Red Team Performance: Summary of Findings

University of Wisconsin-Madison & IDART: Sandia National Laboratories

June 2004

Pascale Carayon and Sara Kraemer
Center for Quality and Productivity Improvement
University of Wisconsin-Madison
610 Walnut Street 575 WARF
Madison, WI 53726

The US Department of Defense is providing funding for this project (PI: Professor S. Robinson; Grant number: DAAD19-01-1-0502; ARO proposal number: 42347-MA-CIP).

For more information:
Dr. Pascale Carayon, Director of the Center for Quality and Productivity Improvement
Tel: +1-608-265-0503
Fax: +1-608-263-1425
carayon@engr.wisc.edu

Research Team

Pascale Carayon & Sara Kraemer
Center for Quality and Productivity Improvement
University of Wisconsin-Madison
610 Walnut Street 575 WARF
Madison, WI 53726
Tel: 1-608-263-2520
Fax: 1-608-263-1425
Email: carayon@engr.wisc.edu/sbkraeme@wisc.edu

Ruth Duggan & John Clem
IDART
Sandia National Laboratories
PO Box, 5800, MS-1375
Albuquerque, NM 87185-1375
Tel: 1-505-844-9320
Fax: 1-505-845-7065
Email:rduggan@sandia.gov/bjbartb@sandia.gov

Summary

Professor Pascale Carayon and her graduate student, Sara Kraemer, have collaborated with Sandia National Laboratories Information Design Assurance Red Team (IDART) program to study red team performance. Ruth Duggan, former red team leader and program coordinator and John Clem, current red team leader and program coordinator at Sandia, are the key collaborators of this study. Funding was provided by the Department of Defense for the project on "Modeling and Simulation for Critical Infrastructure Protection" (#DAAD19-01-1-0502, PI: Professor Robinson, UW-Madison) and by the College of Engineering at the University of Wisconsin-Madison.

Red teaming is an advanced form of assessment that can be used to identify weaknesses in the security of a variety of systems. The red team approach is based on the premise that an analyst who attempts to model an adversary can find systemic vulnerabilities in a computer and information system that would otherwise go undetected. Our study of red team performance encompasses three aims:
(1) identify measures of red team performance
(2) identify factors that contribute to and hinder red team performance
(3) conduct a trade-off analysis comparing red team techniques to automated security methods.

In this summary, findings will be reported in four parts: (1) measures of red team performance; (2) factors contributing to red team performance; (3) factors hindering red team performance; and (4) tradeoffs between red teaming and automated security methods.

For performance measurement, measures were defined and grouped into four categories of different perspectives: (1) individual team members (12 comments); (2) the team as a whole (27 comments); (3) management (12 comments); and (4) customer (30 comments). These perspective categories were further stratified into three dimensions of team measurement: descriptive, evaluative, and diagnostic. The fifth category consisted of comments regarding difficulties in measuring red team effectiveness (14 comments).

Findings on factors that contribute to red team performance will be reported in five parts: organizational context (16 comments), team design (53 comments), team synergy (18 comments), material resources (2), and process criteria of effectiveness (73 comments). Findings on factors that hinder red team performance (28 total comments) will also be reported in five parts: organizational context (6 comments), team design (4 comments), team synergy (3 comments), material resources (1 comment), and process criteria of effectiveness (14 comments).

For tradeoffs comparing red team performance and modeling/simulation methods, findings will be reported in three main areas: (1) weaknesses of simulation methods (16 comments); (2) strengths of simulation methods (6 comments); and (3) strengths of red teaming (12 comments).

Table of Contents

1. Methods

From December 1 to December 5, 2003, Sara Kraemer visited the Sandia Laboratories in Albuquerque, New Mexico. During her stay, she was immersed in the IDART program, interviewing red team members, conducting focus groups with red team members, attending training and technical presentations, and observing the daily activities of Sandia's red teams. Specifically, the study consisted of the following elements: fifteen semi-structured individual interviews and two focus groups with red team members, observation of a red team group training session, attendance of a Sandia technical presentation, personal observation of site surroundings, and analyses of documents pertaining to red team work. The data was collected by Sara Kraemer. Individual interviews and focus groups used the same open-ended interview guide. See Appendix 1 for interview guide. The individual interviews were approximately one hour in length and the focus groups were approximately two and one half hours in length. One focus group and eleven interviews were audio-recorded and one focus group and four individual interviews were not audio-recorded. Personal notes of these interactions were taken by Sara Kraemer and the audio-recordings were transcribed.

The IDART program consists of core, non-core, and matrix red team members. Core red team members are system analysts who regularly participate in red team projects and whose full-time job is within the IDART assessment department. Non-core members are system analysts who semi-regularly participate in red team projects and are not members of the IDART assessment department. Matrix members rarely participate in red team projects. They are accessed for their specific expertise, which is needed for specific systems under consideration. For example, a red team examining a biological and chemical agent detection system could include experts on biological and chemical warfare agents. These members are accessed from the pool of experts within the Sandia organization. Individual interviews were done with eleven core members, three non-core members, and two matrix members. The first focus group included six core members and the second focus group consisted of seven core members.

