The Critical Role of Positive Incentives in Reducing Insider Threat

Andrew P. Moore

Contributors: SEI CERT, SEI Software Solutions Division,
SEI Human Resources,
SEI Organizational Effectiveness Group,
CMU Heinz College/Tepper School of Business

Copyright

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM-0004035

Research Objective

Determine influence of workforce management practices on insider threat behaviors

Negative Incentives

Workforce management practices that attempt to *force* employees to act in the interests of the organization

Employee Constraints, Monitoring, Punishment

Positive Incentives

Workforce management practices that attempt to *attract* employees to act in the interests of the organization

Focus on Employee Strengths, Fair & Respectful Treatment

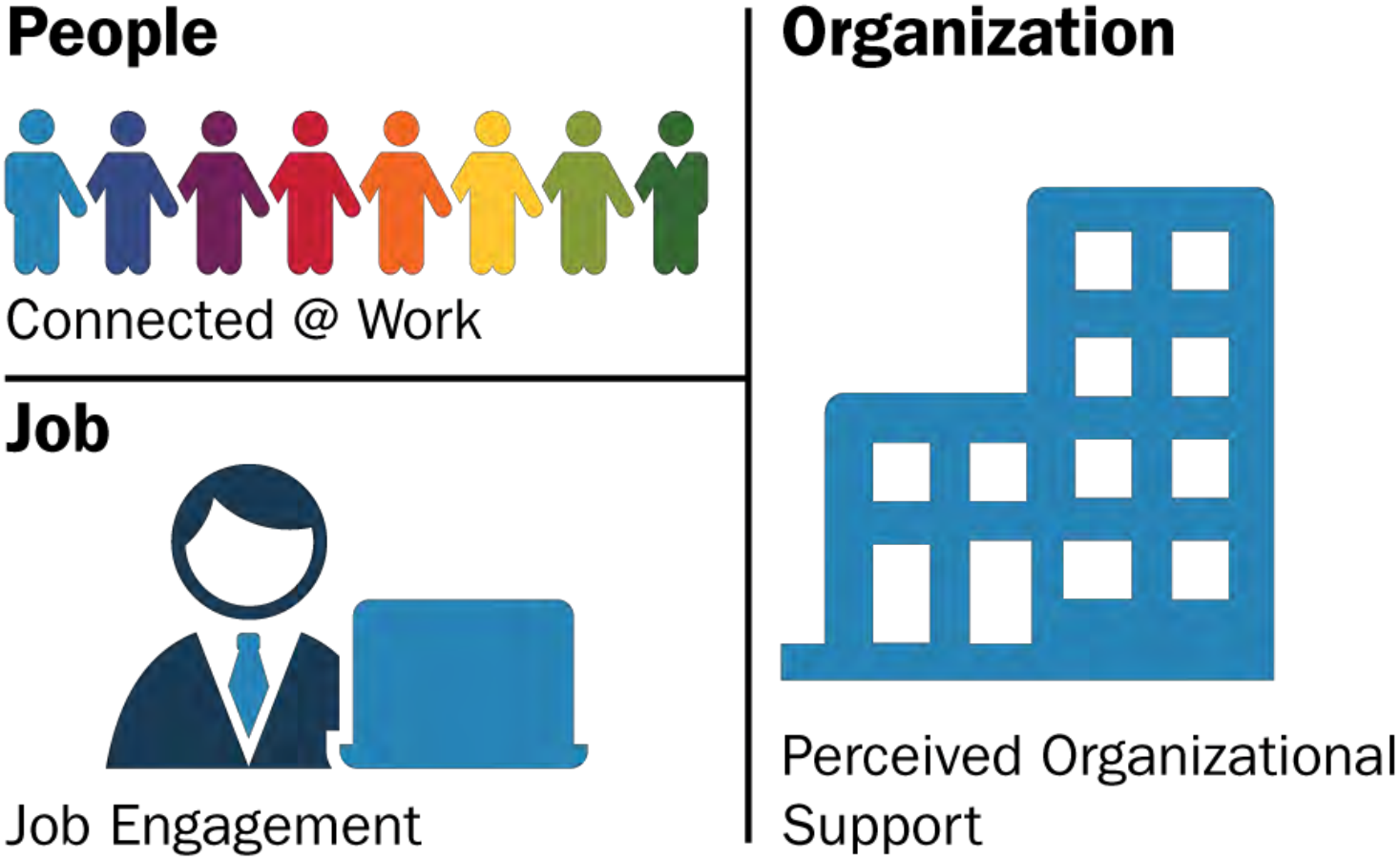
Negative incentives *alone* can *exacerbate* the threat they are intended to mitigate*

Basic Belief: Organizations need to *explicitly* consider a *mix of positive and negative incentives* to build insider threat programs that are a net positive for employees

Initial Scope: Demonstrate value of research in area for insider threat reduction

* See "Effective Insider Threat Programs: Understanding and Avoiding Potential Pitfalls," SEI Digital Library, March 2015.

Three Dimensions of Employee-Organization Alignment



Two-Pronged Exploratory Research Approach

1. *Insider Incident Case Study Analysis*
   - How engaged, connected, and supported are insider threat actors?
2. *Organizational Survey*
   - How much does organizational support influence insider cyber misbehavior?

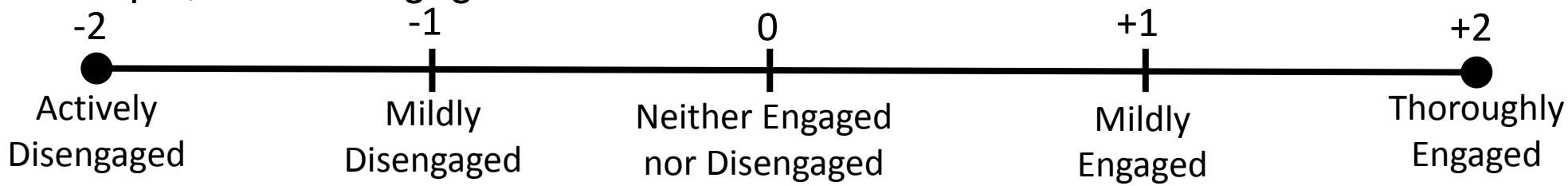
Extension of previous work by focusing on

- Egregious insider threat behaviors
- Organizations actively establishing insider threat programs

Insider Incident Case Study Analysis

How engaged, connected, and supported are insider threat actors?

- **Method:** Rate dimensions on 5-point Likert scales over three time periods
  - For example, for Job Engagement



- **Challenge:** Assessing insider perceptions through observables (w/o interview)
- **Results:** (3 prominent incidents)
  - Dimensions became increasingly negative over time, with some fluctuation
    - *Organizational Support* most strongly negative in all 3 incidents
    - *Job Engagement* negative in 2 out of 3 incidents
    - *Connectedness at Work* negative in 1 out of 3 incidents
- **Initial Decision:** Focus on perceived organizational support as foundation.

Organizational Survey

How much does organizational support influence insider cyber misbehavior?

