



AFRL-RH-WP-TR-2012-0138

**Delineating Cultural Models:
Extending the Cultural Mixture Model**

**Shane Mueller
Applied Research Associates
1750 Commerce Center Blvd. North
Fairborn, OH 45324-6362**

**Andrew Boettcher
Michigan Technological University
1400 Townsend Drive
Houghton, MI 49931**

**Dr. Michael Young
Forecasting Division
Patterns of Life Branch**

**December 2011
Final Tech Report**

DISTRIBUTUION A: Approved for public release; distribution is unlimited.

**AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING,
HUMAN EFFECTIVENESS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RH-WP-TR-2012-0138 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//Signed//

MICHAEL YOUNG, PhD
Work Unit Manager
Patterns of Life Branch

//Signed//

LOUISE CARTER, PhD
Forecasting Division
Human Effectiveness Directorate
711th Human Performance Wing
Air Force Research Laboratory

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 12-20-2011		2. REPORT TYPE Final		3. DATES COVERED (From - To) 10/2010 – 12/2011	
4. TITLE AND SUBTITLE Delineating Cultural Models: Extending the Cultural Mixture Model				5a. CONTRACT NUMBER FA8650-09-D-6939	
				5b. GRANT NUMBER NA	
				5c. PROGRAM ELEMENT NUMBER 62202F	
6. AUTHOR(S) Shane Mueller ¹ Andrew Boettcher ² Michael Young, PhD ³				5d. PROJECT NUMBER 7184	
				5e. TASK NUMBER 0009	
				5f. WORK UNIT NUMBER H073 (7184X25W)	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) ¹ Applied Research Associates 1750 Commerce Center Blvd N. Fairborn OH 45324-6362			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) ³ Air Force Material Command Air Force Research Laboratory 711th Human Performance Wing Human Effectiveness Directorate Forecasting Division Wright-Patterson AFB OH 45433-7022			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-RH-WP-TR-2012-0138		
12. DISTRIBUTION/AVAILABILITY STATEMENT Distribution A: Approved for public release; distribution is unlimited; 88ABW-2012-5648, 31 Oct 2012					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This report describes research and recommendations for describing cultural knowledge. We discuss some of the approaches that have been used in the past, which have considered culture as a factor-analytic problem, as a cognitive style and as a set of shared knowledge. We develop a set of recommendations for extending Cultural Mixture Modeling (CMM) to move beyond the shared-knowledge perspective and enable integration of the factor-analytic perspective as well. We provide a number of examples and recommendations regarding how structural models provide the appropriate representation for extending CMM and discuss issues related to statistical representation and inference that will need to be addressed on this path.					
15. SUBJECT TERMS CMM, cultural models					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include area code)
U	U	U	SAR	35	Michael Young, PhD

Standard Form 298 (Rev. 8/98)

THIS PAGE LEFT INTENTIONALLY BLANK.

TABLE OF CONTENTS

Section	Page
List of Figures	iv
List of Tables.....	iv
Acknowledgments.....	vi
1.0 SUMMARY	1
2.0 INTRODUCTION	2
3.0 METHODS, ASSUMPTIONS, AND PROCEDURES	3
4.0 RESULTS AND DISCUSSION.....	4
4.1 Approaches to Representing and Characterizing Culture	4
4.1.1 Factor-based approaches to Characterizing Culture	4
4.1.2 Cognitive Approaches to Representing Culture	7
4.1.3 Anthropological Approaches to Characterizing Culture	11
4.2 Reconciling Distinct Approaches to Culture using Structural Models	11
4.3 Extending the Cultural Mixture Model	17
4.3.1 Extending to Multiple Choice and Free Responses	18
4.3.2 Relationships Among Groups	18
4.3.3 Finding the Number of Groups	19
4.4 Structural Models	19
4.4.1 Structural Culture Models.....	21
4.5 Mixtures of Directed Acyclic Graphs	22
4.0 RECOMMENDATIONS	25
5.0 REFERENCES	26
6.0 LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS.....	28

LIST OF FIGURES

Figure		Page
1	Example Structural Model	12
2	Example Structural Model Describing Two Independent Cultural Factors Factor.....	13
3	Exam Example Structural Model Describing Two Cultural Factors that Share Some Common Questions.....	14
4	Structural Model Depiction of Cultural Mixture Modeling	25
5	Possible Structural Models that Integrate Dimensional Culture Approach and Cultural Mixture Modeling	16
6	Directed Acyclic Graph.....	19
7	Two Markov Equivalent Graphs and their Essential Graph	25
8	Structural Cultural Model	25
9	Example of an MDAG Model	25

LIST OF TABLES

Table		Page
1	List of the Traits Identified by Different Five-Factor Model of Personality	5
2	Hofstede's Cultural Dimensions	6
3	Other Cultural Influences on Cognition	10

ACKNOWLEDGMENTS

This report prepared under Subcontract No. RQ000746, Prime FA8650-09-D-6939 0009 to SRA, Intl., Project #135599.000 to Applied Research Associates (ARA). This work was partially executed through a subcontract to Michigan Technological University (MTU). The authors wish to thank and acknowledge Dr. Michael Young of the Air Force Research Laboratory (AFRL) for providing guidance and oversight, Scott MacBeth and David Snyder of SRA Intl., Inc., Thomas Drummer of Michigan Technological University Dept. of Statistics and Tom Holzberger of the MTU Office of Innovation and Industry Engagement for support and assistance.

During this project, Shane Mueller was jointly affiliated with ARA and MTU. His contributions to the research reported here were made as an employee of ARA, but he served as an advisor to Andrew Boettcher in his role as Assistant Professor at MTU. A portion of this research will be submitted by Andrew Boettcher for partial fulfillment of a Masters Degree project in the Department of Statistics at MTU.

1.0 SUMMARY

In this report, we examine several approaches to representing cultural knowledge and several types of knowledge that are thought to differ across cultures. These range from the popular factor-based approach inspired by personality research to anthropological approaches that remain highly qualitative characterizations of cultural knowledge. These different approaches initially appear at odds with one another, and seem to provide incommensurate approaches to characterizing culture.

Our research was motivated by the goal of extending Cultural Mixture Modeling (CMM); Mueller & Veinott, 2008; Mueller, 2010), which attempts to understand both agreement and subgroups of disagreement that exist within a culture. CMM itself is built on simpler methods for identifying consensus developed in Cultural Consensus Theory (CCT); Romney, Weller, and Batchelder, 1986). These approaches weren't without their limitations and a primary thread of our research effort was to identify statistical modeling methods for extending CMM in a way that would allow better insight into the nature of cultural knowledge. We began by examining the different approaches to characterizing cultural knowledge and finished by recommending a set of statistical models that would both unify the disparate approaches to modeling cultural knowledge and provide a richer framework upon which to develop theory.

The result of this investigation is a recommendation to adopt structural models in the form of Directed Acyclic Graphs (DAGs) to represent cultural knowledge. This approach built on the simple finite mixture modeling approach taken by CMM, but enabled more interesting and complex structures to be identified, both between knowledge elements and between frames and sub-frames of knowledge. Furthermore, the representational power provided by DAGs enabled a common language for understanding a number of approaches to representing culture, including the traditional factor-analytic approach advocated by Hofstede (1984). We concluded with a set of recommendations for how to employ DAGs to model-free response category norm data, which has previously been a challenge for standard approaches attempting to identify consensus.

2.0 INTRODUCTION

Culture, especially those components of a culture tied to nationality, geographic region, or organizational membership, encompasses many aspects of shared identity. These include (among other things) shared geography, weather, language, vocations, artifacts, history, social groups, behaviors, practices, norms, knowledge, attitudes, beliefs, leaders, celebrities, and stories. The key to a cultural identity centers on the shared understanding of some subset of these aspects. Thus, a topic on which a group does not share a common set of beliefs might not be part of that group's cultural identity. As a consequence, if one understands a cultural identity, one should be able to make inferences about the beliefs, attitudes, and knowledge of an individual within that culture.

A primary goal of much past research on culture has been to characterize the identity of a national or organizational group, in terms of specific attitudes or practices that are typical in the culture but tend to differ across cultures. If the cultural identity can be identified, then one can make inference about individual beliefs and help develop ways to better communicate with, do business with, train, or hire members of that culture. When identifying a cultural identity, it is important to understand whether the members of a culture tend to share that set of beliefs or else the ability to predict individuals from the group identity will fail.

So, for example, one might suggest a hypothetical cultural belief about members of a geographic region - perhaps their political conservatism. The now-ubiquitous categorization of red versus blue state is an example of this as it places each different state along this cultural political spectrum. Yet it remains an untested assumption whether this categorization reflects a true cultural identity because typically, knowing whether a state is red or blue will only give you modest information about one of its residents. Even the most conservative or liberal states have a mixture of individuals along the political spectrum and so mean political conservatism of a state may not indicate that there is a strong consensus regarding those beliefs, only that in a majority-rules society, the winner of elections tends to be in one party or another.

Any method - whether qualitative or quantitative - that attempts to characterize culture should be sensitive to the differences between an (1) individual belief, (2) the mean tendency of a group of individuals, and (3) whether that mean tendency is representative of a shared belief among individuals. Unfortunately, few approaches to culture have been sensitive to these distinctions. Consequently, the research we report here has been conducted in an effort to develop methods for understanding and representing these distinctions.