The transcribed notes and interviews were analyzed by coding the themes of interviews and observations using the qualitative software package, QSR NVivo©. The coding structure consisted of nodes, representing a defined category or dimension of red team performance. When coded, a node held references to passages of text from the observation and interview data.

2. Measures of red team performance

Findings are reported in measurement category and sub-category totals. The coding process resulted in 67 nodes and the total number of comments coded was 95. The nodes were grouped into five major categories. The first four categories were defined four perspectives: (1) individual team members (12 comments); (2) the team as a whole (27 comments); (3) management (12 comments); and (4) customer (30 comments). These categories were further stratified into three dimensions of team measurement: descriptive, evaluative, and diagnostic. The fifth category consisted of comments regarding difficulties in measuring red team effectiveness (14 comments).

2.1 Overall totals for categories of perspectives

The primary types of team performance measures are defined in the following way : (1) descriptive measures (i.e. process), which describe what is happening at any given time and seek to document individual and team behaviors; (2) evaluative measures (i.e. outcome), which judge performance against identifiable standards and serve to answer questions of effectiveness; and (3) diagnostic measures (i.e. process), which seek to identify causes of behavior and question how and why things occurred as they did (Paris, Salas, & Cannon-Bowers, 2000). Table 1 summarizes the number of comments in each perspective category.

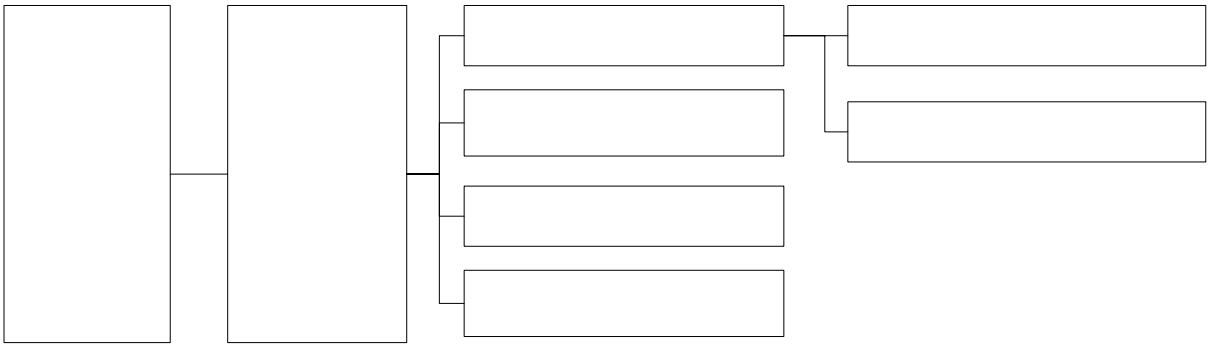
Table 1. Comments on red team performance

Perspective / Measurement Type	Individual	Team	Management	Customer	Total
Descriptive (Process)	0	0	7	22	**29**
Evaluative (Outcome)	0	14	5	8	**27**
Diagnostic (Process)	12	13	0	0	**25**
Total	**12**	**27**	**12**	**30**	**81**

2.1.1 Individual perspective category: Diagnostic dimensions

The individual perspective consists of responses regarding performance measures on the individual red team member. Red team members commented only on the diagnostic dimension of individual measurement (12 of 12 total comments). Figure 1 summarizes the number of responses for the individual perspective category.

