Pacific Northwest National Laboratory
Operated by Battelle for the
U.S. Department of Energy

Violent Intent Modeling: Incorporating Cultural Knowledge into the Analytical Process

AP Sanfilippo
FG Nibbs

August 2007

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights**. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847
fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm

This document was printed on recycled paper.
(9/2003)

Violent Intent Modeling: Incorporating Cultural Knowledge into the Analytical Process

AP Sanfilippo
FG Nibbs

August 2007

Violent Intent Modeling: Incorporating cultural knowledge into the analytical process

This project is currently under contract for research through the Department of Homeland Security until 2011 and classified as FOUA. Therefore all references to the name of the particular group of interest will be referred to as "Group X".

After September 11th, the reality of terrorism threats on American soil motivated the government to find new technologies that would provide a technical advantage in preventing such occurrences from happening again. One area of development has been in advanced information technologies which aim to collect, sort, and analyze massive amounts of data. Some of the challenges posed by such a feat is that information is "massive, complex, incomplete, and uncertain, and it encompasses all data forms, languages, and cultures" (Thomas and Cook 2005:2). Under the leadership of the Pacific Northwest National Laboratory (PNNL), a DHS chartered National Visualization and Analytics Center (NVAC) was established with a research and development agenda was set in this process.

Ideally, terrorist attacks need to be stopped before they occur. Analysts need the ability to piece together information buried in disparate data and spot an emerging attack before it can be executed. Textual data comes from documents, speeches, news, video data, and web pages. As these data are increasing in volume, NVAC's goal is to "support analysis of data volumes growing at a rate of one billion new structured messages or transactions per hour, and one million new unstructured messages or documents per hour" (Thomas and Cook 2005:24). Further, PNNL's goal is to use visual analytics to

"facilitate this reasoning process though the creation of software that maximizes human capacity to perceive, understand, and reason about complex and dynamic data . . . " (Thomas and Cook 2005:6).

This process involves a consideration of natural language techniques, and considering the texts from a meaning-centered, or cultural, approach. Thus far, these nuances of communication through language have only been generated and interpreted by humans, mainly, in the field of linguistic anthropology – a branch of anthropology concerned with describing how people represent the meaning of a word in their mind and how this representation is used in constructing sentences. The incorporation of cultural considerations into automated text analysis has been hampered by the fact that culture-based concepts cannot be quickly identified by computer programs. One goal of PNNL has, therefore, been to explore and design new methods of data transformation that build upon anthropological theories to identify and understand the dynamics of a culture in text analysis. It was PNNL's contention when bringing me into this project that an anthropologist could help shed light on these cultural considerations.

I was assigned to the Group Violent Intent Modeling project (GVIM) that is engaged in developing analytic tools to predict violent intent. The group is presently interested in looking at how group leaders influence their constituents to engage in violent acts. I was assigned research on one particular group known to engage in radical behavior that looked at understanding the specific kinds of language patterns its leaders use in public communications that have influenced their constituents to engage in violent acts. The objective of my research was to see if there were specific linguistic patterns that could be identified within the ideological frame of that group, and whether those

structures could be modeled with any reliability to predict radicalization. For ease of reference, I will refer to the group as "Group X" hereafter.

Theoretical Orientation

My research was rooted in Goffmann's (1974) concept of *frames* as used in the study of social movements. For Goffman, frames denoted blueprints of interpretation that enable individuals to locate, perceive, identity, and label occurrences within their life space and the world at large (Goffman 1974:21). These frames function to organize experience and guide action by simplifying and condensing aspects of the "world out there" (Goffman 1974:25). As such, frames have the intention of mobilizing potential adherents and constituents. These interpretive frames are what are known as "collective action frames" (Goffman 1974:27).

