



Integrity ★ Service ★ Excellence

Trust and Influence

5 MAR 2012

**Joseph Lyons, PhD
Program Manager
AFOSR/RSL**

Air Force Research Laboratory

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the 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 this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 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 a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE 05 MAR 2012		2. REPORT TYPE		3. DATES COVERED 00-00-2012 to 00-00-2012	
4. TITLE AND SUBTITLE Trust and Influence				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Research Laboratory, Wright-Patterson AFB, OH, 45433				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES Presented at the Air Force Office of Scientific Research (AFOSR) Spring Review Arlington, VA 5 through 9 March, 2012					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 22	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			



Outline



- **Program Trends**
- **Portfolio Description**
- **Trust Background & Motivation**
- **Trust Grants - 2**
- **Influence Background & Motivation**
- **Influence Grants - 6**
- **Recent Transitions**
- **Collaborations/Synergies**



Program Trends



- **Trust in Autonomous Systems** 
- **Cross-cultural Trust** 
- **Influence Effects** 
- **Cognitive Mechanisms for Influence** 
- **Computational Methods** 
- **Modeling** 



2012 AFOSR SPRING REVIEW



NAME: Trust and Influence

BRIEF DESCRIPTION OF PORTFOLIO:

Basic research to explore the science of reliance (i.e., how humans establish, maintain, and repair trust of humans and technological systems) and the science of influence (i.e., understanding how to shape the behavior or attitudes of others).

LIST SUB-AREAS IN PORTFOLIO:

Science of Reliance

- Cross-Cultural Trust – Identify the antecedents of trust in different cultures
- Trust in Autonomous Systems/Autonomy – identify the factors that shape reliance in complex human-machine interactions

Science of Influence

- Understanding the behavioral effects of different influence tactics (air strikes, messaging, developmental activities)
- Understanding the cognitive mechanisms that drive influence effects – identify the avenues of influence for different groups



Trust Background



Trust = willingness of individuals to accept vulnerabilities from the actions of others with little ability to monitor their actions (Mayer et al., 1995)

Assumptions:

- Trust is a human phenomenon
- Trust & trustworthiness are independent (Mayer et al, 1995)
- Trust is relational
 - Humans in cross-cultural interactions
 - Complex human-machine interactions
- Trust is dynamic (Levine et al., 2006)
- Trust is only relevant in the context of risk (Parkhe & Miller, 2000)
- Trust has both affective and analytical underpinnings (McAllister et al., 1995; Lount, 2010; Stokes et al., 2011)



Motivation – Trust



AF Tech Horizon's 2010

“In the near to mid-term, developing methods for establishing ‘certifiable trust in autonomous systems’ is the single greatest technological barrier that must be overcome to obtain the capability advantages that are achievable by increasing use of autonomous systems” (p. 42)

Operational Challenges:

- Future battle ground - complex human-machine interactions
- Interactions with other cultural groups – where trust will be critical as HUMINT increases in value – partner capacity service core function (Schwartz, 2011)

Science Challenges: Appropriate reliance is really hard!

- Complacency and or under reliance are common pitfalls
- Automation often has unintended consequences (Parasuraman & Riley, 1997)
 - Automation paradox – reliable automation can lead to catastrophic error
- Interpersonal trust models are based on “Western” data/models
- Humans have trust biases (Lyons & Stokes, 2012)
- Little is known about how human trust principles apply to autonomy

Opportunities:

- Identify human-centric trust vulnerabilities before fielding autonomous systems
- Support AFCLC cultural competencies – trust building



Trust Domain/Scope



Autonomous Systems & Automation



Interpersonal Trustworthiness

- Ability
- Benevolence
- Integrity

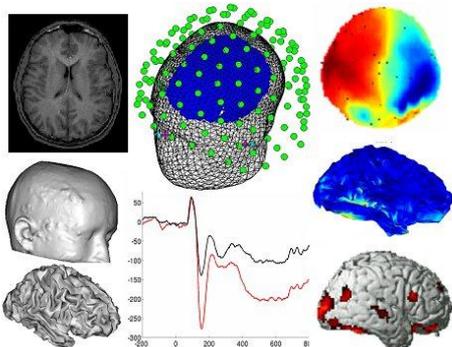
Human-Machine Interactions

Cross-Cultural Trust Issues



Human Trust Biases/Vulnerabilities

Trust Metrics



Human trust





Stokes – Dynamic Trust Model



PI: Charlene Stokes (AFRL) Lab Task
Drs. Lin (Sunway U) and Chen (NICTA)

Objective: Examine contextual factors that influence trust and trustworthiness

- Relationship type, load, culture

Approach: Manipulate trustworthiness and cognitive load in an applicant scenario using US, Malaysian, & Australian subjects

Results: Australian data collected (N=73)

