities simply in order to supplement their income and provide for their families.



Figure 3. Interviewing the young Afghan man through an interpreter (faces have been blurred to protect the identity of the participants); his manner of sitting (see in contrast to the interpreter), along with other elements of his body language and behavior, initially suggested that he may have had significant exposure to Western culture.

The young Afghan man had studied in America, earned a degree in business, and returned to establish a contracting business in Kabul to win construction contracts from the U.S. The economic instability of the village allowed him to step in and obtain additional power beyond what would culturally be expected due to his ability to provide paying employment, while his uncle served as a figurehead to make the situation more palatable within the cultural context. The young man's power, however, depended on the ability to earn contracts. Thus, while the PRT had negotiated with the apparent power brokers, this young man had arranged behind the scenes for the clinic to serve not as a local service, but instead as a display of the young man and his clan's influence with the PRT. The medical clinic was a symbolic demonstration to the village and the traditional tribal hierarchy that his clan could bring money and development to the village.

The PRT's previous mental model of the village and its context, based on a constrained understanding that emphasized culture without taking into account the impact of practical economic and political factors, treated the violence as ideologically motivated. The PRT did not grasp the dynamic relationship between the village's traditional culture and present economic and population stressors, which facilitated the Haqqani network influence and shifted authority within the village. Thus we see that a holistic understanding of PoL, which incorporated not just culture but also these situational exigencies, developed at the local level for this particular village was necessary. In this case, the generalized strategy for "winning hearts and minds" was flawed because it did not meet this population's local needs. A medical clinic, while potentially useful, did not ultimately address the present issue in the village, which was employment for the male population. As these sorts of local, tactical misallocation of efforts occur and accrue, they can result in a large scale strategic failure.

Having identified and filled in the relevant knowledge gaps, and updated the PRT's organizational mental model, the PRT commander was then able to renegotiate with the actual power broker and establish a job development program with the young man's contracting company to address immediately employment needs. During the course of the PRT's term in the area, there was a decrease in violence, which may have been a result of this job program stabilizing the economic situation. This decrease was a part of a 60-70% reduction in kinetic operations within the region as a result of the initial integration of social scientists into PRT operations (*Role of the Social and Behavioral Sciences*, 2008). Following this incident, the PRT was trained by the lead author in this rapid assessment technique. Later reports from this PRT showed that they were better able to conduct negotiations and partner with local populations in developing appropriate aid projects.

CONCLUSIONS

Advisors and liaisons involved in capacity building and stability operations in post-conflict or post-disaster environments face a large number of interconnecting cultural and situational challenges within a complex and evolving human domain. Pre-assignment cross-cultural training, whether culturally or regionally specific or generalized, cannot prepare advisors and liaisons for the full range of situations and people they will need to handle; they must be empowered to quickly but deeply assess the environment, the local status and functioning of the human domain, and identify appropriate stakeholders or influence points. Expanding from an overly rigid and too-broad idea of “culture” to a holistic model of complex Patterns of Life, and performing sensemaking through a methodology like our Ethnographically-informed Sensemaking Protocol can facilitate this flexibility for grasping and operating within complex systems. An understanding of the interrelationship between local factors and overarching contexts will also assist personnel in operating with an orientation toward both tactical and strategic levels. In the future, we hope to expand the current work on PoL and ESP to detail more visual observation-based processes, and work toward the construction of a full program of instruction that could be implemented across operational and subject-matter domains.

ACKNOWLEDGMENTS

This work is supported in part by the Office of Naval Research. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Department of Defense or Office of Naval Research. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon. This work was supported, in part, by the Office of Naval Research project N00014-11-C-0193, Perceptual Training Systems and Tools (PercePTS).