Group leaders have been looked at in this literature as "movement actors," that is, people actively engaged in the production and maintenance of meaning for its constituents, bystanders, and observers (Snow & Benford 1988). Whereas collective action frames are action-oriented sets and meanings that inspire and legitimate the activities and campaigns of social movement organizations, it was our contention that locating patterns within the collective action frames of group leaders may lead to predictive models of group radicalization.

According to research done by Benford and Snow (2000:12) collective action frames are composed of two sets of characteristic features: one refers to the interactive, discursive process that attends to these core framing tasks; the second concerns their action-oriented function, or the "core framing tasks" of the social movement

organization. It is this action-oriented function that I was concerned with in relation to my research question.

My specific approach drew on two different theoretical perspectives in framing. The first consideration is rooted in Klanderman's (1984) idea of the performance of rhetoric. He posits that frames are construct in part as a movement actor negotiates a shared understanding of some problematic condition or situation they define as in need of change, make attributions regarding who or what is to blame, articulate an alternative set of arrangements, and urge others to act collectively to affect change. Snow and Benford (1988) refer to these core framing tasks as "diagnostic," "prognostic," and "motivational" framing.

The second consideration to my approach involved culture, specifically, what Fisher (1984) refers to as "narrative fidelity." This approach looks at the extent that the framings are culturally resonant. In other words, to what extent do the frames resonate with the targets' cultural narrations, or what Campbell (1988) would call its "myths," Gouldner (1970) its "domain assumptions," and Rude (1980) "inherent ideology."

Methodology

Scholarly ideological material was gathered and read on Group X and on religious violence in general (Juergensmeyer 2000; Stern 2003; Kaplan and Bjorgo 1998; Stern 1996; Barkun 1994; Ezekiel 1995). From this, a skeletal frame was devised outlining the group's world view.

Next, twenty-two texts from known leaders of the target group were collected from various publicly accessible internet sources. The communications chosen from these leaders varied in their intent from recruitment to radicalization. The various texts

consisted of transcripts of short-wave radio broadcasts, song lyrics, recruitment brochures, letters written to the general public, sermon transcripts, "training camp" meeting transcripts, newsletters articles, dark website messages, and articles from the organizations official journal magazine.

My approach also employed a critical discourse analysis of the texts to understand how the group's ideology is linguistically transposed by its leaders to its various audiences. In other words, to understand how a leader uses certain phrases, concepts, or certain "codes," to manifest their ideology to its target audience. This involved developing a set of codes used by Group X in describing the phenomenon of faith-based radicalization and building a lexicon of associated words within each of those categories. For example, the code of "*Military*" might include words such as *weapons, war, strategy, officers,* etc,.

The next phase of the research would code all 22 documents from Group X wher the codes were sorted into common categories, summarized, and tested for theoretical and empirical relationships. The following phase of research would have other researches code the same documents using the same system coding scheme to get a sense of inter-rater agreement. In other words, if the coding scheme is correct it should be able to work for more than one coder. Where the intent in direct communications is often implicit, an important aspect of this research is to see whether the coding scheme developed can actually be used to identify violent intent; therefore, the next stage of the project would involve application of this coding scheme to direct communications from leaders in the particular movement of interest to test the utility of the code scheme in predicting violent intent. If predictability can be established, the final phase of this project would run a
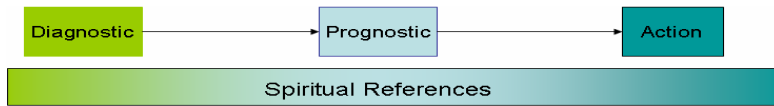
similar battery of tests against a data sample from a broader scope of faith based groups who have also been known to employ violence.

