Modeling Expert Judgments of Insider Threat Using Ontology Structure: Effects of Individual Indicator Threat Value and Class Membership

Frank L. Greitzer
PsyberAnalytix,
Richland, WA, USA
Frank@PsyberAnalytix.com

Justin Purl
Human Resources Research
Organization, Alexandria, VA, USA
jpurl@humrro.org

D.E. (Sunny) Becker
Human Resources Research
Organization, Alexandria, VA, USA
sbecker@humrro.org

Paul J. Sticha
Human Resources Research
Organization, Alexandria, VA, USA
psticha@humrro.org

Yung Mei Leong
Independent Consultant, Hyattsville,
MD, USA
y.leong03@gmail.com

Abstract

We describe research on a comprehensive ontology of sociotechnical and organizational factors for insider threat (SOFIT) and results of an expert knowledge elicitation study. The study examined how alternative insider threat assessment models may reflect associations among constructs beyond the relationships defined in the hierarchical class structure. Results clearly indicate that individual indicators contribute differentially to expert judgments of insider threat risk. Further, models based on ontology class structure more accurately predict expert judgments. There is some (although weak) empirical evidence that other associations among constructs—such as the roles that indicators play in an insider threat exploit—may also contribute to expert judgments of insider threat risk. These findings contribute to ongoing research aimed at development of more effective insider threat decision support tools.

1. Introduction

A serious threat is posed by insiders who seek to destroy, steal, or leak sensitive information, or act in ways that expose their organization to outside attacks. An insider threat is "a current or former employee, contractor, or other business partner who has or had authorized access to an organization's network, system, or data and who intentionally (or unintentionally) exceeds or misuses that access to negatively affect the confidentiality, integrity, or availability of the organization's information or information systems" [1]. Annual industry surveys consistently show that insiders pose the second greatest cybersecurity threat, exceeded only by hackers, and that insider attacks are the costliest to organizations [2][3]. An active research area for development of more effective detection and mitigation approaches is the identification, validation, and integration of cyber and behavioral (sociotechnical) indicators of insider threat risk [4][5][6].

This paper describes continuing work on a comprehensive insider threat ontology [6][7] that supports research to develop more effective decision support tools, facilitates insider threat program evaluation s, and promotes understanding of the complex insider threat domain. A hallmark of the ontology—called Sociotechnical and Organizational Factors for Insider Threat (SOFIT)—is the inclusion of behavioral, social, and organizational factors in addition to the cyber/technical factors traditionally identified with insider threat risk. A general description of SOFIT and its class structure is provided in [7]. While the ontology was based originally on a unidimensional hierarchical taxonomy of factors, relationships have been specified to more fully characterize additional associations among insider threat indicators and related constructs; these associations extend the ontology beyond the simple hierarchical taxonomy from which it was derived. It now represents a collection of taxonomies. Indeed, this paper focuses on how the additional specification of associations among constructs yields a broader ontology that further informs insider threat assessment and mitigation. A primary objective of the current research is to examine how individual indicators and patterns of indicators contribute to judgments of insider threat risk. Though preliminary and requiring further research, results suggest both research and operational implications favoring the inclusion of behavioral and sociotechnical indicators.

URI: https://hdl.handle.net/10125/59756
ISBN: 978-0-9981331-2-6

2. General model

A general context and framework that informs the SOFIT ontology is shown in Fig. 1. This framework depicts presumed underlying factors and processes at work as one progresses along a critical pathway that may culminate in a malicious insider exploit, consistent with the Critical Pathway model described in [8]. Since this framework does not describe processes associated with unintentional (non-malicious) insider threats [9][10], it only partially informs SOFIT.

The model distinguishes personal (individual) factors from external factors, and distal factors from proximal factors. Personal factors include psychological constructs and predispositions, internalized cultural norms and ideology, and capabilities (i.e., knowledge, skills, abilities), which, when combined with external factors, may increase the individual's motivation to act. Personal factors comprise the proclivity or vulnerability to malicious insider activity and include personal predispositions. External factors include stressors, opportunities that present themselves, and actions by the organization that may impact motivation. Distal factors include internal triggering processes, where personal and external factors generate an emotional/cognitive response that culminates in malicious intent. Proximal factors are behaviors that lead to an attack. While proximal factors are the most likely to be identified following the crime, we suggest that distal factors reflecting motivations may be most useful for proactive approaches that attempt to identify individuals who pose greater risks of committing these crimes. Altogether, these processes describe the complex mechanism at play for any potential malicious insider threat. This framework is strongly influenced by earlier works that describe the CMO (capability/motivation/opportunity) model (e.g., [11]), the critical pathway model [8], and numerous behavioral/psychologically oriented works (e.g., [1][3][4][5][7]).

Each element of this framework can be directly or indirectly measured. Following the approaches described in [4] [12], we decomposed the various constructs into a hierarchical set of factors that supports analyses of data to infer observables, indicators and threat behaviors. In this proactive computational approach to insider threat mitigation, *data* are processed to reveal *observables*; collections of *observables* are analyzed to infer *indicators* (collections or patterns of *observables*); *indicators* are examined to infer target *behaviors*. Malicious (threat) behaviors are combinations or sequences of indicators and observables that represent a pattern of actions associated with an exploit. Recognizing target threat behaviors is therefore a complex, model-based classification process that involves inferences about multifaceted combinations or sequences of behavioral, psychological, and technical indicators. This interpretation of the threat assessment process provides a key rationale for related modeling efforts and the design of expert knowledge elicitation studies initially reported in [6] and [7] and extended here.