Challenge: Hard-to-reach population suggests initial exploratory

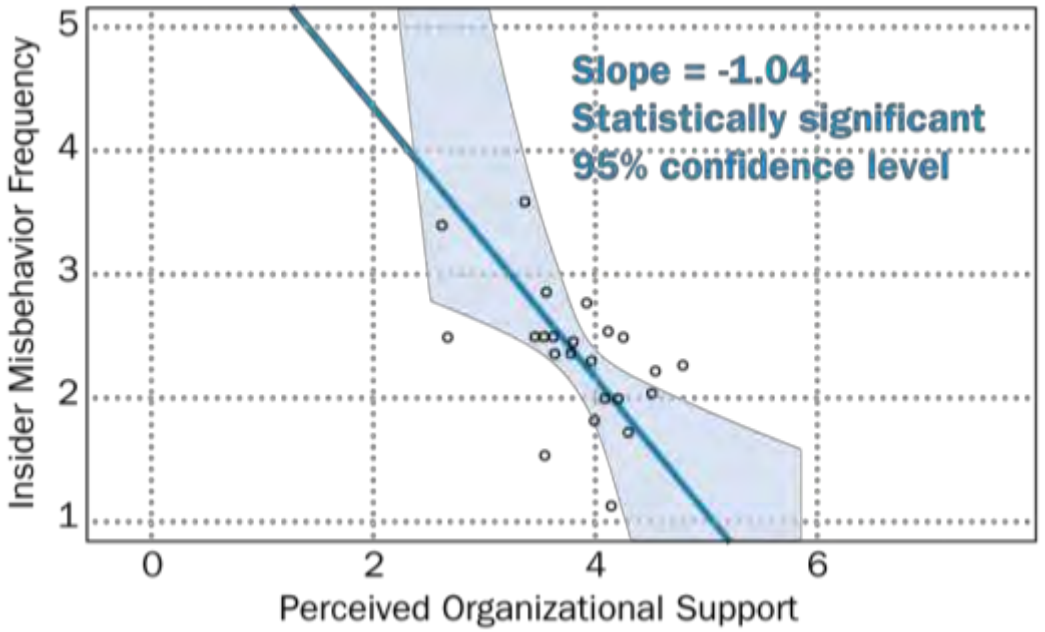
Method: Survey insider threat program managers in an Insider Threat Information Sharing Group

- Independent variable on established 5-point scales
  - *Perceived organizational support* (36 questions)
- Dependent variable on 5-point frequency scale
  - *Cyber misbehavior* from case data (22 questions)

Response:

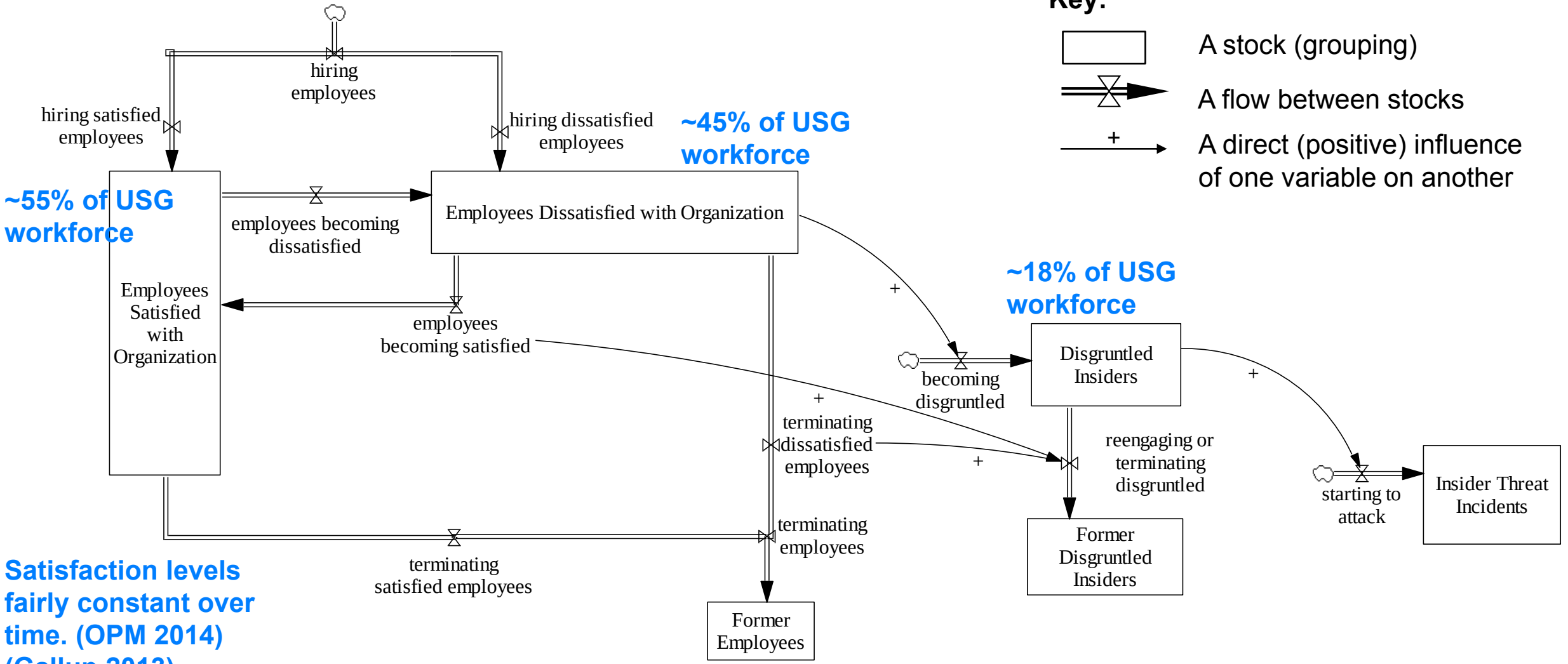
- 25 out of ~90 organizations responded

Results: (23 responses used)*



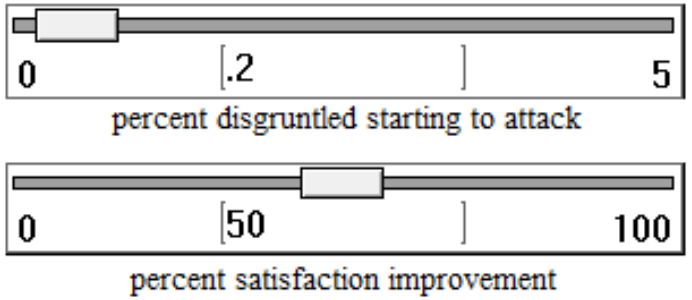
* Analysis used Deming Regression and Multiple Imputation by Chained Equations for missing values.

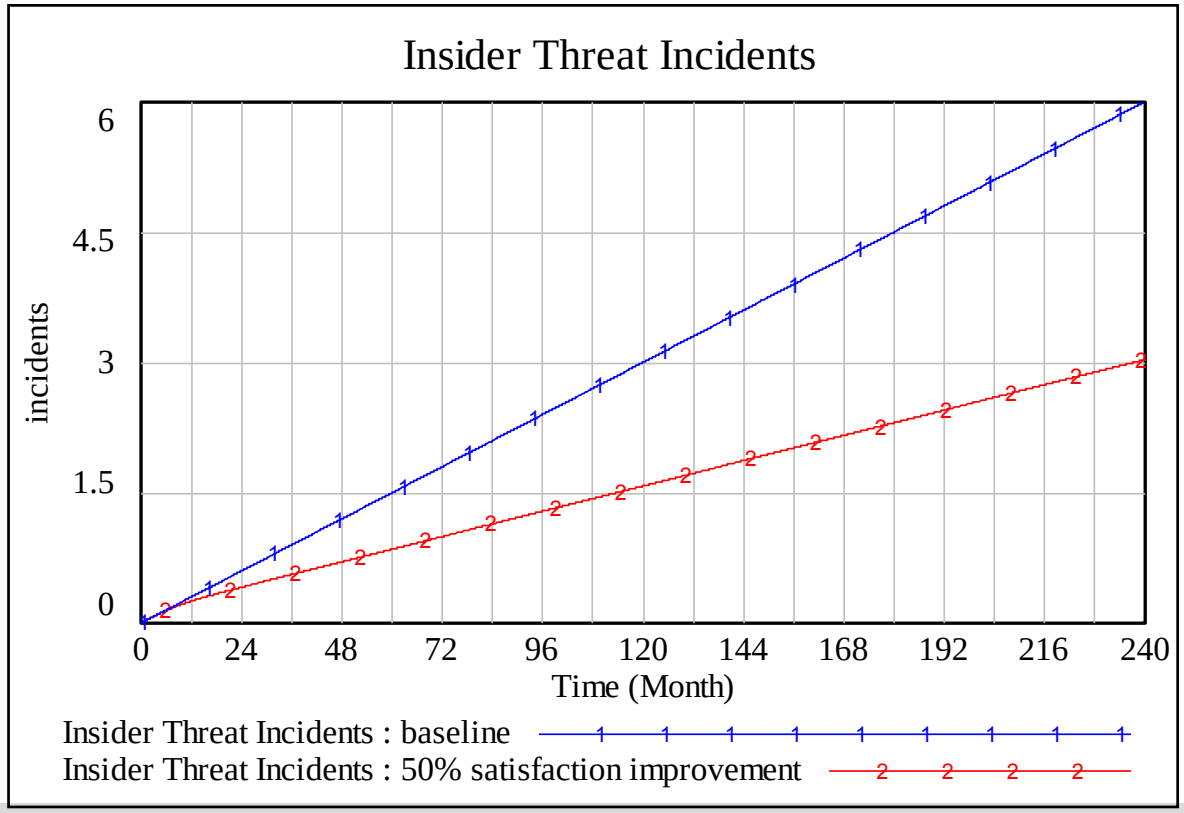
Emerging Physics of Job Satisfaction, Disgruntled Insider Threat