Findings

During the course of analyses specific linguistic patterns of codes emerged. As these code patterns evolved they became the basis for the hypothesis used to test the relevant interpretation and explanation of the texts. These codes, rooted in the ideological material on Group X, and on religious violence in general, were the bones of the skeletal model we used to outline the group's world view. An explanation for the model follows:
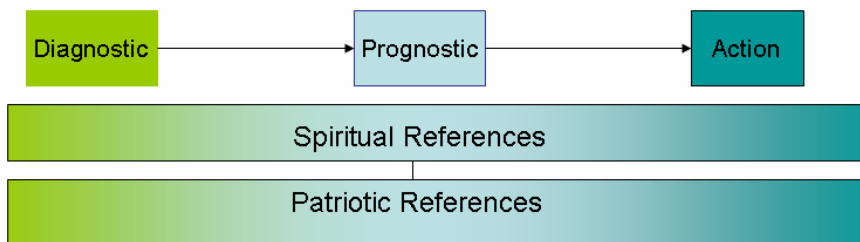
- These leaders of Group X often shroud their activities behind political and legitimate social facades in an attempt to obscure their ideological beliefs and recruitment methods as well as legitimize their message to its receivers. The key overlay used in this movement was identified as Christian fundamentalism. By incorporating what they claim to be fundamental Christian ideology into their communications, these leaders facilitate radicalization by inculcating and converting individuals with their extreme interpretations which function as powerful recruitment tools and make these persons more vulnerable to obscured messages of radicalization. Therefore, the first hypothesis was that movement actors would make extensive use of spiritual references in their communications with reference to all three core framing tasks, diagnostic, prognostic, and motivational, hitherto referred to as "action". This abstraction is illustrated in **figure 1** below:

Figure 1

- According to Group X, members have lost faith in the idea that their nations can be rescued from the clutches of the shadowy conspiratorial entity of the Government (Kaplan and Bjorgo 1998:103). The common discourse of this movement is references that move constituents beyond nationalist loyalties to a world defined by their constructed notions of governance. Religion, broadly conceived, offers a path toward the realization of their conceptualization. It is therefore hypothesized that national and patriotic rhetoric will be heavily mingled with spiritual references at all three core framing tasks, diagnostic, prognostic, and action (motivational). This abstraction is illustrated in **figure 2** below:

Figure 2

After extracting those core framing tasks, the 22 texts were separated into two piles representing what we already knew to be their intent. The communications intended for audiences outside the critical mass of Group X, such as recruitment brochures, introductions to their organization, or interviews with news organizations, and, therefore,
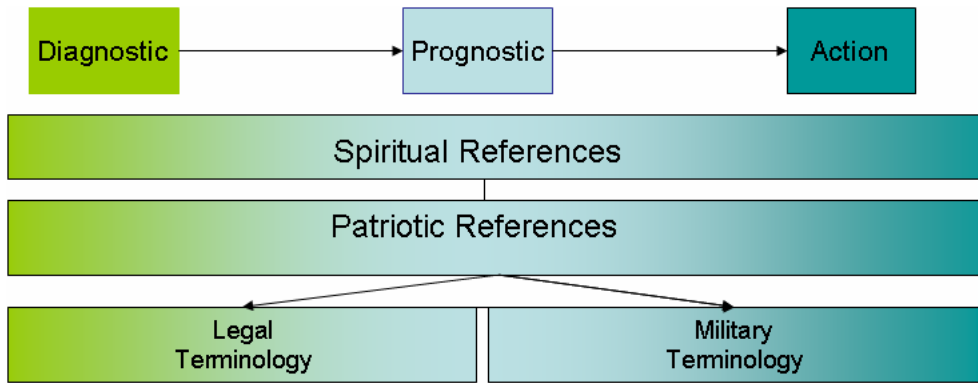
had a non-violent intent, were put in one pile. In the other pile were the other 11 communications directed toward the group's core of believers urging them to carry out a specific violent mission. A critical discourse analysis was done of these two data sets to understand how the leaders linguistically frame these different intents within the ideology of Group X.

For instance, leaders talking to those outside of their core constituents tended to pare spiritual and patriotic terms to legal ones and operate within the diagnostic frame. For example, "Christ wants us to communicate this evidence and judge those guilty of these horrors." Or, "The Bible has been a part of our legal system for almost 400 years," and "God will clearly prove our case to the nation and prosecute the unjust . . ."