Figure 1. Comments on individual team member perspective on red team performance



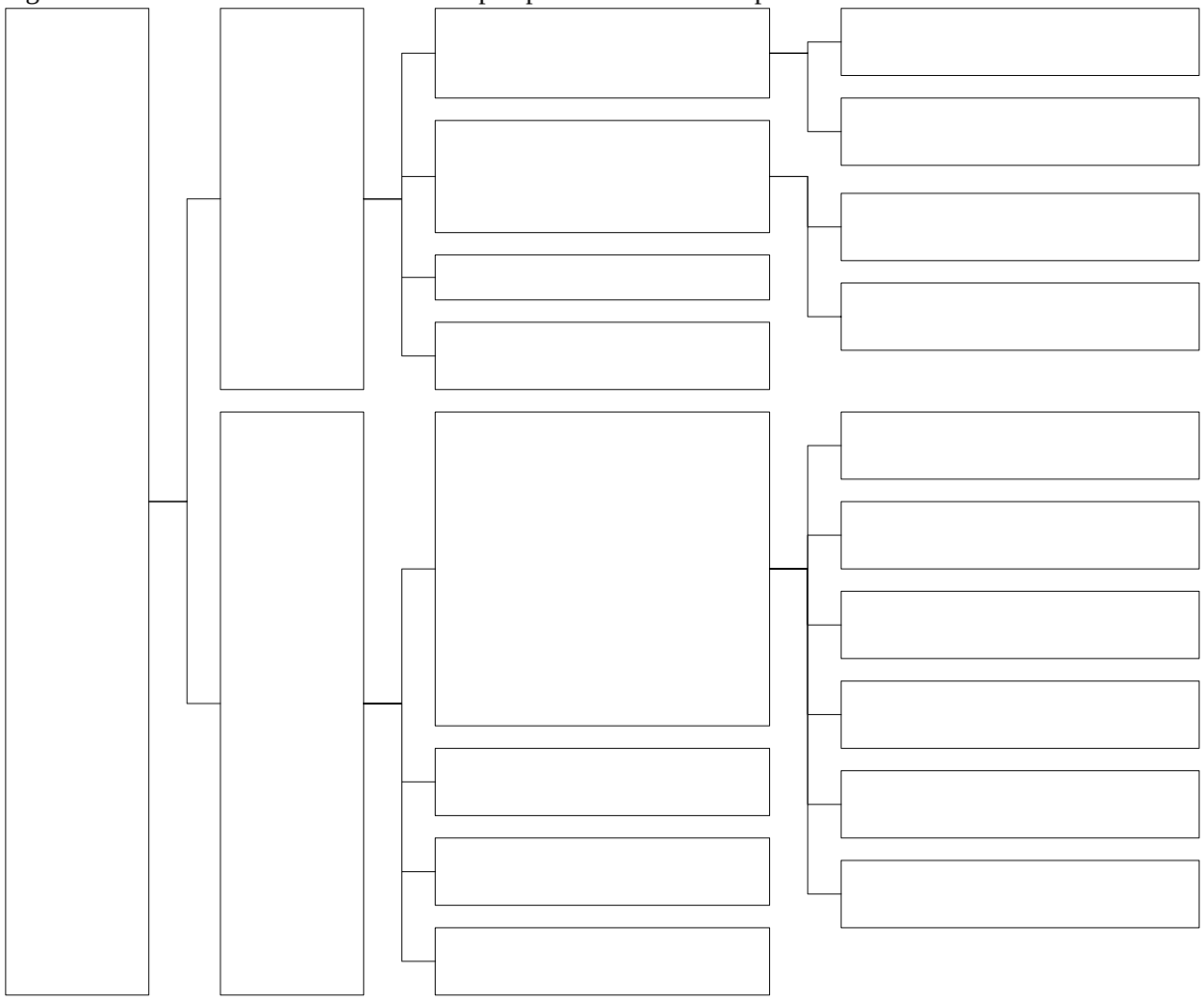
The diagnostic dimensions of the individual perspective category consisted of the following:

- Learning (5 comments) was defined as individual red team member gaining new knowledge either on a project or engagement. Learning was further stratified into two categories.
    - Professional learning (4 comments), which was defined as individual intellectual growth in the job and
    - Individual contributions (1 comment), which was defined as how much each red team member contributed to the overall effectiveness of the red team.
- Fun (4 comments) was defined as whether the team member enjoyed the red team project.
- Degree of creativity (2 comments) was defined as how and the extent to which a red team member exhibits originality or ingenuity in the different phases of projects.
- Satisfaction in work was defined as the individual level of fulfillment with red team work (1 comment).

2.1.2 Team perspective category

The team perspective consists of responses regarding performance measures on the red team as a whole. The red team members reported that team measurement is both evaluative (14 comments) and diagnostic (13 comments), for a total of 27 comments in team perspective category. Figure 2 summarizes the number of comments for the team perspective category.

Figure 2. Comments on the team member perspective on red team performance



2.1.2.1 Team perspective category: Evaluative dimensions

Within the evaluative dimension category (14 of 27 total team perspective category comments), the red team members reported on six sub-categories.

- The ability to understand the system (3 comments) referred to comments regarding how well and to what capacity the red team characterized the system.
- Understanding risks and tradeoffs (2 comments) referred to acknowledging the threats of security and managing them.
- Consistency in results sub-category (3 comments) is comments regarding if multiple red teams targeted the same system, would there be the same findings?
- Expertise of team members (1 comment) referred to comparing teams in terms of categories of expertise across projects.
- Lessons learned (4 comments) referred to the effectiveness of feedback within and across groups and/or projects.
- Collective learning and knowledge (1 comment) referred to the knowledge the red team creates as a whole.

2.1.2.2 Team perspective category: Diagnostic dimensions

Within the diagnostic dimension category (13 of 27 total comments), team dynamics (10 comments) was the largest sub-category of the diagnostic dimension in the team perspective category. There were six sub-categories in the team dynamics category.

- Flexible and adaptive (3 comments), referred to how red teams must accommodate and respond to unplanned issues in projects.
- "Past red team effectiveness" (2 comments) referred to specific examples of red team effectiveness, such as sharing problems and sharing project vision with red team members.
- Shared vision (2 comments) was defined as the team having an understanding of the common goal in each red team project.
- Coming together quickly (1 comment) referred to how fast the red team can organize and mobilize their project efforts.
- Conflict (1 comment) was defined as whether or not red team members "got along with one another".
- Trust (1 comment) referred to the level to which the team had confidence in one another.

