Game-Theoretic Modeling and Analysis of Insider Compliance with Security Policy

Andrew P. Moore
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania
apm@sei.cmu.edu

Stephanie Grzenia
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania
sgrzenia@sei.cmu.edu

Jose A. Morales
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania
jamorales@usna.edu

Cody W. Ickes
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania
wickes@sei.cmu.edu

Joshua E. Fallon
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania
jefallon@sei.cmu.edu

William Casey
U.S. Naval Academy
Annapolis, Maryland
wcasey@usna.edu

ABSTRACT

In this paper, we present our work in progress applying game-theoretic modeling and analysis to our study of the effects of policy compliance requirements on shifting insider motivation. We focus on non-malicious employee non-compliance (possibly intentional) with policy and the potential risks introduced from this non-compliance. We view an employee's decision about whether to comply with policy as a cost-benefit tradeoff and use a compliance budget as the mechanism for modeling those decisions. We demonstrate using game theoretic analysis as a powerful modeling technique to represent how the potentially deleterious effects of requiring employees to follow frequent or burdensome requirements to comply with fixed policy can affect employee decision making. By modeling employee motivation as instance-based learning in a game with players represented by fluctuating Markov decision processes, we can identify conditions where employees are driven to more or less risky behaviors. We calibrated our model execution results to a recent meta-analysis of years of policy compliance research, which provided a level of confidence in the fidelity of our model execution results and our related practice recommendations.

CCS CONCEPTS

• **Security and privacy → Social aspects of security and privacy**

KEYWORDS

game theory, modeling and analysis, simulation modeling, policy compliance, insider risk, unintentional insider threat

ACM Reference Format:

Andrew P. Moore, Cody W. Ickes, Stephanie Grzenia, Joshua E. Fallon, Jose A. Morales, William Casey, 2023. Game-Theoretic Modeling and Analysis of Insider Compliance with Security Policy. *In Proceedings of the Annual Computer Security Applications Conference (ACSAC), Workshop on Research for Insider Threat (WRIT),* December 4, 2023, Austin, TX USA.

1 Introduction

Answering the following questions that developers of workforce security policies face can clarify the challenges that organizations face in creating security policies that actually improve organizational outcomes:

- How much does policy compliance actually reduce organizational risk?
- Do employees understand the organization's goals of achieving the operational mission within acceptable risk tolerances?
- Are employees adequately incentivized to accomplish their part of the organizational mission?
- How do employees' self-interest relate to their policy compliance?

Misalignment among an organization's policy requirements, organizational risk, operational mission, and employees' self-interest creates the potential for insider risk. Although perfect alignment is likely impossible, understanding and making difficult decisions about tradeoffs among competing priorities can help optimize forming policies regarding workforce security. In this paper, we describe an ongoing effort to develop a game-theoretic model of workforce policy compliance that is calibrated with the research literature on organizational security policy compliance and provides insight into security compliance within specific organizational contexts. This work extends our previous game-theoretic modeling in this area [3].

We use game theory as a mechanism to rigorously represent and simulate the actions performed by players (i.e., employees in this case) with certain motivations and investigate the interactions of those actions. These actions are demonstrated in game outcomes to influence the player, the organization, and the organization's ongoing efforts. From these outcomes, we can observe insights that lead to potential informed policy modifications in terms of compliance requirements, cost of compliance, rewards for compliance, and repercussions for non-compliance. The desired outcome of these modifications is a policy that balances productivity and risk with the risk appetite of the organization and supports the development of policies that improve the likelihood of compliance.

The design prototype tool we used for game-theoretic simulation of policy compliance encodes a few high-level assumptions:

1. Policies are designed with the intent that compliance reduces—or at least does not increase—risk.
2. The effort required to comply with policy requirements reduces— or at least does not improve—productivity, compared to what it would have been without compliance.
3. Higher productivity and lower risk are, in general, considered to be preferable.

Insider risk is exhibited through player non-compliance with policy. While we assume players are non-malicious in nature, they may lose the motivation to comply with policy for various reasons. The game prototype provides insight into human behaviors with certain strategic combinatorial modifications of cost, reward, repercussions, and requirements. These insights, in turn, inform the creation and drafting of future policies and procedures to improve employee security policy compliance.