On the other hand, when a leader addressed their own core constituents they switched to operate within action frames, and the legal references were replaced with military ones. For example; "God's chosen elite fighting force. . . " Christ our Savior, our General . . ." We need to be trained in that armor and in using the weapons of warfare given to us by our commander-in-chief, our King, and Master, Jesus Christ." The abstractions of legal and military terminology were added to the skeletal frame, as illustrated by **figure 3**.
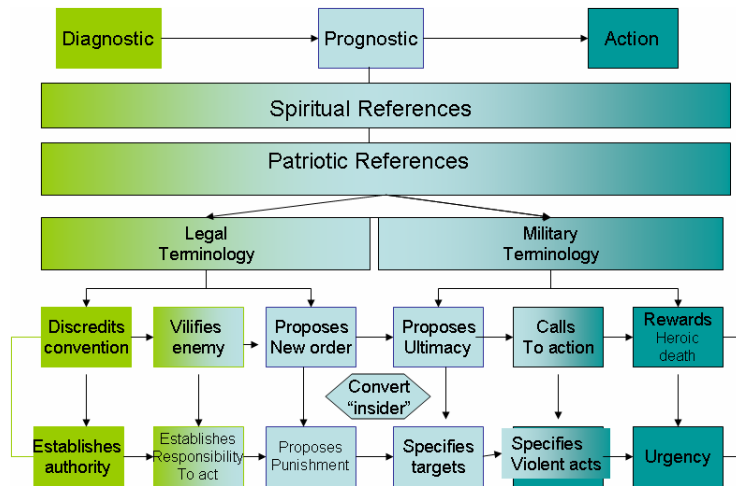
Figure 3

- Within the domains of diagnostic, prognostic, and action, sub-sets of codes within each frame emerged as they were used in conjunction with legal or military terminology. For example, certain linguistic codes seemed more prevalent in rhetoric intended to introduce outsiders to an ideology. These codes were placed in the diagnostic frame. As the intent of a leader's message moved toward proselytization, the discourse changed. These codes fell into the prognostic frame. Once a group of people accepted these beliefs, and were considered "insiders," the rhetoric directed toward them changed again, albeit, still within the prognostic frame. Finally, as the intent became to sway these insiders toward radicalization, the discourse fell within the action frame. These sub-sets of codes were identified as:

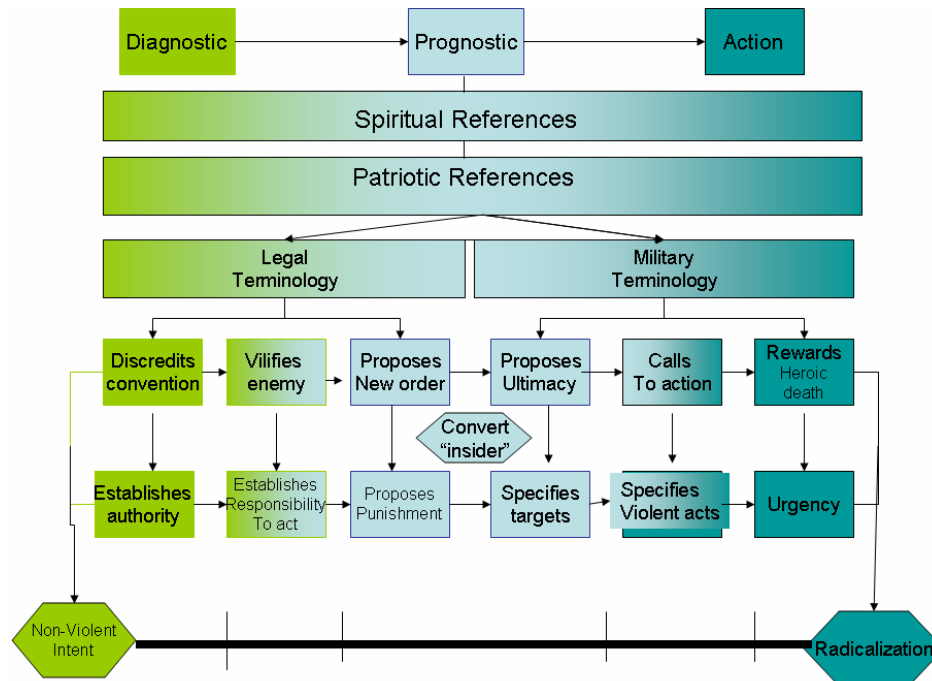| Diagnostic	Prognostic	Action
Discredits Convention	Proposes new order	Calls to action
Establishes authority	Proposes punishment	Specifies violent acts
Vilifies enemy	Proposes ultimacy	Rewards heroic death
Establishes responsibility	Specifies target	Urgency