3.0 METHODS, ASSUMPTIONS, AND PROCEDURES

The Red-State/Blue-State example described in the Introduction illustrates some of the complexity in trying to identify and understand a shared cultural identity or belief. Our basic approach for this research effort was to explore and develop a set of statistical methods that can be used to infer and characterize cultural knowledge in a way that was cognizant of these issues. The path this research has taken is mostly in the form of:

- A review of statistical methods that have previously been used to characterize cultural knowledge
- A qualitative review of the types of knowledge that can be described as cultural.
- An exploration of new mathematical techniques (taking cultural mixture modeling as a starting point) that can be used to describe broader classes of cultural knowledge.
- A recommendation about the best path forward.

The exploration involved implementation of some candidate modeling approaches, but the main outcome of this research is a set of recommendations for extending CCT and CMM, along with a rationale for why reasonable alternative approaches are either insufficient or impractical. The outcomes of this investigation are described in detail in the Results and Discussion section, followed by specific recommendations in the Recommendations section.

4.0 RESULTS AND DISCUSSION

4.1 Approaches to Representing and Characterizing Culture

Culture and cultural knowledge has been studied in a number of distinct ways. The distinctions between approaches are primarily methodological, but they often masquerade as theoretical differences. Furthermore, the theories that are derived from the different approaches are each highly constrained by the methodological choices made.

4.1.1. Factor-Based Approaches to Characterizing Culture

The first example we will describe is the factor-analytic approach to studying culture. This approach begins with the premise that culture can be studied by conducting questionnaire research whose covariance is transformed into a small number of orthogonal dimensions. The resulting theory has a strong correspondence to modern study of personality, which we will discuss next.

Cultural Dimensions as a Group Personality Theory: The factor-analytic approach to studying culture is rooted in the methodologies developed for the study of personality. By understanding the personality profile of an individual (e.g., their answers to a number of questions regarding their attitudes across a wide spectrum of issues), one can understand how they are likely to react in new situations, determine whether they are suitable for certain jobs, identify appropriate therapies or interventions, and provide insight into their behaviors. Personality factors are typically developed through repeated administration of questionnaires to a large participant pool, and the use of factor analytic methods to identify which items cohere. In this sense, coherence means two things: (1) there was considerable variability across the population in a response, and (2) that response co-varied with another response. Items that lack coherence are thought to not be predictive of a factor and are eventually removed from the test body.

The current dominant taxonomy for characterizing individual personality is the so-called “Five-Dactor Model” (e.g., Costa & McRae, 1992). The predecessors to this model date back at least to Tupes & Christal’s (1961) study (funded by the U.S. Air Force) which identified five factors to describe personality (with much overlap to the current five). Norman’s (1963) validation study began the evolution toward the currently-accepted factors. The traits identified by these three models are shown in Table 1.

Table 1: List of Traits Identified by Different Five-Factor Model of Personality

Source		
Tupes & Christal (1961)	Norman (1963)	Costa & McRae (1992)
Surgency	Surgency	Extraversion Openness
Dependability	Conscientiousness	Conscientiousness
Emotional Stability	Emotional Stability	Neuroticism
Agreeableness	Agreeableness	Agreeableness
Culture	Culture	

These personality traits are often considered to be general truths about the behavioral attitudes and perspectives of individuals. However, statistically they can be thought of as latent variables that tend to orthogonally account for a maximal amount of variance across the questions being responded to. The iterative process by which these factors have been identified (which has been going on for more than 50 years) necessarily selects questions that (1) vary across the individuals that are studied (typically westerners), (2) cohere across a number of ways of asking a question and within the factor-topics, and (3) are each primarily related to a single factor. These personality dimensions are the end-point of a process that seeks to find pure dimensions and can only find pure dimensions. In a real sense, the dimensions do not exist in the mind of the personalities being studied, but only in the mind of the researcher who is studying them. Despite this, the outcomes can be very useful and be the basis for diagnoses, treatments, hiring, and other decisions that impact lives.

Interestingly, Tupes and Christal (1961) originally included Culture as one of the factors, but this has been subsumed into other factors in the current versions of the five-factor model. At first examination, the five-factor model (and perhaps the entire factor-analytic personality approach) would seem to have limited application to culture. After all, the factors and the questions are chosen so that responses vary across individuals within a culture, yet account for maximal variance. Thus, questions related to a cultural personality trait - a set of attitudes or beliefs that are consistent for a culture - would never be selected because the questions would not provide discrimination of members within that culture. Furthermore, those questions that are selected will tend to have a large variance across individuals within a culture (as this is how a factor accounts for the most variance) and so would not qualify as a “cultural personality trait,” which should be consistent.

The Dimensional Approach to Characterizing Culture: The solution, therefore, has been to modify this approach so that it can capture culture. In other words, one can take the same factor analytic approach but identify factors that vary across cultures but are consistent within cultures. By taking this approach, one would likely find factors that are distinct from the Big Five personality traits, but have a similar nature: they can be identified by grouping responses on

questionnaires, they involve attitudes, they each relate to unique orthogonal dimensions, and so on. Yet they are still subject to the limitations of factor analytic approaches: they require finding responses to questions that vary across cultures, that co-vary together, and that can account for the greatest proportion of variability (so that a cultural factor on which 95% of the nations that are studied had large agreements would not be a powerful cultural factor, even if a small minority had very different attitudes and beliefs regarding that factor). Furthermore, the factor-analytic approach to culture must pass an even stronger criterion: the responses to a set of questions should cohere within a cultural group. If a factor has large disagreement within a particular culture saying that the culture is moderate on that factor hides the truth, and obscures the possibility that the cultural trait is simply a personality trait rather than any coherent way of describing a culture.

Of course, this factor analytic approach to culture has been under investigation for decades (Hofstede, 1980). The cultural personality traits identified by Hofstede are shown in Table 2.

Table 2: Hofstede’s Cultural Dimensions

Dimension	Description
Individualism vs collectivism	Extent to which individuals are integrated into groups
Power Distance	Extent to which individuals expect power to be distributed unequally.
Masculinity vs femininity	Distribution of roles between genders
Long-term orientation vs. short-term	Focus on future vs. present and past.
Indulgence vs. Restraint	Extent to which hedonistic behavior is permitted or accepted.
Uncertainty Avoidance	Tolerance for uncertainty and ambiguity

As expected, these factors differ somewhat from those personality factors described in Table 1. But just like those factors, cultural dimensions are the result of a process destined to find factors - sets of questions which vary across individuals and co-vary together, are orthogonal between dimensions, and vary maximally. For reasons such as this, we don’t find esoteric traits that might be critical predictors for one or another culture (e.g., attention to time, which could be a critical predictor when comparing Indonesian culture to the rest of the world), or factors that presumably vary across members of each society (e.g., neuroticism, which is likely correlated with the presence of a number of genotypes that vary across cultures).

Despite the fundamental limitations of the approach (and similarly to personality research), the dimensional approach can be useful. For example, it can provide guidance to help corporations understand their multinational operations or to help develop training to allow their corporate cultural identity to embrace and differentiate in different national cultures. Importantly, it distills numerous attitudes and behaviors down to a small set of influences which supposedly govern behavior across a range of situations. However, it must be reiterated that just as with personality theory, the dimensions that come out of the process are a statistical description of survey responses and not necessarily principled psychological trait that influence behavior across a wide range of situations.

4.1.2. Cognitive Approaches to Representing Culture

As pointed out by Nisbett and Norenzayan (2002), psychologists typically assume that attentional, memory, learning, and inference are universal primitives, yet all of these have been found to be influenced by culture. Thus, there is a growing community of researchers who view culture as a fundamentally cognitive phenomenon, or at least an cognitive phenomenon embedded within a cultural context.

The personality-theory approach to characterizing culture implies that cultures and people living in those cultures have certain traits for behaving which govern behavior across a wide spectrum of situations. As we begin describing cognitive approaches, we will consider specific shared declarative knowledge as a carrier of culture. These two views map imperfectly but roughly onto the declarative/procedural or explicit/implicit spectrum, a distinction originally popularized in the study of human memory (cf. Graf & Schacter, 1985). Some cultural knowledge is clearly explicit, declarative information. One can ask a member of a culture to identify family relations, or rules of etiquette, or cultural icons and religious symbols, and these make up an important part of culture. Yet traditional personality-inspired approaches seek to identify styles of behavior, which may be better thought of as procedural or implicit knowledge.

In contrast, using Hofstede's masculinity dimension as an example, a masculinity trait would go beyond simply a listing of gender roles ("This is what women do."), or verbalizable attitudes toward appropriate gender roles ("This is what women SHOULD do") or codified gender roles ("This is what women MUST or ARE PERMITTED to do.") It must be able to predict behavior in new situations and influence behavior over a wide range of situations. Because the dimensional approach seeks to describe culture along a few factors, it is bound to identify these types of traits. At the other extreme, ethnographic and knowledge-based approaches will tend to characterize the explicit shared knowledge of a culture. The cognitive view tends to include both procedural/style aspects and specific knowledge. Some, such as category norms, color names, and possibly even factors such as fatalism and risk tolerance (see below) are essentially embedded within the knowledge one learns from living in a society. These last two could represent specific explicit knowledge to the extent that the cultural norm is embedded within stories, idioms, morals, and laws within that culture, but they could equally-well result from proceduralized reasoning or thinking strategies. Others, such as time understanding, reasoning style, global versus local processing preferences are essentially procedural or implicit knowledge of how various tasks are done, or how systems and practices within a culture should or do work.