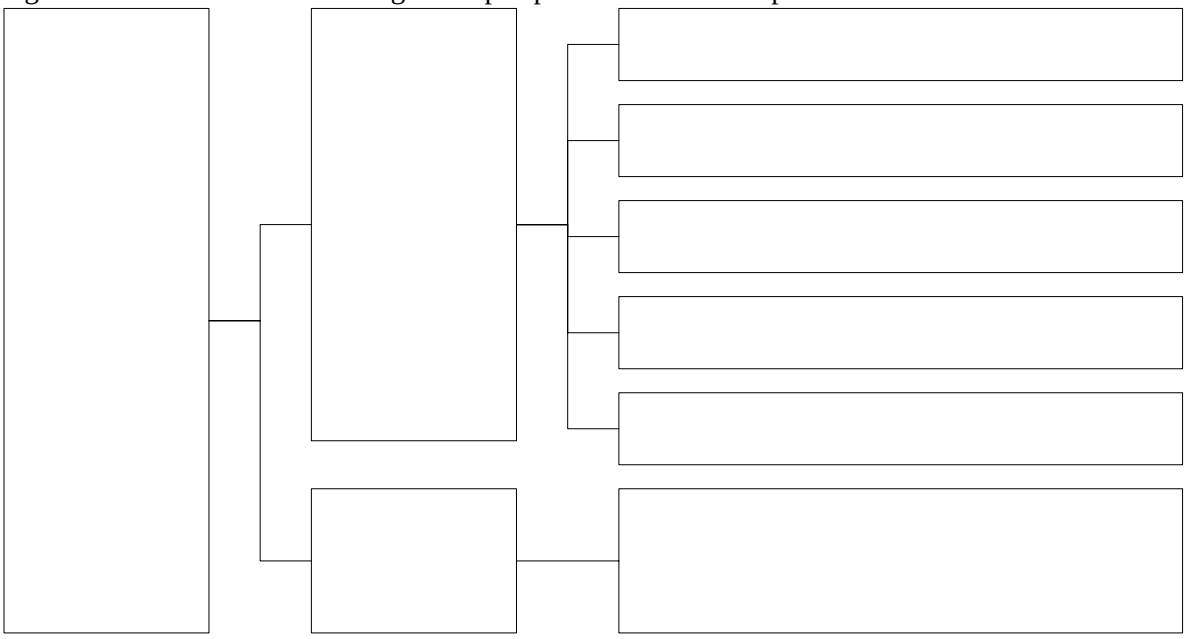
There were three diagnostic dimension unrelated to team dynamics.

- Satisfaction in work (1 comment) referred to the level of team fulfillment in their project efforts.
- Maturity of team (1 comment) referred to the team level of professional development in red teaming.
- Attitude (1 comment) referred to how well (or not so well) the team functions interpersonally.

2.1.3 Management perspective category

The management perspective category consists of responses regarding performance measures by those who manage or hold leader positions within the red team. The red team members reported that measurement in the management perspective category is both descriptive (11 comments) and evaluative (1 comment), for a total of 12 comments in this category. Figure 3 summarizes the number of comments for the management perspective category.

Figure 3. Comments on the management perspective on red team performance



2.1.3.1 Management perspective category: Descriptive dimensions

There were four descriptive dimensions for the management perspective category (11 of 12 comments).

- Capturing the flag (4 comments) referred to obtaining system targets.
- Schedule (3 comments) referred to issues related to adhering to the time deadlines in a project.
- Statement of work goals (2 comments) was defined as meeting the objectives of the project that were outlined in the statement of work document (with cooperation of the customer).
- Budget (1 comment) referred to meeting the monetary constraints of a project.
- Paperwork for project status (1 comment) was defined as completing the necessary documents that track individual and team project work progress and completion.

2.1.3.2 Management perspective category: Evaluative dimension

In the management perspective category, the evaluative dimension consisted of 1 of 12 total comments.