These abstractions were added to the skeletal frame, and are illustrated in the **figure 4.**

Figure 4

- Taken together, this skeleton of collection action frames built from codes derived from the direct communications of leaders of Group X, reveals a nominal scale of progressive intent from non-violence to radicalization. An abstraction of this scale is illustrated below in **figure 5**.

Figure 5

In the next phase of research I built a lexicon of associated words within each of those codes. For example, the code "*Spiritual*" might include in its lexicon an array of Biblical words, phrases, and references, like "angel," Deuteronomy," and "holy." All 22 texts were exhausted for lexical words until no more than one new word per. document was found. Those codes and their lexicons were then entered into the computational linguistic program, Qualrus (http://www.ideaworks.com/qualrus/index.html). Quarlrus is a qualitative content analysis computer program in which researchers can mark segments of a text with qualitative codes then retrieve and analyze those codes, producing useful summaries and overviews.
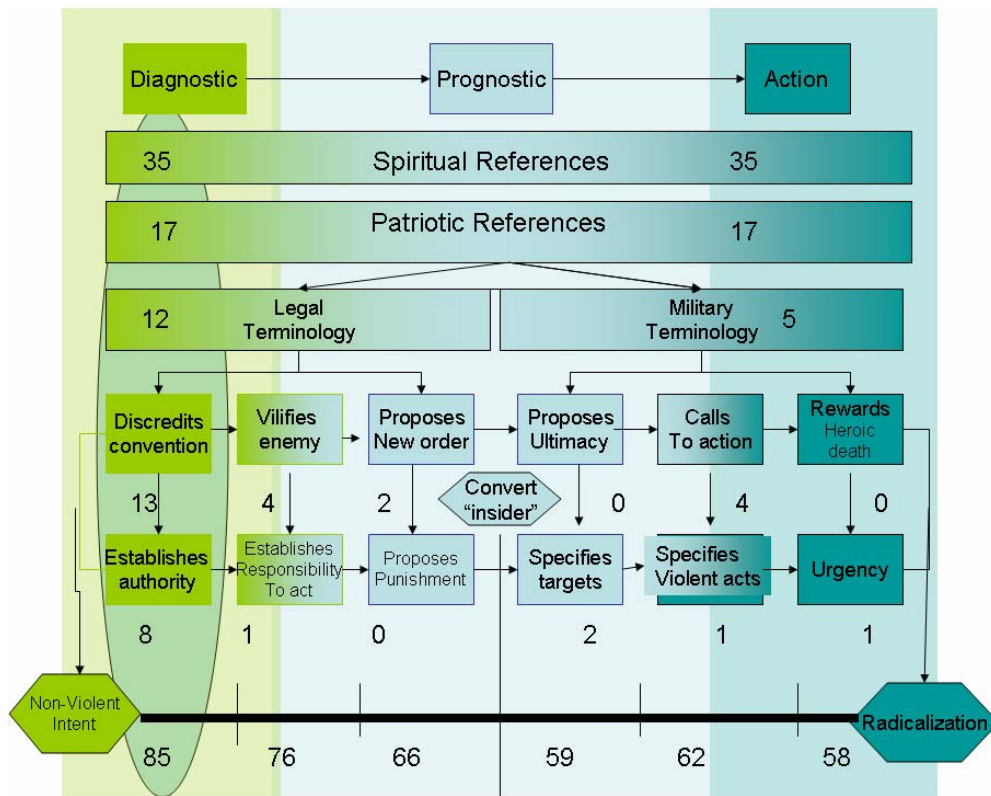
All 22 texts were then loaded into Qualrus omitting any previous designation of intent and then coded by segments (a segment being defined as one sentence in a document, and as one line in a poem or song). The documents along with a description of my coding scheme, lexicons, and examples for coding a segment, were also given to another researcher in Seattle to test for inter-rater agreement. While there was a 100% agreement between the two coders, more raters under more controlled circumstances will need to test the coding scheme to assure its accuracy. This will be taken up on a future phase of the project.