We structured this paper to first describe the background research on security policy compliance. This provides a basis for the game design in the second section and a means for calibrating the game model execution described in the third section. We conclude the paper with a summary of the contributions of our current research and directions for future research and development.

2 Relevant Security Policy Compliance Research

For some time, studies of organizational behavior have shown that employees' decisions on whether to comply with security policies are based on a cost-benefit assessment by the employee. "Employees weighed the perceived need for a specific security policy against the effort required to comply with it" [1]. Recent experiments show that higher required effort (i.e., response cost) to comply with policy hinders even well-intentioned employees [7].

Potential response costs include the time and resources required to comply that detract from the employees' primary duties. Other response costs might include the hassle associated with compliance tasks and uncertainty about the policy's true value. Response benefits might include rewards for complying (e.g., as part of a performance appraisal), avoiding associated punishments for not complying, and positive feelings about protecting organizational assets.

From a purely conceptual viewpoint, some have described employee compliance behavior along an indifference curve plotted on a graph with the x-axis as cost and the y-axis as benefit, as shown in Figure 1. The area above and to the left of the curve is the cost-benefit associated with compliance. The area below and to the right of the curve is the cost-benefit associated with non-compliance. [1]

While the cost-benefit assessment is a useful way to think about how employees make decisions regarding compliance, it is not a calculation that is actually going on in the employees' heads. Employees typically do not explicitly think of costs and benefits when making compliance decisions. Beautement and Sasse have described a conceptual mechanism called the compliance budget as a way of thinking about how employees make compliance decisions [2].

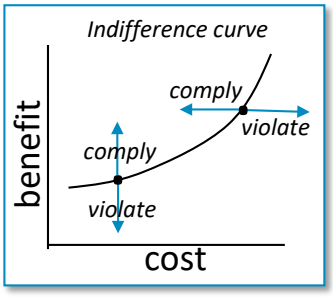

Figure 1: Indifference Curve Representing Employee Compliance with Respect to Cost and Benefit

These researchers lay out the four principles of the compliance budget as follows [1]:

1. *There is a limit to the amount of perceived effort an employee will expend on security tasks.*
2. *This means that any individual security policy associated with significantly more perceived cost than benefits is less likely to be followed.*
3. *When employees expend compliance effort, it accumulates over time, and once an employee's compliance limit is reached, they are less likely to follow any security policy with compliance effort.*
4. *The rate at which the budget is spent matters: the more rapidly they approach their limit, the less likely they are to comply.*

In a nutshell, a person's compliance budget represents the willingness they have, given their current situation, to comply with security policy.

2.1 Insider Risk Types and Motivation

Most of the literature on employee security policy compliance does not assume malicious intent (i.e., an intent to specifically harm the organization). In fact, non-malicious insider risk exhibited through policy non-compliance is the focus of our game-theoretic modeling. Much of the literature applies to this domain for our modeling, and recent advances in insider risk research elaborate the critical path to insider threat in the non-malicious space [8]. As shown in Figure 2, the multiple approach pathways to insider threat (MAP-IT) shows that unintentional, ambivalent, and intentional (i.e., malicious) behaviors all lead to violations of workplace norms (i.e., non-compliance with security policy).

In the MAP-IT framework as illustrated in Figure 2, an insider can exhibit ambivalent behavior when faced with divided loyalties. A divided loyalty can arise for many different reasons. For example, the primary divided loyalty of concern in the security clearance process occurs when a security clearance applicant's loyalty to another nation state or ideology competes with national security interests. Although these divided loyalties may certainly come into play in malicious insider risk, lesser divided loyalties can influence an employee's security policy compliance decisions, such as family needs that compete with work demands or even security policy compliance requests that compete with regular job demands for productive task completion. Ambivalence is viewed as an unstable state in MAP-IT because of the cognitive dissonance that the divided loyalty creates. The discomfort that people naturally associate with

cognitive dissonance moves them toward resolution with either unintentional or intentional behavior.

The unintentional and intentional behavior states are called attractor states in Figure 2; both represent the move away from the unstable state of ambivalence. Unintentional insider behavior is characterized by mental lapses and mistakes that lead to non-compliance. Often unintentional insider behavior is well-intentioned but error prone. But ambivalence also leads employees to commit acts of non-compliance that are impulsive, largely unthinking, and that resolve immediate difficulties in the simplest and least uncomfortable way. Many studies have shown that unintentional insider threat behavior is much more common than intentional (i.e., malicious) insider threat behavior.