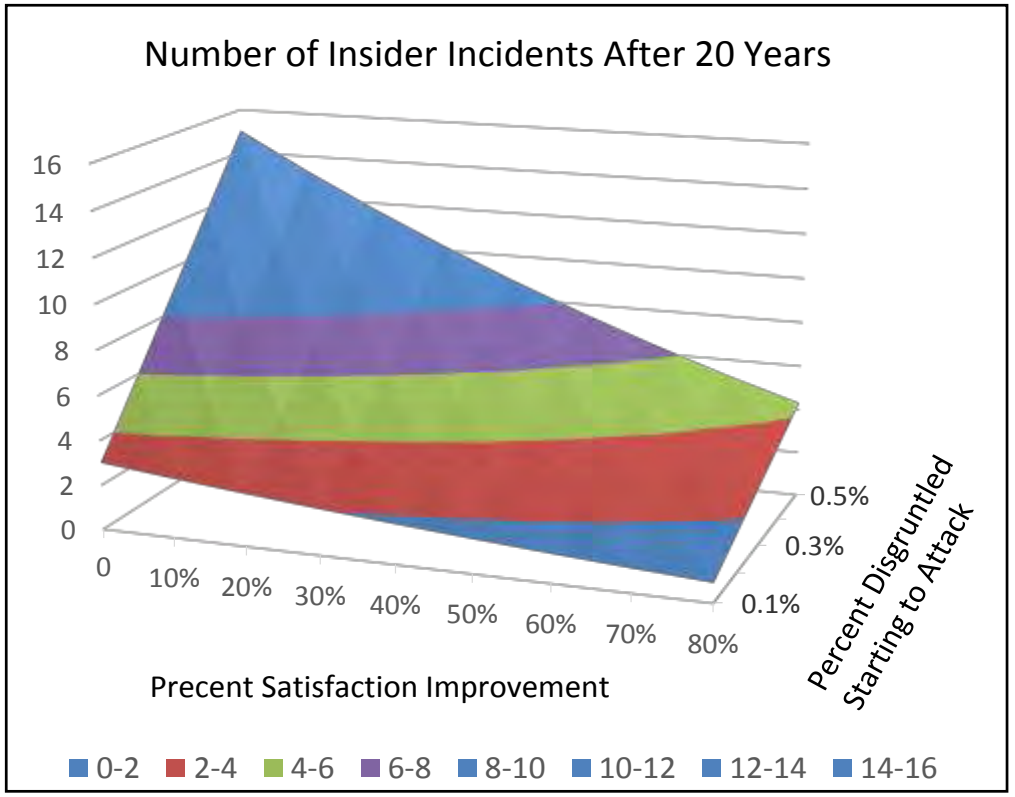
Preliminary Model Simulation Findings

Simulation Controls:





Sensitivity simulation over the two inputs:



Other model uses: Calculate the cost savings from fewer incidents and less counterproductivity

Future Research

Theory Development

- Experiment-based determination of cause-effect relationship between perceived organizational support and insider threat

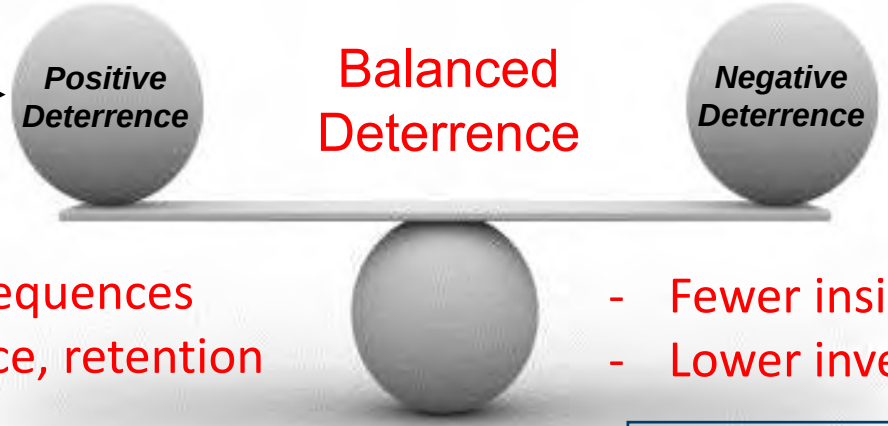
Technology Development

- Detection of insider alienation by identifying at-risk behaviors and indicative changes in insiders' network of workplace relationships

Adoption

- Determine how particular organization can
  - determine an appropriate mix of positive and negative incentives
  - transition to that from their current state

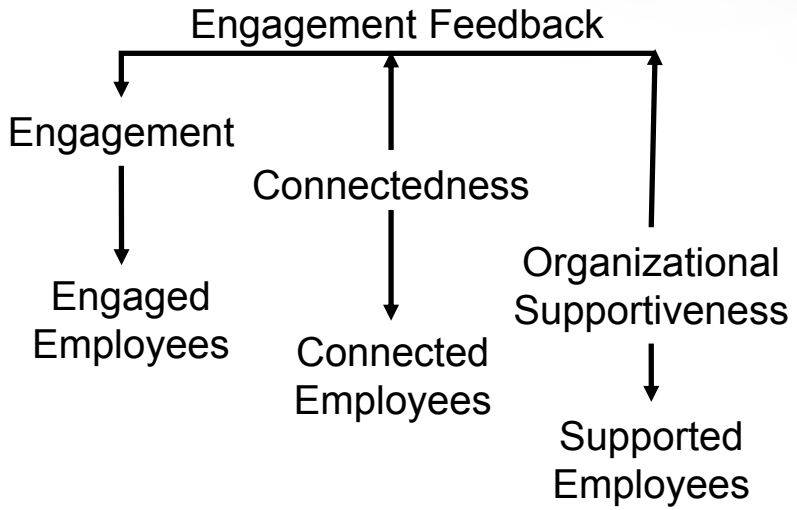
Vision: Extending the Traditional Security Paradigm



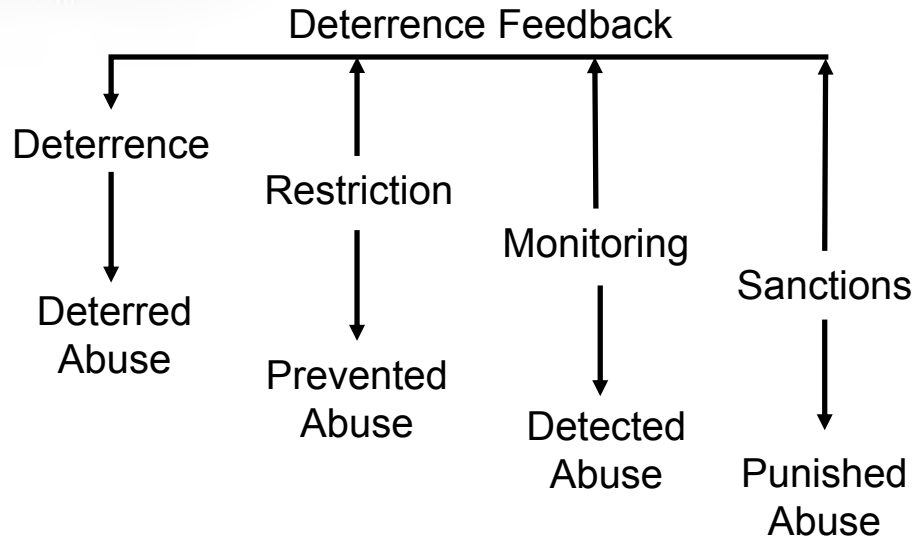
- Fewer unintended consequences
- Satisfaction, performance, retention