3. SOFIT Framework

The SOFIT ontology derives from a systematic review, analysis and synthesis of existing research, case studies, and guidelines by the insider threat research community. It is currently 6-7 levels deep, with 271 constructs defined as individual (human) factors and 49 as organizational factors [7]. Classes, which represent objects with similar structure and properties, are arranged hierarchically: subclasses or members of the classes are referred to as *indicators* [4][5][12].

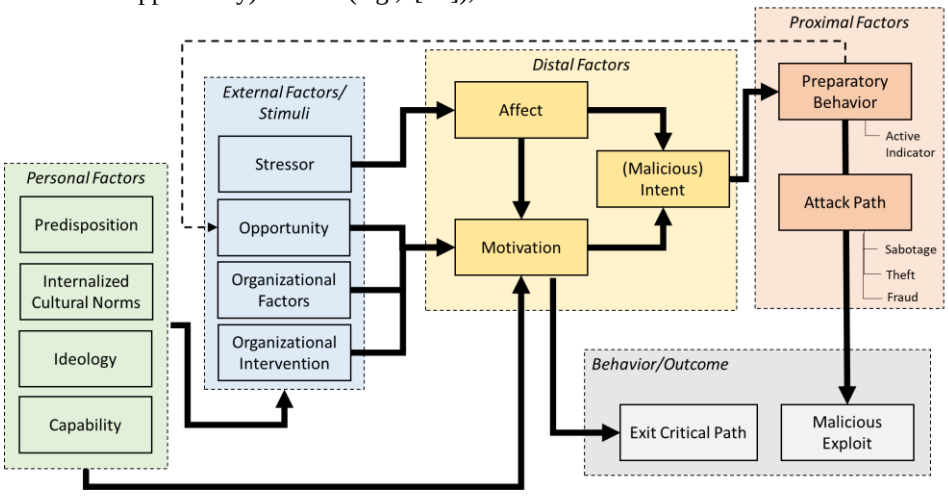

Figure 1. General Model

Fig. 2 shows the main classes comprising the upper levels of the hierarchy. Fundamentally, the ontology attempts to describe individuals and organizations with various characteristics that increase the likelihood that an intentional or unintentional insider threat will occur. The ontology addresses both malicious and non-malicious (unintentional) insider threats, and it distinguishes between actions performed by insiders from those by organizations (e.g., problematic responses to potential threats, poor institutional policies, or security practices). SOFIT is broader and deeper compared to other insider threat ontologies (e.g., [13][14]). The constructs *Factor* (comprising *Individual* and *Organizational* factors), *Actor* (comprising *Person* and *Organization*) and *Intention* (*Malicious* versus *Non-Malicious*) are at the top of the Insider Threat hierarchy. Classes deeper in the hierarchy largely consist of groupings of characteristics at various levels of abstraction. The groups of classes may be related by co-occurrence or cause and effect. Characteristics at the lowest level of abstraction are differentiated by threat type, indicator role, and level of concern.

This paper primarily focuses on the individual factors associated with insider threats; five parent classes and underlying indicator classes are shown in Fig. 3. The ontology accounts for both malicious and non-malicious (unintentional) insider threats, and it distinguishes between actions performed by employees (as insiders) and actions performed by organizations (such as problematic responses to potential threats, poor institutional policies, or security practices). Individual factors reflect behaviors, attitudes, personal issues, sociocultural or ideological factors, and various biographical (life narrative) factors that may indicate increased risk. Protective factors (i.e., those that decrease risk) are not considered in this work. This branch of the taxonomy reflects the substantial body of work by a diverse set of researchers and practitioners focusing on concerning behaviors, sociocultural