Although not as frequent as unintentional insider threat behavior, intentional insider threat behavior is often more severe in its consequences. While we do not explicitly consider intentional insider threat in our current modeling effort, we discuss it here briefly for completeness and as a subject of future model refinement. Schoenherr describes three types of intentional, malicious insider behavior [8]:

- Antisocial—behavior intended to harm the organization
- Prosocial—behavior intended to help people outside the organization
- Asocial—behavior intended to help themselves

All three types of intentional behaviors are purposefully conducted by an individual to the knowing detriment of the organization. These behaviors are ill-intentioned and often characterized by deliberate planning because of the perception of unjust treatment by the organization or its personnel [9].

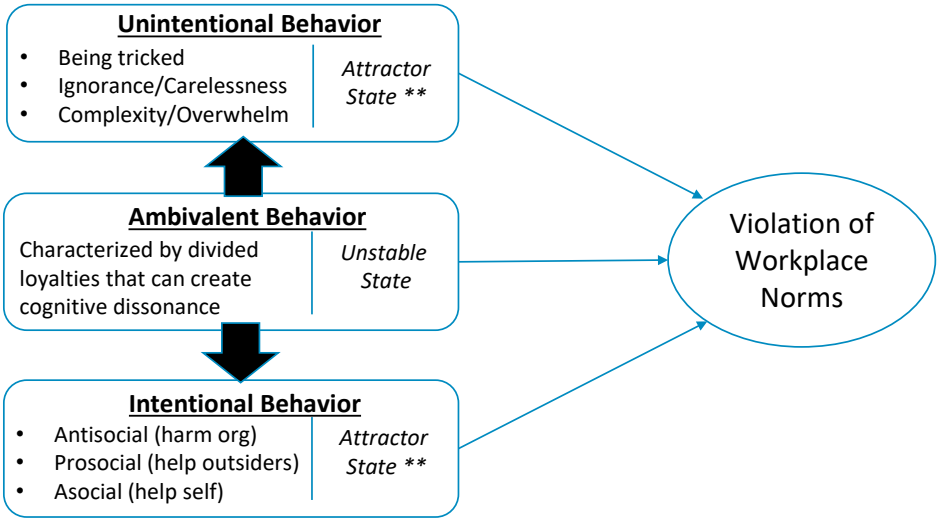

Figure 2: The MAP-IT Framework

2.2 Compliance Control Effectiveness

Consulting scientific organizational behavior literature can be an important way to gain confidence in simulation models, particularly when organizational data is not readily available. It is like the way you might calibrate a measuring device as part of a physics experiment; a simulation model's execution can be calibrated to behave in conformance with the findings in the scientific literature. The most credible research findings are often found in a meta-analysis—a research effort to determine the most important findings across a large number of individual research studies. A meta-analysis "enables [a] researcher to discover the consistencies in a set of seemingly inconsistent findings and to arrive at conclusions more accurate and credible than those presented in any one of the primary studies" [6].

Table 1 shows the primary results from the 2019 Cram meta-analysis. The variables found most significant across the 95 research studies included in their analysis are shown in the leftmost column. All of these variables were found to be statistically significant with effect sizes shown both qualitatively from large to small (based on norms in behavioral science research) and quantitatively (shown as the Pearson correlation coefficient). The number of studies included each variable, and the total sample size across these studies is also shown. We included a rough categorization of the variables to indicate whether they involve a benefit of compliance, or a cost of compliance as discussed earlier.

There are several takeaways from this meta-analysis. Two of the most common practices to improve compliance—compliance rewards and punishments—are at the lower end of effectiveness. This is not to say that they are unimportant but only that their application should be expected to result in only modest gains in compliance. Two factors appear to have a middle-to-large effect on compliance:

- Employee confidence related to security (as reflected in the variables *self-efficacy*, *response efficacy*, and *perceived ease of use*)
- Management support (as reflected in the variables *organizational support*, and *security education & training*)

Organizational culture issues have a major effect on compliance but are often the most difficult to change attitude toward positive compliance, behavioral (personal) norms and ethics, and normative beliefs. While these factors are stubbornly resistant to change, the improvement in the less effective factors might gradually move the culture in a more positive direction.