REFERENCES

- Abbe, A., & Halpin, S. M. (2010). The cultural imperative for professional military education and leader development.
- Appadurai, A. (1990). Disjuncture and difference in the global cultural economy. *Public Culture*, 2(2), 1–24. doi:10.1215/08992363-2-2-1
- Bhawuk, D., & Brislin, R. (2000). Cross-cultural training: A review. *Applied Psychology*, 49(1), 162–191. doi:10.1111/1464-0597.00009
- Byram, M., Nichols, A., & Stevens, D. (2001). *Developing intercultural competence in practice*. Multilingual Matters.
- Byrne, D. S., & Callaghan, G. (Eds.). (2014). *Complexity theory and the social sciences: the state of the art*. London: Routledge.
- Cilliers, P. (1998). Complexity and postmodernism: Understanding complex systems. New York: Routledge.
- Clifford, J., & Marcus, G. E. (Eds.). (1986). *Writing culture: the poetics and politics of ethnography*. Berkeley, CA: University of California Press.
- Department of the Army. (2009). *Tactics in counterinsurgency* (Field Manual No. 3-24.2). Washington, DC: U.S. Government Printing Office. Retrieved from http://armypubs.army.mil/doctrine/DR_pubs/dr_a/pdf/fm3_24x2.pdf
- Earley, P. C., & Ang, S. (2003). *Cultural intelligence: individual interactions across cultures*. Stanford, Calif.: Stanford University Press.
- Eliade, M. (1967). Myths, Dreams, and Mysteries: The Encounter Between Contemporary Faiths and Archaic Realities. (P. Mairet, Trans.). New York: Harper & Row.
- Eliade, M. (1971). *The myth of the eternal return, or, Cosmos and history*. (W. R. Trask, Trans.). Princeton, N.J.: Princeton University Press. (Original work published 1954)
- Eliade, M. (1998). *Myth and Reality*. Prospect Heights, IL: Waveland Press. (Original work published 1963)
- Erez, M., & Earley, P. C. (1993). *Culture, self-identity, and work*. N.Y.: Oxford University Press.
- Garfinkel, A. (1987). The Slime Mold Dictyostelium as a Model of Self-Organization in Social Systems. In F. E. Yates, A. Garfinkel, D. O. Walter, & G. B. Yates (Eds.), *Self-Organizing Systems* (pp. 181–213). Springer US. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4613-0883-6_11
- Gertsen, M. C. (1990). Intercultural competence and expatriates. *International Journal of Human Resource Management*, 1(3), 341–362.
- Knipper, M. (2013). Joining ethnography and history in cultural competence training. *Culture, Medicine, and Psychiatry*, 37(2), 373–384. doi:10.1007/s11013-013-9315-1
- Leung, A. K. -y., Lee, S. -l., & Chiu, C. -y. (2013). Meta-knowledge of culture promotes cultural competence. *Journal of Cross-*

- Cultural Psychology*, 44(6), 992–1006. doi:10.1177/0022022113493137
- Littrell, L. N., Salas, E., Hess, K. P., Paley, M., & Riedel, S. (2006). Expatriate Preparation: A Critical Analysis of 25 Years of Cross-Cultural Training Research. *Human Resource Development Review*, 5(3), 355–388. doi:10.1177/1534484306290106
- McSweeney, B. (2002). Hofstede's Model of National Cultural Differences and their consequences: A triumph of faith - a failure of analysis. *Human Relations*, 55(1), 89–118. doi:10.1177/0018726702551004
- Mitchell, D. (1995). There's no such thing as culture: Towards a reconceptualization of the idea of culture in geography. *Transactions of the Institute of British Geographers*, 20(1), 102. doi:10.2307/622727
- Osland, J. S., & Bird, A. (2000). Beyond sophisticated stereotyping: Cultural sensemaking in context. *The Academy of Management Executive*, 14(1), 65–77. doi:10.5465/AME.2000.2909840
- Puck, J. F., Kittler, M. G., & Wright, C. (2008). Does it really work? Re-assessing the impact of pre-departure cross-cultural training on expatriate adjustment. *International Journal of Human Resource Management*, 19(12), 2182–2197. doi:10.1080/09585190802479413
- Queensland Health. (2010). Five cross cultural capabilities for non-clinical staff. Division of the Chief Health Officer, Queensland Health.
- Rasmussen, L. J., Sieck, W. R., Crandall, B. W., Simpkins, B. G., & Smith, J. L. (2011). *Data collection and analysis for a cross-cultural competence model* (Final Report). Applied Research Associates, Inc. Retrieved from http://www.defenseculture.org/research/emergeresearchfiles/decisionsupportandmodeling/data_collection_analysis.pdf
- Role of the Social and Behavioral Sciences in National Security: Hearings before the Subcommittee on Terrorism & Unconventional Threats, of the House Armed Services Committee and the Subcommittee on Research and Education, of the House Science & Technology Committee*, 110th Cong. (2008) (testimony of COL Martin P. Schweitzer). Retrieved from <https://science.house.gov/hearing/subcommittee-research-and-science-education-hearing-role-social-and-behavioral-sciences>
- Said, E. W. (1978). *Orientalism*. New York: Pantheon Books.
- Sands, R. G. (2012). Cultural relativism and the convergence of ethnography and 3C. *Military Intelligence Professional Bulletin*, 38(1), 13–20.
- Sieck, W. R., Smith, J. L., & Rasmussen, L. J. (2013). Metacognitive strategies for making sense of cross-cultural encounters. *Journal of Cross-Cultural Psychology*, 44(6), 1007–1023. doi:10.1177/0022022113492890
- Simpson, G. G., Gerard, R. W., Goodenough, W. H., & Inkeles, A. (1961). Comments on Cultural Evolution. *Daedalus*, 90(3), 514–533.
- Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks: Sage Publications.
- Zbylut, M. R., Metcalf, K. A., McGowan, B. D., Beemer, M., Vowels, C. L., & Brunner, J. M. (2008). An analysis of cross-cultural behaviors for military advisors in the middle east. Proceedings of the 26th Army Science Conference, Orlando, FL, December 1-4, 2008

Basic Study on Prevention of Human Error - Debiasing Method of Cognitive Biases in Decision Making -

