2009

Trustworthiness in Virtual Organizations

Shuyuan Mary Ho
Syracuse University, smho@acm.org

Trustworthiness in Virtual Organizations

Shuyuan Mary Ho
Syracuse University
smho@syr.edu

ABSTRACT

This study examines perceptions of human *trustworthiness* as a key component in countering insider threats. The term *insider threat* refers to situations where a critical member of an organization behaves against the interests of the organization, in an illegal and/or unethical manner. Identifying how an individual's behavior varies over time - and how anomalous behavior can be detected - are important elements in the preventive control of insider threat behaviors in securing cyber infrastructure. Using online team-based game-playing, this study seeks to re-create realistic insider threat situations in which human "sensors" can observe changes in a target's behavior. The intellectual merit of this socio-technical study lies in its capability to tackle complex insider threat problems by adopting a social psychological theory on predicting human *trustworthiness* in a virtual collaborative environment. The study contributes to a theoretical framework of trustworthiness attribution, and game-playing methodology to predict the occurrence of malfeasance.

Keywords

Trustworthiness, insider threats, virtual organizations, socio-technical study.

INTRODUCTION

According to 2007 CSI Survey, financial losses caused by computer crime have been increasing dramatically to $67 millions in 2007, up from $52.5 millions in 2006 (Richardson, 2007). Among those losses, nearly 37 percent of respondents attributed more than 20 percent of losses are caused by insiders. It indicates an increase of insider abuse of network resources from 42 to 59 percent when compared with 2006 CSI/FBI Survey. Insider misuse of authorized privileges or abuse of network access has caused significant damage and loss to corporate internal information assets. While employees are essential to the productive operation of an organization, their inside knowledge of corporate resources can also threaten corporate security as a result of improper use of information resources. Such improper uses are often termed by security experts as the "insider threats." The capability to anticipate the occurrence of such improper uses would be valuable.

Every individual's behavior varies over time; and unusual or unexpected changes in an individual's behavior that are detected by human observers may possibly provide clues to the imminence of an insider threat activity. A shorthand term that some might use to characterize and assess the vulnerability of an insider is "trustworthiness." This research explores basic mechanisms of how to detect changes in the trustworthiness of an individual who holds a key position in an organization through the observation of overt behavior – including communication behavior – over time. The class of individual whose behavior is the focus of this investigation is someone granted access to and authority over important information within the organization. This access and authority is what is meant by "holding a key position" in the organization. Employees with access and authority have the most potential to cause damage to that information, to organizational reputation, or to the operational stability of the organization.

CONCEPTUALIZING INSIDER THREATS

The phenomenon of insider threats is a social, human behavioral problem (Martinez-Moyano, 2006, Ho, 2008a, Ho, 2008b, Ho, 2009). *C4ISR*[1] Joint Chiefs of Staff (2007) believe that in addition to integrating technology, policy, management, and procedure, "human factors" have greatly threatened and caused vulnerability to the chains of defense (General Pace, 2007). In the *Insider Threat Study by CERT (2004-2005),* US DoD[2], DHS[3], and Secret Service investigated various insider threat cases and discovered that embedded in a mesh of communications, a person given high social power but with insufficient

[1] C4ISR stands for Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance. It is an architecture framework used by the US DOD.

[2] US DOD stands for the US. Department of Defense.

[3] DHS stands for the Department of Homeland Security.

trustworthiness can create a single point of trust failure (Randazzo, 2005, Keeney, 2005, Kowalski, 2008). Thus, "insider threat" as an organizational problem is defined as *situations where a critical member of an organization with authorized access, high social power and holding a critical job position, inflicts damage within an organization*. In a way, *this critical member behaves against the interests of the organization, generally in an illegal and/or unethical manner.* As Mitnick noted, bolstering against the weakest linkage - the human factor - becomes critical in the chain of security defense (Mitnick, 2002).

Insider threats can be analyzed in an abstract view as in Figure 1. Actor A depicts someone who holds a critical position and has authorized access to intelligence or assets in an organization or a community. Observer B_n depicts a group of peers and subordinates that work closely with actor A in which B_n forms the social network of A. B_{org} represents the interests of the organization. One pre-requisite in this social network is that B_n is dependent on A. The communications between actors A and B_n builds the meaning to their social interactions. The group containing actors A and B_n may be goal-oriented and team-based with those whom they work together to achieve certain goals. In order for actors A and B_n to work together, a certain level of trust relationship between actors A and B_n must exist. In order to achieve goals and enhance corporate performance and productivity, trust can be built when each actor displays willingness and competence[4] to work together.