- Fewer insider incidents and misbehaviors
- Lower investigative costs, productivity loss

Security through Positive Incentives



Traditional Security Approach (Negative Incentives)



Contact Information

Presenter / Point of Contact :

Andrew Moore

Lead Insider Threat Researcher

Telephone: +1 412.268.5465

Email: apm@cert.org

Contributors :

SEI CERT:

Samuel J. Perl
Jennifer Cowley
Matthew L. Collins
Tracy M. Cassidy
Nathan VanHoudnos

SEI SSD:

William Novak
David Zubrow

Contributors :

SEI Directors Office:

Palma Buttles

SEI Human Resources:

Daniel Bauer
Allison Parshall
Jeff Savinda

SEI Organizational Effectiveness Group:

Elizabeth A. Monaco
Jamie L. Moyes

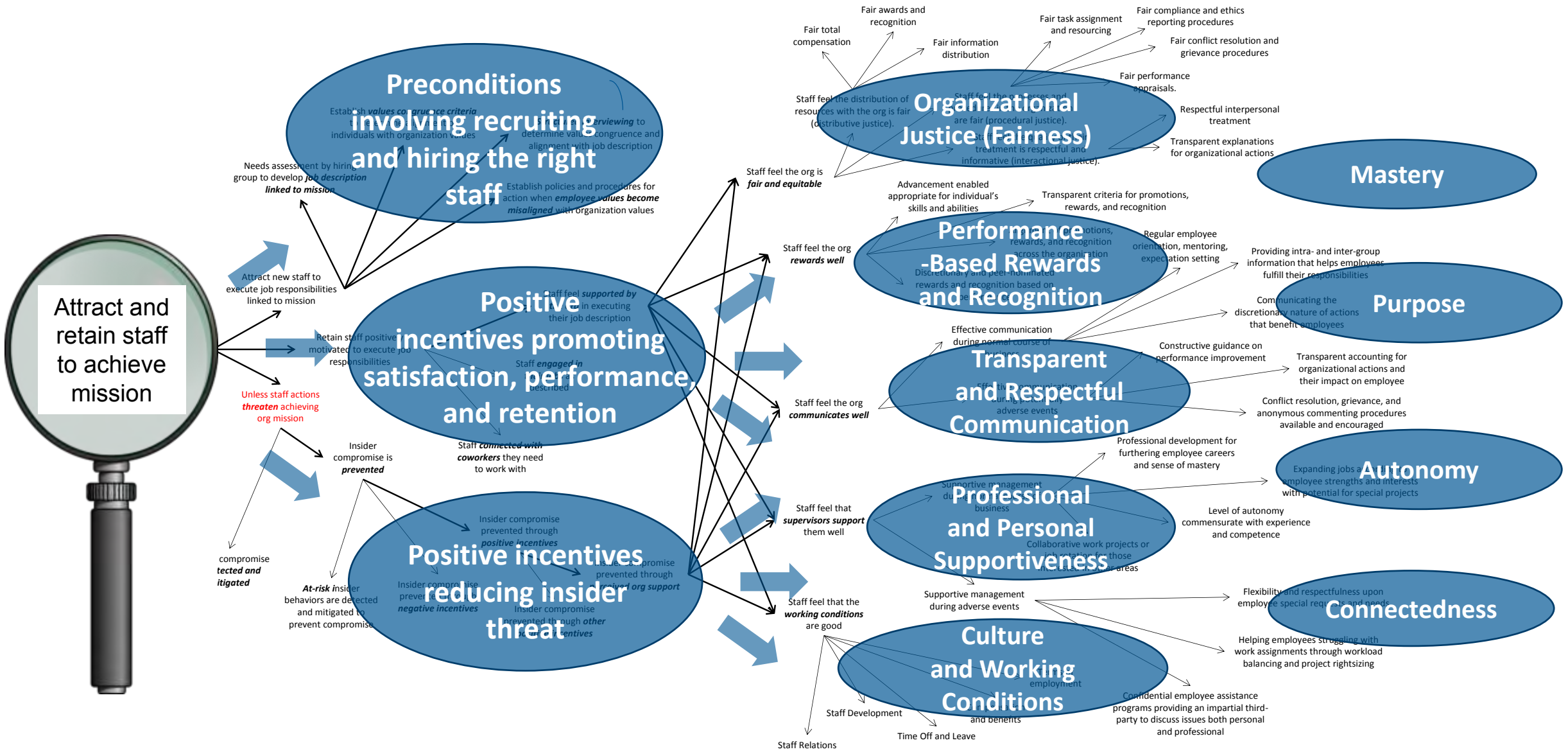
CMU Heinz College and Tepper School of Business:

Professor Denise M. Rousseau

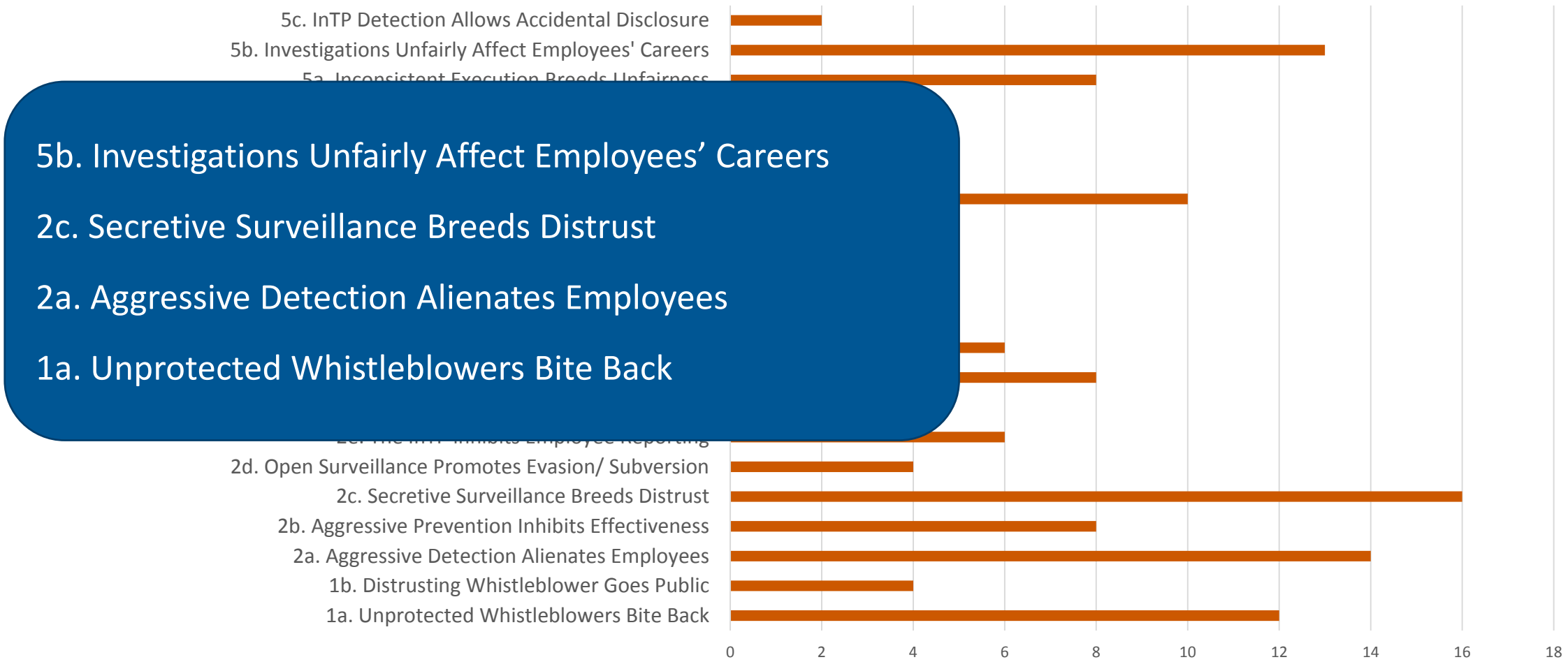
Special thanks to the Open Source Insider Threat (OSIT) Information Sharing Group for their responses to our survey.