A number of cognitive approaches have identified ways in which cognition appears to differ across cultures. Some of these have been collected by H. Klein (2004) into a theory called the

Cultural Lens, which we will discuss next. Following that, we will describe several other cognitive findings that are influenced by culture, and follow this with a discussion of how such information might be represented in a formal system.

The Cultural Lens: A Macrocognitive Factors Approach for Describing Culture Klein's (2004) Cultural Lens model is a descriptive taxonomy for understanding how cultural factors influence cognition, developed from a naturalistic perspective. Rather than relying strictly on questionnaires to identify coherent sets of attitudes, Klein examined individual results from both laboratory and naturalistic studies that showed consistent differences across cultures. She identified eight main factors, which include:

- Time Horizon
- Achievement vs. Relationship
- Mastery vs. Fatalism
- Tolerance for Uncertainty
- Power Distance
- Hypothetical vs. Concrete Reasoning
- Attribution
- Differentiation vs. Dialectical Reasoning

These have some overlap with the dimensional approach advocated by Hofstede, but attempt to place cultural knowledge and cognitive styles within a naturalistic "Macrocognitive" setting. Thus, these dimensions must be thought of in their social and work contexts, rather than simply as either personality traits or primary cognitive functions.

Like Hofstede's dimensions, these factors are primarily framed as cultural personality traits, but related to macrocognitive issues such as reasoning and decision making. For example, tolerance for uncertainty relates to risk and planning. Klein (2004) described how the US military culture permits much more tolerance for uncertainty than even United Kingdom (UK) military planning. One might ask how uncertainty tolerance is represented and thus why it might differ across cultures.

One way to conceptualize uncertainty tolerance is fundamentally cognitive - we might assume that decision makers evaluate plans along a number of dimensions, such as probability of success, value of outcomes, flexibility, and thoroughness, and make decisions using some expected utility combination rule that incorporates all of these factors. A culture that is tolerant of uncertainty would simply weigh the thoroughness dimension less than another culture. Cultures may differ in how they weigh these different aspects of planning in decision making, and a simplistic cognitive view might suggest that those weights could be adjusted via feedback or reward/penalty structure. A number of more nuanced views are possible, which might place the locus of this weight outside the individual, which might explain why two cultures may differ on their decision making styles.

In contrast, a personality-based perspective might consider personality traits related to risk-seeking or risk-aversion as primitives (cf. LeJeuz et al., 2002), and hypothesize that risky situations may elicit an emotional or autonomic response that influences decision making and

planning. In this perspective, cultures could differ in the extent to which they permit intuitive argumentation, or to the extent that they create individuals who are risk-seeking or risk-averse because of societal conditions such as violence, poverty, and the like.

Finally, an ecological approach might suggest that the cognitive and personality factors themselves are irrelevant, and uncertainty tolerance is an institutionalized practice that exists 'outside' of the head of any individuals. A culture with low tolerance for uncertainty might have emerged because of cultural practices of blame-shifting after accidents or mistakes. If a group has had high-profile failures in the past which led to investigations that laid blame on poor planning, it may have led to practices which make the thoroughness of a plan unimpeachable. This practice could persist long after the events that produced it are forgotten. This would have little to do with individual cognitive or personality style, but would be a practice to prevent similar reprovals in the future.

These alternative views have different perspectives on why cultures may differ on the tolerance for uncertainty and different predictions about whether or how this tolerance is communicable. For what was termed the cognitive approach, cultural differences would lie in a learned decision making strategy that could presumably be re-learned or influenced by different reward and outcome structures. For the personality approach, the style may be much more pernicious and may not even be changeable on an individual level. The ecological perspective would suggest that the style is the result of macro-level phenomenon and might be difficult to change, but change would be possible should these macro-level pressures change.

Other Cognitive Approaches to Characterizing Culture: Much of the Cultural Lens model brings together research in cognition and cognitive style which has been shown to differ across cultures with special attention to naturalistic functions. There are a number of other specific functions that have often been thought of as cognitive primitives but which have also been shown to depend on culture. In general, this research goes back to Linguistic Relativity Theory, also known as the Sapir-Whorf hypothesis (Whorf, 1956). We won't offer a comprehensive review of this literature, but Table 3 provides several examples of how culture has been shown to impact fairly primitive cognitive operations.

Table 3: Other Cultural Influences on Cognition

Topic	Citation
Category Norms	Yoon, et al. (2004)
Color Name Categories	Robeson, Davies, & Davidoff (2000)
Color Preference	Palmer & Schloss (2010)
Global/Local	Masuda & Nisbett, (2001)
Time Understanding	Boriditsky (2001)
Risk Preference/Tolerance	Hsee & Weber (1999)

As a brief review, Yoon et al. (2004) conducted a study looking at category norms for Chinese and American respondents, characterizing categories which were both common across the cultures and ones that were distinct. It is obvious that some category norms must differ across cultures and this type of study simply establishes that the knowledge we have is dependent on the context in which we live. A controversial related phenomenon was established by Robeson, Davies, & Davidoff (2000), in which they showed how the perception of color spaces and color similarities were indeed impacted by culture, most likely because of the language used to describe and label color. Slightly different is Palmer & Schloss's (2010) WAVE model of color preference, in which they found that different cultures had different associations with, and consequently preferences for, different colors. Here, when asked which of two colors one preferred, responses could be accounted for by identifying associations with that color and using secondary positive and negative associations to predict the valence for that color. Cultures differ in their color palette because of geography, technology, wealth, tradition, and fashion and so it is not surprising that these associations have secondary effects on arbitrary ratings of color. Each of these examples represents an explicit knowledge categorization that happens to differ across cultures consistent with a view of culture as shared knowledge.

Other phenomena are more procedural. For example, Masuda & Nisbett (2001) established that visual encoding styles may differ across cultures, with Japanese participants attending to the background and contextual cues of a scene more so than American participants. Similarly, Boriditsky (e.g., 2001) has shown how reasoning about time differs across cultures. Finally, Hsee & Weber's (1999) findings are somewhat related to the Tolerance for Uncertainty dimension identified by Klein (2004), and relate to cultural differences in a level of risk that is accepted or preferred. These phenomena go beyond establishing difference in what people know, and impact how people in different cultures think, reason, or act. These may be

proceduralized knowledge, which control behavior in limited situations, and are impacted by repeated practice of a particular behavior norm. The behavior norm may be an instantiation of a ubiquitous philosophy (e.g., a holistic world-view, a fatalistic views of causality), or it could be related to more prosaic practices (practice with different styles of video games or puzzles popular in different cultures; verb tense systems in different languages). To the extent that such phenomenon stem from procedural knowledge, it may be possible to change individual behavior via deliberate practice with different reasoning modes. However, these behaviors are likely a consequence of some other shared practice or belief, rather than a central aspect of culture. Thus, if the shared practice or belief can be changed, the procedural skill may be flexible as well.

4.1.3. Anthropological Approaches to Characterizing Culture

The anthropological and ethnographic approach to characterizing cultural knowledge centers on developing qualitative narratives for understanding a culture (e.g., Watson & Huntington, 2008). Thus, this approach is closely related to the knowledge-based cognitive characterizations discussed earlier. However, the earlier cognitive approach focused on fairly simple associations, between a category and its members, or between colors and color names or objects that have such a color. A narrative encompasses a much more complex type of declarative knowledge.

Little work has been done identifying methods for translating such narratives into data structures that would permit identifying consensus narratives, but the development of methods for identifying shared narratives could have a number of applications. Current computational and Artificial Intelligence (AI) approaches to representing narrative (Van Den Braak et al., 2007; Richards, Finlayson, & Winston, P. H., 2009) rely networks of nodes and relationships to represent narrative.

Thus, a solution to representing cultural narratives may not need to establish shared knowledge on the textual description of narratives produces as an output of a typical ethnography. Rather, a method that can perform cultural consensus inference on directed graphs may be sufficient, insofar as the narrative can be mapped onto that graph. We will describe such an approach in subsequent sections of this report.

4.2 Reconciling Distinct Approaches to Culture using Structural Models

Initially, the different approaches that fall along a spectrum from personality trait to declarative knowledge and narrative appear to incommensurable. The factor-based approach is interested in placing a culture in an attitude space, whereas the epidemiological approach is interested in identifying specific aspects of knowledge that are embedded and transmitted within a culture. However, both approaches can be understood from a generic structural models approach. Figure 1 shows a typical representation for a single factor, which could be Hofstede's Power Distance. Power distance essentially describes a set of practices, beliefs, and attitudes regarding the power hierarchy of a society, inferred via the responses (typically on a 1-5 scale) of questions such as the following:

In most situations managers should make decisions without consulting their subordinates.