Atsuo MURATA, Tomoko NAKAMURA, and Saki KUBO

*Dept. of Intelligent Mechanical Systems
Graduate School of Natural Science and Technology, Okayama University
Okayama, Japan*

ABSTRACT

Cognitive biases potentially and unexpectedly induce crucial disasters such as Three Mile Island disaster and Challenger space shuttle disaster. This study explored how cognitive biases can be eliminated by paying attention to the characteristics or properties of each bias. The following cognitive biases were used to discuss the effectiveness of debiasing method of cognitive biases: ignorance of base rate, regression to mean, conjunction fallacy, framing effect, illusion of covariation, and overconfidence. In other words, the effectiveness of debiasing methods of these biases in decision making was experimentally discussed. The debiasing methods presented in this study were effective to some extent for suppressing the biases (conjunction fallacy, ignorance of base rate, and regression to mean) to some extent. However, for some cognitive biases (framing effect, fallacy of covariation, and overconfidence), the debiasing methods in this study were not necessarily effective. Some implications for the prevention of crucial human errors and accidents were given from the viewpoints of cognitive biases included in disastrous accidents.

Keywords: Fallacy of Conjunction, Ignorance of Base Rate, Regression to Mean, Framing Effect, Illusion of Covariance, Overconfidence

INTRODUCTION

Due to bounded rationality, we generally cannot make decision rationally, and thus suffer from cognitive biases (Kahneman, 2011, Tversky and Kahneman, 1974, Kahneman and Tversky, 1984, Altman, 2012, Angner, 2012 and Bazerman and Moore, 2001). Kahneman (2011) shows that our cognitive information processing is conducted by System1 or System2. While System2 requires us to conduct effortful, demanding and deliberate mental activities, System1 operates quickly, automatically, without time consuming, and intuitively with little or no efforts. Although heuristic approaches that we adopt when we have no time to deliberate are based on System1, and are very simple and intuitive, such approaches are suffering from cognitive biases.

According to Bazerman and Moore (2001), we summarized how cognitive biases are induced. It is hypothesized that the heuristics such as availability, representativeness, confirmation, or affect cause the biases such as confirmation biases, and anchoring and adjustment. In Figure 2, not only heuristics but also overconfidence and framing is shown as causes of biases. Moreover, it is assumed that our bounded awareness and uncertain (risk) situations forms the basis of heuristics, overconfidence, and framing. Due to such bounded ability, it is valid to judge that we cannot behave rationally but irrationally.

As mentioned above, we frequently tend to behave irrationally, and are in most cases unaware of how and to what

extent these irrational behaviors influence us. Such irrational properties are sure to distort our decisions, and in the worst cases lead to crucial accidents. Without consideration of our bounded rationality (irrationality), we cannot prevent crucial accidents from occurring and analyze the true (genuine) cause (source) of accidents.

Representativeness heuristic induces representative biases such as ignorance of base rate, regression to mean, and conjunction biases. Overconfidence also leads to biases such as illusion of control, fallacy of planning, and optimistic biases. Such biases lead to crucial accidents such as the Challenger space shuttle disaster (Reason, 1990 and Vaughan, 1997), Three Mile Island nuclear reactor meltdown (Lusted, 2012 and Osif et. al., 2006) or the collision accident (occurred in 2008) between the Japanese Aegis-equipped destroyer “Atago” belonging to Ministry of Defense and the fishing boat “Seitokumaru.” Therefore, the elimination of cognitive biases must be one of the effective and promising counter measures for preventing crucial accidents. It is also possible that representative biases distort the evaluation of risks.

This study explored how cognitive biases can be eliminated. The following cognitive biases were used to discuss the effectiveness of debiasing method of cognitive biases: ignorance of base rate, regression to mean, conjunction fallacy, framing effect, illusion of covariation, and overconfidence. In other words, the effectiveness of debiasing methods of these biases in decision making was experimentally discussed.

METHOD

Eighty participants took part in the experiment. The participants allocated to each experiment below were equally allocated to either biasing representation or debiasing representation. In other words, half were allocated to each condition. The answering time limitation differed among questions. On the basis of the answer for each type of bias, the performance data such as the percentage correct of each question was compared between biasing and debiasing conditions.

Conjunction fallacy

Linda is 31 years old, single, outspoken, and bright. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations. Which of the following two alternatives is more probable ?

Probability-based (biasing) representation:

- (a) Linda is a bank teller.
- (b) Linda is a bank teller and active in the feminist movement.

Frequency-based (debiasing) representation:

Imagine 200 females like Linda. Which of the following two alternatives is more frequent?

- (a) Linda is a bank teller.
- (b) Linda is a bank teller and active in the feminist movement.

The experimental factor (between-subject factor) was whether debiasing was tried or not. The percentage correct was compared between the biasing and the debiasing conditions.

Ignorance of base rate

Probability-based (biasing) representation:

A device has been invented for screening a population for some disease. Although the device is a very excellent one, it is not perfect. Therefore, if someone is not a sufferer, there is a 1% chance that he or she will be judged positive. Roughly 1% of the population has the disease. Mr.Suzuki was tested, and the result was positive. Answer the chance (probability) that he is a sufferer.