The last part of the project that I participated in used the statistical overviews of the texts provided by Quarlrus to see if the model could accurately identify violent intent. My hypothesis of group leader violent intent was based on simple quantitative principles. Code occurrences in a document are aggregated according to specific combinations in which they appear. Those combinations of interest are based on where they appear on a nominal scale of frames from "diagnostic" to "action." The highest aggregate segment of

occurrences indicates whether the particular communication being analyzed was intended to radicalize its listeners or carries a non-violent intent. Two tests of this hypothesis are illustrated below.

The code boxes in **Figure 6** are filled in with their corresponding number of occurrences in one document.
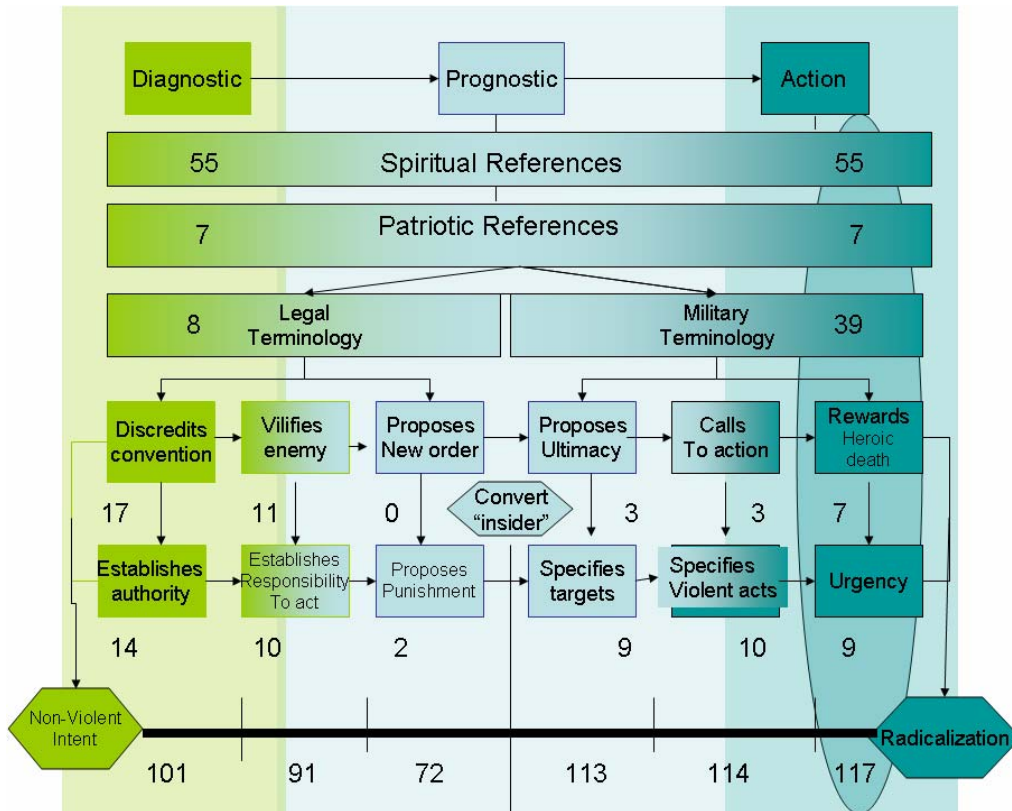
Figure 6