A survey will contain a number of like questions (depicted as rectangles on the right side of Figure 1). These questions were validated in the sense that they have been shown to co-vary, different cultures differ in the extent to which individuals tend to agree with the statements. On

the left, the latent factor (labeled “power distance”) represents a causal source for the responses. For simplicity, we will consider power distance and the responses all two-state random variables, so that individuals will either have high or low power distance attitudes, and agreement or disagreement with the question. In general, these assumptions can be relaxed with only minor additional complexity. The arrows connecting the power distance latent variable with observable responses to questions describe the two conditional probabilities of giving a positive response to the question (the probability of “Yes” given high power distance, and the probability of “Yes” given low power distance).

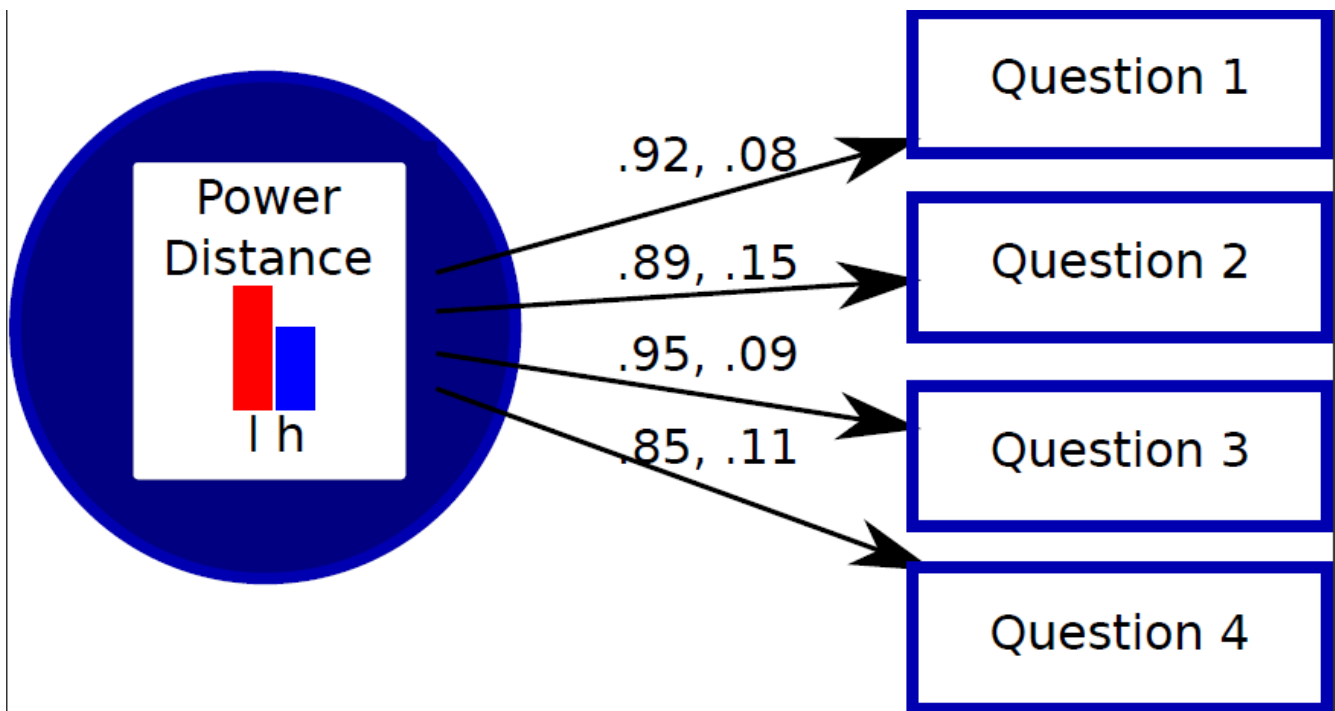


Figure 1: Example Structural Model Describing a Single Cultural Factor

The structural model depicted in Figure 1 frames the analytic problem in terms of a Markov network, rather than the eigenfactor decomposition utilized by typical factor-analytic approaches. However, they are closely related and similar in spirit. The Markov framework treats each node as a random variable, with both unconditional and conditional probability distributions at each state. The ultimate distribution of responses can be simulated via Monte Carlo simulation if all the probability distributions are known, and the main inference process is to identify a best estimate for those probabilities given a set of data.

In a typical survey, many questions will be asked, and they will be explained best by some set of latent variables. Through an iterative testing and selection process, the factor analytic approach selects questions that are relatively independent, and identifies the major themes which describe these questions. Figure 2 illustrates with two of Hofstede’s dimensions. Ideally, responses to

each bank of questions would be highly determined by a knowledge of the latent node, as shown in Figure 2.

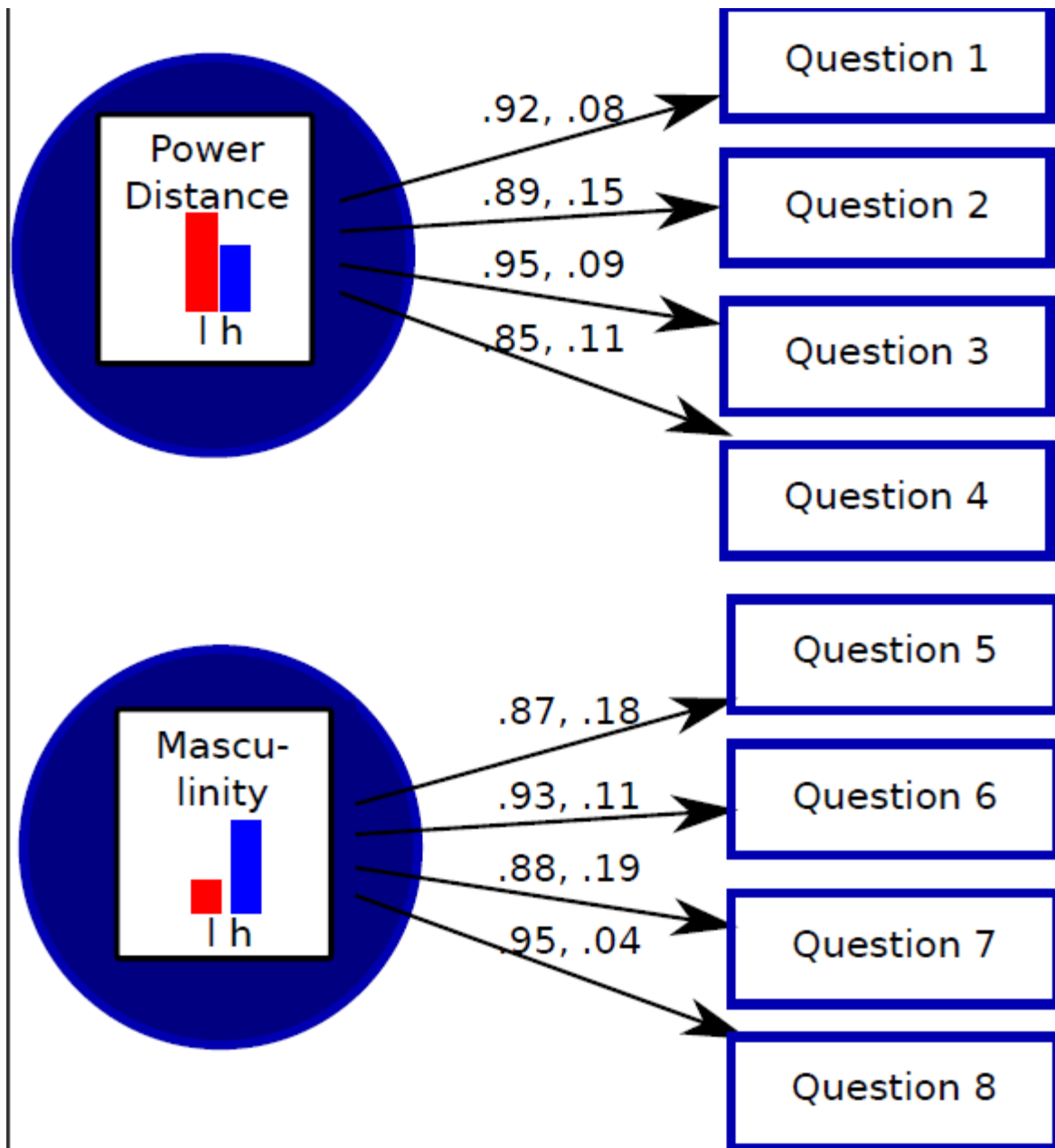


Figure 2: Example Structural Model Describing Two Independent Cultural Factors Factor

Without explicit care to select questions and factors that are conditionally independent, one is more likely to find a situation like that shown in Figure 3. Here, some questions are predicted by both latent states. However, both of these situations essentially described the potential attitude space of an individual, rather than the attitudes of a group. So, for the example in Figure 2, we suppose that each factor has two levels (high and low), and that knowing whether an individual is high or low on these factors can tell you what their responses are likely to be on the different

questions. A culture will have a distribution of these latent states, so that perhaps (as depicted) most individuals will have low power distance and high masculinity within a culture. The distribution across these latent states would provide the location in the dimensional state implied by Hofstede's dimensional analysis.