- (A) 85%
- (B) 50%
- (C) 20%

Frequency-based (debiasing) representation:

Roughly one out of 100 has the disease. A device has been invented for screening a population for some disease. Although the device is a very excellent one, it is not perfect. The device judges that one of 100 is a sufferer in spite of being healthy. How many of 100 are judged positive and how many of these persons is actually a sufferer?

- (A) One out of 100 is judged positive, and one is actually a sufferer.
- (B) Two out of 100 are judged positive, and one is actually a sufferer.
- (C) Five out of 100 are judged positive, and one is actually a sufferer.

The experimental factor (between-subject factor) was whether debiasing was tried or not. The percentage correct was compared between the biasing and the debiasing conditions.

Regression to mean

The participants were required to predict the winning average of one NPB (Nippon Professional Baseball Association) pitcher on the basis of data for past two years. The participant was also required to predict the batting average of one NPB player on the basis of data for past two years.

Biasing representation:

In this condition, as shown in Figure 1(a), the prediction of next year's winning average was conducted with no reference of the mean winning average for past two years. The participant was also ordered to predict next year's batting average of three players A-C on the basis of Table 1(a) with no reference of the mean winning average of three players A-C.

Debiasing representation:

In this condition, as shown in Figure 1(b), the prediction of next year's batting average in 2013 was conducted with a reference (horizontal line) of the mean batting average for past two years. The participant was also ordered to predict next year's batting average of three players A-C on the basis of Table 1(b) with a reference of the mean winning average of three players A-C.

The experimental factor (between-subject factor) was whether debiasing was tried or not. The mean difference between the predicted and the actual batting average was calculated and compared between the biasing and the debiasing conditions.

Framing effect

The following Asian disease problem (refer to Kahneman, 2011) was used. Imagine that US is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative measures have been proposed to combat the disease. Assume that the exact scientific estimates of the effectiveness of the measures are as follows when using the positive frame.

If measure A is adopted, 200 people will be saved.

If measure B is adopted, there is a one-third probability that 600 people will be saved.

According to Kahneman (2011), a substantial majority of respondents chose measure A, and preferred the certain option (measure A) to the gamble (measure B).

The negative frame can be expressed as follows.

If measure C is adopted, 400 people will die.

If measure D is adopted, there is a one-third probability that nobody will die, and a two-thirds probability that 600 people will die.

If we look closely and compare the positive and the negative versions, we easily notice that measures A and C are identical, and that measures B and D are identical. In spite of this, a large majority of people tended to prefer the gamble (measure D) to measure C. This corresponds to the reversal of preferences, and the framing induces such a cognitive bias.

Therefore, it was explored whether such a bias could be deleted or not using the following experimental procedure.

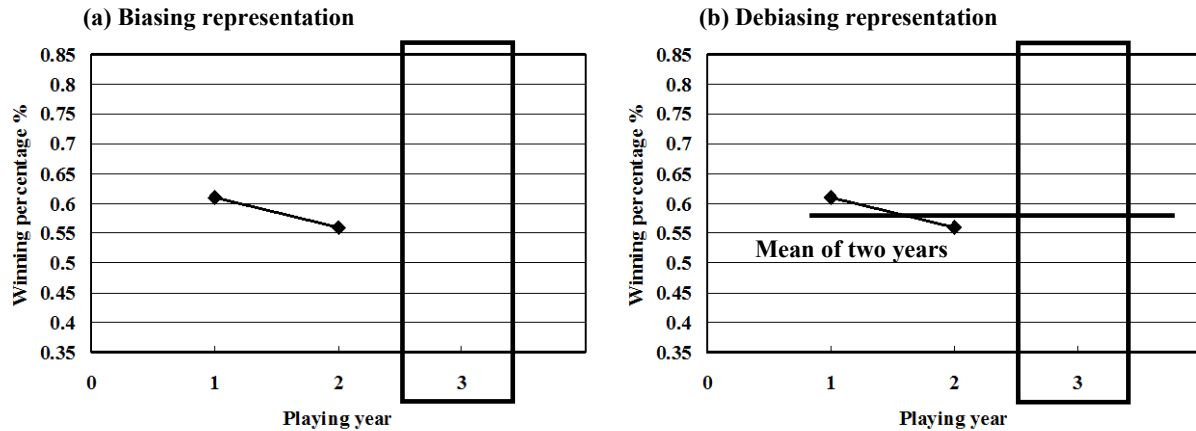


Figure 1. Graphic representation of winning percentage for the past two years. (a) Biasing representation, (b) Debiasing representation.

Table 1. Table representation of batting average for three players. (a) Biasing representation, (b) Debiasing representation.

(a) Biasing representation

Player	Batting average(this year)	Batting average(Next year)
A	0.34	
B	0.258	
C	0.239	

(b) Debiasing representation

Player	Batting average(this year)	Batting average(Next year)
A	0.34	
B	0.258	
C	0.239	
Mean	0.279	

Biasing condition:

The participant was provided with the positive and the negative frames above, and required to select an alternative A or B, or C or D. No other instructions were given to the participant.