Backups

Positive Incentive-Based Principles and Practice Areas



Potential Unintended Consequences of Traditional Insider Threat Management Practices*

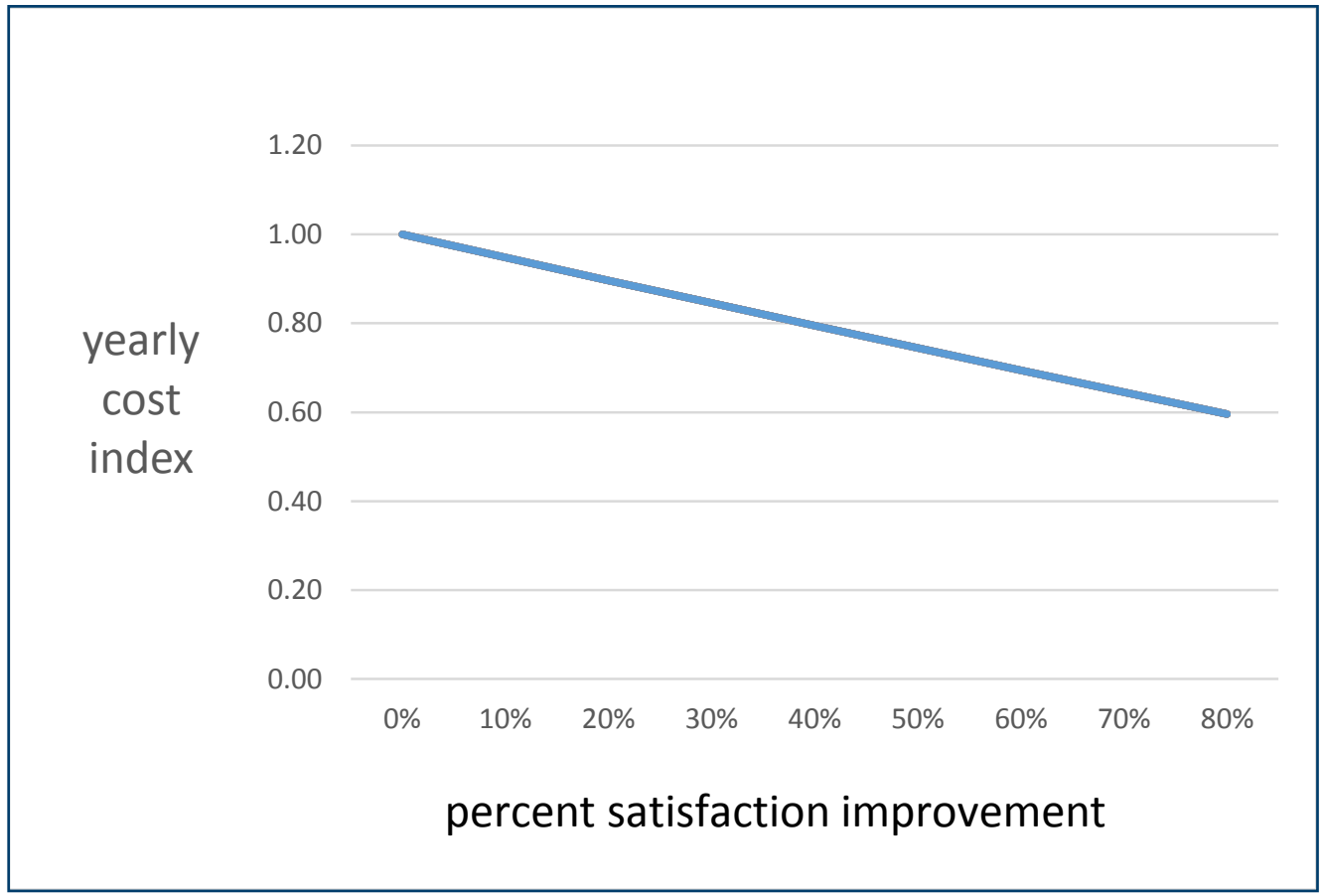


* See "Effective Insider Threat Programs: Understanding and Avoiding Potential Pitfalls," SEI Digital Library, March 2015.

Cost Benefits Due to Fewer Counterproductive Behaviors (CWB) and Insider Threat Incidents

Assumptions:

- Satisfied employees engage in one CWB every 2 months
- Dissatisfied employees engage in two CWBs every month (4 times the rate of satisfied)
- Average cost of a CWB is $500
- Average cost of an insider incident is $1M
  - Includes data loss, IP loss, investigation costs, etc.
- Cost index calculated by dividing the costs associated with CWBs and incidents by the cost in the baseline run



Workplace Violence/IT Sabotage: Two Sides of the Same Coin?

Presenter: Michael C. Theis

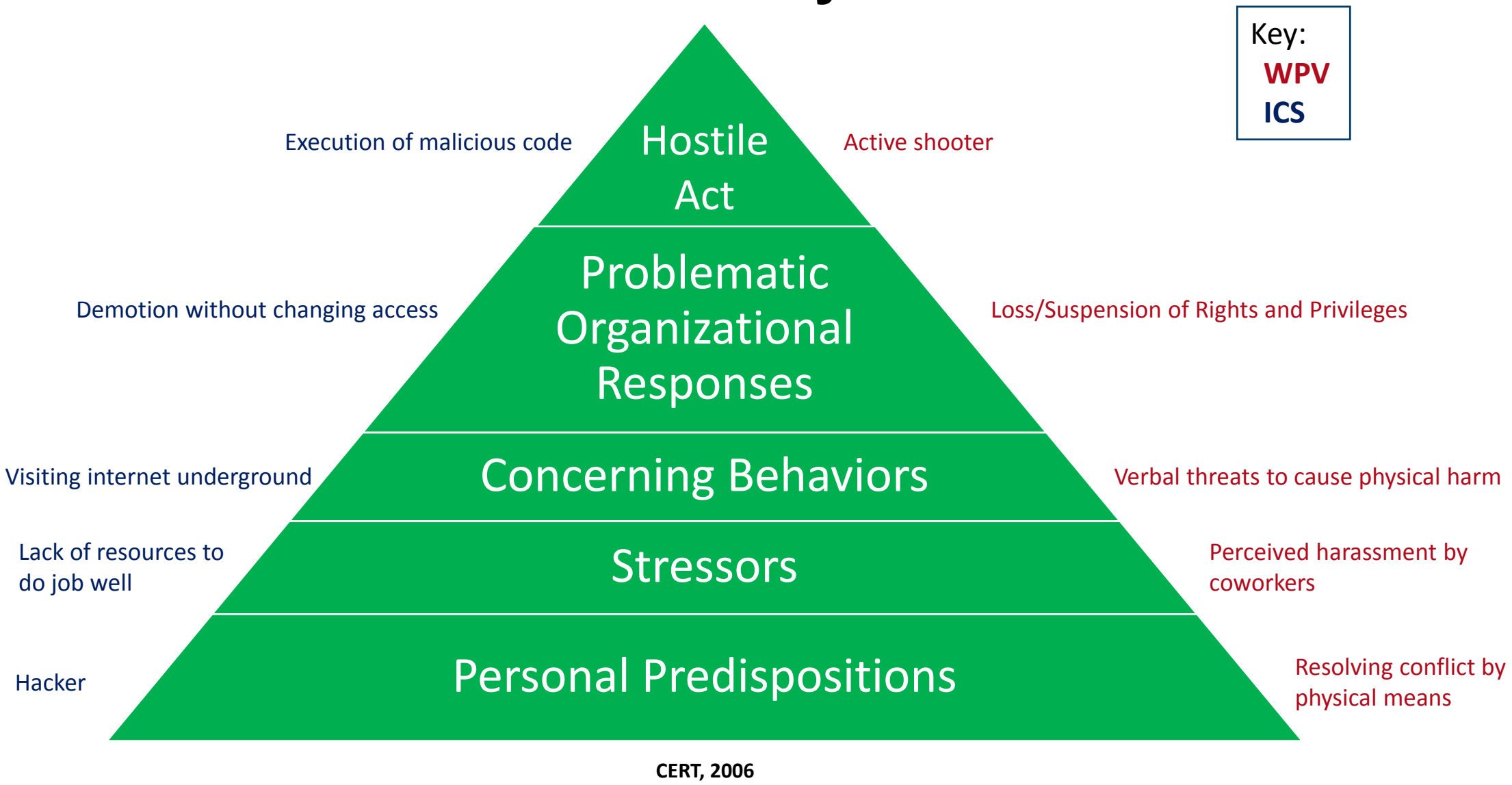
Research Objective and Approach

Objective: Determine if coherent, integrated, and validated indicators for Insider Workplace Violence (WPV) and Insider Cyber Sabotage (ICS) can be identified.