One factor that has a major effect on actual compliance behavior (rather than merely the intention to comply) is response cost. In experiments that researchers conducted, higher required effort of policy compliance hinders even well-intentioned employees [7]. Guidance provided by prominent researchers recommends not prioritizing efficiency of security policy design [5]. This guidance suggests the need to consider response costs when fine-tuning policies (or rather the implementation of policies) for different groups within an organization based on the nature and needs of the group. In other words, fine-tuning policy implementation to reduce response costs can be a major factor in increasing compliance, but that fine-tuning may vary across groups within the organization.

Table 1 presents study results that are averaged across organizations of varied cultures. If more information is known about the culture of the organization targeted as the object of the simulation, that information may be useful in customizing the model to represent that organization more accurately. Studies may help in this customization if relevant organizational distinctions are accounted for in the studies. For example, Table 2 shows the primary differences in the effect sizes for compliance-related variables for three different geographic regions: North America, Europe, and Asia-Pacific. These regions differ in their individualism vs. collectivism, and their perceptions on freedom, hierarchy, and organizational control/power. The rest of this paper develops a model that is calibrated to the average organization represented in the studies (as in Table 1) rather than a customized model representing a specific organizational culture.

Table 1: Meta-Analysis of Previous Compliance Research [4]

Variable Studied	Effect Size Magnitude	Overall Effect Size	Number of Studies	Total Sample Size	95% Conf. Interval
Perceived Usefulness (BC)	Large	0.651	7	1955	[0.60,0.70]
Personal Norms & Ethics	Large	0.579	20	4,970	[0.55,0.61]
Response Cost (CC)	**Large**	**-0.568****	**25**	**5,271**	**[-0.63,-0.51]**
Attitude	Large	0.564	37	10,975	[0.54,0.59]
Normative Beliefs	Large	0.531	43	12,416	[0.51,0.55]
Organizational Support	Large	0.518	12	2,749	[0.48,0.56]
Self-Efficacy	Medium	0.447	57	14,014	[0.43,.47]
Response Efficacy (BC)	Medium	0.442	24	6,019	[0.41,0.47]
Perceived Benefits (BC)	Medium	0.432	11	2,274	[0.39,0.48]
Security Education & Training	Medium	0.418	30	8,398	[0.39,0.44]
Detection Certainty (BC)	Medium	0.416	20	6,520	[0.39,0.44]
Perceived Ease of Use (CC)	Medium	0.381	7	1,788	[0.32,0.44]
Threat Severity (BC)	Medium	0.342	22	5,700	[0.31,0.37]
Punishment Severity (BC)	**Medium**	**0.323**	**27**	**8,010**	**[0.30,0.35]**
Punishment Expectancy (BC)	**Medium**	**0.317**	**29**	**9,979**	**[0.29,0.34]**
Resource Vulnerability	Small	0.218	20	6,061	[0.19,0.25]
Rewards (BC)	Small	0.09	10	4,812	0.06,0.12]

BC= Benefits of Compliance; CC = Cost of Compliance

Table 2: Meta-Analytic Differences across Cultures [4]

Variable Studied	Moderator Group	Weighted Effect Size	Number of Studies	Total Sample Size	95% Conf. Interval
Normative Beliefs	Asia-Pacific Europe	0.696 0.577	R 5	1.980 1.903	[0.65,0.74] [0.52,0.63]
	N. America Asia-Pacific	0.568 0.696	19 7	4,644 1,980	[0.54,0.60] [0.65,0.74]
Response Cost	**Asia-Pacific Europe**	**-0.129 -0.896**	**8 6**	**1,282 976**	**[-0.19,-0.07] [-0.97,-0.82]**
	Europe N. America	**-0.896 -0.266**	**6 10**	**976 3,113**	**[-0.97,-0.82] [-0.30,-0.23]**
	N. America Asia Pacific	**-0.266 -0.129**	**10 8**	**3,113 1,282**	**[-0.30,-0.23] [-0.19, -0.07]**
Self Efficacy	Asia-Pacific Europe	0.571 0.407	8 10	1,879 2,873	[0.52,0.62] [0.36,0.45]
	Europe N. America	0.407 0.434	10 28	2,873 6,092	[0.36,0.45] [0.41,0.46]
	N. America Asia-Pacific	0.434 0.571	26 8	6,092 1,879	[0.41,0.46] [0.52,0.62]
Response Efficacy	Europe N. America	0.349 0.514	12 7	3,384 1,691	[0.31,0.39] [0.46,0.57]
Security Education & Training	Asia-Pacific N. America	0.486 0.432	10 11	2,548 3,217	[0.44,0.53] [0.39,0.47]
Detection Certainty	Asia-Pacific N. America	0.576 0.342	7 10	1,580 3,876	[0.52,0.63] [0.34,0.41]
Threat Severity	Europe N. America	0.385 0.167	13 5	3,941 1,048	[0.35,0.42] [0.10,0.23]
Punishment Severity	**N. America Asia-Pacific**	**0.285 0.374**	**12 7**	**3,521 1,797**	**[0.25,0.32] [0.32,0.42]**
Punishment Expectancy	**Europe N. America**	**0.251 0.244**	**8 12**	**2,081 3,744**	**[0.20,0.30] [0.21,0.28]**
Resource Vulnerability	Europe N. America	0.109 0.268	8 7	3,088 1,815	[0.07,0.15] [0.22,0.32]

