

Changing Discriminatory Norms Using Models of Conceptually-Mediated Cognition and Cultural Worldviews

Daniel Olsher (dan@intmind.com)

National University of Singapore

Abstract

Cognitive modeling can provide immense social benefits, especially in the case of unconscious processes causing significant psychological and societal distress. One such process, discrimination against minorities, is typically grounded in unconsciously held stereotypes reflecting deep ignorance of the realities minorities face. Another, suicide terrorism, results in significant suffering for many communities. Related beliefs often arise from unquestionable values, norms and worldviews, however, making direct/conscious change appeals unworkable. Building upon cultural knowledge representation and cognitive modeling, this paper shows how change could be effected via cognitive operations on conceptual worldviews. After introducing a novel framework for modeling conceptually-mediated belief systems and techniques for guided dissonance reduction and network change, the paper applies these to anti-discrimination and terrorism reduction. Such work holds great potential for better aligning perception of minorities with the realities they face, reducing suffering, and disrupting psychological processes dependent on improper views of stigmatized minorities and other phenomena.

Keywords: Culture; Cognitive Models; Nuance; Knowledge Representation; Norm Change; Discrimination Reduction; Terrorism Reduction

Introduction

Cognitive Science has the potential to provide immense social benefits, especially in shedding light on processes that are normally unconscious but which tend to cause significant psychological and societal distress.

One such process is discrimination, often grounded in stereotypes reflecting deep ignorance of the realities of what minorities are and what they face. Another is the process by which people decide to undertake violent action during conflict.

Because the beliefs underlying such processes often arise from norms, cultures, values, and other strongly unconscious, often nearly unquestionable, belief systems, direct conscious appeals are often unworkable and insufficient for effecting change.

In this paper, we consider how to construct procedures capable of achieving change in beliefs by modeling the beliefs people already hold (and the connections between them) and designing procedures tailor-made to effect change.

Incremental changes are most effective. The use of familiar concepts lowers fear, and as the status quo already appears to be 'reality', small changes to it are more likely to succeed. Direct/conscious attempts to change such beliefs often fall afoul of cognitive and social defense mechanisms, with excessive change causing dissonance, negative emotional outcomes, and ultimately, rejection.

This paper introduces a framework for modeling conceptually-mediated belief systems, a novel change

strategy entitled Directed Dissonance Reduction (DDR), a conceptually-mediated formulation of inoculation theory, and a brief discussion of techniques for modifying belief networks. Examples (including sample networks) are then given demonstrating how to apply these tools in the prejudice and terrorism reduction domains.

Beyond norm change, this research seeks to advance nuanced knowledge representation of worldviews, cultures, and norms as well as to advance our understanding of conceptually-mediated cognition.

Core Representation Frameworks: COGVIEW and INTELNET

In this paper, the author's COGVIEW and INTELNET formalisms are used to model nuanced conceptual worldviews. Both formalisms are the subject of ongoing development, with this paper providing an early demonstration of their combined capabilities.

COGVIEW is a network-based framework for representing complex worldviews through 'fields' of interconnected concepts and energy flows. It serves as a base for simulating important psychological, conceptual, and cognitive semantics processes mediated by worldviews. COGVIEW reflects the insight that information originating from human cognitive processes is of a different type than that typically considered in the 'hard sciences'; namely, it is *deeply nuanced* and *inherently distributed*.

COGVIEW uses networks of concepts to store information. Networks are defined as collections of nodes connected with edges (or links). Energy originates from energy sources and flows across edges (capable of modifying the energy that flows across them). Each edge has an indicated direction in which energy traversing it will flow.

COGVIEW supports two levels of cognitive processing, the C(onscious)-Level and the U(nconscious)-Level. U-Level processing is not accessible to conscious awareness and is dominated by the effects of associational memory in that the primary activity taking place at this level is the spreading of energy between concepts and concept fields and the associative detection of congruence/similarity between various portions of the extended concept universe. Critically, 'rational' or 'logical' thought does not take place at this level and is therefore incapable of acting as a filter for change. Growing evidence (see Shermer, 2002) suggests that at this level, mere understanding may be cognitively tantamount to acceptance, requiring an active act of 'disbelief' to be overcome. 'Raw' concepts and a wide range of energies can be accessed directly here and immense amounts of energy easily shifted and deployed, as well as emotions triggered, all without conscious or rational intervention.