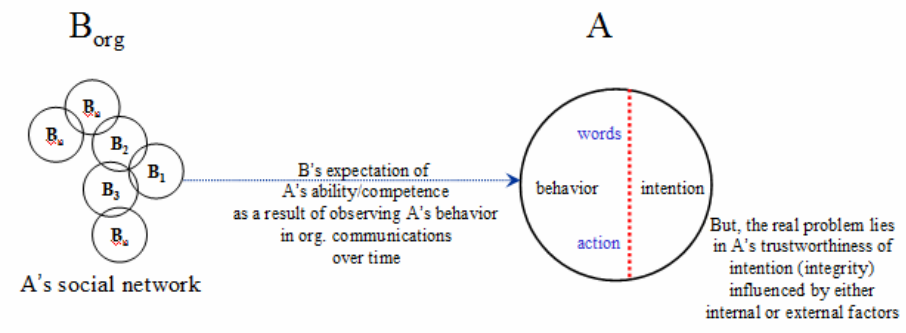

Figure 1. Abstract View of Insider Threats

In normal and regular situations, actors A and B_n trust one another and cooperate to achieve pre-determined goals. Within that pattern of organizational communication, a norm of social interactions between actors A and B_n is formed. However, when actor A's integrity is compromised by either internal causes, such as improper thoughts to gain illegal profits - or external causes, such as being persuaded by others to commit crime, actor A's behavior would somehow reflect what his or her intentions and willingness are. In some way, the behavior of actor A might change to reflect his unethical thoughts when his or her integrity has been compromised. While the reasons that cause actor A to behave against the interests of the organization, B_{org}, are unknown, such behavior of actor A causes the organizations or community to suffer loss whether tangible or intangible. A low competent, inefficient, actor A, can be scrupulously honest; this actor A, by the definition of insider threats, would not typically cause significant insider threats. One critical differentiation in this insider threats phenomenon is that actor A may continue to act normally in his or her daily job performance which fulfills observer B_n's expectation of actor A's competence and willingness to work with B_n as a team. But the real problems lie in actor A's compromised integrity - that his or her "character," or "quality," of having strong moral principles is downgraded and corrupted.

RESEARCH QUESTION

This study examines basic mechanisms for detecting changes in the trustworthiness of an individual who holds a key position in an organization, by observing overt behavior – including communication behavior – over time. Since Steinke (1975) suggests that it is possible to detect cheating behavior without directly observing the individual, the overarching question is: *What changes of behaviors can reflect a downward*[5] *shift in the trustworthiness of a critical member in a virtual organization which might signal possible insider threats?* A virtual organization refers to a group of individuals whose members and resources may be dispersed geographically, but function as a coherent unit through the use of cyber-infrastructure. This group of individuals is team-based and goal-oriented, where leaders and subordinates work together to achieve pre-determined goals. Derived from the question, I hypothesize that the downward shift in a person's trustworthiness can be reflected in his

[4] Same as *ability*.

[5] Same as *unethical*, or *illegal*.

or her behavior. And, the inconsistency and unreliability in this actor's unexpected behaviors when compared to his or her communicated intentions can be detected by the observers' subjective perceptions over time. The observers refers to the members of his or her close social network.

LITERATURE REVIEW

In order to understand how people observe a target's behavior over time, make inferences about changes in behavior that signify something abnormal, and be able to predict a likelihood of a downfall shift of the target's intention as reflected in his or her behavior, attribution theory is adopted to look at a basic human relationship, trust. Attribution theory explains how an observer perceives a target's behavior and gives it an explanation (Kelley, 1973, Heider, 1944, Heider, 1958). Moreover, the observer assigns the target's behavior to a cause and may suggest a possible threat if the attributed cause is abnormal. The trustworthiness of the target's intention is perceived, attributed and assigned with meaning by his or her social network. In other words, the social network assigns meaning to the perceived trustworthiness in the target's intentions according to the observed behavior.

The theoretical framework of this research is introduced by first reviewing trust theories, which sets general understanding for trustworthiness, then differentiating how trustworthiness is from trust. Then, attribution theory is discussed on the ground of a trust relationship. The overall syntheses of the theory of trustworthiness attribution are given at the end.

Trust

Trust can exist within a relationship of reliance between two individuals or groups. Trust is an estimation, an opinion, an evaluation, or a belief, that trustor B has concerning trustee A (Meyerson, 2006). Moreover, Shoda, Mischel and Wright stated that trust can refer to both a belief and a behavior of trustee A (Shoda, 1994). Hosmer characterizes trust as "the expectation… of ethically justifiable behavior—that is, morally correct decisions and actions based upon ethical principles of analysis" (Hosmer, 1995).

A basic trust relationship involves two parties: a trustor, B, who may (or, may not) trust a trustee, A. To continue the previous analogy, I use the term 'trustee A' to refer to the target subject who is studied and may (or, may not) be trusted by B. I use the term 'trustor B,' to refer to a trusting party, in this case, members of A's social network such as his or her supervisees. Trustworthiness refers to as an inferred quality of A. The establishment of B's beliefs about the trustworthiness of A is an antecedent to a trust relationship between B and A. B must have a vulnerability to A for a trust relationship to matter. If B has no vulnerability to the actions of A, then it does not matter whether B trusts A. B's trustfulness represents his or her tendency to trust A based on limited evidence of A's behavior (Mayer, 1995, Mayer, 1999).