Debiasing condition:

In this condition, the participant was required to select an alternative A or B, or C or D, and to state the reason why they selected measures A or B, or C or D. This will be called elaboration condition.

The experimental factor (between-subject factor) was whether we made an attempt to delete a cognitive bias or not.

Illusion of covariation

The participant was provided with the data in Table 2(a)-(b) and required to answer whether there existed causality between dizziness symptom and brain tumor, and between drinking habit and liver cancer. No causality exists in Table 2(a), while there exists causality in Table 2 (b).

Table 2. Cases (a) without causality and (b) with causality.

(a) No causality		
Dizziness symptom	Brain tumor	
	Yes	No
Yes	160	40
No	40	10

(b) Causality		
Drinking habit	Liver cancer	
	Yes	No
Yes	180	20
No	30	170

Biasing condition:

The participant was provided with no information necessary for scientifically verifying the causality.

Debiasing condition:

The participant was provided with information (hint) necessary for scientifically verifying the causality.

The experimental factor (between-subject factor) was whether we made an attempt to delete a cognitive bias or not.

Overconfidence

Scientific questionnaires which consist of 20 questions were prepared. Examples of scientific questions were as follows.

- (1) Which atomic weight is larger, Ag (silver) or Au (gold)?
- (2) Which condition is $(x-2)(x+3)=0$ for $x=2$, necessary condition or sufficient condition?
- (3) Which condition is inequality $x>1$ for $x>-2$, necessary condition or sufficient condition?
- (4) Which is larger, 2^5 or 3^3 ?
- (5) Which wave length is shorter, infrared rays or ultraviolet rays?
- (6) Which corresponds to the period of 500 Hz, 0.12s or 0.002s?
- (7) Which probability is higher when drawing a card from the deck of 13 cards of heart 1 to heart 13, (A) the probability of drawing multiples of 5, or (B) the probability of drawing two even numbers in a row?

The participants were required to answer one of the following questions after having finished answering all of 20 questions.

Biasing condition:

The participant was required to rate his or her confidence to the answer using 0-100%.

Debiasing condition:

The participant was required to predict how many correct answers out of 20 problems he or she reached.

The experimental factor (between-subject factor) was whether we made an attempt to delete a cognitive bias or not. It was discussed whether overconfidence can be deleted by asking the participant about frequencies of correct answer instead of asking him or her confidence of the answer.

Table 3. Number of participants trapped into conjunction fallacy compared between biasing and debiasing conditions (conjunction fallacy).

	Bias group	Debias group	Comparison of two population proportions
Conjunction fallacy	35	5	**($p < 0.01$)

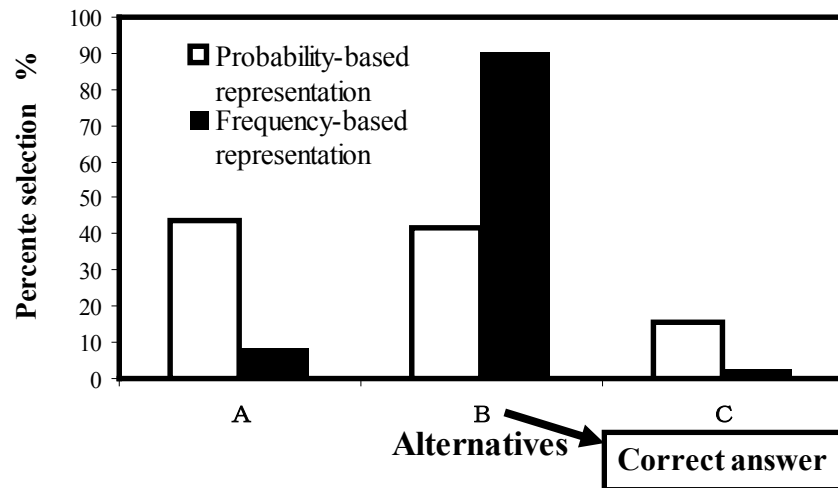


Figure 2. Percentage selection for each alternative (A-C) compared between probability-based and frequency-based representations (ignorance of base rate). In this case, the correct answer is B.

RESULTS

Conjunction fallacy

The comparisons of percentage correct between biased and the debiased versions is summarized in Table 3. From Table 3, it is clear that the debiasing method is effective for deriving correct answers and eliminating conjunction fallacy. The result means that the definite distinction between single-event probabilities and frequencies was important for the bias related to conjunction fallacy.

Ignorance of base rate

In Figure 2, the frequency-based representation produced more correct answers than the probability-based representation. As hypothesized, the frequency-based representation was effective for eliminating representativeness bias (ignorance of base rate).