Clashes In COGVIEW, a *clash* occurs when energy flowing in one direction meets energy flowing in the other, an especially important occurrence when the positive or negative valences of clashing energies are opposed to one another.

Clashes represent the point at which conceptual incompatibilities become manifest, and when the ability and/or need to take some sort of meaningful action has been identified, that is, when the model makes a certain type of 'conclusion'.

The subnetwork within which a clash occurs will naturally be activated most when a result propagates to consciousness, and thus is the subnetwork from which the conscious knower will consider the 'insight' of the clash to have originated. Critically, this is true even (as is the common case) when the energy which caused the clash came from other concept fields. This phenomenon allows for persuasive processes in which the semantic/conceptual content of a clash is removed from its original context (usually one in which conscious processing would have caused the desired conceptual understanding to be rejected) and shifted to another context (i.e. another subportion of the conceptual cloud) wherein which the attribution of the clash output to that context is 'safer' (or more desired).

The sites of clashes often coincide with the most relevant and important (moral) 'issues' that a human would identify within particular conceptual fields. This phenomenon provides one source of evidence that specific COGVIEW networks are accurately reflecting certain real world semantics.

Arguably, clashes have a neural basis; for example, Lieberman, Schreiber, and Ochsner (2003, p. 689-690) suggest that "The C-system [conscious] (named for the "c" in reflection), consisting of the prefrontal cortex, anterior cingulate cortex, and medial temporal lobes, is recruited when the X-system [unconscious] fails to create coherent outputs from the different sources of input. ... According to this model, the C-system is usually involved only to the extent that the X-system fails to resolve the current set of inputs into a coherent output ... [I]f the conflict between different considerations is too large, the C-system will detect this tension in the X-system and become involved (Botvinick, Braver, Barch, Carter, and Cohen, 2000)."

Tools for Effecting (and Inoculating Against) Belief Change

In this section we introduce tools for use both at the macro and COGVIEW network levels.

Network-Level Tools

This section provides a very high-level listing (due to space limitations) of techniques for adjusting COGVIEW networks. While useful on their own, a high-level understanding of these techniques is also helpful in understanding Directed Dissonance Reduction.

1. Break Link Between Two Concepts: Disconnects energy sources from downstream concepts (can be accomplished via DDR).

2. Introduce new connections to compete for energy with pre-existing connections.
3. 'Attraction': Prime concepts likely to attract energy and thus create more desirable energy flows. Priming of 'incompatible' concepts (those causing dissonance with other salient concepts) may cause confusion, weakening previous connections.
4. Use 'attractor cognitions', that is, new cognitions more pleasant than undesired ones. Prototypical examples: 'hot-button' concepts like 'family'. Energy from these may 'overwhelm' that of undesired concepts or paths.
5. 'Blocking cognitions' capable of blocking flow along undesired paths. To accomplish, associate original path with something unpleasant, confusing, or fear-inducing. Highlight path nonoptimality by pointing out logical inconsistencies or suboptimal conclusions arising from the path.
6. Create Links: Repeatedly reference pairs of concepts together, point out semantic similarities, and/or create scenarios where concepts are always associated or required to be used together to accomplish a goal.
7. Emphasize / De-emphasize Concepts: De-emphasize through neglect - emphasize through repetition.
8. Redirect Link: Prime desired concepts and paths. Reinforce through Repetition and Need-To-Use. Repetition: use narratives or linguistic constructions continually invoking and refreshing the salience of desired concepts or paths. Need-To-Use: cognitions become more salient when they are necessary to accomplish important goals.
9. Change Polarity (what was once sending negative energy now sends positive, and vice versa): Represents significant changes in perspective. Overwhelm energy of one polarity with energy from a source of the opposite polarity.
10. Re-Normalize: Show that concept 'X' is actually much more like concept 'Y' than was originally believed to be the case.
11. Stigma Disconnect: Stigma is a highly potent source of negative energy. Stigma Disconnect breaks energy flows associated with stigma as early as possible and/or refits them with more positive energy sources.
12. Concept Implantation and Reconstruction of Incorrect Concept Fields: Entire concept fields may be incorrect, either in terms of content or in connections.

Associations may be reinforced if they cover entire concept chains - that is, from energy input nodes to output nodes, as this provides cognitive 'closure' and makes chains easier to process.