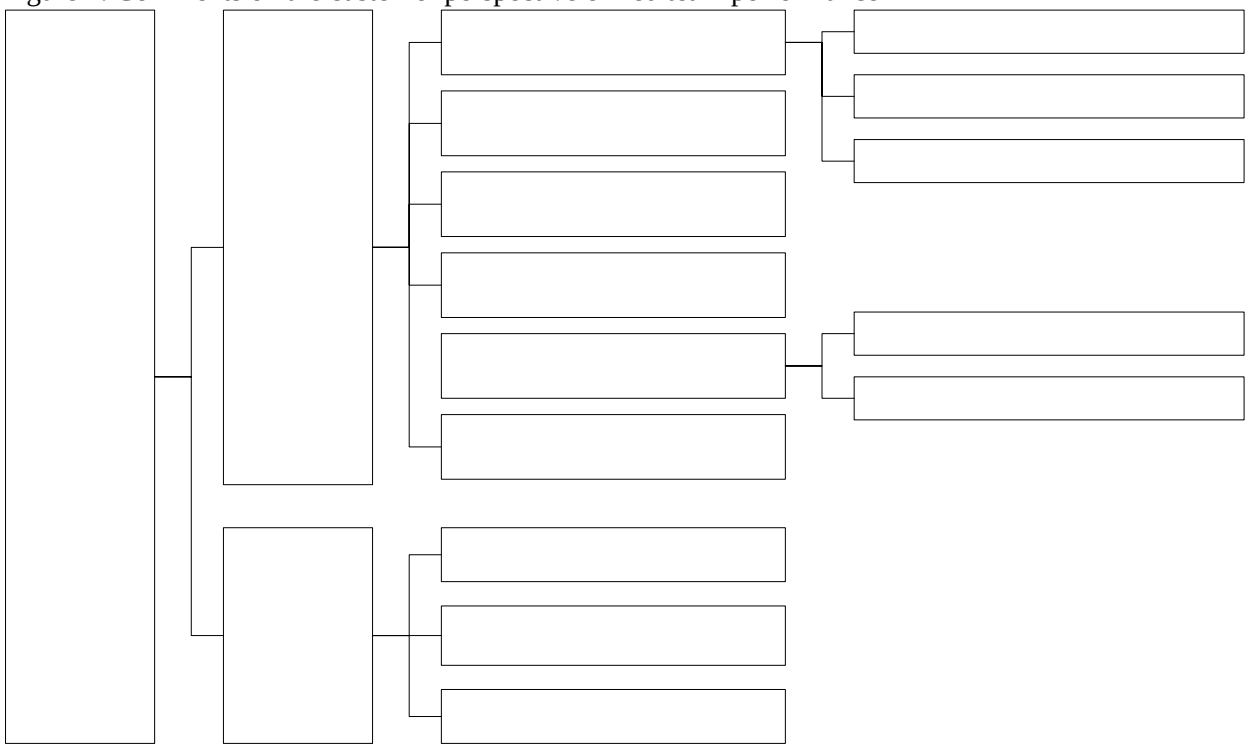
- Lessons learned (1 comment) referred to ideas, solutions, or problems that were relayed to the management at the end of a project, so that management may integrate solutions or information into red team improvement.

Management perspective category: 12 total comments

2.1.4 Customer perspective category

The customer perspective consists of responses regarding performance measures by red team members from the perspective of their customers. The red team members reported that customer perspective measurement is both descriptive (22 comments) and evaluative (8 comments), for a total of 30 comments in customer perspective category. Figure 4 summarizes the number of responses for the customer perspective category.

Figure 4. Comments on the customer perspective on red team performance



2.1.4.1 Customer perspective category: Descriptive dimensions

In the customer perspective category, descriptive dimension consisted of 22 of 30 total comments.

- Feedback to customer (3 comments) referred to how well the red team communicates its findings to the customer.
- Useful information (6 comments) was defined as how helpful or valuable the findings of the red team are to their customers.
- Quality of final report (5 comments) referred to the overall value of the report to the customer.
- Goals, budget, and timeliness (2 comments) is defined as meeting these requirements with the time specifications to the customer.
- Disrupt network service (1 comment) referred to the ability to interrupt service on the customer's network.
- Follow-up recommendations (1 comment) described an instance where a customer would request that a red team do a follow-up exercise after completing a project.
- The description of "informal customers" (2 comments) referred to customers who want to learn about the security of their systems and be interactive with the team. They give the red team latitude in their approach and methods.
- The description of "formal customers" (1 comment) described customers who are usually mandated a red team analysis, have a low level of trust for the red team, and do no allow the red team to stray from the plan of the project.
- Problems with implementation (1 comment) referred to if the customer was successful in implementing the red team recommendations.

Descriptive dimension: 22 of 30 total comments

Customer perspective category: 30 total comments

2.1.4.1 Customer perspective category: Evaluative dimensions

The evaluative dimension consisted of 8 of 30 total comments.

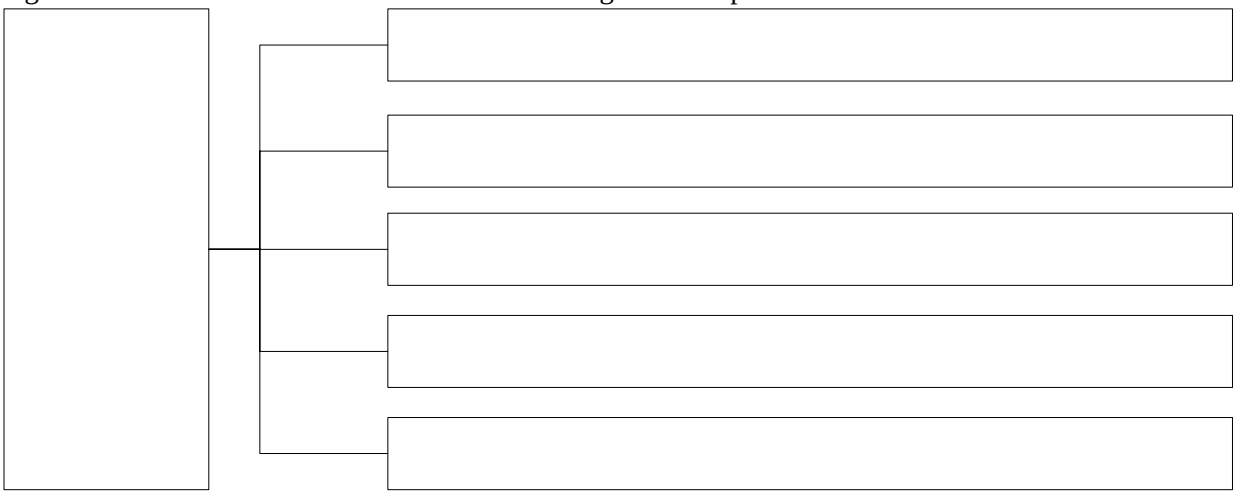
- Customer needs (4 comments) referred to customer project assessment expectations.