Regression to mean

The difference between the predicted and the actual winning percentage (or batting average) was calculated for each participant. In Figure 3(a), the difference between the mean and the prediction is compared between biasing (without baseline as shown in Figure 1(a)) and debiasing (with baseline as shown in Figure 1(b)) conditions. In Figure 3(b), the difference between the mean and the prediction is compared between biasing (without mean value as shown in Table 1(a)) and debiasing (with mean value as shown in Table 1(b)) conditions. Comparing Figure 3(a) and Figure 3(b), it is clear that the table representation is more effective than the graphic representation. Adding the mean value to the table as a reference was effective for eliminating the bias related to the regression to mean.

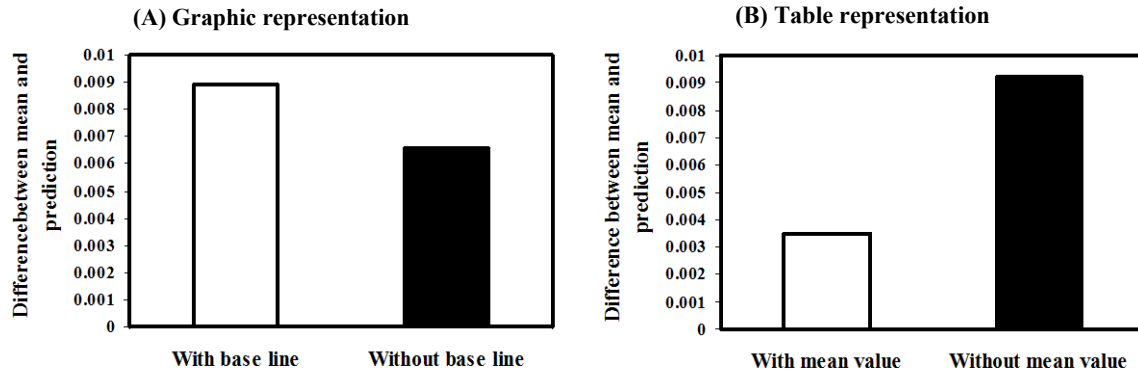


Figure 3. Difference between mean and prediction as a function of debiasing condition (with base line or without base line or mean value) and representation method ((A) graphic representation and (B) table representation) (regression to mean).

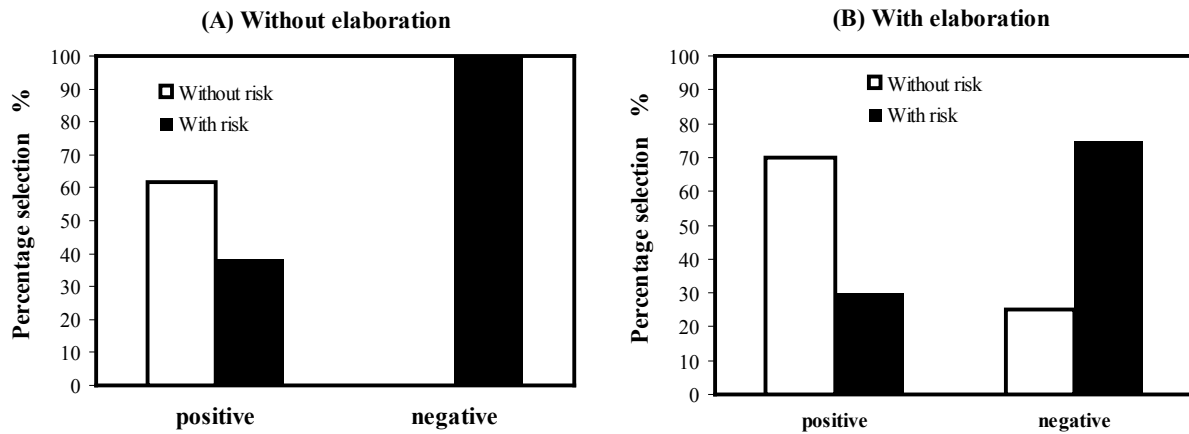


Figure 4. Percentage selection as a function of frame (positive and negative), risk condition (with risk and without risk), and debiasing condition (without elaboration and with elaboration) (Framing effect).

Framing effect

In Figure 4(a) and (b), the debiasing effect by elaboration is summarized. In the elaboration condition (debiasing condition), the participant was required to explain why the alternative was selected. The reversal of preference (Alternative A is selected for positive frame, while alternative D is selected for negative frame) was not removed by the elaboration instruction (explanation of why they selected the alternative).

Illusion of covariation

The percentage of correct and wrong answers for both biasing and debiasing conditions is shown in Figure 5(a) and (b). The debiasing measure (providing the participant with elaborated explanation for scientifically verifying the causality) did not work properly for the deletion of cognitive bias in capturing the causality. This shows that it is not so easy for us to recognize the causality correctly, and that we suffer from cognitive biases when understanding the relationship between the cause and effect.

Overconfidence

Figure 6 shows the relationship between the percentage correct and the prediction of number of correct answers (expressed as percentage correct answer) as a function of the way how the participant expressed his or her confidence (frequency representation: asking the participant about frequencies of correct answer out of 20 problems).