Directed Dissonance Reduction

Dissonance reduction (Festinger, 1957) represents a powerful mechanism for effecting psychological change, capable of marshaling significant emotional and mental energies. In this section we reformulate this mechanism in terms of COGVIEW networks, and later demonstrate how it may be repurposed in service of both belief change and 'inoculation'.

DDR redefines dissonance reduction as a process of *positive energy maximization*, suggesting that people seek to maximize the most positive energy obtainable within their belief networks. Dissonance is understood as a threat to current energy maximization, and dissonance reduction as an attempt to create a new equilibrium. Motivation arises from discomfort, with motivation levels proportional to the potential amount of threatened energy change.

Stages of Directed Dissonance Reduction

Phase I. Design

Phase I analyzes subjects' COGVIEW networks and identifies clash nodes, important positive energy sources, and potential psychological defenses.

COGVIEW allows putative belief changers to identify where concepts and energy sources lie in relation to others and to determine what concepts to prime and interconnect in order to 'reach' desired energy sources.

Phase II. Initiation (Build Initial Context)

Priming of concepts identified in Phase I takes place here. This phase may also seek to add new useful concepts or links, though strategies built significantly on pre-existing networks are likely to be more effective, reliable, and easier to implement in practice.

If subjects perceive that large negative changes in energy balance may result, denial may occur. In order to avoid this defense strategy, significant positive energy sources (such as FAMILY) may be identified and primed at this stage.

In some cases, negative energy sources may be primed, for example, if subjects believe that there is 'no problem' or that a potential dissonance-inducing stimulus is of little or no significance. A canonical example is the use of fear by terrorists to cause target populations to take their goals seriously. Another involves health appeals, in which subjects may not be convinced of the reality of possible threats.

Phase III. Dissonance Introduction

In this stage, a stimulus - a single or small set of new links whose existence threatens the existing positive energy balance - is introduced, causing dissonance. These must be significant enough to threaten belief network energy maximization.

Care must be taken here to avoid resort to denial as a clash defense strategy; if the target is able to simply ignore the stimulus altogether, it will not be effective. Thus, stimuli should be factually accurate, important, and otherwise impossible to ignore.

Phase IV. Directed Dissonance Reduction

Here, energy is introduced and links are adjusted so that dissonance will be resolved as desired. This stage attempts to direct the choice of energy maximization equilibrium; energy placed at clash sites is most beneficial.

Phase V. Solidification

This phase solidifies energy at important clash points in order to enhance satisfaction and attitudes towards outcomes, as well as make it less likely that changes will be reversed or re-examined.

The more likely it is that the new energy balance will persist in future, the more anxiety will be reduced during the dissonance reduction process.

Directed Dissonance Reduction Examples

We now provide two detailed examples of how DDR can be used in diverse domains.

Coming Out

The first example illustrates the case of a son who comes out to his father, whose view of the situation is initially quite inaccurate. The goal of DDR is that the father maintain his positive associations and attitudes about his son.

Figure 1 provides a diagram of the father's COGVIEW network before DDR has taken place. Energy flows from the energy source nodes, following the arrows indicated in the diagram, with two properties: magnitude (amount) and valence (positive or negative). Whenever energy crosses a graph edge labeled 'NEG', the valence of that energy is reversed.

Energy derives from three sources. *Biological* energy arises from parental love, the need to guide children, parent-child and father-son bonding, the desire to raise well-adjusted children capable of participating in the wider world, and the psychological need to view oneself as a good parent. *Societal* energy arises from in-built needs to be seen positively by the community as a good parent and as someone who has raised a well-adjusted child. Lastly, *Cultural* energy draws on local imperatives to guide children, act as good parents, and raise moral children.

Phase I (Design): DDR goals are as follows: *break* the link between INCORRECT-BELIEFS-ABOUT-GAYS and GAY (Link *IncorrectBeliefs*) and *maintain* the link between IDEAL-SON and THE-SON (Link *IdealSon*). The unfavorable outcome is the opposite, that is, the breaking of link *IdealSon* and the maintenance of the *IncorrectBeliefs* link. Another undesired outcome could be the breaking of the link between THE-SON and FATHER - in other words, the father disowning the son.