3 The Prototype Game Design

Our goal for this work is to produce a prototype tool that can be used to simulate performer interaction with an organizational policy. Simulation results reflect the effect a policy has on motivating users to comply with or deviate from policy requirements. These results also provide a way for the organization to explore features of policy and organizational culture that can be adjusted to influence performer actions. Users of the prototype tool can experiment with encoding policy to explore unintended negative effects or compare the effects of various compliance requirements when exploring new or updated policies.

The game prototype primarily revolves around three player role types: a (principal) subject, a reviewer, and an adjudicator. These role types act in sequence:

1. A *subject* requests to take an action.
2. The request is passed to a *reviewer* who considers and determines whether the action is allowable by some set of policies and decides to approve or deny the action.
3. The *adjudicator* allows the subject's action to take effect or pursues intervention with the subject.

This three-step sequence (3S) of events is a high-level description of several everyday scenarios that follow the process of comply, review, respond.

The 3S process starts with *comply*, where an individual desires some action to be taken for one of several reasons, and it is in compliance with or in opposition to a policy. In the second step, *review*, the reviewer considers policy conformance and the budget available for auditing, chooses to audit the signaled compliance-related subject action, and indicates if the action should be taken or not. In the final step, *respond*, the adjudicator potentially punishes the subject for detected non-compliance. The person carrying out this action can fully or partially implement this step or not implement it at all; the implementation is chosen resulting from a diverse set of motivations.

Ideally, 3S would always occur in a form that provides an equilibrium of maximum positive benefit to all involved entities, though reality may diverge with both positive and negative benefits to each of the various parties. Since human behavior has critical influence on 3S, the results can be drawn from a large and complex matrix of possibilities. Game theory provides the mechanics of generating the matrix of possibilities and traversing it in a guided process that is driven primarily by the motivations of the players. The player's motivations are further influenced by considering the potential impacts of compliance with or divergence from a policy. Considering and rigorously representing the diverse influence of human behaviors, in terms of motivation, on 3S is a key interest of this work because it can provide insights on policy creation, modification, and sustainment.

The prototype game implements three motivation types of the subject player: compliant, unintentional, and ambivalent. Two of these types are as indicated in the MAP-IT framework [8] shown in Figure 2, and they represent the two primary types of non-malicious insider threat: unintentional and ambivalent. The compliant subject is simply a player that currently tries to do their job while complying with the organization's security policies. The intentional (malicious) behavior state from MAP-IT is outside the scope of the current game, but it could be implemented in the future as an extension or enhancement.

Figure 3 shows the motivational state transitions among these three motivation types. A transition to the unintentional subject motivation state occurs when job stress grows past a certain threshold while compliance stress holds steady. While unintentional subjects are largely well-intentioned in nature, their job stress causes them to fail more often in complying with policy, at least in part due to mental lapses and mistakes made because of job stress. If job stress relaxes, the unintentional subject may in time return to being compliant. However, if the subject's compliance stress starts to increase because of unintentional compliance failures and continued job stress, the subject may enter the ambivalent motivational state.