Rotter defined trust as a generalized expectancy - held by an individual or a group - that the communications of another (individual or group) can be relied upon regardless of situational specificity (Rotter, 1967, Rotter, 1980). The term "expectancy" refers to a subjective probability - learned from previous experience - that an action will lead to a particular outcome (Rotter, 1980). A generalized expectancy, in the eyes of the beholders, is one that holds broadly across a variety of situations. Interpersonal trust, defined by Rotter, is "an expectancy held by an individual or a group that the word, promise, verbal or written statement of another individual or group can be relied upon" (Rotter, 1967, Rotter, 1980).

Trustworthiness

Rotter asserted that "trust and trustworthiness are closely related," but trust depicts a relationship among two or multiple parties or actors while trustworthiness is an attribute or a quality of a person (Rotter, 1971). Trustworthiness is defined as a generalized expectancy concerning a target person's degree of correspondence between communicated intentions and behavioral outcomes that are observed[6] and evaluated, which remain reliable, ethical and consistent, and any fluctuation between target's intentions and actions does not exceed the observer's expectations over time (Hardin, 2002, Hardin, 2003, Hosmer, 1995, Rotter, 1967).

Tyler discovered that trustworthiness can serve as a strong predictor of the fairness in authorities' decision-making (Tyler, 1988). Tyler and Degoey examined how a group attributes their authorities' trustworthiness will shape the willingness of group members in accepting authoritative decisions (Tyler, 1996). This research has two interesting points in understanding the authorities' trustworthiness. First, Tyler and Degoey studied the authorities' ability, which is the same as competence (Tyler, 1996). It helps to clarify the role of ability in trustworthiness. Second, this research looks at authorities'

[6] Same as *perceived*.

trustworthiness from the group's perspectives. Tyler and Degoey proved that trustworthiness of the authorities is determined by both relational and intentional concerns, rather by instrumental concerns, from the group members (Tyler, 1996). In other words, the competence of the authority is not found to be significantly correlated to the subordinate's attribution of the authority's trustworthiness.

Trust lies in the existence of a relationship of reliance between the two (individuals or groups). To further differentiate trustworthiness from trust, we may use an example to describe one key difference between these two constructs. To illustrate, a criminal, B, can trust a criminal-partner, A, to conduct a jointly orchestrated crime, and there is no moral or ethical notion involved. It simply takes a belief or reliance that the criminal-partner, A, can deliver what was promised. Here, the consistency between words and actions of this criminal, A, are what is important. However, if B requires A to act de facto on B's behalf but B cannot fully count on A's acting in B's interest, this adds a complexity of decision in the trust relationship. In this situation of complexity, B reserves to trust in A because B may not find A trustworthy (Hardin, 1996, Hardin, 2001, Hardin, 2003). Overall, we may say that the cause of this ruptured trust relationship lies in the fact that either B is not trustful or A is not trustworthy.

Trust Models that Elucidate Trustworthiness

The findings of Tyler and Degoey proposed psychological models of trustworthiness. Specifically, the findings showed subordinates' attribution[7] about authorities' trustworthiness and whether this attribution would shape subordinates' willingness to accept authority's decision (Tyler, 1996). Although the decision to trust can be calculated by a trustor as described in Lewicki and Bunker's work (Lewicki, 1996), trust can simply be a response to "environmental contingencies" (Tyler, 1996, Williamson, 1993). When a trustor trusts a trustee, it simply means that there is a belief in this trustor that the consequences from the action performed by the trustee will be positive and beneficial. In other words, this trustor finds the trustee trustworthy and can have certain percentage of assurance that the trustee will perform an action that is beneficial to the trustor.

Mayer, Davis and Schoorman defined three factors of perceived trustworthiness to be ability (competence), benevolence (kindness) and integrity (goodwill/ethics) (Mayer, 1995). Furthermore, Mayer and Davis enhanced these factors by using field quasi-experience to validate a trust model that competence (ability), benevolence (kindness) and integrity (goodwill/ethics) are antecedent factors in trustworthiness (Mayer, 1999). Mayer and Davis differentiate trustworthiness from trust in the effective use of a performance appraisal system (Mayer, 1999). Mayer and Davis found a significant impact of the appraisal system's acceptability on trust for management, which was mediated by the factors of perceived trustworthiness (Mayer, 1999). The implication of this finding was that trust is constantly influenced by the combination of competence, benevolence and integrity. In addition, trustor's propensity to trust influences the perceived trustworthiness of a trustee.

Overview of Attribution Theory

Attribution theory has been adopted to understand how people attribute (or assign) the causes of others' behaviors (Heider, 1944, Heider, 1958). The perception given by observers can vary depending on the interpretations of different observers, the target individual being studied, and different situational settings (Heider, 1958, Weiner, 1985, Weiner, 1986, Weiner, 2006). The attribution of the target's (A's) behavior by observers (B_n's) is determined by B's judgment that A intentionally or unintentionally (Heider, 1958) behaves in a way that is attributable to either external (situational) causality or internal (dispositional) causality (Kelley, 1973).