Clash points occur at FAMILY, RAISING CHILDREN, and THE-SON. At FAMILY, positive energy (+50) comes from SOCIETY and significant negative energy from COMBINED-ENERGY (initially positive but reversed after crossing the NEG edge near INCORRECT-BELIEFS-ABOUT-GAYS). Without mitigation this will flow into IDEAL-SON, ultimately making this path one that contravenes DDR goals. If positive energy is introduced at this node through priming or statements such as 'your son wants to raise children and start a family', 'gay people want to start loving families', and/or associating FAMILY with significant positive energy, positive energy may 'overwhelm' the negative energy at the clash point. A similar technique may be applied at RAISING CHILDREN.

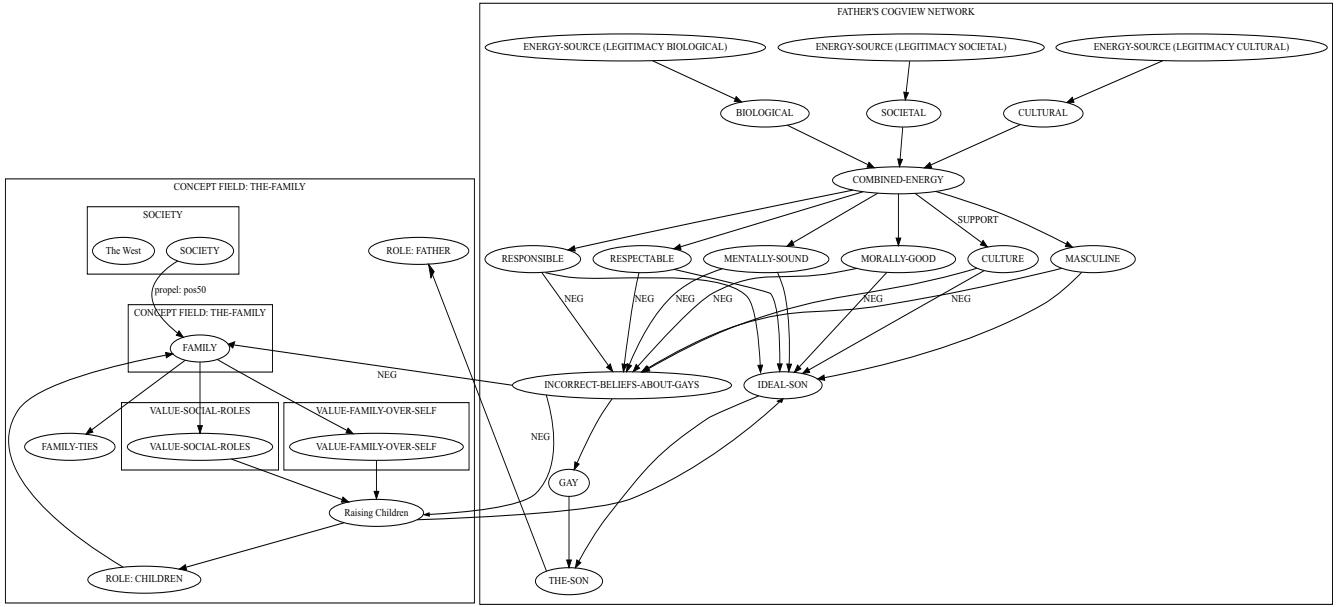


Figure 1: Father's COGVIEW Network before DDR

The clash located at THE-SON is particularly interesting, as it points to the exact dilemma faced by the father in this situation: either accept his son and break links to his incorrect beliefs or vice versa. Such correspondences can help verify how well COGVIEW networks fit the situations they represent.

Energy does not actually flow from INCORRECT-BELIEFS-ABOUT-GAYS to THE-SON until the dissonance-inducing cognition (linking GAY to THE-SON - the *ComingOut* link) is introduced, matching the reality of the problem being modeled.

Because THE-SON connects to FATHER, reminding the father that his son is a 'good' son will introduce positive energy into this node and make it more likely that he will not want to disown his son so that he may remain associated with this positive energy source.

Other positive energy sources include caring/compassion, parental attachment, the desire for enhanced well-being, the desire to focus on strengthening family and not break it apart, and so on. Other potential strategies could include creating new links between desirable nodes or changing the polarity of existing links between personal attributes (RESPONSIBLE, RESPECTABLE, etc.) and GAY or THE-SON.