- Load manipulation effective
High *M*: 3.625, Low *M*: 3.037; $t(72)=5.201, p<.01$
- Trustworthiness dimensions uniquely influence selection
- The unique influence diminish under high cognitive load
 - Dispositional influences stronger

LowCL Condition		% of time selected for position of ..	
Applicant	Supervisor	Other's Supervisor	Co-Worker
Ability	0.250	0.175	0.575
Benevolence	0.513	0.238	0.250
Integrity	0.238	0.575	0.175
Neutral	0.000	0.013	0.000
Total	1.000	1.000	1.000
HiCL Condition		% of time selected for position of ..	
Applicant	Supervisor	Other's Supervisor	Co-Worker
Ability	0.250	0.275	0.450
Benevolence	0.350	0.325	0.313
Integrity	0.400	0.388	0.200
Neutral	0.000	0.013	0.038
Total	1.000	1.000	1.000



Atkinson – Role of Benevolence in Trust of Autonomous Systems



PI: David Atkinson (IHMC)

Objective: Experimentally evaluate the Impact of benevolence within a human-machine interaction

- Interpersonal trust and trust in automation models have been treated as orthogonal

Approach:

- Operationalize “benevolence” in an autonomous system scenario
- Evaluate the impact of benevolence using experimental methods

Collaborators:

- Peter Hancock (UCF)
- Deborah Billings (UCF)
- Robert Hoffman (IHMC)





Motivation – Influence



Lt Gen Flynn (Nov 2011)

Need to understand the “Precursors of war” – what are the triggers for attitudes and behavior in different parts of the world

Operational Challenges:

- DoD lacks precision in cultural-based influence – a science base is needed!
- Military Information Support Operations (AF IFO Roadmap, May 2008)
 - Target audience analysis is a top MISO requirements
- Air Force Targeting Center - quantification of behavioral effects for influence

Science Challenges:

- Manipulation of influence tactics not plausible in practice – but maybe in the lab
- Rational actor models may not generalize beyond the laboratory – need field research
- Data mining and modeling tools have outpaced theory in social media research

Opportunities:

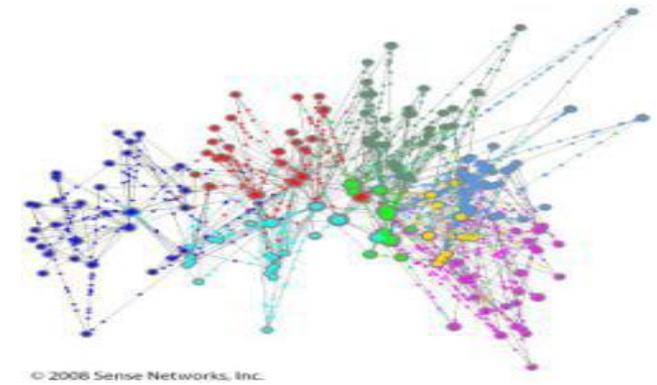
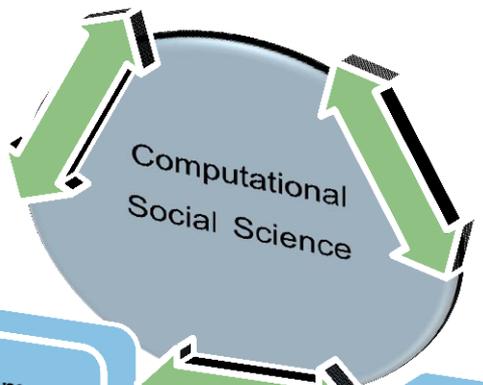
- Revolutionize “targeting” within the AF – quantify hidden costs
- Support AFSOC/ACC need for cultural awareness (e.g., AFRICOM)



Influence Domain/Scope



Target Audience
-Values/Beliefs
-Behavioral Patterns
-Narrative



© 2006 Sense Networks, Inc.

Enabling Factors
-Social Media

Military Action
-Kinetic/Non-Kinetic
-Stability Operations
-Trust Building





Lyall: Terrorism, Governance, and Development



PI: Dr. Jason Lyall (Yale)
Lead: Jacob Shapiro (Princeton)

Objective: Evaluate the impact of Kinetic versus Non-Kinetic actions on subsequent violence

Approach: Working with CENTCOM, identify unclassified data and used a matching algorithm to compare villages in Afghanistan before and after an event

•Matched on: geography, pre-existing violence, economic indicators, etc.