Kelley suggested that attribution theory explains how people answer "why" questions in causal situations (Kelley, 1973). It deals with people's social perceptions by assigning causes to an observed behavior. For example, if a person behaves in 'rushing to get things done' quickly, we tend to question whether this behavior is a result of his inclination, or a type of pressure from an external source. Because all human beings are of the same species and born with similar types of features and functions, man should "know" and be able to "sense" from his own perceptions and with his judgment of how the world operates (Kelley, 1973, Heider, 1958) despite the fact that sometimes those attributions may not be accurate or valid.

Theoretical Framework

The main construct of this framework is human perceived trustworthiness, particularly, of those members who hold critical positions in an organization. The intentions of the target subject are reflected in the information behavior (Ajzen, 1991, Beck, 1991) through perceived trustworthiness in a virtual organization or an online community. A multi-level analysis of mixed

[7] Same as, judgment.

"lenses" is adopted in looking at how behavior of a target is influenced by his or her intention within existing organizational norms. Perceived trustworthiness, as a primary construct, is communicated with his or her social network (Holmes, 1989a, Holmes, 1989b, Rempel, 1985).

This study focuses on what behavioral changes of the target might trigger a downward shift in his or her perceived trustworthiness. Thus, observations by members of the target's social network are where the behavioral anomaly detection occurs. Social actors communicate and construct *meaning* within an organization or a community (Giddens, 1979). Through communciation[8], peers in a social network can judge the trustworthiness of the target subject. The target subject's trustworthiness is judged by the perceptions of his or her peers from the communications within his or her community. Since the target's behavior is observed within his or her social network, this level of analysis is at the group level where the behavior is the unit of analysis. However, the perceptions of the target subject are collected from the peers at the individual level.

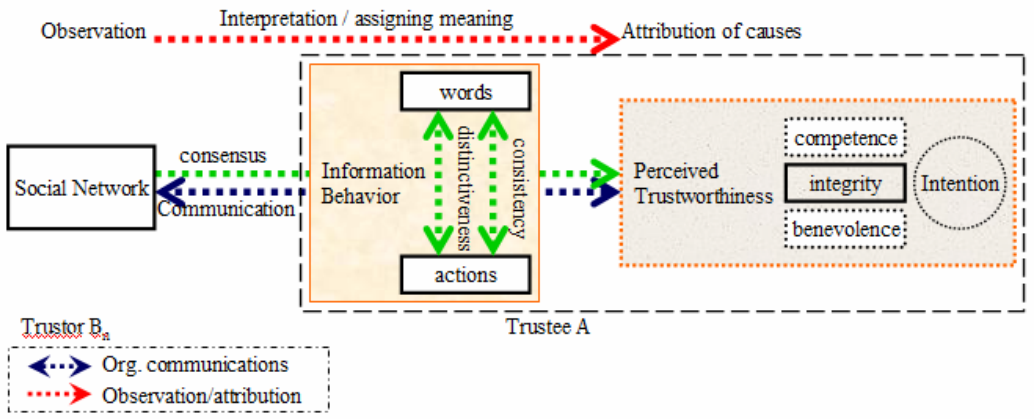

Figure 2. Theoretical Framework

In Figure 2, the relationship of the constructs is represented by arrows. Blue arrows represent communication between the target and the social network. Red arrows represent the observation and attribution by members of the social network regarding the target's behavior. Green arrows represent three attribution principles depicted by Kelley (Kelley, 1973). In this framework, the communications among the target's social network (including communications to and from the target) sheds light on the target's perceived trustworthiness. In other words, members of the social network attribute (or assign) meaning to the target's trustworthiness level based on their observations of the target's behavior.

Syntheses of Trust and Attribution Theory

Trust denotes a basic human relationship and it is so in a workplace. Trust represents predictability and correspondence between behavior and words. Unexpected behaviors – actions that do not fit the usual pattern – trigger B's attributional analysis of A. The following hypothesized theoretical mechanism is at work. With *external attribution* to plausible circumstances, B attributes any anomalous behavior of A to circumstances beyond A's control. As a result, B's perception of A's trustworthiness is not adversely affected (at least with respect to integrity; trust in A's competence may be affected, but that is beyond the scope of this study). However with an *internal attribution* by B of A's anomalous behavior, B concludes that A is performing the anomalous behavior by choice, and therefore B will downgrade her perception of A's trustworthiness. The empirically well-established reasons for attributing behavior to external or internal causes (e.g., distinctiveness, consensus, and consistency) apply to B's thinking processes about A's behavior. It is B's attributions about A's anomalous behavior that influence perceived changes to trustworthiness.

In short, the behavioral anomaly detection involves in a two-staged recursive process. There is a relational dependence from B to A. B is dependent on A and B trusts A. The first-stage of this process takes place when B builds a mental model of A's "normal" behavior in standard situations. The second-stage of this process takes place when a situation happens through which B assigns meaning to A's "abnormal" behavior that is incongruent with the existing model of A's "normal" behavior.

[8] Communication behavior is interchangeable with information behavior. However, more precisely it refers to communication behavior that appears in exchanged information.

B finds A's anomalous behavior non-significant if the causes are attributed to A's competence, as an external attribution. However, B finds A's anomalous behavior significant if the causes are attributed to A's trustworthiness - particularly with regards to integrity - as an internal attribution.