Phase II (Initiation): In Stage II, the priming-related components of the Phase I design are carried out.

Phase III (Dissonance Introduction): In this stage, the link (from GAY to THE-SON), labeled as *ComingOut* above, is introduced by the son coming out to the father. Denial must

be avoided, which would appear as refusing to believe the son is gay.

Phase IV: (Directed Dissonance Reduction) At this stage, the full strategy designed at Phase I is implemented. Positive energy is inserted at FAMILY, RAISING CHILDREN, THE-SON, and so on. This stage makes use of link change strategies 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, and 12 (see above).

The goal here is to make it more painful for the father to configure his belief network in a manner contrary to the goals of DDR - that is, if he breaks or maintains non-preferred links, the maximum energy balance (equilibrium) he will achieve will ultimately be lower than that if he configures his network as desired. As an example, if the father breaks the link from IDEAL-SON to THE-SON, he will cease to receive the benefit of the positive energy entering the node THE-SON from IDEAL-SON. A similar outcome will arise if he breaks the link between IDEAL-SON and THE-SON.

Phase V (Solidification): Operations remain at Phase IV until sufficient change has been effected. To test, we may ask the father how he feels about his son, tallying the total number of positive and negative qualities.

Further positive energy may be added to THE-SON by focusing on the father's need that his son and other family members be associated with positive attributes, and re-priming may be performed on FAMILY, CHILDREN, and so on.

Terrorism Reduction

DDR also has significant application in the domain of terrorism reduction. DDR can support public communications de-