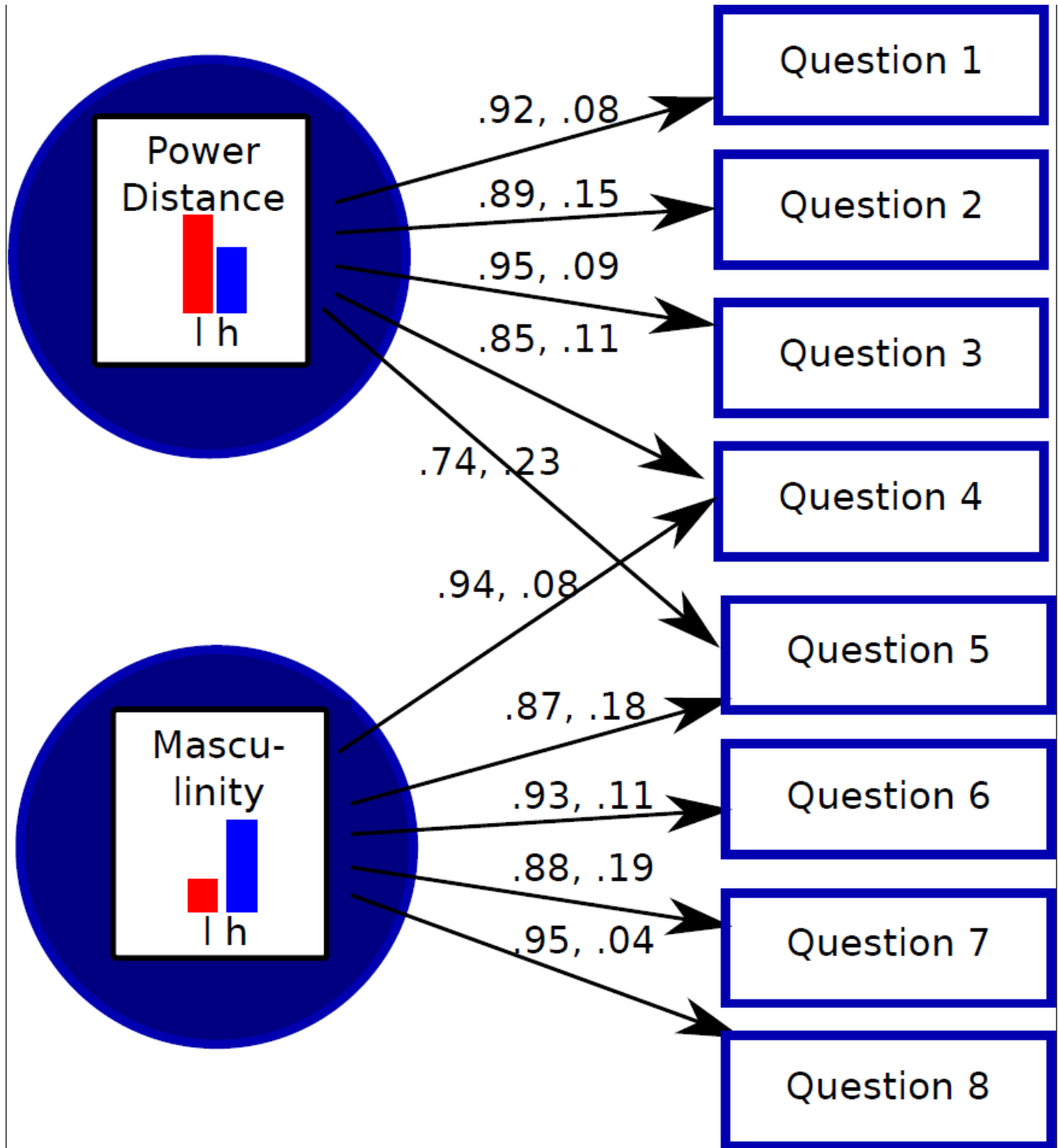


Figure 3: Example Structural Model Describing Two Cultural Factors that Share Some Common Questions

Of course, the dimensional approach assumes that each of these variables is not a categorical value, but rather some continuum. The general Markov network can be extended to capture continuous-valued random variables as well to handle such a situation. But the important limitation of the standard dimensional approach is that it essentially assumes the pooled distribution describes the individual practice or experience. This may or may not be the case in general.

Mueller & Veinott (2008) introduced CMM to address that limitation. They examined single-node latent variables with multiple states, as shown in Figure 4. Essentially, a node might have multiple categories representing a group of individuals. Now, multiple conditional edge probabilities exist for each state of the latent node, but the strong consensus model they proposed assumed that each edge must have a value of either α or $1 - \alpha$, for some small value of α (e.g., 0.05). By restricting the latent nodes to be categorical, and the conditional probabilities to be close to 0 or 1, this model attempts to find groups of consensus.

The advantage of this approach is that it allows one to characterize a culture as a distribution of beliefs, rather than as a single value. Thus, if there really is strong agreement about a moderate position on the power distance scale, one can describe this consensus; if rather there is disagreement, so that one group holds beliefs highly consistent with power distance and another group holds attitudes consistent with equal power distribution, that would be represented as a mixture of two belief groups.

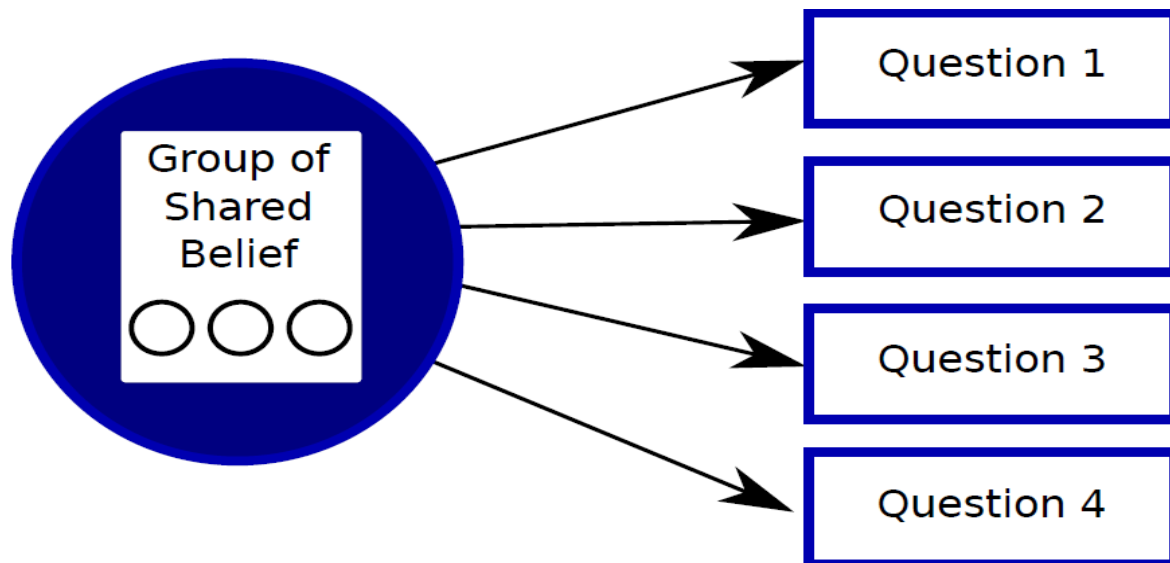


Figure 4: Structural Model Depiction of Cultural Mixture Modeling

The focus on finding groups of agreement has, however, led to a severe limitation in how CMM has so far been applied. Unlike the factor analytic approach, it places no structure on the set of ideas it is looking at. The structural model vocabulary can easily handle extensions like this, and enable different theories about shared belief to be tested on data. Figure 5 depicts several possible ways this integration might happen, using for concreteness two of Hofstede's dimensions as example latent variables. It should be recognized that the basic approach we are advocating does not rely on the existence of these dimensions; it simply allows for framing a model that incorporates dimensions such as this.