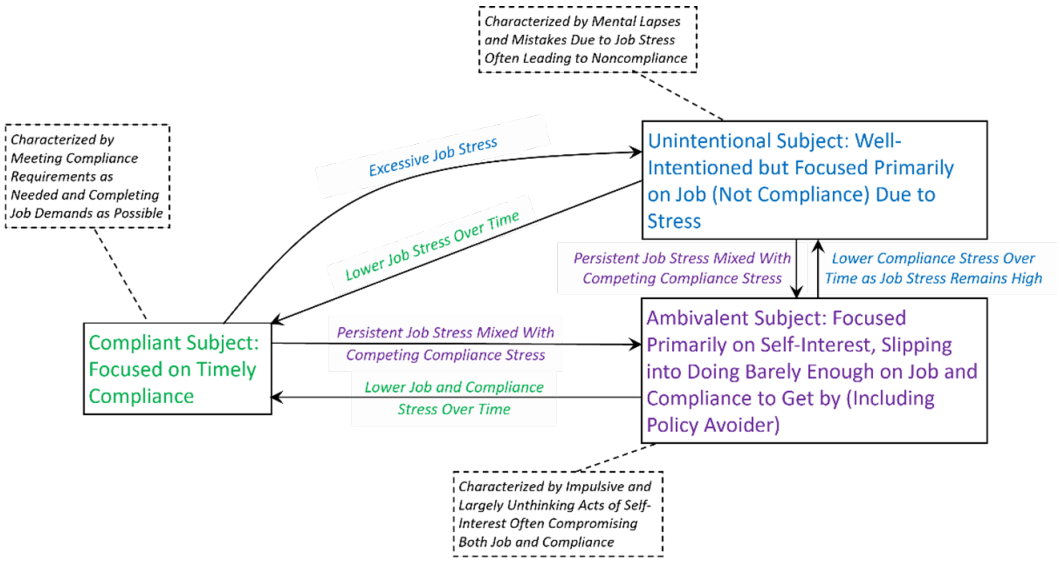

Figure 3: Illustration of the Factors Influencing Shifts Among Motivations

An ambivalent subject may move back to the unintentional state if compliance stress subsides or straight back to the compliant state if both job stress and compliance stress subside. Both transitions likely involve a time lag between entering the state to transitioning to a new state, since subject motivations take time to adjust.

The thresholds at which a subject will transition between any two states are set during game configuration. They are dependent on the context where policy compliance is required, the perceived demand on the subject, and the sensitivity of the subject to both job-related and compliance-related stress.

Figure 4 shows how the prototype tool implements the state transitions depicted in Figure 3. Transitions among the motivational states occur as a function of compliance stress and job stress. A compliant subject remains mostly compliant until job stress increases above a threshold. Above that threshold, a compliant subject becomes an unintentional subject if their compliance stress is below the compliance stress threshold. The compliant subject becomes an ambivalent subject if their compliance stress is above the compliance stress threshold.

An additional insight derived from this work is the discovery of unintended impacts resulting from policy compliance. It is broadly accepted that policy non-compliance impacts negatively on one or more entities. What is less often overtly obvious is that a negative impact can arise from policy compliance. This phenomenon is apparent where the policies being evaluated apply only to some components of an entity.

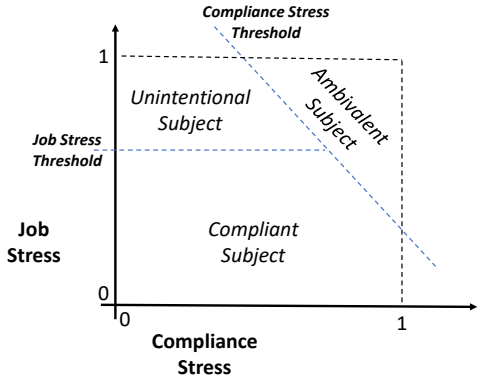

Figure 4: A Model of Player Motivations as a Function of Stress

A familiar example is in the tension between organizational components that focus, respectively, on security compliance and resource availability. For instance, a security operator tasked with vulnerability remediation may be required to install a patch to comply with a software update policy. This compliance may result in the unintended consequence of denying service to internal users of the software.

In our work, we address the discovery of unintended impacts by creating several scenarios in the form of a game. This game contains multiple organizational components that evaluate requests against a policy pertaining only to one component. We used general organizational policies for request adherence, and both the organization and its efforts with its customer base are required to sustain certain operational thresholds. We examine, via game theory, the hypothesis that abiding by organizational policies can have unintended consequences negatively impacting the organization's customer-related efforts. We address this in the game prototype by assigning risk and productivity measures as the basis of determining impacts to the player role types, the organization, and customer-related efforts.