Results:

•Kinetic actions led to higher retaliation but the effects were limited to 5km radius

Type	15-day	30-day	60-day	90-day
Damage	.369*	.542†	1.041*	2.121**
	+9%	+12%	+15%	+19%
No Damage	.182†	.164	.359	.908
	+6%	+5%	+5%	+7%

DV: Changes in mean number of insurgent attacks against ISAF in specified temporal/spatial window
Based on (Lyall, 2011)

Impact: Results transitioned to the Air Force Targeting Center/ACC and briefed to CENTCOM



Matsumoto – Emotions and Intergroup Relations



PI: David Matsumoto (SFSU)
Co-PI: Mark Frank (SUNY Buffalo)

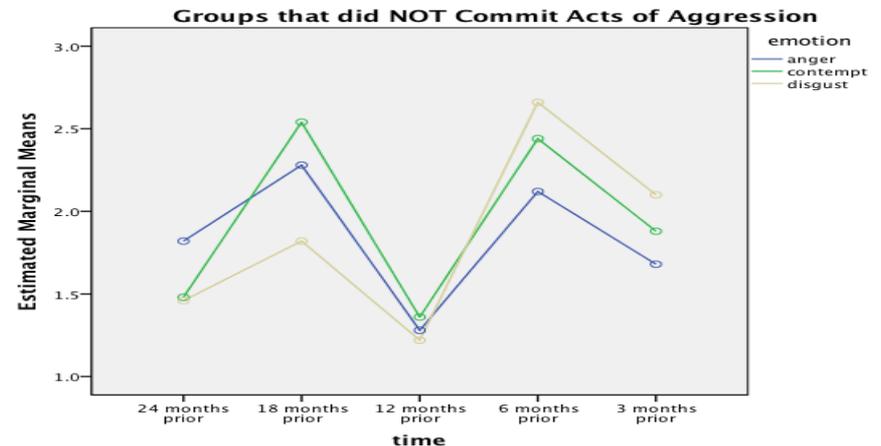
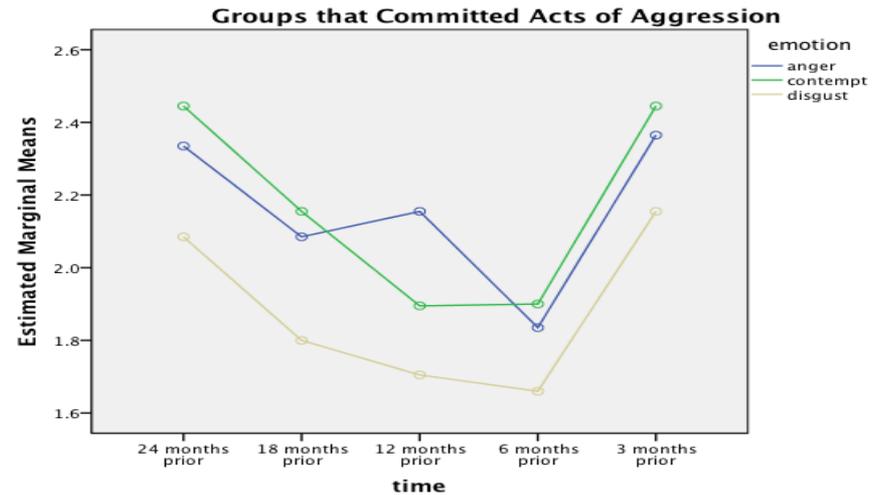
Objective: Evaluate the role of emotions in predicting violent acts

- Emotions trigger action tendencies
- Anger, contempt, disgust

Approach: Examine key leader speeches describing in/out groups prior to acts of violence

- Code emotions in text

Results: Acts of aggression were preceded by increased anger, contempt, and disgust





ARTIS - MUTUAL INFLUENCE OF MORAL VALUES, MENTAL MODELS AND SOCIAL DYNAMICS ON INTERGROUP CONFLICT



PI: Scott Atran, Rich Davis, Jermy Ginges

Objective: To examine how conflict impacts the relationship between religiosity and moral absolutes

- Religious acts may anchor sacred values
- This process could be triggered by threat

Approach: Surveyed Palestinians on attitudes toward Israeli-Palestinian issues

- Recognition of Israel as Jewish State, right of return for refugees, sovereignty, etc.
- Assessed: religiosity, perceived conflict, and moral absolutism

Results: Conflict intensifies commitment of religious devotees to sacred values

Impact: Prior research on sacred values has been presented to Congress, DARPA, House of Lords, Dept of State – and published in top-tier journals (e.g., Science)