METHOD

This section of research method comprises overview, description of actual experiments, factorial design, variables, and hypotheses for the simulated case studies.

Overview of the Research Design

This study seeks to identify indicators of abnormal behavior and the basic criteria of trustworthiness assessment. The "Leader's Dilemma" game is a series of simulated, controlled "honeypot" situations created to test how a leader's trustworthiness is perceived by his team members. This game simulates a virtual organization including the front-end (overt) where normal business operates while the back-end (covert) information is captured regarding how a target is influenced. The "front-end" data includes emails, chats, discussion blogs, surveys, interviews, etc. for team members within the organization, from which we analyze indicators of the "back-end" manifestations of trust within a virtual organization. A virtual organization refers to a group of individuals whose members and resources may be dispersed geographically, but function as a coherent unit through the use of cyber-infrastructure. This group of individuals is team-based and goal-oriented, where leaders and subordinates work together to achieve pre-determined goals. In these experimental settings, the Game-Master (G) has the role of manipulating the dynamics of the virtual competition. The Experimenter takes on the role of a Moderator (M) in these online games. Team-Leader (A) is the target, who is appointed from among the team participants by the Game-Master. Team members are observers (B_n), who work with Team-Leader in achieving their pre-determined goals. Figure 3 depicts how each role in this virtual contest is situated. Figure 4 illustrated how the target and the threat situations were pre-determined.

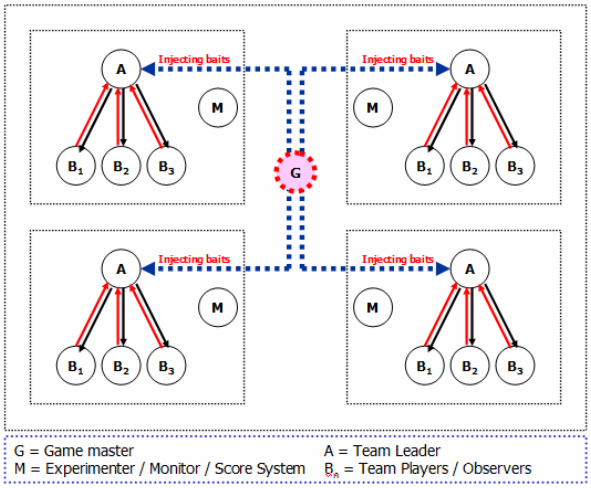

Figure 3. Experiment Control Room

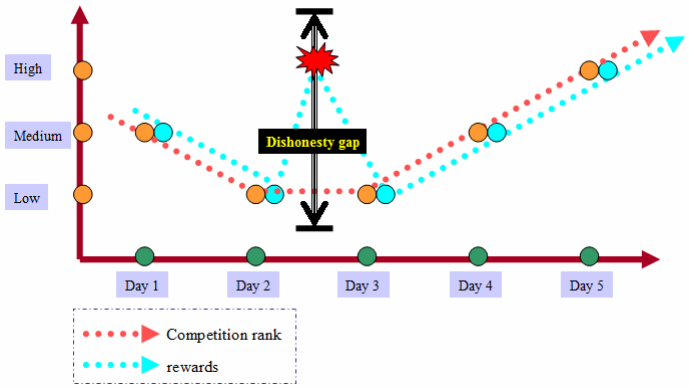

Figure 4. Manipulated Virtual Contest

Factorial Design

A 2×2×3 factorial design for this full-scale experiment includes four sets of simulated case studies of online games (Figure 5). In other words, 12 sets of group observations will be obtained. Sample size for each group is indicated in Figure 5, where n_A, n_B, n_C, n_D refer to the number of observers, and n_{TAR} refers to the number of target.

Treatment / Time / Setting	without treatment/bait; (B_0)			with treatment/bait; (B_1)		
	Time (T_1)	Time (T_2)	Time (T_3)	Time (T_1)	Time (T_2)	Time (T_3)
Increase group sensibility (S_1)	Group A Average (n_A=6; n_{TAR}=1)			Group C Average (n_C=6; n_{TAR}=1)		
Decrease group sensibility (S_2)	Group B Average (n_B=4; n_{TAR}=1)			Group D Average (n_D=5; n_{TAR}=1)		

Figure 5. 2×2×3 Factorial Design

While the dependent variable (response) is target's perceived trustworthiness, major independent variables (factors) include: the bait (B_0 and B_1) as the treatment, a mole that increases or decreases group sensibility (S_1 and S_2) by either encouraging or discouraging conversations about the team-leader, and time (T_1, T_2 and T_3) representing measurement obtained from each day, in particular, after conflict of interest between the team-leader and the team members is created.

Research Model

My hypothesis is that the downward shift in a person's trustworthiness can be reflected in his or her behavior. And, the inconsistency and unreliability in this actor's unexpected behaviors when compared to his or her communicated intentions can be detected by the observers' subjective perceptions over time (Figure 6).