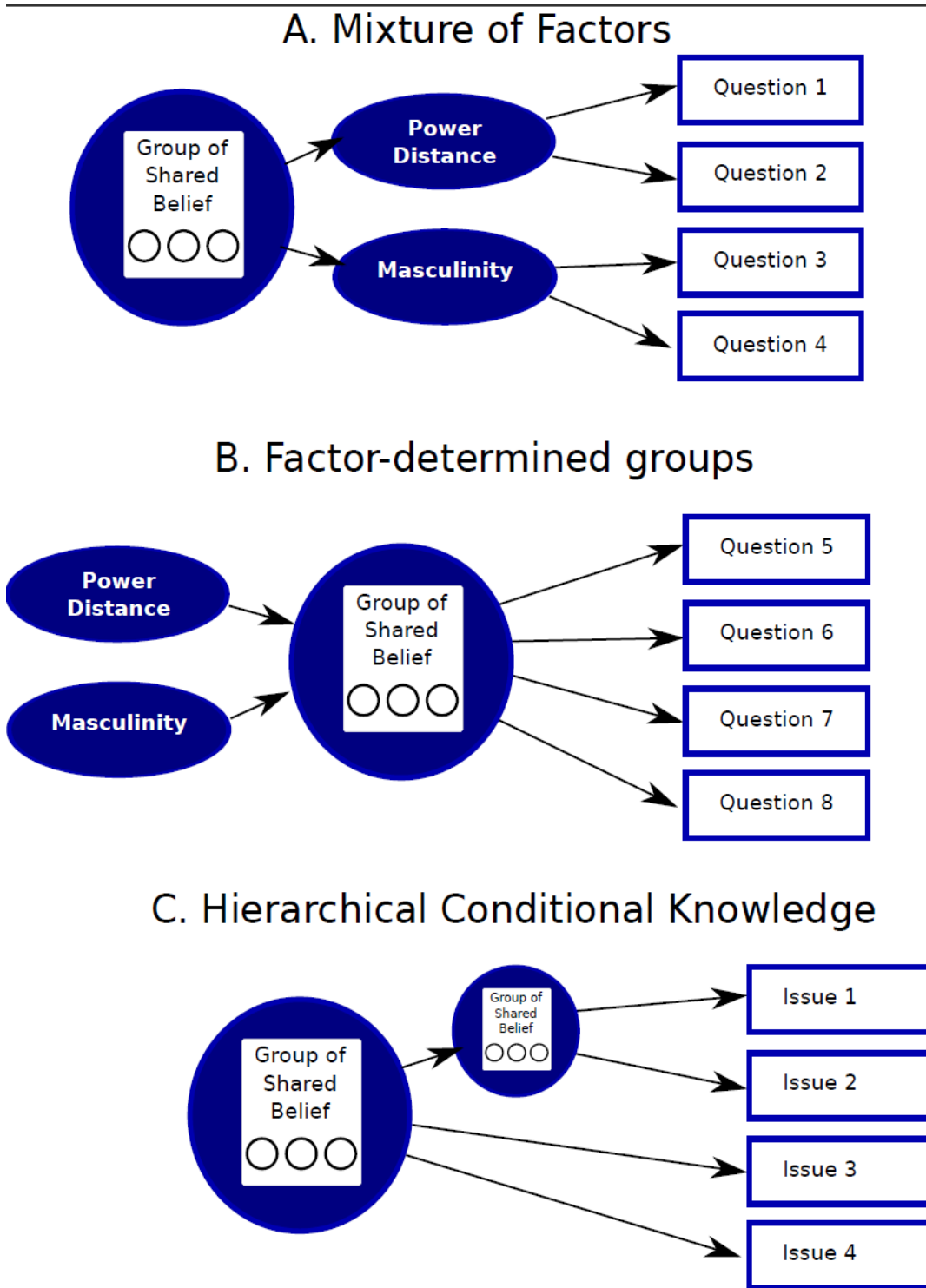


Figure 5: Possible Structural Models that Integrate

Figure 5 shows three hypothetical structures that a structural model permits, which enable inference about shared knowledge within a culture. The top Panel A shows one of the simplest

ways to combine these factors. Suppose that a set of orthogonal factors were chosen already. If you can determine an individual's state on those two factors, the responses to all questions can be determined with high reliability. However, within a culture or across cultures, there may be relationships between those factors. If each factor had two levels (low and high), one might be able to classify a culture as a mixture of groups with just a few of the possible patterns (e.g., only high-high and high-low). Thus, the group identity of an individual would determine their value on a dimension, which would in turn determine their responses to different questions.

The center Panel B reverses this relationship. Here, suppose we ask a new set of questions (along with ones related to masculinity and power distance). We may be able to characterize the entire population based on a small number of groups that cross cultures, but a respondent's group membership can be determined by his or her responses on the masculinity and power distance dimensions.

Finally, the bottom Panel C shows a hierarchical knowledge group structure. The left most node may define a large-scale category of belief (e.g., Political Party) which determines the views on a number of issues (e.g., 3 and 4). That is, the node might specify that a politician is either a Republican or a Democrat, and knowing only this can allow one to predict votes on Issues 3 and 4. But a subdivision of one party splits on another set of issues (1 and 2), but the probability of holding this sub-view is highly dependent on the primary party, but determined even more-so by the subgroup membership.

These three examples provide some initial examples of how different approaches to representing cultural knowledge can be unified. Both the factorial-based personality approaches and the knowledge-based ethnography approaches are amenable to these representations, especially as the relationships between knowledge (in the form of causal reasoning and narrative) can be made explicitly using this approach. In the remainder of this report, we will discuss in technical detail the advantages and limitations of this approach.

4.3 Extending the Cultural Mixture Model

As originally described by Mueller & Veinott (2008), CMM is a simplistic way to allow inference about groups and subcultures of agreement from survey-style data. There are a number of ways in which the model limits the types of inference that can be made about knowledge. In the current research effort, we have examined ways in which CMM can be extended to provide a better account for data and better understanding of shared knowledge.

In this research, we have identified three primary ways in which CMM can be extended. Even if we retain the basic framework of mixture modeling, we can consider a few new alternatives. These include: (1) we can expand beyond binary questions, which would allow CMM to be extended to free-response and multiple choice data. This change could be incorporated into the current CMM with minimal work, mostly related to developing a multinomial likelihood model for the data. Alternatively, (2) we can consider relationships among the groups. These types of relationships, which were discussed at a high level in the previous section, will enable a much richer characterization of cultural knowledge, by extending the model from simple mixtures to structural models. Finally, (3) we could improve upon the fitting method itself. This might entail using means to consider the number of groups itself a part of the model, making the search for

the model a single iterative process.

4.3.1. Extending to Multiple Choice and Free Responses

When we extend the model to multiple choice questions we must consider exactly what entails a consensus. Before we considered binary questions with a strong consensus model, where for each question there is one parameter per group, γ_i . This parameter was restricted to be α or $1 - \alpha$, where α was set close to 0. With m choices there are $m - 1$ parameters. Consider one multiple choice response with three choices: A, B, and C. For each group in the data there is a pair of parameters, $\gamma_{A,i}$ and $\gamma_{B,i}$. The probability that a member of group i gives the response x is

$$p_i(x) = \gamma_{A,i}^{I_{\{x=A\}}} \gamma_{B,i}^{I_{\{x=B\}}} (1 - \gamma_{A,i} - \gamma_{B,i})^{I_{\{x=C\}}}, \quad (1)$$

Where I is 0 if the condition is false, and 1 if it is true. A similar restriction to the strong consensus model would require at least one of $\gamma_{A,i}$, $\gamma_{B,i}$, or $1 - \gamma_{A,i} - \gamma_{B,i}$ is α or $1 - \alpha$. It is interesting to note that there could be a consensus that one response is not correct, even if there is no agreement about which one is correct, (e.g. $\gamma_{A,i} = .01$ and $\gamma_{B,i} = \gamma_{C,i} = .495$). Here there is an agreement against A, but an even split between B and C. However, inference about a consensus against an option must be made with care, a fact that proven by Arrow's (1950) Impossibility theorem.

For free responses we could consider a strong consensus based on a threshold $1 - \alpha$, which might be far less than 0.95 if there are many differing responses, mostly with very low response levels. A reasonable value for α might be selected by making an estimate of the size of the total response pool related to the number of responses typically given. An alternative would be to consider a mixture of structural models, discussed later.

4.3.2. Relationships Among Groups

In the mixture model we assumed that the groups are independent of each other. This assumption does not fit when we consider a population where an individual can belong to multiple groups. A well-studied extension to consider is that of hierarchical groups.

There is a vast literature and multiple software packages available for hierarchical mixture models, making it easy to extend CMM to allow for hierarchies among groups. This also over-constrains the group structure. Political affiliations offer an example: A person might be a Republican or a Democrat overall, however this is not the only dimension to political affiliation. Suppose a person generally affiliates Democratic because he or she is socially liberal, yet they are also fiscally conservative. Another individual might be socially conservative and fiscally conservative, identifying as a Republican. Yet another person is socially conservative, fiscally liberal, and identifies as a Republican. If fiscal and social affiliations are nested in the overall party affiliations, the assumption of a hierarchy does not allow one to be a fiscal conservative and either a Democrat or a Republican. If we relax the independence imposed by the hierarchy, we can allow for almost any relationship among groups. These types of relationships among groups can be represented with a structural model.

The calculation of the mixture model under the independent groups and hierarchy assumptions is computationally tractable because the likelihood factorizes and can be optimized very easily. In the more general case of a structural model we must work with specialized algorithms that are much more computationally intensive.

4.3.3 Finding the Number of Groups

The original CMM used the Expectation-Maximization (EM) algorithm to find the best set of parameters under a fixed number of groups, which was repeated for different numbers of groups, using the Bayesian Information Criterion (BIC) to select the best model. There are methods for combining the search for the number of groups with the iteration to optimize the parameters, thereby reducing the amount of computation involved in repetitively iterating over the data.

The most common of these methods is Reversible Jump Markov Chain Monte Carlo (RJ-MCMC). In this method the model can jump between parameter sets of varying sizes. Since there is a set of parameters for each question in each group, the number of parameters in the model depends on the number of groups. Moving the model fitting problem from the EM algorithm to Bayesian MCMC methods allows more flexibility in the model itself. The cost is the introduction of more algorithmic details, such as the label switching problem.

When we use the EM algorithm to find the mixture model, we start out with random assignments, and run multiple different sets at the same time. Comparing these different sets directly poses a problem because the same groups might have different labels in each different set. Since we previously used the BIC to compare models, this was not a problem for the standard CMM. However, in RJ-MCMC label switching becomes a problem when the number of groups jumps (increases or decreases). This can be rectified by imposing some form of well ordering on the groups. Problems like this and other algorithmic details make RJ-MCMC more difficult to implement. Other methods for considering the number of groups as a variable similarly require moving to a Bayesian framework. In particular RJ-MCMC has been studied in the context of structural models, lending an aid to implementation.