Reason: If there are common indicators organizations may be able to develop socio-technical controls that prevent, detect, and help respond to both threats without identifying which crime will eventually be committed.

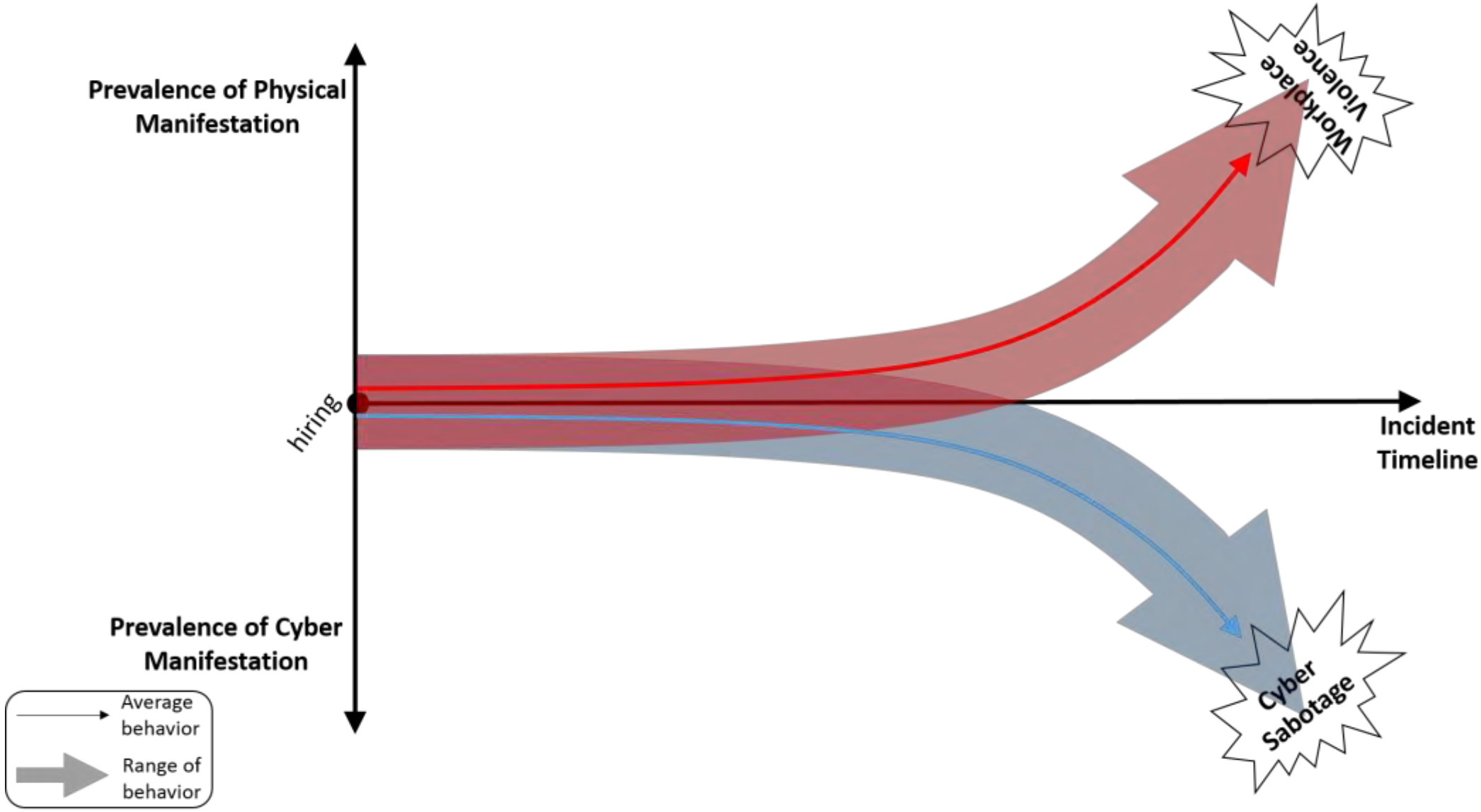
Approach: Collect, code, and analyze cases of WPV and compare them to cases of ICS in the CERT Insider Threat Center's corpus.