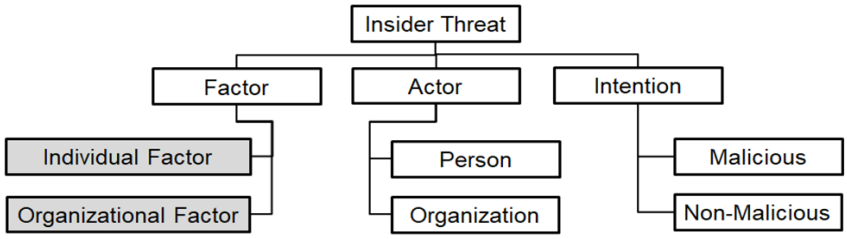

Figure 3. SOFIT Ontology Higher-Level Classes

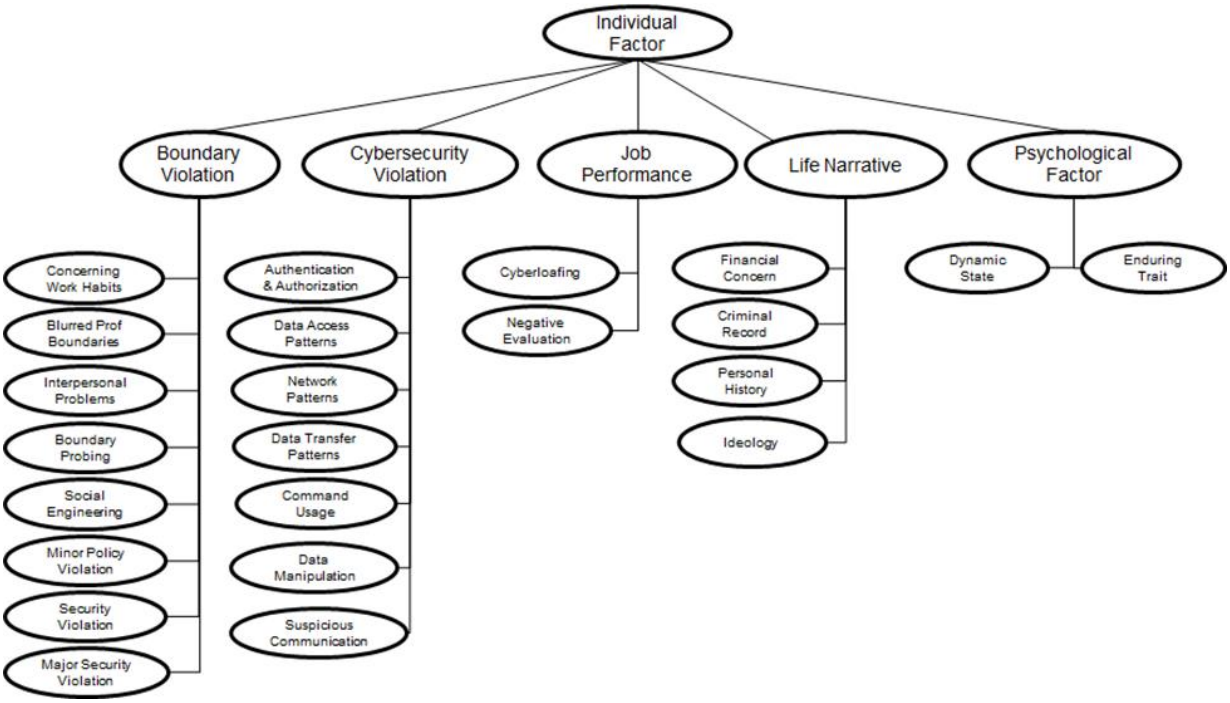

Figure 2. Individual Factor Branch of SOFIT Ontology

factors, and psychological factors underlying insider threats, as shown in Fig.3. Examination and discussion of works relating to psychological constructs (especially [5][15]) led us to differentiate enduring psychological traits from dynamic states, consistent with findings that these two constructs are reliably distinct despite their admitted overlap (e.g., [16][17]) and with the diverse body of psychological research that hinges on (e.g., [18][19][20]) or capitalizes on (e.g., [21][22]) that distinction. Finally, the inclusion of personal history and sociocultural factors derives from research and case studies (e.g., [23][24][25][26]). We adopted a "life narrative" factor construct based on the notion that certain sociocultural factors may be discerned from life narratives of individuals [27].

Fig. 4 depicts lower-level constructs within the *Individual Factor* branch: *viz*., a decomposition of the Job Performance class into two deeper-level subclasses, "Cyberloafing" and "Negative Evaluation." Within each of these subclasses are observables (such as "Excessive Personal Use of Work Computer"); lower-level constructs are defined but not shown in the figure. Measuring observables requires specifying and implementing detectors associated with these constructs (e.g., a detector for excessive personal use of work computer might be number of visits to non-work-related websites). SOFIT stops short of specifying detectors, since these are organization-specific and their specification would likely increase the size of the ontology by an order of magnitude.

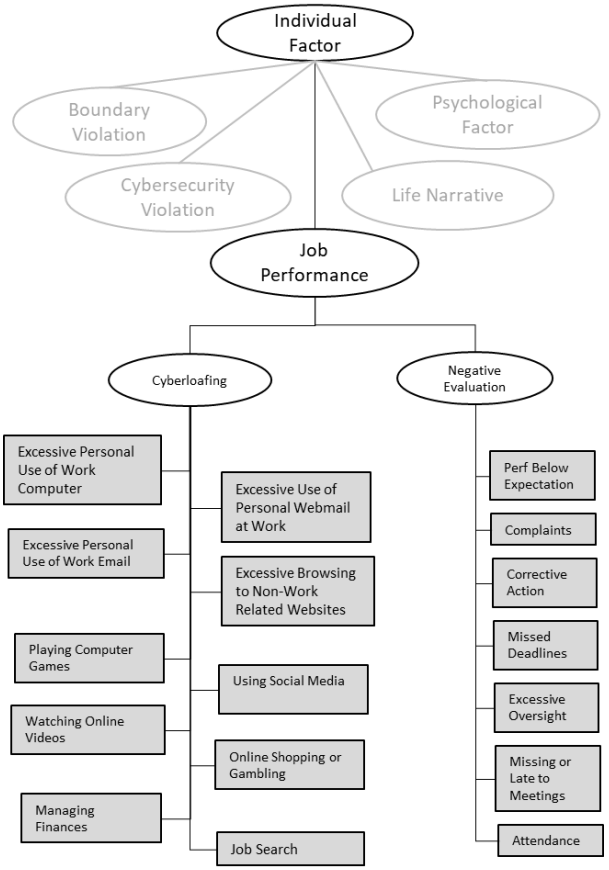

Figure 4. Job Performance branch of Individual Factor hierarchy

4. Associated Constructs

Our research team relied upon research literature and our collective expert judgments to examine numerous associations among the factors represented in the ontology and relationships between these factors and other relevant constructs. We considered possible associations of the insider threat indicators with the following threat types [1]:

- ***Insider Sabotage:*** An act by an insider to direct specific harm toward an organization or its assets.
- ***Insider Data Theft/Exfiltration*****:** Theft of sensitive information by an insider.
- ***Insider Fraud*****:** Modification, addition, deletion, or theft, of an organization's data for personal gain, leading to an identity crime (e.g., identity theft, credit card fraud).
- ***Unintentional Insider Threat (UIT):*** An act or failure to act by an insider, without malicious intent, that causes harm or substantially increases the probability of future harm to an organization or its assets.
- ***Workplace Violence*****:** Any act or threat of physical violence, harassment, intimidation, or other threatening disruptive behavior that occurs at the work site.

Based on our own judgments, the coauthors (FLG, JP, DEB, YML) individually and by consensus identified associations of individual factors with these six threat types. Considered preliminary until validated against independent expert knowledge or empirical evidence, these additional relationships among ontology constructs can support queries to generate lists of factors associated with insider threat types.

Using a similar individual-and-consensus procedure, we considered possible relationships between insider threat indicators and certain constructs that help to describe an indicator's role in the insider threat exploit: Precipitating Event, Predisposition, Behavioral Precursor, Technical Precursor, Access Path, and Contextual Factor (these constructs are defined in Table 1). Table 2 shows output from a query listing factors associated with the role, Predisposition. These factors come from different ontology classes (e.g., *Boundary Violation, Job Performance, Psychological Factor/Enduring Trait*).

Table 1. SOFIT Constructs Characterizing an Indicator's Role in Insider Threat Exploits

Construct	Definition	Examples
Precipitating Event	An event that triggers or motivates the insider to carry out an insider crime	1.1.4.3.4.4.5 Disciplinary Action 1.1.4.3.4.4.6. Passed over for promotion
Personal Predisposition	A characteristic historically linked to a propensity to exhibit malicious insider behavior	1.1.5.2.1.5. Low Honesty-Humility 1.1.5.2.2.1.1. Manipulative 1.1.5.2.2.2. Narcissism
Behavioral Precursor	An individual action, event, or condition that involves personal or interpersonal behaviors and that precedes and is associated with insider activity	1.1.5.1.2.5. Disgruntlement 1.1.5.1.2.6. Overly Critical
Technical Precursor	An individual action, event, or condition that involves computer or electronic media and that precedes and is associated with malicious insider activity	1.1.3.6.3. Delete or edit audit logs 1.1.3.7. Suspicious Communication
Access Path	Sequence of one or more access points along a critical path (also known as "attack vector" or "kill chain")	1.1.3.4.4. Unauthorized storage device 1.1.3.4.1. Attempts to access prohibited file-sharing websites
Contextual Variable	Factor that adds context (not necessarily predictive)	1.1.4.2.1.3. Unexplained affluence 1.1.4.3.1. Age 1.1.4.3.2. Gender

Table 2. Factors associated with the role "Predispositions"

1.1.1.	**Boundary Violation**	1.1.5.2.1.3.2.	Rebellious- Nonconforming
1.1.1.2.	Blurred Professional Boundaries	1.1.5.2.1.4.	Excitement-seeking
1.1.1.2.1.	Excessive Socialization	1.1.5.2.1.5.	Low Honesty- Humility
1.1.2.	**Job Performance**	1.1.5.2.2.	Dark Triad
1.1.2.2.	Negative Evaluation	1.1.5.2.2.1.	Machiavellianism
1.1.2.2.6.	Missing or Late To Meetings	1.1.5.2.2.1.1.	Manipulative
1.1.5.2.	**Enduring Trait**	1.1.5.2.2.2.	Narcissism
1.1.5.2.1.	Personality Dimensions	1.1.5.2.2.2.1.	Self- Centered
1.1.5.2.1.1.	Emotional Instability/ Neuroticism	1.1.5.2.2.2.2.	Grandiosity
1.1.5.2.1.2.	Low- Conscientiousness	1.1.5.2.2.2.3.	Rejects Criticism
1.1.5.2.1.2.1.	Unreliable	1.1.5.2.2.2.4.	Lack of Empathy
1.1.5.2.1.2.2.	Impulsivity	1.1.5.2.2.3.	Psychopathy
1.1.5.2.1.2.3.	Poor Time Management	1.1.5.2.2.3.1.	Callousness
1.1.5.2.1.3.	Disagreeableness	1.1.5.2.2.3.2.	Lack of Remorse
1.1.5.2.1.3.1.	Socially Averse	1.1.5.2.2.3.3.	Sadism

The primary associations of the ontology describe the hierarchical nature of the indicators. The associations between parent classes and child classes such as those illustrated in Table 2 are the major organizing principle for the indicators. Role type associations are considered secondary because indicators within a role are more heterogenous and less closely related. However, this perspective is based on narrative evidence alone. Indeed, it is possible to develop an alternative ontological structure based on roles instead of the class structure described in Section 3. It is therefore appropriate to ask: What aspects of the ontology (e.g., individual indicators, their roles, and parent class relationships) might best account for expert judgments of insider threat? We conducted an expert knowledge elicitation study to address this and related questions.