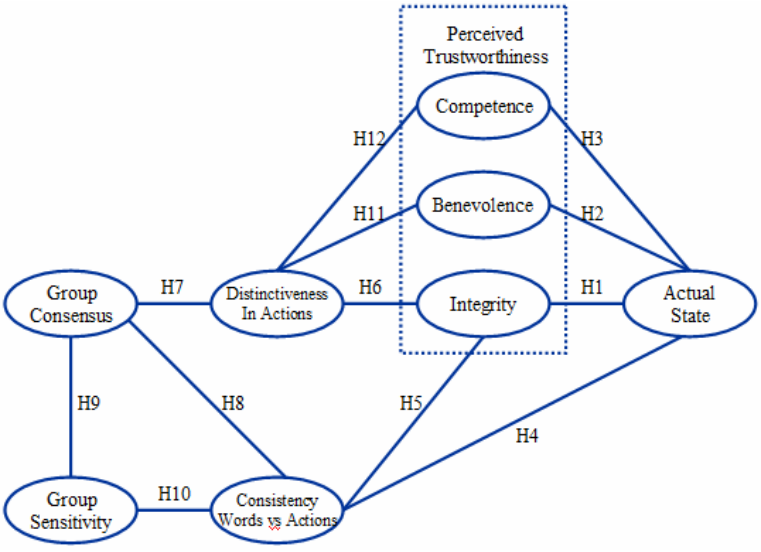

Figure 6. Research Model

Specific hypotheses include:

- *Hypothesis 1* (H_1): There is a negative relationship between the target's actual state and the target's perceived trustworthiness[9], in terms of his or her integrity.

[9] When target's actual state is positive (meaning that he or she has taken the bait), his or her perceived trustworthiness is found positive. When target's actual state is negative (meaning that he or she has not taken the bait), his or her perceveid trustworthiness is found negative.

- *Hypothesis 2* (H2): There is a negative relationship between the target's actual state and the target's perceived trustworthiness, in terms of his or her benevolence.
- *Hypothesis 3* (H3): There is a negative relationship between the target's actual state and the target's perceived trustworthiness, in terms of his or her competence.
- *Hypothesis 4* (H4): When target's actual state is positive (meaning that he has taken the bait), the target's inconsistency between communicated intentions (words) and information behavior (actions) can be found.
- *Hypothesis 5* (H5): When target's perceived trustworthiness, in terms of integrity, is relatively low, the target's inconsistency between communicated intentions (words) and information behavior (actions) can be found.
- *Hypothesis 6* (H6): When target's perceived trustworthiness, in terms of integrity, is low, the target's distinctive behaviors are likely to be re-evaluated by observers using internal causality. Likewise, when target's perceived trustworthiness, in terms of integrity, is high, the target's distinctive behaviors are likely to be re-evaluated by observers using external causality.
- *Hypothesis 7* (H7): When target's behavior is distinctive, groups tend to have low consensus about the target's trustworthiness.
- *Hypothesis 8* (H8): When target's words and actions are found inconsistent, groups tend to have low consensus about the target's trustworthiness.
- *Hypothesis 9* (H9): when target's social network is sensitive; this group tends to have fewer consensuses about the target's trustworthiness
- *Hypothesis 10* (H10): The group sensitivity has a significant influence on the perceived trustworthiness, in terms of integrity, of the target. The higher the group sensitivity is, the more likely for the group to detect inconsistency between target's words and actions. However, the group sensitivity is less likely to be influenced by false accusation. And, the strength of the questions from the group is less likely to influence the perceptions of the target's trustworthiness.
- *Hypothesis 11* (H11): When target's perceived trustworthiness, in terms of benevolence, is relatively low, the target's distinctive behavior can be found. However, this hypothesis (H11) is subject to the establishment of Hypothesis 2 (H2).
- *Hypothesis 12* (H12): When target's perceived trustworthiness, in terms of competence, is relatively low, the target's distinctive behavior can be found. However, this hypothesis (H12) is subject to the establishment of Hypothesis 3 (H3).

PRELIMINARY DISCUSSION

Activities in Day 1 and Day 2 were designed for the purpose of enhancing group rapport. The bait offered at the end of Day 2 game stirred up a conflict of interest for the target team-leader. The results of participant observations confirmed these hypothesized situations during the full-scale experiment.

Specifically, these experimental situations can be explained as follows. The team-leaders for Group 1 and 2 were not presented with bait. The group sensitivity toward the team-leader's perceived trustworthiness was increased in Group 1 through questioning of the leader's behavior. As a result, the Group 1 team-leader suffered from being questioned about his integrity and fairness by his team members throughout the game. He allocated more micro-payments to his team members than to himself. In Group 2, the group sensitivity was decreased toward the team leader's perceived trustworthiness. The general perceptions of the leader's trustworthiness remained high. The results showed that Group 2 had the most harmonous teamwork atmosphere.

The team-leaders in Groups 3 and 4 in the full-scale experiment were all presented with monetary bait. All three team-leaders took the bait without letting the team members know. As for Group 3, bait was offered to the team-leader and the group sensitivity was enhanced through encouraging discussion about the target's questionable behavior. The attribution toward the perceived trustworthiness dropped in Group 3. The target team-leader in Group 4 was also given, bait but group sensitivity was reduced through discouraging discussion. The attribution toward the target's perceived trustworthiness dropped accordingly. However some team members could sense the team-leader's behavioral change. One of the survey results, as illustrated in Figure 7, indicates groups attribution over target's integrity in the dimension of steadfastness (Mayer, 1999).