WPV and ICS Incident Pathway

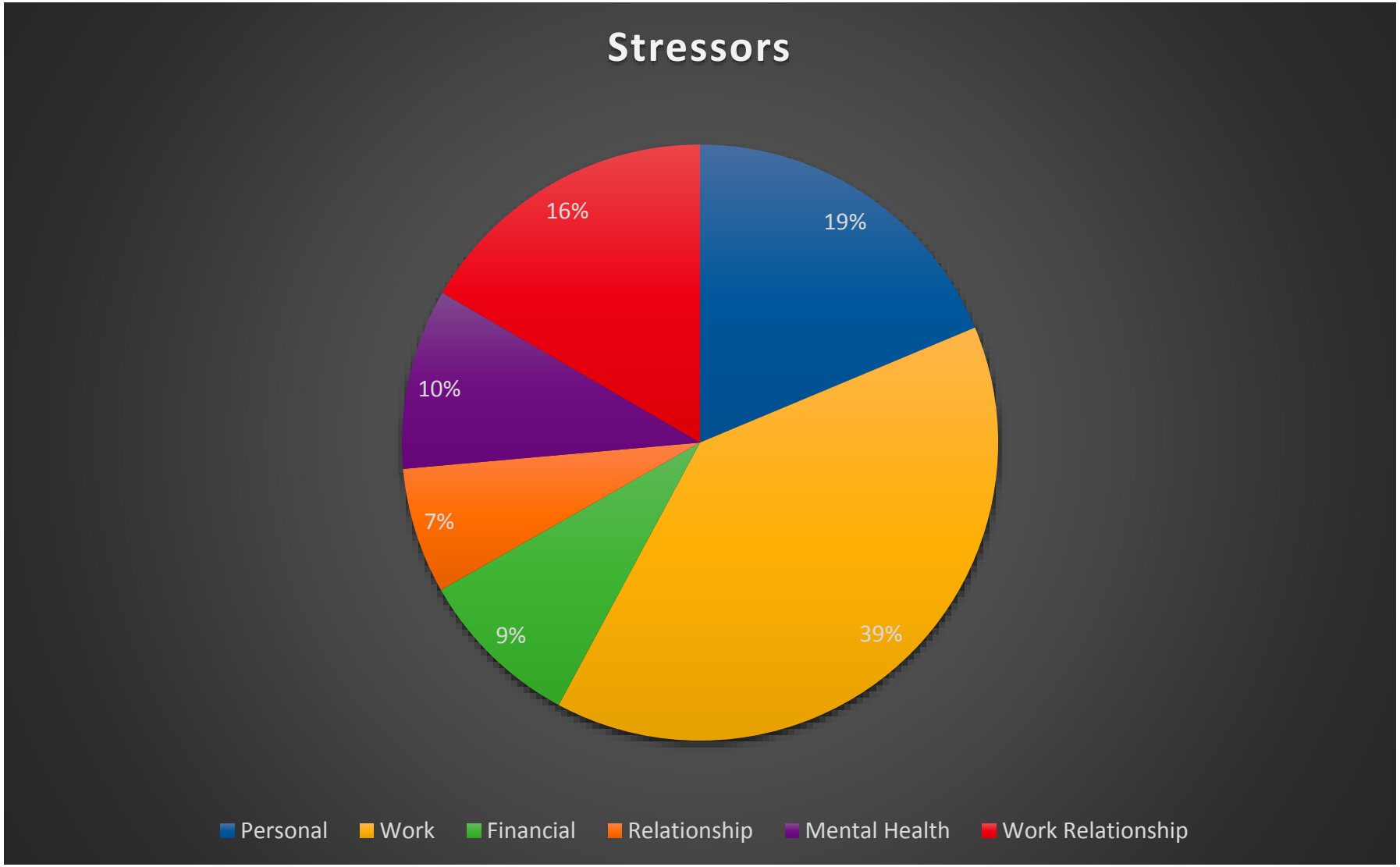


CERT, 2006

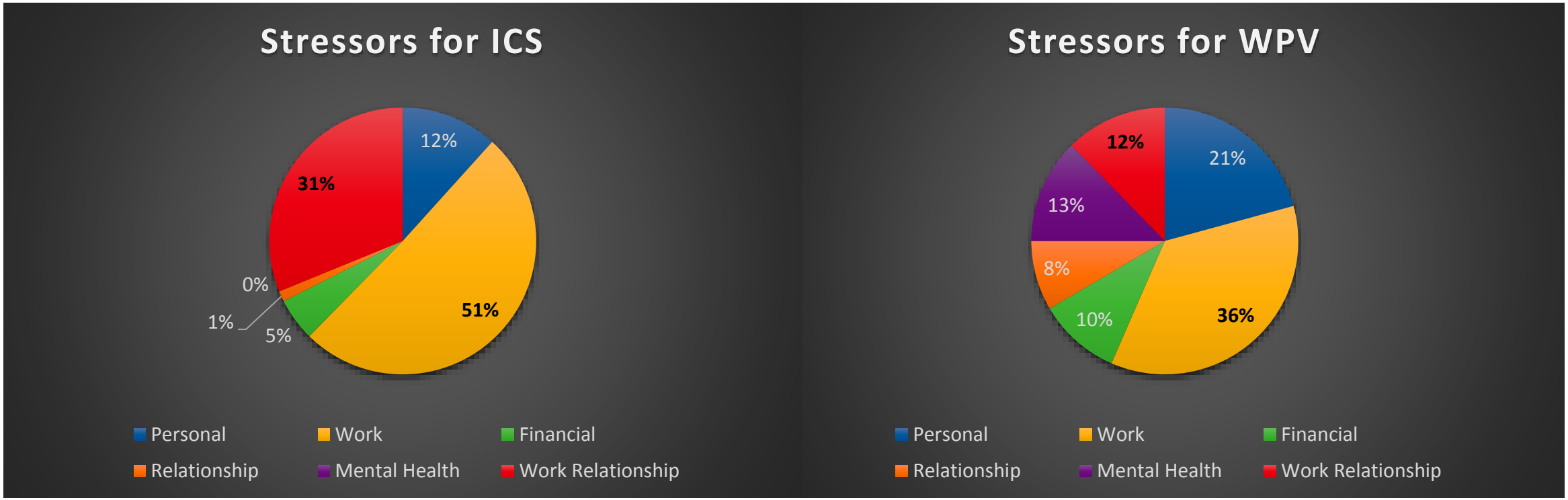
Hypothesis: Common Path Before Divergence