5. Expert Knowledge Elicitation Study

The ontology class structure is a framework for describing the domain of insider threat indicators. Further, SOFIT's structure encapsulates an inherent schema that we hypothesize analysts use to make judgments about insider threat risk. For example, an

analyst may not be gravely concerned about a case within an organization unless an indicator from the data manipulation class is present. Likewise, the role of an indicator may provide valuable information about how an analyst interprets a case. For example, an analyst may deem a case less worthy of further investigation because there is no precipitating event or other factor indicating a motive. Therefore, the class structure and role structure may provide explanatory power for modeling expert judgments of potential insider threat cases.

5.1. Research Questions

Our research questions center on the hypothesis that expert judgments of the threat/risk rating (level of concern) for an indicator in isolation will provide a relatively powerful way to predict judgments of combinations of indicators (i.e., threat/risk rankings of cases). This is suggested by previous findings [5][7], but here we focus on more specific model-based questions:

1) Do the threat/risk ratings for individual indicators predict the threat ranks of cases? (**Sum-of-Risk** model)
2) Does a count of the number of indicators from each of the parent classes in a case predict the rank of the case? (**Class-Count** model)
3) Does a count of the number of indicators from each of the 6 possible roles predict the rank of the case? (**Role-Count** model)
4) Do the class-count and role-count models contribute independently to the prediction of case rank?
5) Do the class-count and role-count models predict case rank above and beyond sum-of-risk model?

5.2. Method and Procedure

Thirteen experts from at least five participating organizations representing both research and operational experience participated in the study. Participants were recruited via a snowball recruitment method (seeded from our contacts across the research/operational communities). All had more than 5 total years of experience in insider threat or related fields, and 12 experts had 11 or more years of experience. In Part I of the study (survey open for 3 months), each of the participants provided ratings of level of concern for 202 indicators (out of the 271 individual indicators) in the ontology [ratings were on a 0-100 scale, where *0 = no concern at all* and *100 = gravest concern about an actual exploit or strong inclination/likelihood of committing an exploit*). Seven of these experts, all with 11 or more years of experience, opted to go on to Part II of the study (open for 2 months after Part I). In Part II, participants ranked 45 stratified random cases presented as combinations of 2 to 5 individual indicators. The cases were constructed to balance the number of indicators and degree of concern as considered by the experimenters who judged cases as low, medium, or high concern. Except for five cases that were given to everyone as "anchors", the remaining 40 cases varied across participants. An example of a case (considered by the experimenters to represent high concern) is the following, with indicators enclosed in brackets:

[Resigned] [Extreme Discontent] [Establish Backdoor] [Transfer Large Amount of Data] [Strong Reaction to Organizational Sanctions]

Cases comprised simply the list of indicator labels (as above); definitions of indicators were available for review. Instructions for this ranking task were to sort cases into five "bins" corresponding to increasing levels of concern (Low, Low-Moderate, Moderate, Moderate-High, Extreme), and then to rank-order the cases in each bin from highest concern to lowest concern. This results in a rank-ordering for the set of cases. The 315 cases ranked by the 7 experts were the unit of analysis for all regression analyses.

5.3. Results

5.3.1. Quantitative Models. We examined five models that attempt to predict the expert's ranking of the level of threat in the cases, based on the indicators present in them. We use the variable, *R*, to denote the predicted level of threat or risk specified by each of these models.

- **Counting** model. $R = \Sigma x_i$, where x_i has the value of 1 if indicator *i* is present, otherwise 0. Thus, if there are *n* indicators in a case, the risk will be *n*, irrespective of any differences in threat level for individual indicators.
- **Regression** model**.** $R = \Sigma b_i x_i$**,** where b_i is the regression weight for indicator *i*. The regression model estimates many empirically-derived weights to predict the case rankings.
- **Sum-of-Risk** model. The risk for a case is the sum of the ratings of concern (r_i) for the individual indicators contained in the case, i.e., $R = \Sigma r_i x_i$, where the r_i represents the rating of concern for indicator *i*.
- **Class-Count** model sums the weights based on the parent class for each indicator represented in a case, i.e., $R = \Sigma c_{j(i)} x_i$, where $c_{j(i)}$ is an

empirically-derived weight for the parent class associated with indicator *i*.

- **Role-Count** model counts the number of roles represented, or $R = \Sigma r_{k(i)} x_i$,, where $r_{k(i)}$ is an empirically-derived weight for the role associated with indicator *i*.

5.3.2. Modeling Results. Consistent with previous findings [5][7], the predictive power of a **Counting** model (using only the number of indicators observed) provides a logical lower bound on our measure of predictive strength (R^2 = 0.26); whereas, a **Regression** model freely estimating the weight of each indicator on rank provides a logical upper bound (R^2 = 0.76).

We compared the performance of the other three models with the upper and lower bound alternatives presented by the **Counting** and **Regression** models. The **Sum-of-Risk** model predicted the rankings nearly twice as well as the **Counting** model (R^2 = 0.48). The **Class-Count** model predicted case ranks (R^2 = 0.54) slightly better than the **Sum-of-Risk** model. Table 3 shows the beta weights and significance levels for the parent class indicators in the **Class-Count** model[1]. Note that because lower ranks represent higher concern, negative regression weights are expected.