- Degree of difficulty (3 comments) was defined as the level of effort required to achieve the project goals.
- Creative solutions (1 comment) referred to the type of imaginative and innovative solutions that cannot be found by other easier methods, such as downloading tools from the internet.

2.2 Difficulties in measuring red team performance

Red team members reported some difficulties in measuring red team performance (14 total comments). Refer to Figure 5 for the number of comments regarding the difficulties in measuring red team performance.

Figure 5. Comments on the difficulties in measuring red team performance

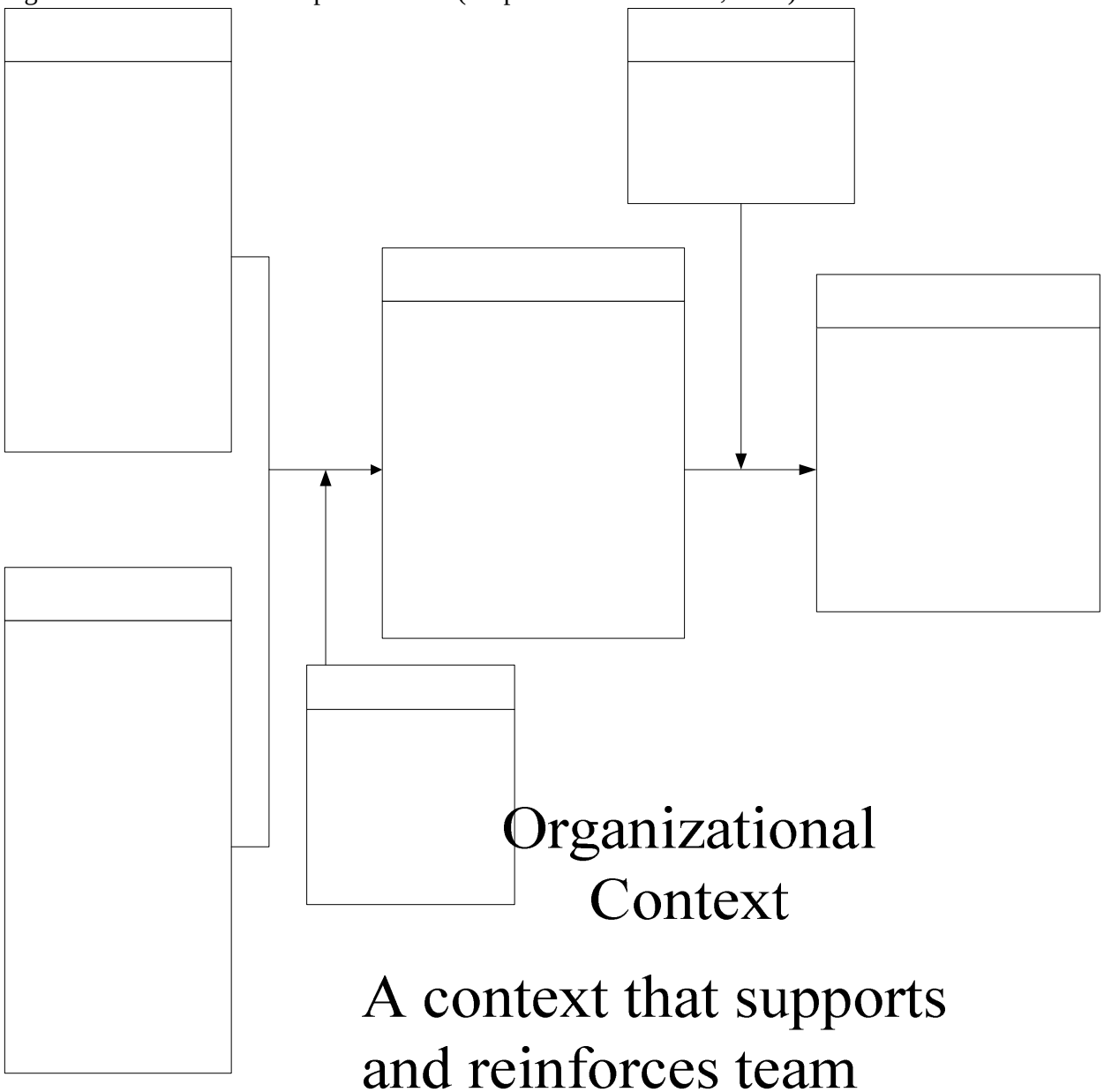


- Difficulty in measuring the customer's perception of effectiveness (9 comments) referred to how the customer reacts to a red team assessment.
- Measuring the unknown (2 comments) was defined as the difficulty in measuring the "defense" of a system, since there certain constraints and assumptions about a system are made during an assessment.
- Quantification effectiveness issues (1 comment) referred to the difficulty to express "helping people in security" quantitatively.
- Simulated environment (1 comment) accounts for the fact that red teams often have to engage systems in a simulated, not live, environment. In this environment, they may not be able to account for the realm of possibilities in a live environment.