Since the Diagnostic column has the largest cumulative total (85), this indicates that the communication by the group leader was within a diagnostic frame, and since it falls to the extreme left of the scale, would be said to carry a message of non-violent intent.

The next illustration in **Figure 7** shows the results of another coded document that fell at the other end of the spectrum. With the highest aggregate total (117) falling within

the Action frame, and to the extreme right of the scale, it indicates that the intent of this

leaders communication is to radicalize its listeners.

Figure 7

At the end of this phase of research, the null hypothesis was rejected as the

placement on the scale of both documents did in fact correctly indicate what we knew to

be their intent. The hypothesis was then tested on the remaining 20 texts from leaders of

Group X where the coding scheme applied to this model continued to accurately reflect

their known intent 100% of the time. Equally important, the results indicated the

usefulness of cultural considerations and anthropological theories in the analytic process

of data transformation.

Limitations and Future Possibilities

We now turn to a brief examination of the implications of the outcomes of this research in relation to the previously outlined goals of the Group Violent Intent Modeling project and DHS.

While this particular hypothesis of radical intent and relative set of collective action frames appears most promising, one of its key constraints, as in other collective action frames associated with particular movements, is its limitations to the interest of the particular group or to this particular set of related problems. As one stated goal of GVIM is to define a broader set of master frames that constrain the orientations and activities of a variety of movements, its applicability to a wider variety of groups would have to be tested.

Another limitation is that literature suggests that frames are not static. In one ethnographic study of framing processes, Snow & Miller (unpublished data) posited that collective action frames are continuously reconstituted during the course of interaction that occurs in the context of gatherings, and that the key to understanding the evolution of frames resides in the articulation and amplification processes rather than in the topics or issues comprising the frames. This leads to questions of shelf-life for such a model. Researchers would have to address how often the model should be updated by cultural specialist to account for relevant shifts in a groups discourse.

Moreover, before this research can be transitioned into practice, it must undergo further evaluation within the contexts of its intended use. Progressively, more sophisticated evaluations will be required with more representative, but still unclassified, test data to evaluate and adjust the coding scheme and levels of predictability. The

resulting theory must be evaluated to ensure that it represents a significant advance over current practice and to ensure that it operates successfully.

To truly leverage the successful research results described in this paper, these results must eventually be moved into practice. Ultimately, that means a piece of software will be developed and tested in the lab that directly replicates the target production environment and will require close collaboration among government and research communities. The hopeful outcome will be that the resulting research matures into software that finds broad usage for DHS.

On Working at PNNL

Working at PNNL on this project has allowed me several exceptional opportunities for academic and professional growth.

Firstly, the PNNL atmosphere is set up so that every project is handled by a group of diverse experts. In graduate school, although there is plenty of opportunity for group discovery, the groups are often esoteric. This internship has allowed me to benefit from a "think tank" environment that was both energetic and challenging, requiring that I always pull my weight by bringing my best scholarship to the table.

Secondly, it allowed me to experience the integration of theory and practice. In academia, often the end result of research is pure knowledge. At a national lab the goal is usable product. I was able to witness first hand how theoretical outcomes were integrated into this process.

Thirdly, the timing of this particular project allowed me the opportunity to present my research to its intended client. During my third week of the internship, DHS came for a scheduled visit to the lab to inquire on the progress of their visual analytics contract.