The **Role-Count** model (R^2 = 0.42) was somewhat less predictive of case ranks than the **Sum-of-Risk** model. Table 4 shows the beta weights and significance levels for the **Role-Count** model. Notably, and surprisingly considering that all factors are assumed to reflect some degree of risk/threat, the role type of Personal Predisposition attained a positive weight (β = 0.15).

Including both the class-structure and the role-structure in the analyses provides explanatory power; however, there is a high potential for overlap. The relative incremental validities of parent class and role type were examined to determine if the contributions are independent. Role type predicted significantly beyond parent class, $F(1,312) = 6.237$, $p = .013$, $\Delta R^2 = 0.010$. Similarly, parent class predicted significantly beyond role type, $F(1,312) = 58.447$, $p < .001$, $\Delta R^2 = 0.094$. These results supported the notion that parent class and role type provide independent contributions to the expert judgments of threat rank. Although statistically significant, the incremental validity of role type over parent class was small, which implies parent class accounts for most of the variance and that the independent variance contributed by role type is relatively limited.

Table 3. Weights for Count of Parent Class

Parent Class	β
Boundary Violation	
Concerning Work Habits	-0.13***
Blurred Professional Boundaries	0.00
Interpersonal Problems	-0.16***
Boundary Probing	-0.21***
Social Engineering	-0.05
Minor Policy Violation	-0.12***
Security Violation	-0.28***
Major Security Violation	-0.31***
Job Performance	
Cyberloafing	-0.05
Attendance	0.03
Negative Evaluation	-0.13**
Cybersecurity Violation	
Authentication/Authorization	-0.21***
Data Access Patterns	-0.23***
Network Patterns	-0.32***
Data Transfer Patterns	-0.33***
Command Usage	-0.22***
Data Manipulation	-0.27***
Suspicious Communication	-0.21***
Life Narrative	
Criminal Record	-0.07
Financial Concern	-0.07
Personal History/Major Life Changes	-0.08
Behavioral Health Issues	-0.07
Disloyalty	-0.21***
Radical Beliefs	0.00
Suspicious Foreign Travel	-0.18***
Psychological Factor	
Affect	-0.09*
Attitude	-0.23***
(Concerning) Personality Dimensions	-0.02
Dark Triad	-0.02

*$p<.05$ **$p<.01$ ***$p<.001$

Hierarchical regression was conducted to determine how much variance the **Class-Count** and **Role-Count** models account for beyond the **Sum-of-Risk** model. In the first step, the **Sum-of-Risk** model was considered independently to determine the baseline variance accounted for, R^2 = 0.48. The other two models were added in the second step and any increase in model fit was attributed to the added predictors, ΔR^2 = 0.04. As expected, given the small independent prediction, parent class predicted beyond the sum of ratings of concern (β = -0.34, $p < .001$);

[1] The 29 parent classes in Table 3 differ slightly from the most current representation (Fig 3); subsequent to the study, some indicator classes (e.g., attendance, affect) were placed lower in the hierarchy and therefore do not appear in the figure. We replicated the analyses using the current class structure, obtaining similar weights and an R^2 of 0.52.

whereas, role type does not (β = -0.09, p = 0.25) when both structures are in the same model (see Table 5).

Ultimately, the variance in expert ratings of concern in isolation predicted level of concern for a case relatively well. However, the judgment process for cases is not solely explained as a simple summation of the individual risks. The additional prediction provided by parent class implies that the adjustment from a simple sum involves the parent class structure in some fashion. The lack of additional prediction from role type implies that the adjustment either does not involve role type or that the role type overlaps with parent class or ratings of concern.

Table 4. Weights for Count of Role Type

Role Type	β
Precipitating Event**	-0.13**
Personal Predisposition***	0.15***
Behavioral Precursor***	-0.19***
Technical Precursor***	-0.30***
Access Path***	-0.34***
Contextual Variable*	-0.14*

*$p<.05$ **$p<.01$ ***$p<.001$

Table 5. Incremental Prediction of Case Rank over Sum-of-Risk

Predictor	R^2	β
Step 1***	0.484***	
Sum-of-Risk***		-0.70***
Step 2***	0.525***	
Sum-of-Risk***		-0.33***
Parent Class***		-0.34***
Role Type		-0.09

*$p<.05$ **$p<.01$ ***$p<.001$

6. Discussion

The indicator level-of-concern values and number of indicators from each parent class predict expert judges' ranking of insider threat cases. As expected, the indicators for major security violation, network pattern, and data transfer pattern were the most predictive of the threat ranks for cases. Indicator role type was a substantially weaker predictor of case rankings, with, technical precursors and access path indicators showing the strongest relationship.

When paired with ratings of concern to make predictions about case rankings, the number of parent classes represented in the indicators (class structure) outweighed the number of role types (role structure).

Relative to the **Counting** model and the freely estimated **Regression** model, the predictive strength of the **Class-Count** model lends support for the ontological structure that was built for SOFIT. These results support the notion that the ontological structure aligns with the internal schema used by experts to make judgments about the relative concern of insider threat cases. In contrast, the mixed results of the **Role-Count** model may reflect the fact that judgments of level of concern at the individual indicator level matter more than judgments of indicator role. This suggests that the raters considered both psychological and technical indicators in rough proportion to their risk. Future research should investigate the conditions in which indicator role might have a more substantial impact on judgments of insider threat risk.