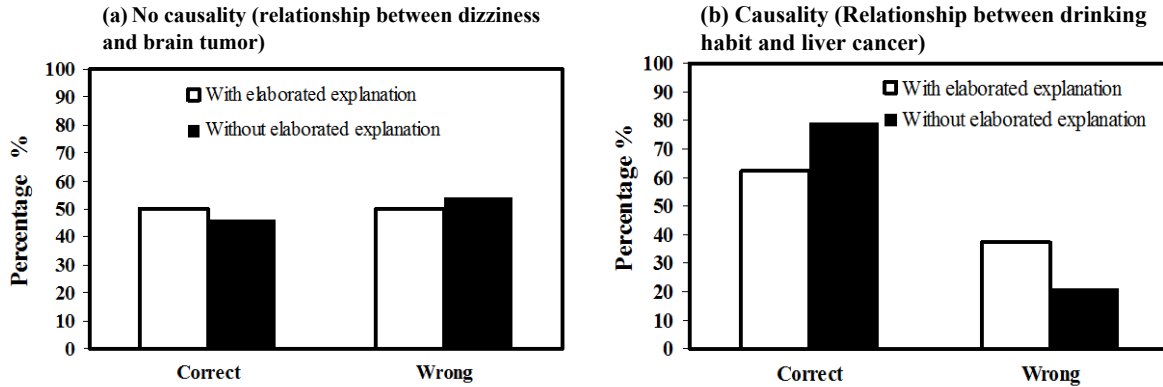


Figure 5. Percentage of correct and wrong answers compared between biasing and debiasing conditions (illusion of covariation). (a) without causality, (b) with causality.

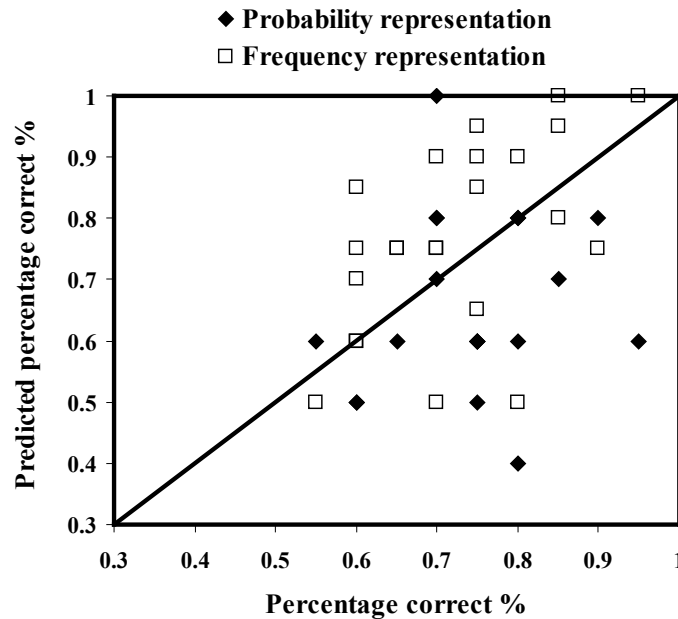


Figure 6. Relationship between percentage correct and predicted percentage correct (overconfidence). Data above the diagonal line represents overconfidence. Data below the diagonal line represents underconfidence.

or probability representation: asking him or her confidence of the answer). If overconfidence is dominant, the greater part of data exists above the diagonal line in Figure 6. The following tendencies were as a whole observed. Overconfidence was dominant for the frequency representation, while underconfidence was dominant for the probability representation. The result might show that overconfidence is deleted by probability representation. However, it must be noted that not correct evaluation but underconfidence appeared as a result of deleting overconfidence.

DISCUSSION

Conjunction fallacy

From Table 3, it was confirmed that the debiasing method by frequency-based representation of sentences is effective for deriving correct answers and eliminating cognitive biases. Frequency-based representation can prevent us from trapping into a celebrated reasoning error, that is, misjudgment that the conjunction of two facts is larger

than the conjunction of one fact.

The finding that the frequency-based expression contributes to delete conjunction fallacy might be effective for deleting cognitive bias especially when evaluating the risks for a variety of events.

Ignorance of base rate

As hypothesized and shown in Figure 2, the frequency-based representation was found to be effective for eliminating representativeness bias (ignorance of base rate). Generally, we are not so good at understanding and recognizing probabilistic phenomena, especially risk. In order to overcome this, the frequency-based representation and explanation of probabilistic phenomenon must be promising. The result means that the definite distinction between single-event probabilities and frequencies is important for deleting the cognitive bias related to the ignorance of base rate. The frequency-based representation helped eliminate bias of ignorance of base rate due to the probability-based representation.

Taking such characteristics into account, it is desirable to use frequency-based representation when we need to assess some risk so that cognitive biases are prevented.

Regression to mean

Comparison of Figure 3(a) and Figure 3(b) made it clear that the table representation is more effective for representativeness bias (regression to mean) than the graphic representation. Adding not the mean line to the graphic representation but the mean value to the table as a reference was effective for eliminating the bias related to the regression to mean.