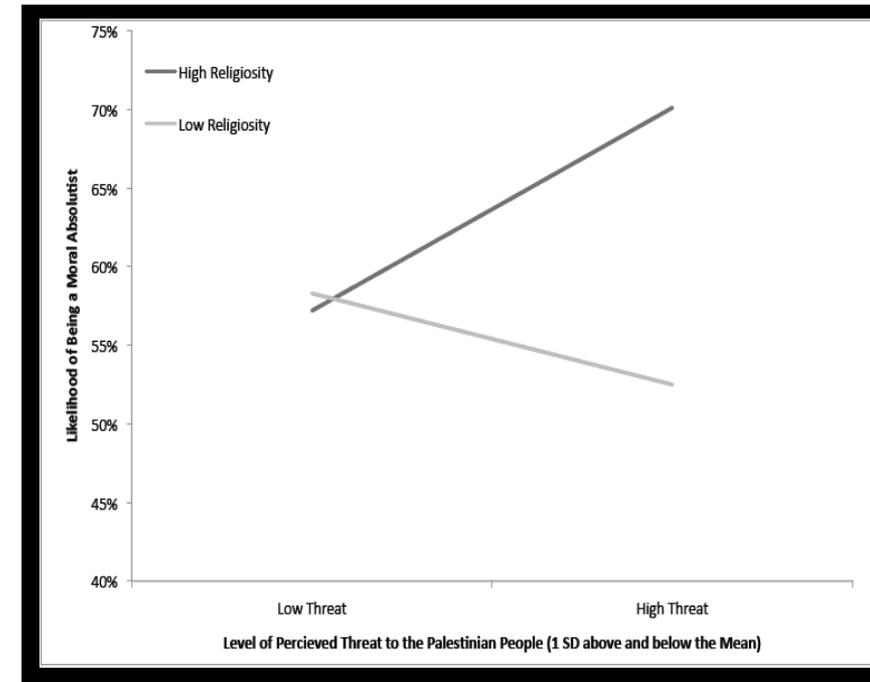


Figure 1. Likelihood Of Being a Moral Absolutist (%) Depending On Religiosity And Perceived Threat to the Palestinian People (+/- 1 SD).



Burns – Neurobiology of Violence



PI: Greg Burns (Emory)

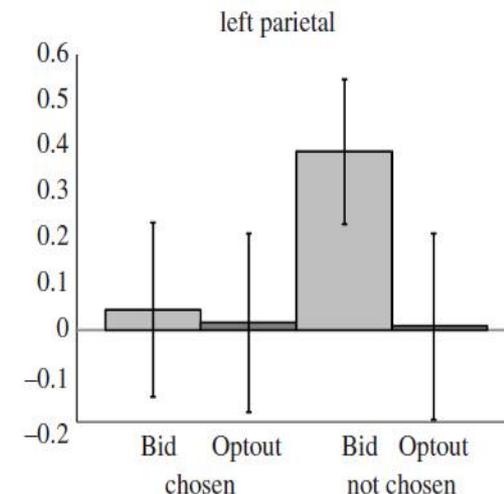
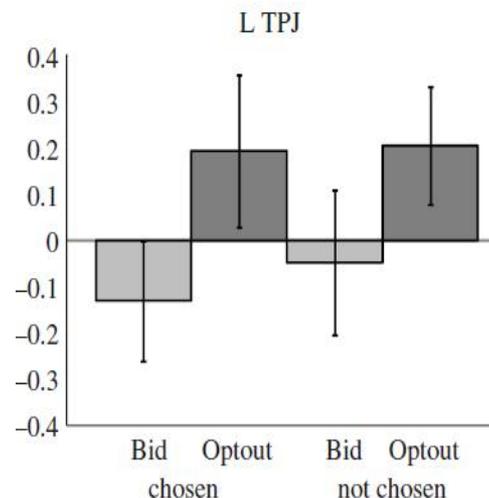
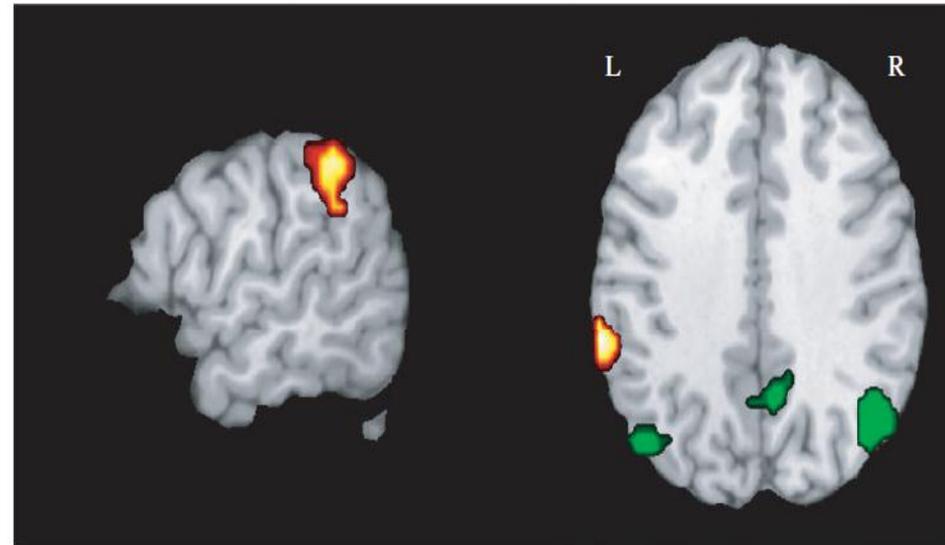
Objective: To assess the cognitive processing sacred decisions