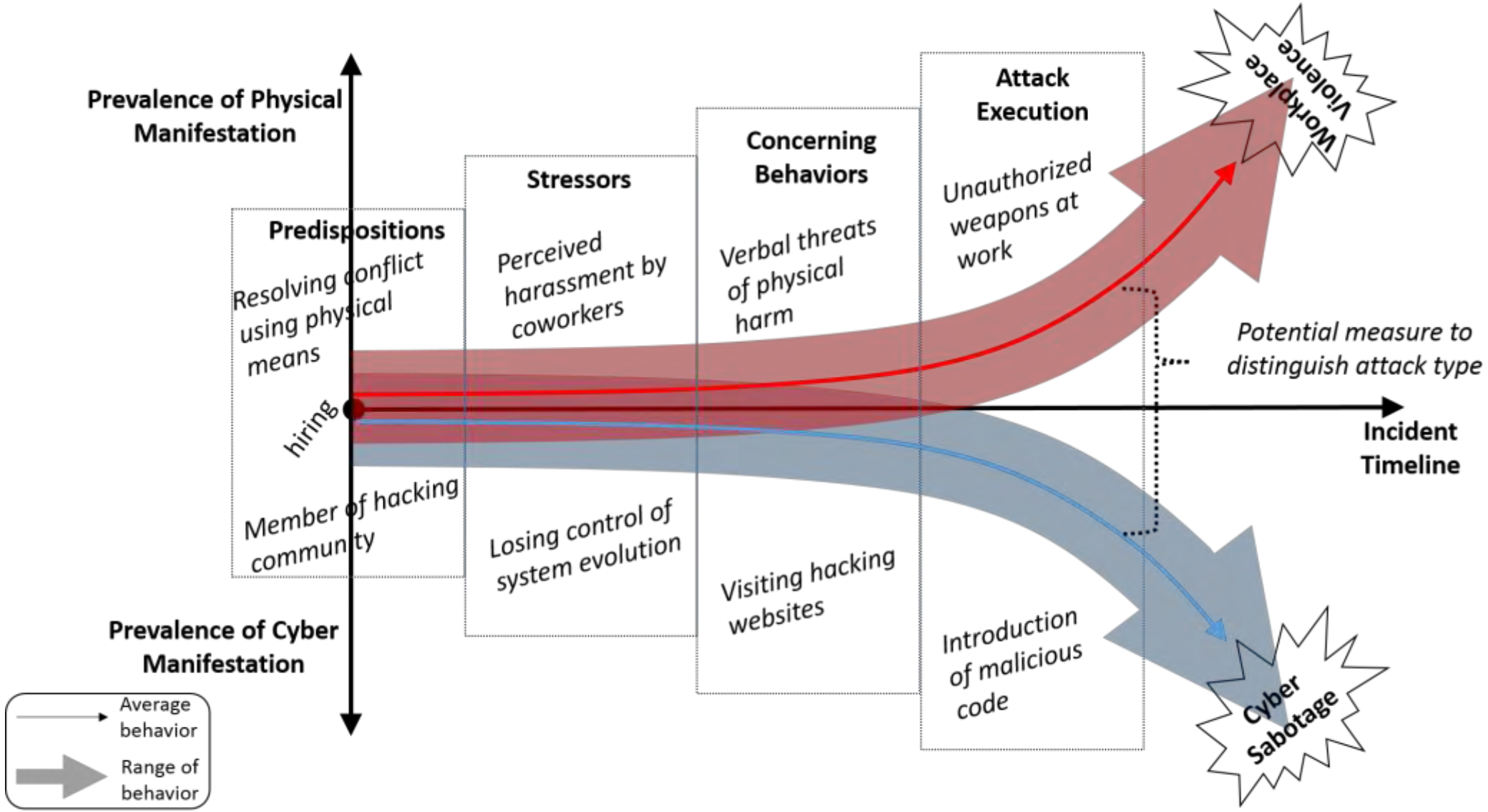
Aggregation of Stressors



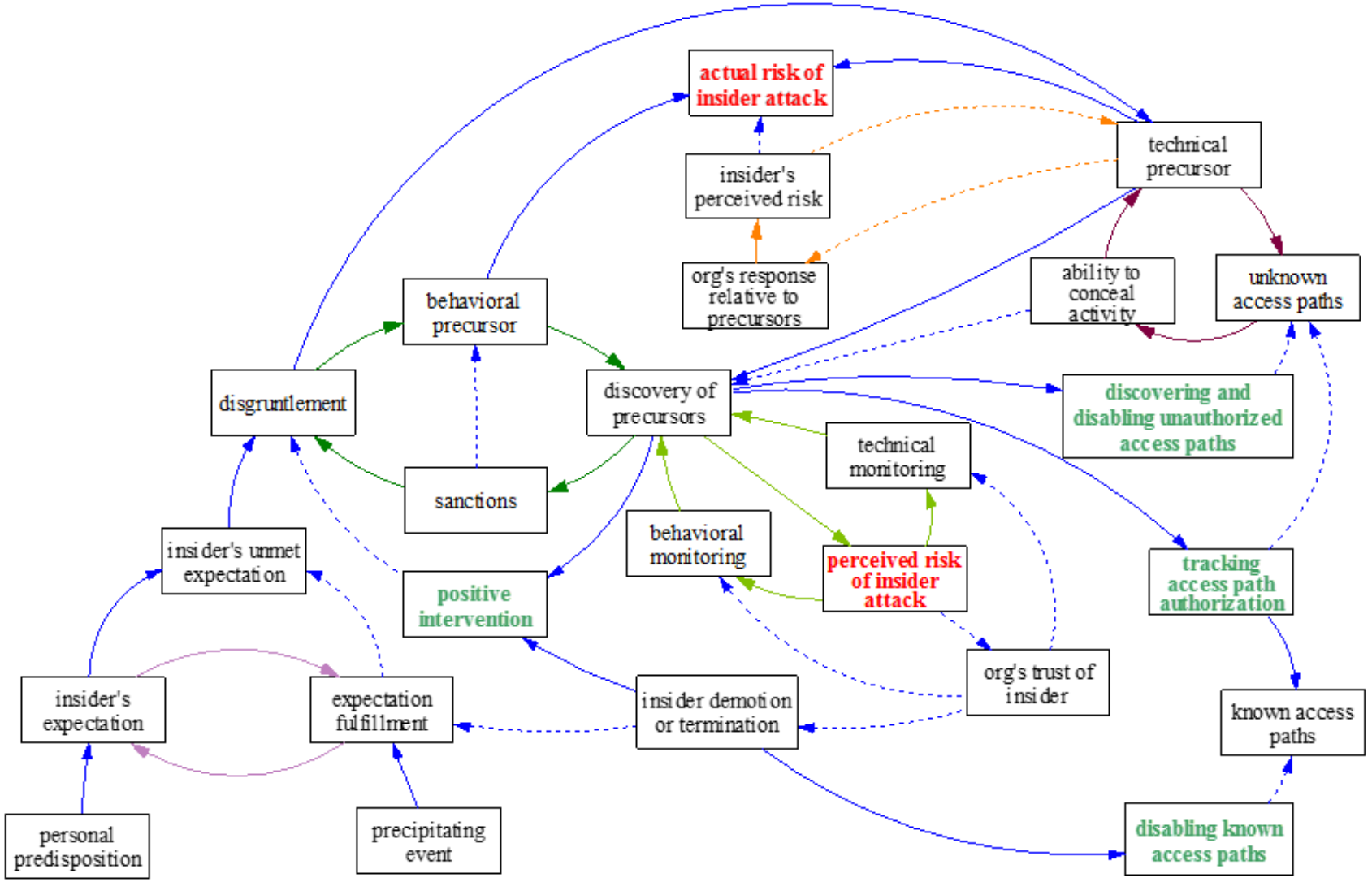
Stressors by ICS and WPV



Distinguishing the WPV and ICS Pathways

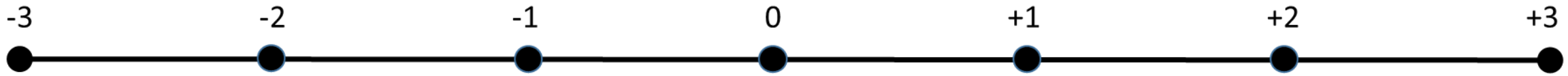


ICS Causal Loop Diagram



Backups

A Cyber-Physical Scale for Assessing Observables*



-3	-2	-1	0	+1	+2	+3
Direct, Active Cyber Aggression	Indirect, Active Cyber Aggression	Passive (Direct or Indirect) Cyber Aggression	Non-aggressive behaviors whether physical or cyber	Passive (Direct or Indirect) Physical Aggression	Indirect, Active Physical Aggression	Direct, Active Physical Aggression
Examples: Execution of malicious code to delete critical files	Disabling the code that creates the backup tapes	Not submitting online reports in a timely manner	Other observables that do not reach the level of aggression	Not coming to meetings scheduled	Spreading false rumors	Verbal assault, bullying, shooting coworkers

* Note: combined cyber-physical observables may be broken down into their constituent components for measurement. See the Reality-Virtuality Continuum for a loosely related construct applied to virtual reality technologies. https://en.wikipedia.org/wiki/Reality%E2%80%93virtuality_continuum

Operational Definitions (from Buss and Parrot)

Aggression – intentional behaviors that can cause significant harm to a victim (person or organization) who wishes to *avoid* the act. (note: definition excludes desired harm (sadomasochism, going to dentist) and unintentional harm (stepping on foot))

- **Direct Aggression –** person-to-person interactions (but not necessarily face-to-face) in which the perpetrator is easily identifiable by the victim (e.g., Active: Shooting, email a threat; Passive: intentionally not write a letter of recommendation and harming victim's application for new job).
- **Indirect Aggression–** circuitous interactions in which the perpetrator may remain unidentified, possibly to avoid accusation, direct confrontation, and/or counterattack by the victim (e.g., Active: (anonymously) spreading false rumors; Passive (rare): (anonymously) not coming to the defense of someone being criticized).
- **Active Aggression–** an act of commission by the perpetrator, which involves active engagement in harming the victim (e.g., Direct: shooting; Indirect: (anonymously) spreading harmful rumors)
- **Passive Aggression –** an act of omission by the perpetrator, which involves a lack of active responding that causes harm to the victim (e.g., Direct: intentionally not write a letter of recommendation and harming victim's application for new job; Indirect (rare): (anonymously) not coming to the defense of someone being criticized)

Physical - intentional acts involving personal or interpersonal interaction that does not involve cyber

Cyber - intentional acts involving interaction with computers, computer networks, or electronic media

Hasan, Fort Hood – 2009: Concerning Behaviors

Major Period	Sub-Period	Direct-Active Cyber Aggression (-3)	Indirect Active Cyber Aggression (-2)	Passive Cyber (Indirect or Direct) (-1)	Center of Scale (0)	Passive Physical (Indirect or Direct) (+1)	Indirect Active Physical Aggression (+2)	Direct Active Physical Aggression (+3)	Sub-Period Concerning Behaviors (non-zero)	Major Period Concerning Behaviors (non-zero)
'92-97		0	0	0		1	2	0		3
'98-03		0	0	0		0	1	0		1
'04-09		2	3	0		1	5	3		14
Sub-Periods of Last Major Period	'04-05	0	0	0		0	2	0	2	
	'06-07	0	0	0		0	2	0	2	
	'08-09	2	3	0		1	1	3	10	
Major Period Totals		2	3	0		2	8	3		18

Alexis, WNY – 2013: Concerning Behaviors

Major Period	Sub-Period	Direct-Active Cyber Aggression (-3)	Indirect Active Cyber Aggression (-2)	Passive Cyber (Indirect or Direct) (-1)	Center of Scale (0)	Passive Physical (Indirect or Direct) (+1)	Indirect Active Physical Aggression (+2)	Direct Active Physical Aggression (+3)	Sub-Period Concerning Behaviors (non-zero)	Major Period Concerning Behaviors (non-zero)
3/04-3/07		0	0	0		1	0	2		3
4/07-12/10		0	0	2		1	0	1		4
’1/11-9/13		0	0	0		1	3	0		4
Sub-Periods of Last Major Period	2011	0	0	0		0	0	0		
	2012	0	0	0		0	0	0		
	2013	0	0	0		1	3	0	4	
Major Period Totals		0	0	2		3	3	3		11

7-Point Scale Analysis of Results