- Variation in projects (1 comment) referred to differences across projects. It is difficult to take into account all of those variations when accounting for the overall effectiveness of red team performance.

3. Factors associated with red team performance

The literature on team effectiveness has proposed various models. We have adapted the Hackman's (1987) Normative Model of Group Effectiveness to red team performance (see Figure 6). The model defines a range of organizational factors that can affect red team performance. We used this model to categorize the factors that contribute to and hinder red team effectiveness.

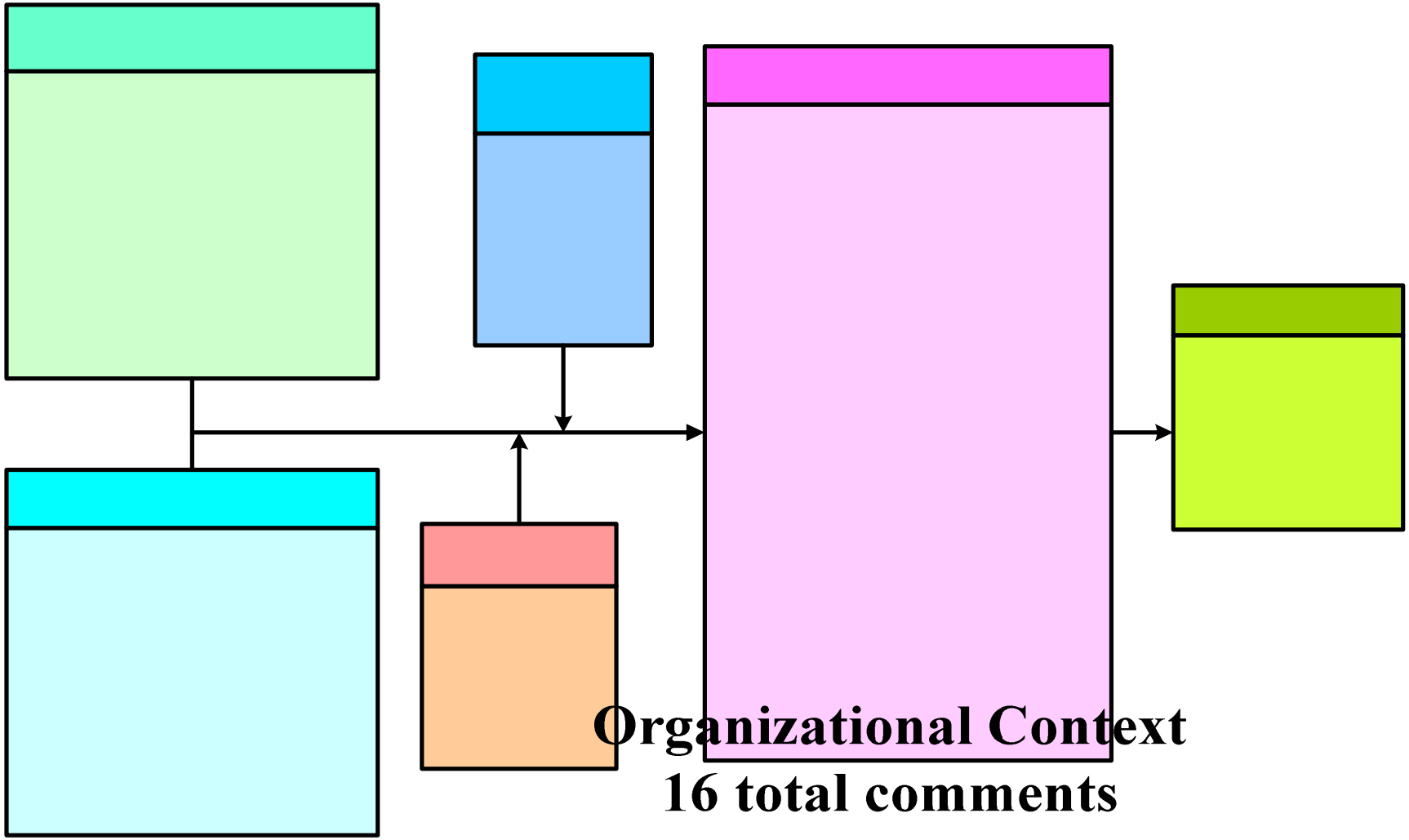
Figure 6. Model of red team performance (adapted from Hackman, 1987)



3.1 Factors that contribute to red team performance

Findings on factors that contribute to red team performance will be reported in five parts: organizational context (16 comments), team design (53 comments), team synergy (18 comments), material resources (2), and process criteria of effectiveness (74 comments). Refer to Figure 7 for a summary of the comments on the factors in the model of red team performance.