4.4 Structural Models

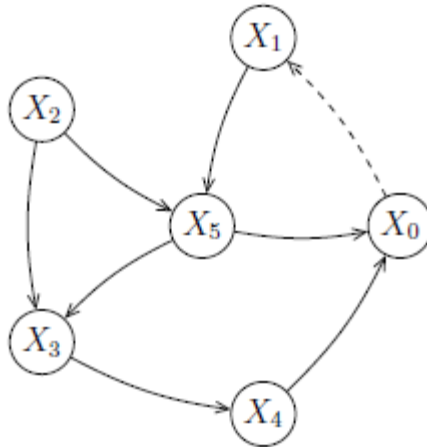


Figure 6: Directed Acyclic Graph

The parents of a node are all nodes that point toward it, e.g. the parents of X_5 are X_1 and X_2 , while the parents of X_3 are X_2 and X_5 . If there were an edge from X_0 to X_1 , there would be a cycle with X_5 , so this edge cannot exist.

Several potential improvements to the standard CMM could be achieved by adapting a structural model to represent the data. A structural model is a model where the joint distribution of a set of variables can be factored into a DAG which represents the conditional independence of the variables. If $X = \text{class}=\text{"MathClass-open"}(X_1, \dots, X_n)$ are the random variables and b is the structure of the variables,

$$p(X = \mathbf{x}) = \prod_{i=1}^n p(x_i | \text{pa}(\mathbf{b})_i), \quad (2)$$

where $\text{pa}(\mathbf{b})_i$ is the structure of the parents of X_i .

The structure is built from information about conditional independence. We say Y is conditionally independent of Z given W if given any W , Y is independent of Z . This is denoted $Y \perp\!\!\!\perp Z | W$. We can use the information about conditional independence to find subsets of X which are not conditionally independent. We give those subsets edges in the structural graph, b . The directions of the edges and further properties of the graph are determined from the pairwise, local, and global Markov properties:

The Pairwise Markov Property: for any pair (X_i, X_j) of non-adjacent vertices,

$$X_i \perp\!\!\!\perp X_j \mid \mathbf{X} \setminus \{X_i, X_j\},$$

The Local Markov Property: For any vertex X_i with the set of its parents and neighbours denoted $\text{bd}(X_i)$,

$$X_i \perp\!\!\!\perp \mathbf{X} \setminus \{X_i, \text{bd}(X_i)\} \mid \text{bd}(X_i),$$

The Global Markov Property: Given any triple, (A, B, S) , of disjoint subsets of \mathbf{X} such that S separates A from B ,

$$A \perp\!\!\!\perp B \mid S.$$

With these additional restrictions and a density which is always positive, we can guarantee there is a factorization of the density that has such a graphical structure. An important note is that this graphical structure is not necessarily unique. When more than one graph represents the same set of conditional independence relation they are said to have Markov equivalence. This requires careful consideration since we are searching for a single structure that describes our data. All Markov equivalent graphs have the same skeleton, that is the underlying graph without directions. The essential graph can be used to characterize the information in the data. The essential graph is the graph where there is an arrow if at least one Markov equivalent graph has that arrow, and none have the reverse arrow.

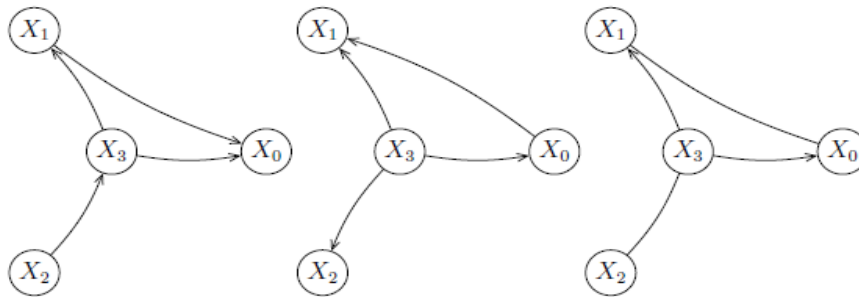


Figure 7: Two Markov Equivalent Graphs and their Essential Graph

(In both graphs X_3 is a parent of X_0 and X_1)

This leads us to the need to be careful about how we interpret the structure of these graphs. There are two viewpoints we can take. We can seek the ‘true’ underlying graph that represents the unique relation between the variables, and gives us the ability to explain them as well as casually model and predict. The other point of view is that we are just drawing an approximate model to predict future data, but not necessarily find a truth in it. In this point of view, we allow

room for multiple possible models, while in the original viewpoint there can be only one. If we seek a single truthful answer then we might view the variance that would admit other models as the uncertainty in our answer. In this way the essential graph represents the part we are certain about.

4.4.1. Structural Culture Models

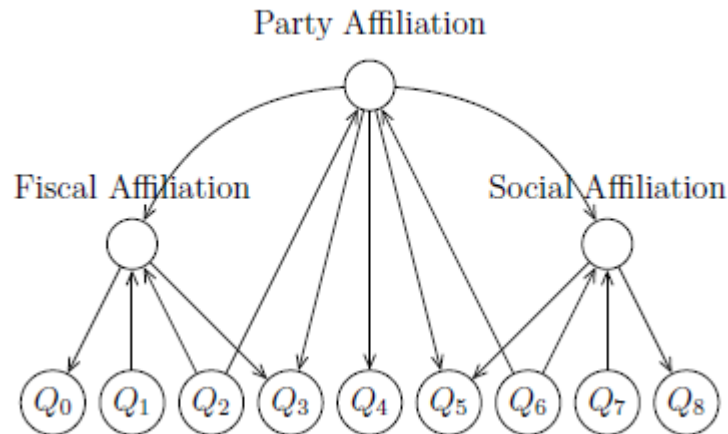


Figure 8: Structural Cultural Model
(with structured latent variables and independent observed data)

If we consider affiliations either liberal or conservative, the answers to Q2 and Q4 together might determine affiliation overall, but the structure caused by the moderates is not eliminated. Someone could be liberal fiscally and conservative socially, their party affiliation might come from their answers to questions that are more aligned by that status.

We would like to infer the structure, b , from the data. This is particularly challenging as the space of possible structures is super exponential on the number of variables. Searching the space of structures combinatorially is quickly infeasible even for small n , so we must sample from the space of structures to find a best fit. In the case of a structural model where there is no missing data we can find explicit solutions, but even the size of this problem grows fast. We are interested in in models with latent variables, the unobserved cultural groups. This compounds the computational difficulties because calculating the marginal likelihoods cannot be done analytically and is difficult computationally.

Chickering and Heckerman (1997) give a comparison of methods in approximating the marginal likelihood for models with incomplete data, and arrive at the Cheeseman-Stutz (1997) method. We can combine this with the Structural EM algorithm to search for our model. As the structural EM algorithm iterates it can either improve the structural model or the parameter estimates. It always converges on a local minimum.

We can simplify the space of structures if we are only interested in the relationships between the groups and the questions, and not the questions among themselves. We consider the groups a set of unobserved random variables with an unknown structure between them, and a structure

between the groups and the questions. We consider the questions independent of each other structurally, which vastly simplifies the search space.

This model is attractive as it gives us information about the relationships between different cultures as well as relationships between different questions and cultures. We can use this information to predict what latent classes a person might belong to, or we can use partial information about a person's responses to predict their responses to the child nodes of the information we have.

4.5 Mixtures of Directed Acyclic Graphs

If we wish to extend the model to free responses, there is a broader structure that might be more appropriate, a multi-DAG (MDAG). An MDAG is a mixture of DAGs, each latent class has its own DAG structure between variables. Here we would be intimately curious about the relations between individual responses, and so we would not consider the questions independent. However, we will once again be without information about the relationship between populations. Implementing the structural search is very difficult, but Thiesson et al. (1998) propose some heuristics to make it tractable.

In an MDAG model there is a distinguished random variable, C , that is the latent class of the observations. Once again the model is like our original model, but we look for structural relationships among responses. We can write the density as

$$p(C = c, \mathbf{X} = \mathbf{x}) = \pi_c \prod_{i=1}^n p(x_i | pa(\mathbf{b}_c)_i), \quad (3)$$

where \mathbf{b}_c is the structural model under group c and π_c is the probability of being in group c under structure \mathbf{b}_c . If we assume C has a multinomial distribution then this is just a mixture of DAG models.

What we lose in information about the group structures is gained in information about the relationship among responses. If we consider the model with free responses, our only option before was to attempt to select a multinomial subset of the responses and use our previous methodology on that. Now we can allow for each response to be an indicator, so a single question can have a set of responses. Each different cultural group might see a different set of responses and substructure information might be present as relationships among the responses themselves.

Consider a very small survey with two questions. There are no fixed answers, the questions ask to list as many ideas as one can think of related to a certain subject. Figure 9 gives a potential graph of the relationship among the responses. We obtain information about which responses imply other responses, essentially factorizing the set of responses for each different cultural group.

An example is if we ask people what their vegetable and flavor of ice cream are. Perhaps the group of people who like vanilla ice cream and broccoli with a few spurious answers is large.

Knowing someone is in this group and that they like cauliflower might give us information on if they like radishes too. Perhaps there is another group of people who predominantly like chocolate ice cream and carrots. Knowing someone from this group likes cauliflower might not tell us anything about if they like radishes.