As constructed, the SOFIT ontology specifies *concerning* indicators, and consistent with this construction, our study showed that expert ratings of concern for all individual indicators were at least somewhat concerning on average (i.e., >20 on a 100-point scale). Indicators relating to the personal predisposition role type tended to reflect cases that were of lesser concern (positive β weight in Table 4). Since role type does not add validity to level of concern, this result likely reflects the fact that the personal predisposition role has the lowest average concern. Because the number of indicators in a case is limited, the roles do not occur independently in the sample of cases. This means that a case that includes one or more personal predisposition indicators would be likely to include fewer indicators from more concerning roles, such as technical precursors or access paths.

7. Conclusions

The results reported here support the inclusion of behavioral/social indicators of insider threat in modeling expert judgments, and further suggest that operational contexts relying on analyst judgment may benefit from decision support tools that use the SOFIT ontology. More specifically, we conclude that decision support tools for insider threat assessments should take account of the class structure. Further research is warranted to assess the possible impact of indicator role types.

7.1. Limitations

The results of the current and previous expert knowledge elicitation studies [7] are proxies for empirically investigating the predictive strength of indicators in an operational setting. Validity of proposed approaches and models cannot be faithfully

determined without testing in real operational settings. The present line of research on insider threat indicator structure seems to warrant further investigation using real data with ground truth to validate the models beyond the prediction of expert judgments.

7.2. Future Work

To better understand relationships among constructs and their influence on insider threat judgments, additional expert knowledge elicitation studies should be conducted.

We continue to implement SOFIT functionality to support qualitative and quantitative insider threat assessment approaches. The intended solution will provide an interface to explore information beyond a simple list of indicators for a case being evaluated, so that the analyst may be provided additional information that serves to explain or justify the assessment. With this knowledge, the analyst can make an informed decision about forwarding the case for further investigation.

Our team has also begun to implement functionality within the ontology to help an organization assess its insider threat monitoring approach. This is based on a comparison of an organization's indicator portfolio with the domain of insider threat indicators specified in the ontology. This could be used in the design of a web-based assessment tool such as reported in [28], and it would support the mandate of the National Insider Threat Task Force to conduct technology maturity level assessments of US departments/agencies capabilities for detecting insider threats [29].

The present research offers two contributions. First. the results advance research on insider threat indicators. Second, results reported here can facilitate the development of ontology-based operational tools for both insider threat assessment and technical maturity level assessments of organizational insider threat program portfolios. This research will lead to the development and use of more effective, knowledge-based decision support tools for insider threat assessment.

8. Acknowledgments

This research was supported under IARPA contract 2016-16031400006. The content is solely the responsibility of the authors and does not necessarily represent the official views of the U.S. Government. The research study was reviewed and approved by the HumRRO IRB (Memorandum for the Record dated September 1, 2017). The authors gratefully acknowledge contributions from team members KB Laskey, J Lee and A. Zaidi in developing the ontology.

9. References

[1] Cappelli, D.M., A. P. Moore, & R. F. Trzeciak. (2012). *The CERT guide to insider threats: How to prevent, detect, and respond to information technology crimes (theft, sabotage, fraud)*. Addison-Wesley.

[2] *CSO Magazine*. (2011). US Secret Service, Software Engineering Institute CERT Program at Carnegie Mellon University, and Deloitte. "2011 cybersecurity watch survey." January.

[3] Keeney, M., E. Kowalski, D. Cappelli, A. Moore, T. Shimeall, & S. Rogers. (2005). *Insider threat study: computer system sabotage in critical infrastructure sectors.* Washington, DC (National Threat Assessment Center), U.S. Secret Service and Carnegie-Mellon University, SEI/CERT Coordination Center.

[4] Greitzer, F. L., & D. A. Frincke. (2010). Combining traditional cyber security audit data with psychosocial data: towards predictive modeling for insider threat. In *Insider Threats in Cyber Security*. vol. 49, C. W. Probst, *et al.*, Eds., Springer US, 2010, 85–114.

[5] Greitzer, F. L., L. J. Kangas, C. F. Noonan, C. R. Brown, & T. Ferryman. (2013). Psychosocial modeling of insider threat risk based on behavioral and word use analysis. *e-Service Journal*, *9(1)*, 106-138.

[6] Greitzer, F. L, M. Imran, J. Purl, E. T. Axelrad, Y. M. Leong, D. E. Becker, K. B. Laskey, & P. J. Sticha. (2016). Developing an ontology for individual and organizational sociotechnical indicators of insider threat risk. *Eleventh International Conference on Semantic Technology for Intelligence, Defense, and Security (STIDS 2016)*, Fairfax, VA, November 15-16, 2016, 19-27.

[7] Greitzer, F. L., J. Purl, Y.M. Leong, & D.E. Becker. (2018). SOFIT: Sociotechnical and Organizational Factors for Insider Threat. *IEEE Symposium on Security & Privacy, Workshop on Research for Insider Threat (WRIT)*, San Francisco, CA, May 24, 2018.

[8] Shaw, E. D. & L. Sellers. (2015). Application of the critical-path method to evaluate insider risks. *Studies in Intelligence*, *59(2)*, 41-48.

[9] Greitzer, F. L., J. Strozer, S. Cohen, J. Bergey, J. Cowley, A. Moore, & D. Mundie. (2014). Unintentional insider threat: contributing factors,

observables, and mitigation strategies. *47th Hawaii International Conference on Systems Sciences (HICSS-47)*, Big Island, Hawaii.