Approach: Used fMRI, exposed subjects statements varying in sacredness

- “I’m a dog person; I believe in God”
- Offered subjects money to sign a document contradicting their value

Results: Statements identified as sacred evidenced rule-based activity whereas bid statements showed utilitarian activity

Impact: Transitioned to DARPA





Axelrod – Case-based Influence



PI: Robert Axelrod (Michigan)

Objective: Examine the use of analogies and exemplars within different cultural groups

- Examine cultural sensemaking/narrative – historical references
- Salience, similarity, compellingness

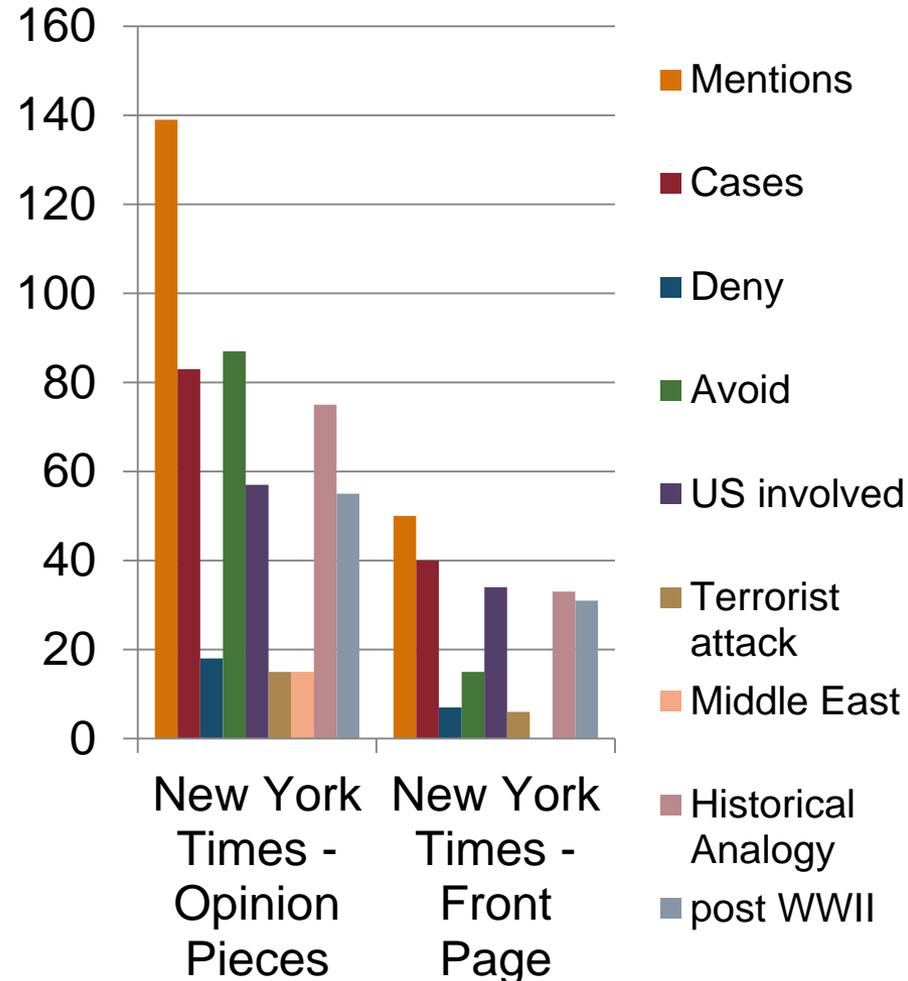
Approach: code newspapers after significant events to examine use of historical cases

- 9/11; Arab Spring; Osama Letters

Results:

- Wide variety of historical analogies used
- Varies by geography
 - Cases more salient when more proximal