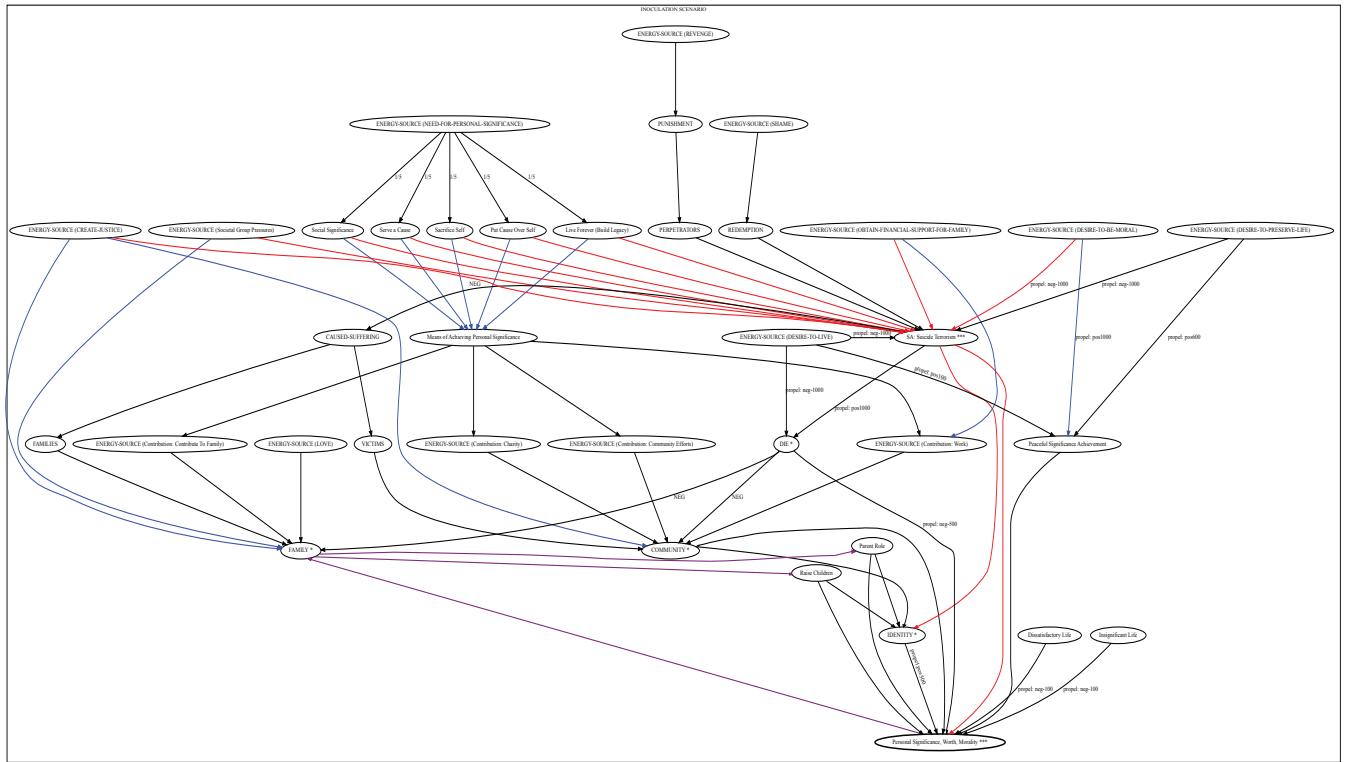


Figure 2: Citizen COGVIEW Network before Persuasion

signed to slow the adoption or 'taking root' of radical ideologies in susceptible individuals (termed 'inoculation'). Focus can also be placed on interrupting the flow from conception to the carrying out of acts.

We focus here on 'inoculation' against suicide terrorism, as it is a highly effective strategy capable of blocking entire event chains from occurring. Suicide terrorism's reliance on interlocking societal and personal value systems is an excellent fit for the COGVIEW+DDR approach.

We draw on Kruglanski et al. (2009)'s view of suicide terrorism as a quest for personal significance. We also draw on 'inoculation theory' (McGuire, 1961; Pfau et al., 1997), which suggests that subjects may be 'inoculated' against later attempts to change their beliefs by being exposed to weakened forms of counterarguments to their current positions. 'Beliefs' are understood here as COGVIEW links relevant to suicide terrorism and 'attempts to change beliefs' as attempts to convince others to undertake suicide terrorism.

Inoculation seeks to protect existing beliefs against change, assuming that people initially believe that suicide is wrong and that they should seek to make their life meaningful through appeal to family and activities rather than through terrorism.

Inoculation consists of two phases: *Threat* and *Refutational Preemption* (Pfau et al., 1997, 188-189). Threat motivates belief protection by demonstrating that beliefs (and energy maximization) can be threatened. Refutational Preemption inoculates against future threats by presenting weak-

ened forms of counter-arguments (which may ultimately be the same or different as what are eventually encountered, as the effect works in both cases.) (Pfau et al., 1997, 188-189) (McGuire, 1961, 330).

DDR assists (and induces) participants to 'create their own' counterarguments against persuasion attempts. As Pfau (1997, 189) suggests, "[i]t is the motivation provided to generate answers to potential counterarguments, as opposed to the specific information provided, that is responsible for [conferring resistance]."

By highlighting anxiety related to energy equilibrium loss, DDR can help explain findings related to distraction and differential performance of inoculation as a strategy (see, for example Keating & Brock, 1974).

Applying DDR Threat occurs at Phases II and III of DDR, and Refutational Preemption at Phases IV and V.

Phase I (Design - Threat): Identify nodes expected to lose significant positive energy if persuasion occurs: FAMILY, COMMUNITY, IDENTITY, PERSONAL SIGNIFICANCE, WORTH, AND MORALITY and PEACEFUL SIGNIFICANCE ACHIEVEMENT. Also threatened is the *loop* between FAMILY and PERSONAL SIGNIFICANCE, WORTH, AND MORALITY. Loops such as these serve to simulate stability; the loss of stability represented by energy flowing within stable feedback loops is a significant loss in and of itself.

Simulate the introduction of negative energy into threatened

nodes (and loss of stability). This can be accomplished by calling attention to the following potential outcomes of energy loss:

FAMILY: Children without parents, not close to loved ones, loss of significance, judged as immoral, not supportive, broken family (leading to suffering, shame), ostracism, discrimination, prejudice, family suffering from revenge taken by others.

COMMUNITY: Chaotic communities, crime out of control, suffering, no rule of law, fear, bad example, children hurt, not growing up well, bad publicity for cause, martial law.

IDENTITY: Unclear about role in life, no personal meaning, unsure how to live life, deep instability, lose identity as good parent, good spouse, good member of community.

PERSONAL SIGNIFICANCE, WORTH, AND MORALITY: loss of face/worth to family, community, and friends, seen as wrong and immoral, family members feel societal shame, ashamed, unhappy due to lack of access when in need.