[10] Greitzer, F. L., J. Strozer, S. Cohen, A. Moore, D. Mundie, & J. Cowley. (2014). Analysis of unintentional insider threats deriving from social engineering exploits. *IEEE Security and Privacy Workshop on Research for Insider Threat (WRIT)*, San Jose, CA, May 17-18, 2014.

[11] Schultz, E. E. (2002). A framework for understanding and predicting insider attacks. *Computers & Security*, *21*, 526–531.

[12] Greitzer, F. L., D. A. Frincke, & M. M. Zabriskie. (2011). Social/ethical issues in predictive insider threat monitoring. In: MJ Dark (Ed.), Information Assurance and Security Ethics in Complex Systems: Interdisciplinary Perspectives. Hershey, Pennsylvania: IGI Global. Chapter 7, pp.132-161.

[13] Costa, D. L., M. Collins, J. S. Perl, J. M. Albrethsen, J.G. Silowash, & D. Spooner. (2014). An Ontology for Insider Threat Indicators. In K. B. Laskey, I. Emmons and P C.G. Costa (Eds.), *Proceedings of the Ninth Conference on Semantic Technologies for Intelligence, Defense, and Security* (STIDS 2014), 2014, 48–53.

[14] Costa, D. L., M. J. Albrethsen, M. L. Collins, S. J. Perl, G. J. Silowash, & D. L. Spooner. (2016). *An Insider Threat Indicator Ontology*. Pittsburgh, PA: TECHNICAL REPORT CMU/SEI-2016-TR-007.

[15] Gelles, M. (2005). Exploring the mind of the spy. In Online *Employees' Guide to Security Responsibilities: Treason 101*. Retrieved from Texas A&M University Research Foundation website: http://www.dss.mil/search-dir/training/csg/security/Treason/Mind.htm

[16] Chaplin, W. E., O. P. John, & L. R. Goldberg. (1988). Conceptions of states and traits: Dimensional attributes with ideals as prototypes. *Journal of Personality and Social Psychology*, *54(4)*, 541-557.

[17] Steyer, R., A. Mayer, C. Geiser, & D. A. Cole. (2015). A theory of states and traits—Revised. *Ann. Review of Clinical Psychology*, *11*, 71-98.

[18] Roesch, S. C., A. A. Aldridge, S. N. Stocking, F. Villodas, Q. Leung, C. E. Bartley, & L. J. Black. (2010). Multilevel factor analysis and structural equation modeling of daily diary coping data: Modeling trait and state variation. *Multivariate Behavioral Research, 45(5)*, 767-789.

[19] Van Gelder, L., & R. E. De Vries. (2016). Traits and states at work: Lure, risk and personality as predictors of occupational crime. *Psychology, Crime & Law, 22(7)*, 701-720. DOI 10.1080/1068316X.2016. 1174863

[20] Grös, D. F., L. J. Simms, M. M. Antony, & R. E. McCabe. (2007). Psychometric properties of the State–Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State–Trait Anxiety Inventory (STAI). *Psych Assessment*, *19(4)*, 369–381.

[21] Douglas, K. S., S. D. Hart, C. D. Webster, & H. Belfrage. (2013). *HCR-20V3: Assessing risk of violence – User guide.* Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

[22] Meloy, J. R, S. G. White, & S. Hart. (2013). Workplace assessment of targeted violence risk: The development and reliability of the WAVR-21. *J. of Forensic Sciences, 58(5)*, 1353-1358.

[23] Shaw, E. D., & L. F. Fischer. (2005). *Ten Tales of Betrayal: The Threat to Corporate Infrastructures by Information Technology Insiders. Report 1—Overview and General Observations.* Technical Report 05-04, April 2005. Monterey, CA: Defense Personnel Security Research Center.

[24] Shaw, E. D., J. M. Post, & K. G. Ruby. (1999). Inside the mind of the insider. *Security Management*, *43 (12)*, 34-42.

[25] Band, S. R., D. M. Cappelli, L. F. Fischer, A. P. Moore, E. D. Shaw, & R. F. Trzeciak. (2006). *Comparing insider IT sabotage and espionage: a model-based analysis*. Carnegie-Mellon University, SEI/CERT Coordination Center. CMU/SEI-2006-TR-026.

[26] Parker, D. B. (1988). *Fighting computer crime: A new framework for protecting information*. New York, NY: John Wiley & Sons, Inc.

[27] Greitzer, F. L. & B. Zadeh. (2010). *Behavioral and Cultural Factors in the Workforce: A Framework for Insider Threat Modeling Research and Development (OUO).* Technical Report. Richland, WA: Pacific Northwest National Laboratory.

[28] Mylrea, M., S. N. G. Gourisetti, C. Larimer, & C. Noonan. (2018). Insider threat cybersecurity framework webtool & methodology: Defending against complex cyber-physical threats. *IEEE SPW Workshop on Research for Insider Threat (WRIT), San Francisco, CA.*

[29] The White House. (2012). *Presidential Memorandum—National Insider Threat Policy and Minimum Standards for Executive Branch Insider Threat Programs.* November 21, 2012. https://obamawhitehouse.archives.gov/the-press-office/2012/11/21/presidential-memorandum-national-insider-threat-policy-and-minimum-stand