4 Calibrating the Model to the Literature

Our strategy for calibrating model behavior to the research literature is to calibrate based on the relative sizes of the effects of different factors. Calibration to absolute values would have little meaning since the absolute values of compliance within organizations for the factors we studied are not provided in the literature. However, calibration to the effect of the factors relative to each other is both possible and a reasonable first step in fine-tuning the model simulation to improve its credibility.

For example, in the meta-analytic results provided in Table 1, we consider the relative effects of commonly applied punishments for non-compliance on two dimensions: punishment expectancy (to what extent the subject expects to be punished for non-compliance) and punishment severity (the strictness of the punishment for non-compliance). To further explore the cost-benefit aspects of subject compliance decisions, we also consider the impact of response cost (the productivity lost due to compliance). The model simulation should, on average, show that there is about twice the improvement from a decrease in response cost compared to a commensurate increase in punishment expectancy or punishment severity. The effect size of -0.568 for response cost is very roughly two times the effect size of either 0.317 for punishment expectancy or 0.323 for punishment severity. The goal of the simulation is to be in "the right ballpark" rather than excessively precise.

The goal in this calibration effort is to use the relative effect sizes from the literature as general guideposts for refining the model rather than as precise targets. Comparing effect sizes is based on the magnitude of the effect size, disregarding the polarity of that size. Finally, determining what a "commensurate" increase is can be estimated by translating the factors into a common unit of analysis; dollars is often a convenient unit to use.

We calibrated the model for a "moderate" compliance environment. We characterize a moderate compliance environment as one that has average policy compliance rates between 60% and 75%. In contrast, we characterize a "strict" compliance environment as one that has compliance rates greater than 75% on average. Using a moderate compliance rate environment for calibration seems reasonable since the typical organizations studied in research are going to be academic or corporate where the negative consequences of non-compliance are generally more limited than other organizations that may have negative national security implications associated with non-compliance.

The calibrated moderate compliance environment does show in execution that average risk decreases about twice as much for reduced response cost compared to a commensurate increase in punishment expectancy or punishment severity. Interestingly, in addition to reducing risk to a greater extent, reducing response cost also increases average productivity about six times as much as increasing the audit rate. While the compliance research literature does not report on productivity effects, this "side benefit" of reducing response cost is to be expected since response cost is defined as the negative impact on productivity due to policy compliance.

We also tested the impact of reducing response cost and increasing punishment severity and expectancy in a strict

compliance environment. In our simulation, the strict compliance environment had an average compliance rate of 77% while the moderate compliance environment had an average compliance rate of 63% in the baseline run. The scenario we ran was to test which of two options the strict compliance organization should choose in order to best improve their functioning:

1. invest in a training package on compliance that will reduce a player's response cost by 5%
2. or, invest in a user tracking/monitoring tool that will increase the reviewer's audit rate by 5%.

We assume, for simplicity, that these two options cost about the same. Simulation of the strict compliance environment shows that investing in the compliance training package to reduce response cost (option 1) reduces average risk by a third, while increasing the audit rate through the user tracking/monitoring tool leads to virtually no reduction in average risk. Also, as expected, average productivity is also improved much more in option 1 than option 2.

5 CONCLUSIONS

Human decision making is an inherently volatile and complex process, and game theory frequently assumes strictly rational actors who make objective decisions based on numeric inputs. Examples of these inputs include the actors' outcome utility for themselves and their beliefs about the values assigned by other (also perfectly rational) players. However, humans are not perfectly rational, and they are flawed in their assessment of their own preferences and values. Human cybersecurity professionals, for example, execute significantly different activities in a simulation environment than they report in surveys asking them hypothetical questions about the very same computing context.

With that understanding, we caution that the most useful analytical results will derive from games that are calibrated against baseline behaviors that are justified by rigorous support (e.g., observed behaviors in related contexts or research on relevant actors). For example, in a cybersecurity context, the open research on insider threats offers useful statistical insights into the behavior of a professional population. However, these statistical descriptions will need to be adapted to improve model utility in governmental contexts, where organizational structures and individual expectations differ from the corporate workplace.