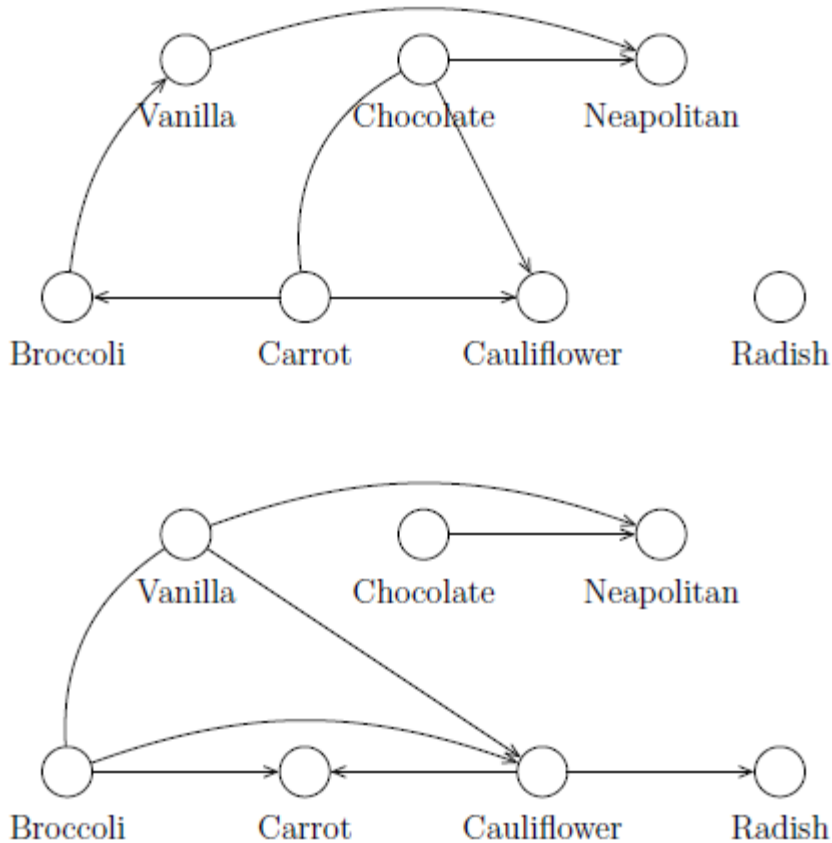


Figure 9: Example of an MDAG Model

(Different DAG structures on the same data for each latent class. The top half and bottom half describe a different relationship over the same responses.)

5.0 RECOMMENDATIONS

The limitations of the original CMM offer a number of opportunities for improvement in such a way that we believe can provide not only a unified statistical methodology, but bring additional theoretical coherence to the field. The key to this progress is developing a structural model approach to characterizing cultural knowledge, and identifying ways to restrict and interpret these models to allow the greatest insights about cultural knowledge.

Consequently, we recommend adopting a structural model approach using independent multiple choice questions and unobserved groups, where the goal is to infer the structure among the groups. Using a combination of the structural EM algorithm and the Cheeseman-Stutz (1997) approximation would be a good first step.

For the more general free response questions, an implementation of MDAGs as in Theisson et al. (1998) should be implemented. The casual interpretation of responses across questions provides a model of not only which cultural groups exist, but also relationships among opinions within those groups.

6.0 REFERENCES

- Arrow, K. J. (1950), "A Difficulty in the Concept of Social Welfare," *Journal of Political Economy* 58(4), pp. 328 - 346.
- Atran, S., Medin, D., Ross, N., Lynch, E., Vapnarsky, V., Ucanek, E., Coley, J., et al. (2002). *Folkecology, Cultural Epidemiology, and the Spirit of the Commons: A Garden Experiment in the Maya Lowlands, 1991-2001*. *Current Anthropology*, 43(3), pp. 421- 450.
- Boroditsky, L. (2001). Does Language Shape Thought: Mandarin and English Speakers' Conceptions of Time. *Cognitive Psychology*, 43(1), pp. 1 - 22.
- Cheeseman, P., Stutz, J. (1997). *Bayesian Classification (Autoclass): Theory and Results*. *Advances in Knowledge Discover and Data Mining*. Menlo Park, CA: AAAI Press, pp. 153-180.
- Chickering, D. M. & Heckerman, D. (1997). Efficient Approximations for the Marginal Likelihood of Bayesian Networks with Hidden Variables. *Machine Learning*, 29(2) pp. 181-212.
- Costa, P.T., Jr. & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual*. Odessa, FL: Psychological Assessment Resources.
- Cowell, R.G., Philip Dawid, A., Lauritzen, S.L., Spiegelhalter, D.J. (2007). *Probabilistic Networks and Expert Systems*. New York, NY: Springer.
- Friedman, N. (1998) The Bayesian Structural EM Algorithm. *Proceedings of the 14th International Conference on Uncertainty in Artificial Intelligence*. San Francisco, CA: Morgan Kaufmann, pp. 129-138.
- Graf, P., & Schacter, D. L. (1985). Implicit and Explicit Memory for New Associations in Normal and Amnesic Subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, pp. 501-518.
- Hofstede, G. (1980). *Culture's Consequences: International Differences in Work- Related Values*. Newbury Park, CA: Sage.
- Hsee, C. K., & Weber, E. U. (1999). Cross-National Differences in Risk Preference and Lay Predictions. *Journal of Behavioral Decision Making*, 12, pp. 165-179.
- Klein, H. A. (2004). Cognition in Natural Settings: The Cultural Lens Model. *Advances in Human Performance and Cognitive Engineering Research*, 4, pp. 278.
- Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a Behavioral Measure of Risk- Taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, 8, pp. 75-84.
- Masuda, T., & Nisbett, R. E. (2001). Attending Holistically versus Analytically: Comparing the Context Sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology*, 81(5), pp. 922-934.
- McLachlan, G., Peel, D. (2000). *Finite Mixture Models*. New York, NY: Wiley.

- Mueller, S. T., & Veinott, E. S. (2008). Cultural Mixture Modeling: Identifying Cultural Consensus (and Disagreement) Using Finite Mixture Modeling. In Proceedings of the 30th Annual Conference of the Cognitive Science Society, pp. 64 - 70.
- Mueller, S. T. (2010). Delineating Cultural Models. Technical Report for AFRL Subcontract No. RQ000746, Prime FA8650-09-D-6939 to SRA, Intl., Project #135599.000 TO 009 to Applied Research Associates. DTIC Document #ADA542838.
- Norman, W. T. (1963). Toward an Adequate Taxonomy of Personality Attributes: Replicated Factor Structure in Peer Nomination Personality Ratings. *Journal of Abnormal and Social Psychology* 66 (6): pp. 574 - 583. doi:10.1037/h0040291. PMID 13938947.
- Palmer, S. E., & Schloss, K. B. (2010). An Ecological Valence Theory of Human Color Preference. *Proceedings of the National Academy of Sciences of the United States of America*, 107(19), pp. 8877 - 8882. doi:10.1073/pnas.0906172107
- Richards, W., Finlayson, M. A., & Winston, P. H. (2009). Advancing Computational Models of Narrative. CSAIL Technical Report No. 2009-063. Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory, Cambridge, MA.
- Robeson, D., Davies, I., & Davidoff, J. (2000). Color Categories are not Universal: Replications and New Evidence from a Stone-Age Culture. *Journal of Experimental Psychology: General*, 129, pp. 369-398.
- Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as Consensus: A Theory of Culture and Informant Accuracy. *American Anthropologist*, pp. 313-338.
- Theisson, B., Meek, C., Chickering, D.M., Heckerman, D. (1998). Learning Mixtures of DAG Models. *Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence*, pp. 504-514.
- Tupes, E. C., & Christal, R. E. (1961). Recurrent Personality Factors Based on Trait Ratings. USAF ASD Tech. Rep. No. 61-97, Lackland Air Force Base, TX: US Air Force.
- Van Den Braak, S. W., Vreeswijk, G. A., & Prakken, H. (2007). AVERs: An Argument Visualization Tool for Representing Stories About Evidence. *Proceedings of the 11th International Conference on Artificial Intelligence and Law*, pp. 11-15; ACM.
- Watson, A., & Huntington, O. H. (2008). They're Here—I can Feel Them: The Epistemic Spaces of Indigenous and Western Knowledges. *Social & Cultural Geography*, 9(3), pp. 257-281.
- Yoon, C., Feinberg, F., Hu, P., Gutchess, A. H., Hedden, T., Chen, H. M., Jing, Q., et al., (2004). Category Norms as a Function of Culture and Age: Comparisons of Item Responses to 105 Categories by American and Chinese Adults. *Psychology and Aging*, 19, pp. 379 - 393.

LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

AI	Artificial Intelligence
AFRL	Air Force Research Laboratory
ARA	Applied Research Associates
BIC	Bayesian Information Criterion
CCT	Cultural Consensus Theory
CMM	Cultural Mixture Modeling
DAG	Directed Acyclic Graph
EM	Expectation-Maximization
MCMC	Markov Chain Monte Carlo
MDAG	Mixture of DAGs
MTU	Michigan Technological University
RJ-MCMC	Reversible Jump Monte Carlo Markov Chain
UK	United Kingdom