Following 9/11

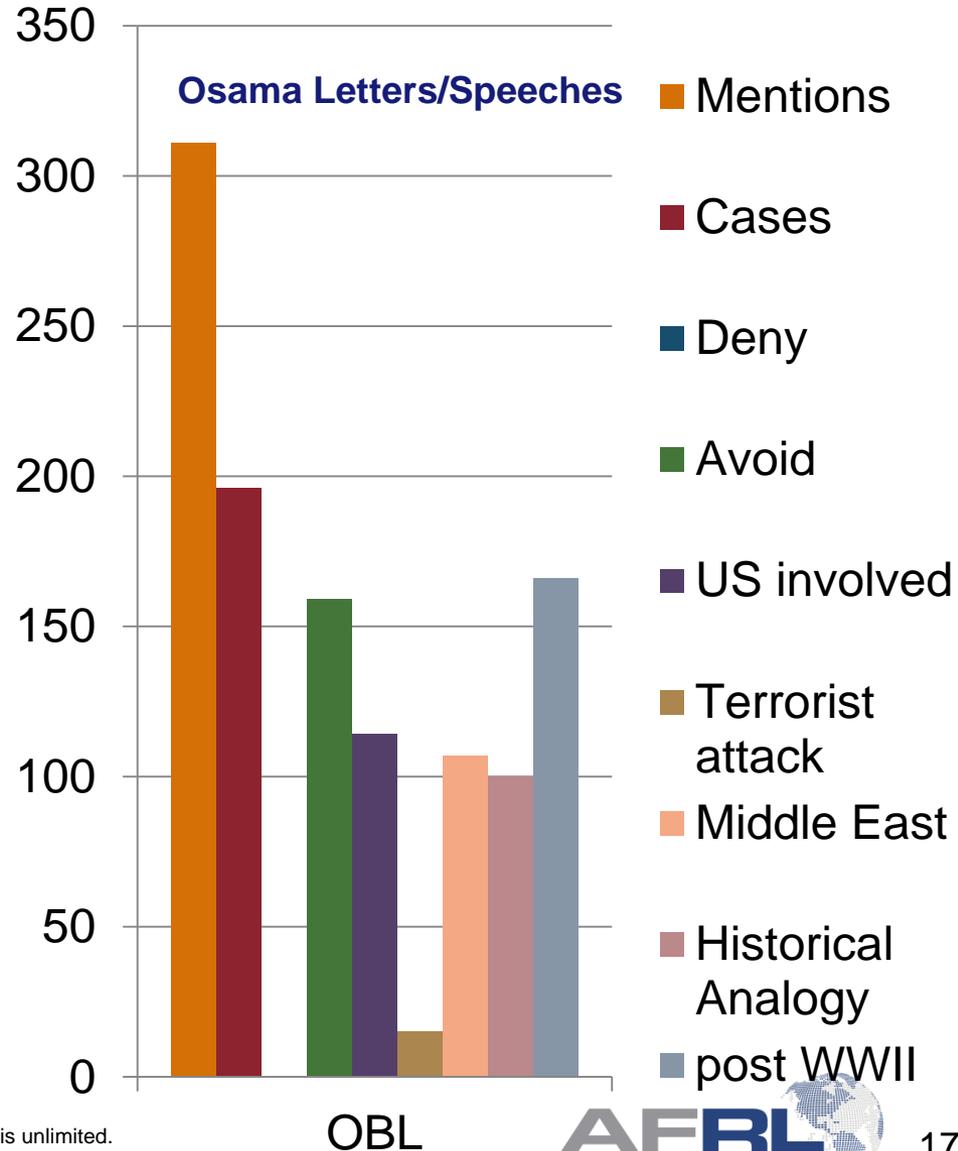
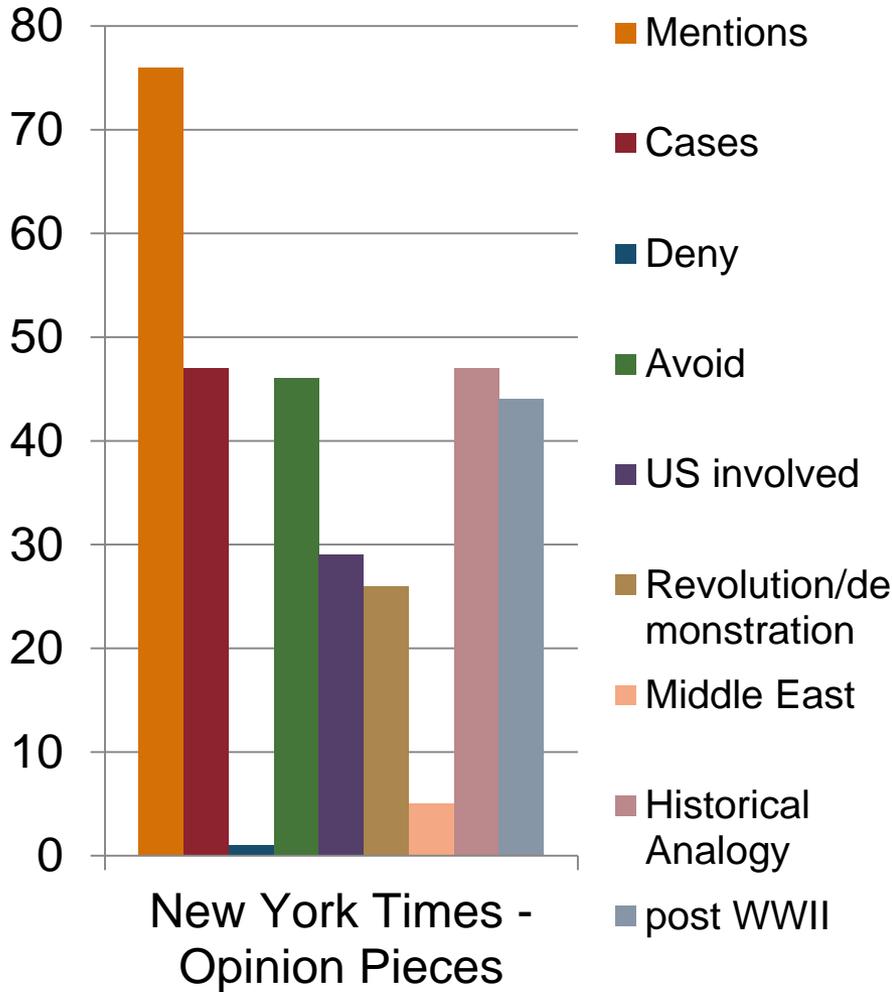