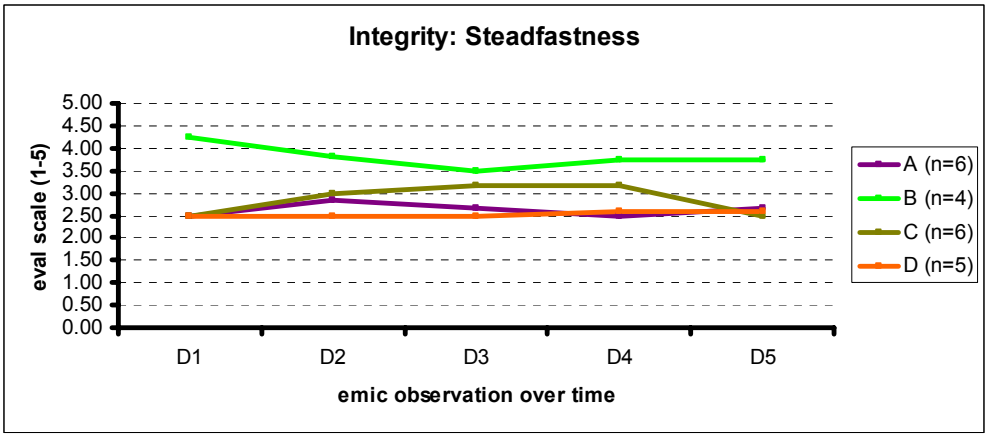

Figure 7. Group Attribution of Target's Integrity in the Dimension of Steadfastness

In all these experiments, the participants did not know that the target Team-Leader was being manipulated by the Game-Master – which occurred in the background. Since insufficient evidence existed regarding the target, the resulting perceptions depended primarily on whether the target's behavior was generally reliable or ethical, and on the outcome itself. The perception of all targets' behavior was nearly all positive from Day 1 through Day 3 during the experiment. This inferred that the targets' competence in leading the team was found to be satisfactory, and their communicated intention was found to be consistent in terms of his information behavior. In this case, the "anomalous" behavior was not found to be significant. However, the outcome of the targets' behavior for Group 3 and 4 turned negative on Day 5, and the level of the target's integrity dropped as a result of taking the bait. Thus, the target's anomalous behavior was found to be significant.

CONTRIBUTION AND CONCLUSION

This study explains and has capacity to predict insider threats by attributing trustworthiness of individuals in virtual organizations. However, it obviously has some limitations. Human perception is not fully reliable due to the fact that not all information is made transparent to the perceivers. Humans attribute their perception of people's trustworthiness based on limited social interactions; however, group consensus overcomes the limitations of fallible individual observations. Most of the attributions are context-specific, time-dependent and combined with judgment regarding the target's capability. Basic struggles of personal gain, selfishness and greediness remain. Ethical values and moral standards are vaguely defined by society and therefore vaguely adopted by individuals.

The findings demonstrate that it is possible to trace and detect anomalous information behavior of an insider leader with a high degree of accuracy, although the leader's change in behavior is subtle. It is believed that this framework of trustworthiness attribution can be generalized through future understanding of human conversational acts and logic. The intellectual merit of this socio-technical study lies in its rigorous capability in building the theory of trustworthiness attribution that tackles complex insider threat problems. By adopting a social psychological theory to predict human trustworthiness in a virtual collaborative environment, findings of this study may contribute to research in geographic dispersed virtual teams, online communities, virtual organizations, and virtual worlds. The broader impact of this theory of trustworthiness attribution can be generalized and formalized to build a social computing model, and socio-technical systems, for insider threat prediction.

ACKNOWLEDGMENTS

I thank Jeffrey M. Stanton, my advisor, for his constant support and insight. I also wish to thank Conrad Metcalfe for his helpful comments and editing assistance.

REFERENCES

1. Ajzen, I. (1991) *Organizational Behavior and Human Decision Processes,* **50,** 179-211.
2. Beck, L., & Ajzen, I. (1991) *Journal of Research in Personality,* **25,** 285-301.
3. General Pace, P. (2007) In *C4ISR Joint Chief of Staff*(Ed, Staff, U. S. J. C. o.) United States Joint Chiefs of Staff.