Consequently, I was asked to prepare a short presentation for them on the research I was doing. Over the two days that DHS was here I was able to participate in several meetings, gave my presentation, answered questions, got directional feedback, and enjoyed informal networking at dinner parties held in their honor. I received incredibly positive feedback from both DHS as well as my mentor on my presentation and ideas. It was very empowering.

Finally, this lab and my mentor were wonderfully hospitable making this internship a joy and pleasure. PNNL hosts quite a few interns over the summer from different programs and therefore puts together a plethora of intern activities, tours, trips, cookouts, etc., for us to enjoy. While here in Washington I hiked on Mount Reiner, went on a wine vineyard tour, biked along the Columbia River, ate at least three company sponsored picnics, went on a tour of the Hanford Site, listened to host of special topic lectures form lab directors, and got sent to Seattle on a three day business trip. This lab highly valued its DHS interns and we were given special status and additional "treats" by the office of Student Internships throughout the summer. On top of that, my mentor and his wife took me out and entertained me at their house on numerous occasions. Three of us DHS fellows were able to find each other early in the summer and spent many occasions together developing what I hope will be lifelong friendships.

Most importantly, this experience has stretched me to consider working in a national lab doing this kind of work after I receive my PhD. Had I not come here, I would not have had any idea of the kinds of work done in these facilities nor the environment in which it is conducted. As my hypotheses has proven to have promise, PNNL has asked me to consider continuing this research for them on a 10 hour p/week basis through a

contract they are developing with out university for this purpose. Moreover, I have been invited by my mentor to co-author a journal article with him on this research and to present it at a professional conference next spring.

Overall I found my internship to be a very rewarding experience that allowed me to apply my anthropological education in numerous ways, the least of which was the analysis of overwhelming amounts of information that could potentially predict emerging threats.

Works Cited

Barkun, Michael. 1994. Religion and the Racist Right.: The origins of the Christian Identity Movement. Chapel Hill: University of North Carolina Press.

Benford, RD, David Snow. 2000. Framing Processes and Social Movements: An overview and assessment. *Annual Review of Sociology* 26: 11-39.

Brent, Edward, Pawel Slusarz and Alan Thompson. 2002. Qualrus: The intelligent Qualitative Analysis Program. Thousand Oaks, CA: Idea Works Inc. http://www.ideaworks.com/qualrus/index.html

Campbell J. 1988. The Power of Myth. New York: Doubleday

Ezekiel, Raphael. 1995. The Racist Mind: Portraits of American Neo-Nazis and Klansmen. New York: Viking Press

Fisher WR. 1984. Narration as a human communication paradigm: the case of public moral argument. *Communication Monogram*. 51:1-23

Goffman E. 1974. Frame Analysis: An Essay on the Organization of the Experience. New York: Harper Colophon

Gouldner AW. 1970. The Coming Crisis in Western Sociology. New York: Basic Books

Juergensmeyer, Mark. 2000. Terror in the Mind of God: The Global Rise of Religious Violence. Berkeley: University of California Press.

Kaplan, Jeffery and Tore Bjorgo. 1998. Nation and Race: The developing Euro-American Racist Subculture. Boston: Northeastern University Press.

Klandermans B. 1984. Mobilization and participation: social-psychological expansions of resource mobilization theory. Am. Soc. Rev. 49:583-600

Rude G. 1980. Ideology and Popular Protest. New York: Knopf

Snow DA, Benford RD. 1988. Ideology, frame resonance, and participant mobilization. Int. Soc. Mov. Res. 1:197-218

Stern, Jessica. 2003. Terror in the Name of God: Why Religious Militants Kill. New York: Harper Collins.

Stern, Kenneth. 1996. A Force Upon the Plain: The American Militia Movement and the Politics of Hate. Oklahoma: University of Oklahoma Press.

Thomas, James and Kristin Cook. 2005. Illuminating the Path: The Research and Development Agenda for Visual Analytics. National Visualization and Analytics Center. Los Alamitos, CA: IEEE Computer Society