Axelrod – Case-based Influence



Demonstrations in Egypt 2011





UCLA MURI - Inferring Structure and Forecasting Dynamics on Evolving Networks



PI: Kristina Lerman (USC)

MURI Lead: Jeff Brantingham (UCLA)

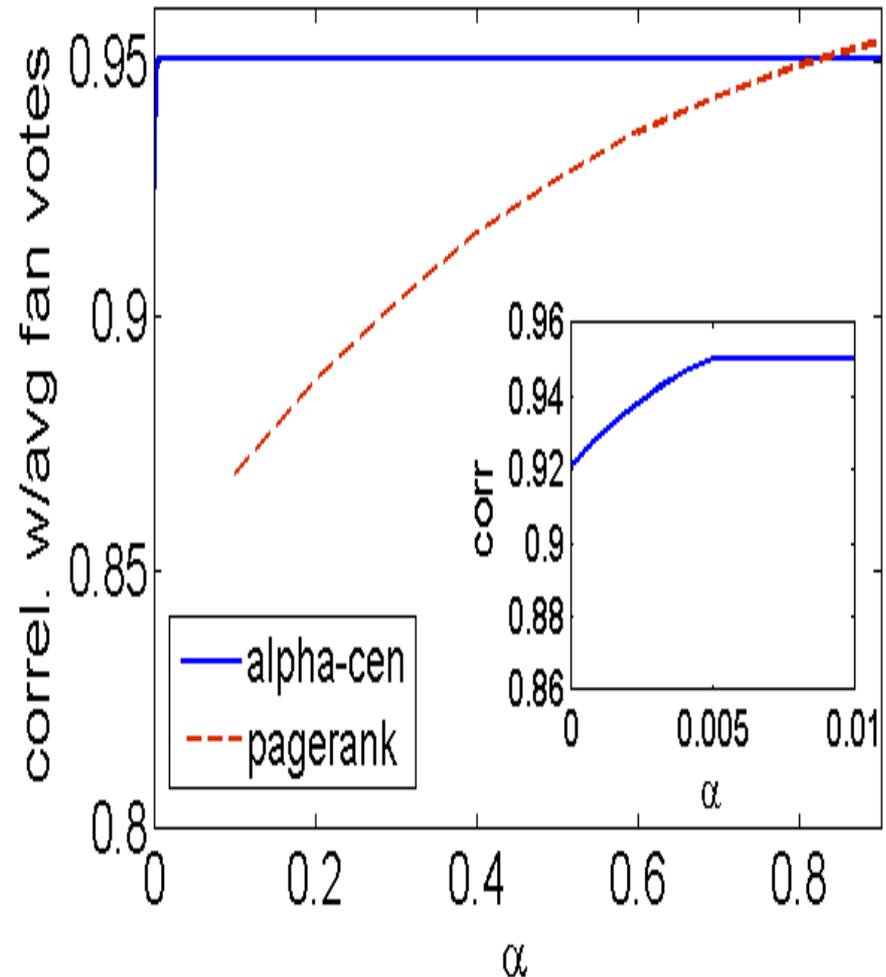
Objective: Examine the utility of social network analysis metrics for understanding influence

Approach: (sample)

- Examine utility social networks using DIGG
- Influence = re-posting
- Alpha Centrality vs. Page Rank

Results:

- Alpha Centrality better predictor of influence than Page rank for more local network structures





UCLA MURI - Inferring Structure and Forecasting Dynamics on Evolving Networks



PI: Jeff Brantingham (UCLA) & team!