PEACEFUL SIGNIFICANCE ACHIEVEMENT: never-ending violence throughout the community, loss of life, family, and friends, international condemnation, loss of resources, great danger, great instability, fear throughout.

Loss of stability loop between FAMILY and PERSONAL SIGNIFICANCE, WORTH, AND MORALITY: great uncertainty, chaos, lack of predictability surrounding life.

Refutational Preemption: For this component we identify clashes at the following nodes: SIGNIFICANCE ACHIEVEMENT: SUICIDE TERRORISM (3 clashes), PERSONAL SIGNIFICANCE, WORTH, MORALITY (3 clashes), FAMILY, COMMUNITY, DIE, and IDENTITY.

Clash nodes are those most contested from a good-vs-bad moral perspective; at these sites worldviews may be modified significantly and immense shifts may occur. Clash points are marked with asterisks in Figure 2; the number of asterisks denotes the number of clashes occurring at any given node.

Phases II and III (Initiation and Dissonance Introduction) involve execution of the Threat component. Threat primes appropriate concepts and the existence of a threat in and of itself creates dissonance.

In Figure 2, blue links mark those that would be threatened and red links those that would be newly introduced should indoctrination succeed.

Phase IV (DDR and Refutational Preemption): The goal of DDR is to cause maintenance of blue links (those that are threatened) and rejection of red links (those that are indoctrination-related).

This can be accomplished by priming the clash nodes and significant energy sources that feed into them as follows:

PERSONAL SIGNIFICANCE, WORTH, MORALITY: Prime RAISE CHILDREN and PARENT ROLE by highlighting importance of personal roles and positive emotions for all. Instinctual protecting, providing for. Prime IDENTITY through critical non-violent roles played by subject in the community.

FAMILY: Children with supportive parents in a stable home with a good reputation worthy of respect. The importance of the family as a unit *with the individual present to participate*

within it.

COMMUNITY: Stable communities conducive for leading a good, moral life, international support, and so on. Link strengthening, performed by simultaneously invoking JUSTICE and COMMUNITY, affecting the link between CREATE-JUSTICE and COMMUNITY.

DIE: Prime DESIRE-TO-LIVE by cultural allusions to the importance of life, stories about what can be achieved during one's lifetime, meaning behind life, and so on. Mortality itself may be primed, but with undesirable follow-on effects.

Attempts may be made to introduce negative energy into SIGNIFICANCE ACHIEVEMENT: SUICIDE TERRORISM, though this is risky as it may be seen as a clumsy attempt to persuade, inducing change in an undesired direction.

The stability loop between FAMILY and PERSONAL SIGNIFICANCE, WORTH, MORALITY, will be automatically strengthened by introduction of positive energy to any constituent node, serving as further support for the status quo during DDR.

In the above, link-changing strategies 4,5, and 7 are employed.

Conclusion and Next Steps

This paper has introduced strategies for modeling and changing beliefs and demonstrated examples of their application in the domains of prejudice and terrorism reduction. Further work will involve extended validation of the COGVIEW modeling framework and DDR technique, as well as development of further rubrics for using these techniques in critical real-world scenarios.

References

Festinger, L. (1957). *A theory of cognitive dissonance*. Row, Peterson.

Keating, J., & Brock, T. (1974). Acceptance of persuasion and the inhibition of counterargumentation under various distraction tasks. *Experimental Social Psychology*(10).

Kruglanski, A. W., Chen, X., Dechesne, M., Fishman, S., & Orehek, E. (2009). Fully committed: Suicide bombers' motivation and the quest for personal significance. *Political Psychology*, 30(3), 331–357.

Lieberman, M. D., Schreiber, D., & Ochsner, K. N. (2003). Is political cognition like riding a bicycle? How cognitive neuroscience can inform research on political thinking. *Political Psychology*, 24.

McGuire, W. (1961). Resistance to persuasion conferred by active and passive prior refutation of the same and alternative counterarguments. *Abnormal and Social Psychology*, 63(2).

Pfau, M., Tusing, K., Koerner, A. F., Lee, W., Godbold, L. C., Penaloza, L. J., ... Hong, Y.-H. (1997). Enriching the inoculation construct: The role of critical components in the process of resistance. *Human Communication Research*, 24(2), 187–215.

Shermer, M. (2002). *Why people believe weird things* (2nd ed.). Holt.