The reason might be inferred as follows. The graphic representation in Figure 1 makes the respondent (participant) more conscious of the decrease of winning average for the past two years, and thus led to the rebound of the predicted value in the third year. On the other hand, the numerical expression in Table 1 might help make the respondent (participant) conscious of the mean value.

Framing effect

The reversal of preference (Alternative A is selected for positive frame, while alternative D is selected for negative frame) was not removed by the elaboration instruction in this study. Although Kahneman (2011), Tversky and Kahneman (1974), and Kahneman and Tversky (1984) insist that framing effect is robust property, it is not certain whether framing effect is in fact robust or not. As shown in Figure 4(a) and (b), irrespective of debiasing procedure (elaboration by requiring the respondent (participant) to explain the reason for his or her selection), framing effect (reversal of preference) was not removed properly.

Simon et. al. (2004), on the other hand, insisted that framing effect is not robust as Kahneman (2011), Tversky and Kahneman (1974), and Kahneman and Tversky (1984) insist. The elaboration (debiasing) was tried through the explanation why the respondent (participant) selected the alternative. Other elaboration (debiasing) technique such as making the respondent clearly recognize the difference of positive and negative frames might be more effective for deleting such a cognitive bias.

The debiasing of framing effect can be summarized as follows. Trapped into such a framing effect might lead to the distorted decision making, and thus the judgment of events from a specific frame. This might potentially lead to crucial accidents. Viewing events from multiple frames will be effective for preventing us from being trapped into a framing effect.

Illusion of covariation

Figure 5(a) and (b) shows that it is not so easy for us to recognize the causality (relationship between dizziness and brain tumor, or relationship between drinking habit and liver cancer) correctly, and that we suffer from cognitive biases when understanding the relationship between the cause and effect. When identifying the cause of failures or accidents, such biases are ubiquitous around us.

Failing to identify the genuine cause of failures, malfunctions, or accidents frequently induces crucial accidents such

as Three Mile Island nuclear power plant accident (nuclear reactor meltdown) (Lusted, 2012 and Osif et. al., 2006). In the range of our experiment, only a brief explanation for identifying causality does not work effectively for deleting fallacy of covariation. The consistent and scientific training for identifying causality with high reliability might be effective for deleting such a bias.

Overconfidence

The result in Figure 6 might show that overconfidence is deleted by probability-based representation. However, it must be noted that this debiasing measure does not lead to correct evaluation. In the range of this experiment, underconfidence appeared as a result of deleting overconfidence. According to Bazerman and Moore (2001), overconfidence leads to illusion of control, fallacy of plan, or optimistic biases. These cognitive biases potentially become the trigger of failures of business such as bankruptcy, crucial accidents or disasters. Moreover, underconfidence produces or induces passive and negative activities, leads to the insufficiency of a variety of systems, and eventually induces failures of business such as bankruptcy, crucial accidents or disasters. Future research should explore how overconfidence or underconfidence should be deleted.

CONCLUSIONS

In conclusion, the debiasing methods presented in this study were effective to some extent for suppressing the biases (conjunction fallacy, ignorance of base rate, and regression to mean). However, for some cognitive biases (framing effect, fallacy of covariation, and overconfidence), the debiasing methods were not necessarily effective. Future research should further continue to explore whether cognitive biases such as framing effect, fallacy of covariation, and overconfidence can be deleted or not. Moreover, we must systematically investigate the mechanism how such cognitive biases are linked to the occurrence of crucial accidents.

REFERENCES

- Altman, M. (2012), *"Behavioral Economics for Dummies"*, John Wiley & Sons Canada, Ltd.
- Angner, E. (2012), *"A Course in Behavioral Economics"*, Palgrave Macmillan.
- Bazerman, M.H. and Moore, D.A. (2001), *"Judgment in Managerial Decision Making"*, Harvard University Press.
- Kahneman, D. (2011), *"Thinking, Fast and Slow"*, Penguin Books.
- Kahneman, D. and Tversky, A. (1984), *"Choices, Values, and Frames"*, American Psychologist, Vol.34.
- Lusted, M.A. (2012), *"The Three Mile Island Nuclear Disaster"*, Essential Library.
- Osif, B.A., Baratta, A.J. and Conkling, T.W. (2006), *"TMI 25 Years Later - The Three Mile Island Nuclear Power Plant Accident and Its Impact"*, The Pennsylvania State University Press.
- Reason, J. (1990), *"Human Error"*, Cambridge University Press.
- Simon, A.F., Fagley, N.S. and Halleran, J.G. (2004), *"Decision framing: Moderating effects of individual differences and cognitive processing"*, Journal of Behavioral Decision Making, Vol.17, pp.77-93.
- Tversky, A. and Kahneman, D. (1974), *"Judgment under Uncertainty: Heuristics and Biases"*, Science, Vol.185.
- Vaughan, D. (1997), *"The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA"*, University of Chicago Press.