4. Giddens, A. (1979) *Central Problems in Social Theory: Action, Structure and Contradiction in Social Analysis,* University of California Press, Berkeley, CA.
5. Hardin, R. (1996) *Ethics,* **107,** 26-42.
6. Hardin, R. (2001) In *Trust in Society*(Ed, Cook, K. S.) Russell Sage Foundation, New York, NY, pp. 3-39.
7. Hardin, R. (2002) *Trust and Trustworthiness,* Russell Sage Foundation, New York, NY.
8. Hardin, R. (2003) In *Trust and Reciprocity: Interdisciplinary Lessons From Experimental Research*(Ed, Ostrom, E., & Walker, J.) Russell Sage Foundation, New York, pp. 80-101.
9. Heider, F. (1944) *Psychological Review,* **51,** 358-374.
10. Heider, F. (1958) *The Psychology of Interpersonal Relations,* John Wiley & Sons, New York, NY.
11. Ho, S. M. (2008a) In *Cyber Warfare and Cyber Terrorism*(Ed, Janczewski, L. J., & Colarik, Andrew M.) Information Science Reference (an imprint of IGI Global), Hershey, PA, pp. 206-215.
12. Ho, S. M. (2008b) In *Social Computing, Behavioral Modeling, and Prediction*(Ed, Liu, H., Salerno, John J., & Young, Michael J.) Springer, Tempe, AZ, pp. 129-140.
13. Ho, S. M. (2009) In *Social Computing, Behavioral Modeling, and Prediction*(Ed, Liu, H., Salerno, John J., & Young, Michael J.) Springer, Tempe, AZ, pp. 113-122.
14. Holmes, J. G., & Rempel, J.K. (1989a) In *Review of Personality and Social Psychology*, Vol. 10 (Ed, Hendrick, C.) Sage, Beverly Hills, CA.
15. Holmes, J. G., & Rempel, J.K. (1989b) In *Close Relationship*(Ed, Hendrick, C.) Sage, Newbury Park, CA, pp. 187-220.
16. Hosmer, L. T. (1995) *Academy of Management Review,* **20,** 379-403.
17. Keeney, M., Kowalski, E., Cappelli, D., Moore, A., Shimeall, T., & Rogers, S. (2005) (Ed, Service, U. S. S.) National Threat Assessment Center & CERT Program.
18. Kelley, H. H., Holmes, J.G., Kerr, N.L., Reis, H.T., Rusbult, C.E., & Van Lange, P.A.M. (1973) *American Psychology,* **28,** 107-128.
19. Kowalski, E., Conway, T., Keverline, S., Williams, M., Cappelli, D., Willke, B., Moore, A. (2008) (Ed, Service, U. S. S.) National Threat Assessment Center & CERT Program.
20. Lewicki, R. J., Bunker, B.B. (1996) In *Trust in Organizations: Frontiers of Theory and Research*(Ed, Kramer, R. M., & Tyler, T.R.) Sage, Thousand Oaks, CA, pp. 114-139.
21. Martinez-Moyano, I. J., & Rich, Eliot H., Conrad, Stephen H., & Andersen, David F. (2006) In *Proceedings of the 2006 Winter Simulation Conference*(Ed, Perrone, L. F., Wieland, F.P., Liu, J., Lawson, B.G., Nicol, D.M., & Fujimoto, R.M.) IEEE, pp. 562-568.
22. Mayer, R. C., & Davis, J.H. (1999) *Journal of Applied Psychology,* **84,** 123-136.
23. Mayer, R. C., Davis, J.H., & Schoorman, F.D. (1995) *Academy of Management Review,* **20,** 709-734.
24. Meyerson, D., Weick, K.E., Kramer, R.M. (2006) In *Organizational Trust: A Reader*(Ed, Kramer, R. M.) Oxford University Press, New York, NY, pp. 415-444.
25. Mitnick, K. D., & Simon, W.L. (2002) *The Art of Deception: Controlling the Human Element of Security,* Wiley, Indianapolis, Indiana.
26. Randazzo, M. R., Keeney, M., Kowalski, E., Cappelli, D., & Moore, A. (2005) (Ed, Service, U. S. S.) National Threat Assessment Center & CERT Program.
27. Rempel, J. K., Holmes, J.G., & Zanba., M.D. (1985) *Journal of Personality and Social Psychology,* **49,** 95-112.
28. Richardson, R. (2007), Vol. 2008 Computer Security Institute.
29. Rotter, J. B. (1967) *Journal of Personality,* **35,** 651-665.
30. Rotter, J. B. (1980) *American Psychologist,* **35,** 1-7.
31. Rotter, J. B., & Stein, D.K. (1971) *Journal of Applied Social Psychology,* **1,** 334-343.
32. Shoda, Y., Mischel, W., & Wright, J.C. (1994) *Journal of Personality and Social Psychology,* **67,** 674-687.
33. Tyler, T. R. (1988) *Law and Society Review,* **22,** 301-355.
34. Tyler, T. R., & Degoey, P. (1996) In *Trust in Organizations: Frontiers of Theory and Research*(Ed, Kramer, R. M., & Tyler, T.R.) Sage, Thousand Oaks, CA, pp. 331-356.
35. Weiner, B. (1985) *Psychological Review,* **92,** 548-573.
36. Weiner, B. (1986) *An Attributional Theory of Motivation and Emotion,* Springer-Verlag, New York, NY.
37. Weiner, B. (2006) *Social Motivation, Justice and the Moral Emotions: An Attributional Approach,* Lawrence Erlbaum Associates, Inc., Mahwah, NJ.
38. Williamson, O. E. (1993) *Journal of Law and Economics,* **34,** 453-500.