Objective: Use mathematical models and statistics to forecast patterns of violence in LA gangs

Approach:

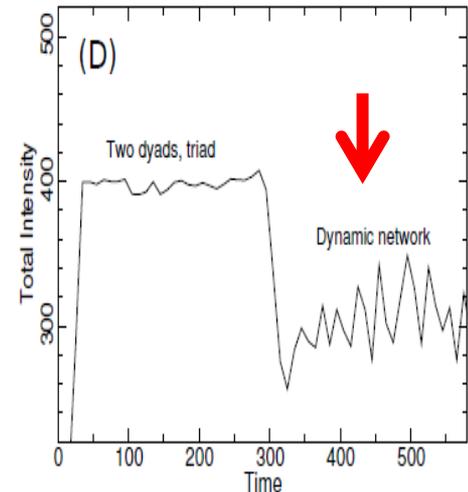
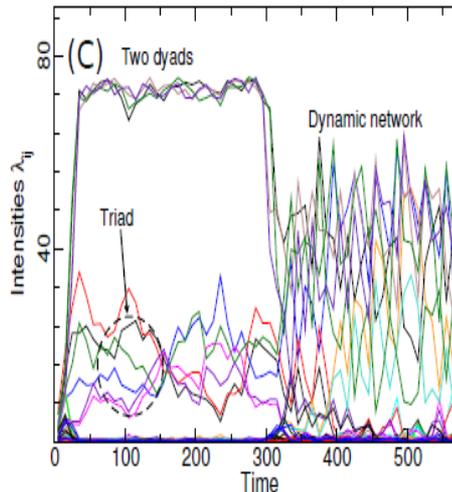
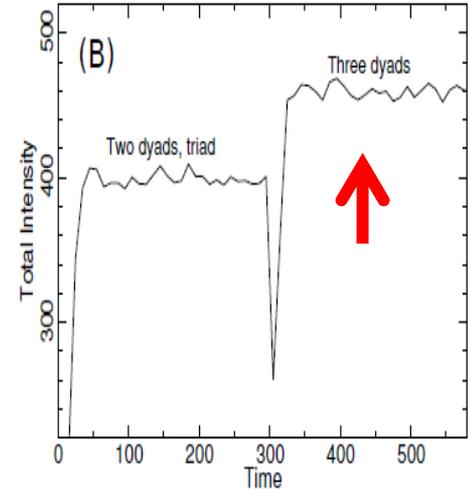
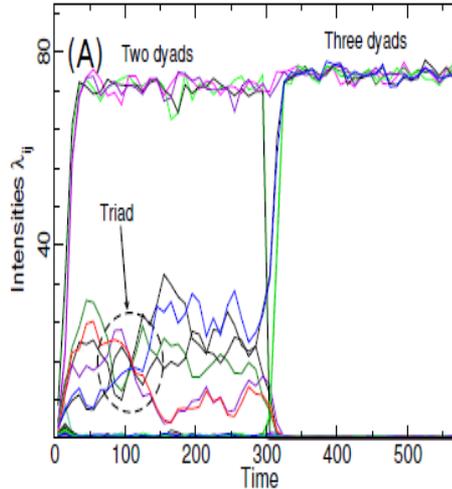
- Apply statistical models to gang activity

Results:

- Model identified dyadic rivalry as most violent – supported by actual crime data

Impact:

- Used by Predictive Police Unit
- Used by 711 HPW Forecasting CTC





Recent Transitions



- AF Minerva Princeton Project:
 - Direct transition to COIN efforts for ISAF, heavily cited in the 2011 Foreign Relations Committee Report on Aid in Afghanistan
 - Briefings to Gen McChrystal, Chairman of the Joint Chief of Staff, ISAF, DARPA
 - Lyall – airpower study transitioned to the Air Combat Command/Air Force Targeting Center
- UCLA MURI used by predictive policing unit in CA (featured on ABC/NBC news)
 - Transitioning to 711 HPW Forecasting CTC
- Yaneer Bar-Yam – Statistical models of geographic/ethnic boundaries & link between food costs and riots presented to Strategic Multi-layer Assessment community – Warfighters in the MISO domain
 - Cited in Top 10 Science Discoveries of 2011 by Wired Magazine!
- Sandy Pentland – computational methods applied in the DARPA Nexus 7 initiative
- Sacred values work has transitioned to the Navy and DARPA impacting national policies -- U.S. Senate, U.S. State Dept., House of Lords, DARPA (Featured in Time magazine article on Gaddafi)
- Greg Burns - Neurobiology of Violence leveraged by DARPA



Collaborations/Synergies



Trust

ARL: Trust work related to interfaces, robotics, networks

ARI: Trust in networked teams

NICTA: Funding technology development, some work on trust and cognitive load

NAVAIR: Trust & culture interests

IARPA: Funding large trust initiative on physiology of trust

ONR: Machine Ethics

AFRL: Trust is a core research area – close collaboration with 6.1 and 6.2

Influence

ONR: Close collaboration, co-funding

ARO: Training, mission rehearsal, face-to-face negotiation/interaction, etc. Focus on near-term: “something for the soldier”

DARPA: Interest in culture but more focused on neuroscience and training

OSD HSCB: Modeling, ops analysis, training (mainly 6.2-6.3)

Dept. of State: Social Media interests, possible co-funding

Air University Culture & Language Center: focus is mainly on training related to language and culture

Defense Equal Opportunity Management Institute: focused on cultural training