The prototype game-theoretic simulation model we describe in this paper is not intended to be used to predict human behavior or render judgement on the goodness or badness of an individual policy. Rather, the outputs of this simulation represent an interpretation of how, overall, policy requirements, enforcement, auditing, and feedback affect performance compliance. Observed concentrations of deviant behavior, decreased productivity or mission capability, or increased risk are intended to be interpreted based on the user's expertise.

The simulation model we describe is a work in progress. Our progress shows that the simulation demonstrates behavior that is consistent with the scientific literature. We will continue to calibrate model execution with other data from the literature. Differences in culture can be accommodated, to a limited extent, when studies distinguish results based on cultural attributes, as Table 2 did for different geographic regions. We also plan to calibrate the model with organizational data as it becomes available. The game-theoretic model can be configured to represent the behavior that can be expected in a variety of organizational contexts.

Calibration with the literature will be accomplished for corporate and academic environments because these contexts are typically the ones covered in scientific studies. Configuration of the model to a governmental context will result in different behavior. Nonetheless, the calibration of the model to the literature should still improve the overall credibility of the simulation results. Integrating risk measures with productivity measures will help establish measures of mission readiness that move beyond one-dimensional thinking to more realistic goodness measures for evaluating security compliance policies and mechanisms.

ACKNOWLEDGMENTS

The authors would like to thank SEI business development personnel – Harold Ennulat and Morgan Farrah; and the technical editors of this paper—Barbara White and Sandy Shrum.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM24-0448

REFERENCES

[1] Adam Beautement and Angela Sasse. 2009. The Economics of User Effort in Information Security. *Computer Fraud & Security 2009,* 10, 8–12. https://doi.org/10.1016/S1361-3723(09)70127-7.

[2] Adam Beautement, M. Angela Sasse, and Mike Wonham. 2008. The Compliance Budget: Managing Security Behaviour in Organisations. In *Proceedings of the 2008 New Security Paradigms Workshop*, *NSPW '08*. Association for Computing Machinery, New York, NY, 47–58. https://doi.org/10.1145/1595676.1595684.

[3] William Casey, Jose Andre Morales, Evan Wright, Quanyan Zhu, and Bud Mishra. 2016. Compliance Signaling Games: Toward Modeling the Deterrence of Insider Threats. *Computational and Mathematical Organization Theory* 22 3, 318–49. https://doi.org/10.1007/s10588-016-9221-5.

[4] W. Alec Cram, Jeffrey Proudfoot, and John D'Arcy. 2019. Seeing the Forest and the Trees: A Meta-Analysis of Information Security Policy Compliance Literature. In *Proceedings of the 50th Hawaii International Conference on System Sciences*. 2007. 4051–4060. http://hdl.handle.net/10125/41649.

[5] W. Alec Cram, Jeffrey G. Proudfoot, and John D'Arcy. 2020. Maximizing Employee Compliance with Cybersecurity Policies. *MIS Quarterly Executive* 19, 3, 183–98. https://doi.org/10.17705/2msqe.00032.

[6] Morton Hunt. 1997. How Science Takes Stock: The Story of Meta-Analysis. Russell Sage Foundation.

[7] Jeffrey Jenkins, Alexandra Durcikova, and Jay F. Nunamaker, Jr. 2021. Mitigating the Security Intention-Behavior Gap: The Moderating Role of Required Effort on the Intention-Behavior Relationship. *Journal of the Association for Information Systems* 22, 1. https://doi.org/10.17705/1jais.00660.

[8] Jordan Richard Schoenherr, Kristoffer Lilja-Lolax, and David Gioe. 2022. Multiple Approach Paths to Insider Threat (MAP-IT): Intentional, Ambivalent and Unintentional Insider Threats. *Counter-Insider Threat Research and Practice* 1, 1. https://citrap.scholasticahq.com/article/37117-multiple-approach-paths-to-insider-threat-map-it-intentional-ambivalent-and-unintentional-insider-threats.

[9] Yucheng Zhang, Xin Liu, Shan Xu, Liu-Qin Yang, and Timothy C. Bednall. 2019. "Why Abusive Supervision Impacts Employee OCB and CWB: A Meta-Analytic Review of Competing Mediating Mechanisms. *Journal of Management* 45, 6, 2474–97. https://doi.org/10.1177/0149206318823935.