Figure 7. Factors that contribute to red team performance



Description of Sandia's organization (11

3.1.1 Factors that contribute to red team performance: Organizational context

In the organizational context (16 total comments), red team members described Sandia's organization (11 comments).

- Sandia's pool of experts (6 comments) was described as Sandia as a large research laboratory, supporting many projects and research initiatives that cross many disciplines. Projects may access experts from those disciplines within Sandia.
- Working at a government laboratory (1 comment) referred to the benefits of working within a government laboratory and working with other government laboratories. This includes the fact that government laboratories are not profit-driven, like private industry.
- Networking within Sandia's organization (1 comment) referred to how Sandians network between departments on projects and initiatives.
- "Sandians have similar backgrounds" (2 comments) referred to the fact that Sandia hires people with similar professional and educational experiences.
- Using new computer and information (CIS) tools (1 comment) referred to using new methods and/or tools in computer science within Sandia.

In the culture category (5 comments), red team members commented on Sandia's organizational culture as well as the culture of the red team.

- In the culture of the red team, a "helpfulness" mindset (1 comment) is associated with the group. This referred to how the red team is organized to help improve the nation's critical systems.
- The culture of Sandia (4 total comments) was referred to in four ways:
    - Hiring employees with a high ethical standard (1 comment).
    - Mentoring to develop new/young employees (1 comment).
    - The culture of initiative at Sandia (1 comment), referred to a "can-do" mentality by both the Sandian employees and management.
    - Rewards and recognition at Sandia (1 comment) referred to the types of programs that Sandia uses to reward employees, such as parties, gift certificates, and certificates of excellence.

3.1.2 Factors that contribute to red team performance: Team design

For team design (53 total comments), components of diversity in team design (28 comments) were cited by red team members.

- Expertise (12 comments) was referred to how team composition needs to have a mix of different skill areas and subject matter experts.
- Culture (1 comment) referred having different cultural backgrounds represented in the team. For example, some cultures have a "linear" mindset to communicate a problem, while others may use a metaphor to describe a problem.
- Experience level (1 comment) referred to how the acquired skill/knowledge levels of team members needs to be different within a project team.
- Thinking type (4 comments) referred to using people with different ways of approaching a problem, such as detailed versus "big picture" thinking.
- Personality type (2 comments) referred to having individuals with different character traits present in team design.
- Communication style (2 comments) referred to how team members convey information to one another.
- Maturity and professional development (2 comments) is defined as having a mix of "old" members and "new" members.
- The importance of diversity (4 comments) addresses why it is critical to incorporate diversity into team design. It is important to design in team diversity, otherwise they may miss vulnerabilities,

or, if there are too many people with the same knowledge domains on the team, it may hinder progress.

For individual team member attributes (25 comments), core and non-core team members (24 comments) and manager (1 comment) perspectives were commented on by red team members.

- Creativity (1 comment) referred to how red team members enjoy being innovative in individual ways within the group.
- Attitude (16 comments) referred to having a strong initiative and commitment to work, no ego, a developed interest in their work, and a competitive disposition.
- Flexibility (4 comments) referred to the ability to "give" in their tasks and duties as well as how they approach, plan, and how to carry out attack plans.
- Design experience (1 comment) referred to the importance of doing security design and implementation work outside of red teaming.
- Staying current in CIS issues (2 comments) referred to keeping up-to-date on the latest hacker issues as well as issues in CIS and IS.
- For project managers, personality (1 comment), was referred to as the personality characteristics of an effective project manager. This includes being able to multi-task, having a certain level of social capability, and the ability to recognize and address problems.

3.1.3 Factors that contribute to red team performance: Team synergy

For the area of team synergy (18 total comments), factors fell in seven areas.

- Traveling (2 comments) referred to the bonding experience of "being on the road" with fellow red team members.
- Collaboration (2 comments referred to team members building off of each other's interests, strengths, and knowledge.
- Trust (1 comment) referred to how the team relies on one another, such as knowing that tasks will get done and duties will be accomplished.
- A common interest (1 comment) referred to how it is enjoyable to work with like-minded individuals on the team.
- Social interests (1 comment) described being friendly and communicative with other team members.
- Honesty (1 comment) among team members referred to the ability and desire to communicate thoughts and feelings frankly and with respect.
- Creativity (5 comments) referred to how innovation and imagination contributes to the dynamics of the team.
- Fun (5 comments) included sharing food with one another and being excited about red team work.

3.1.4 Factors that contribute to red team performance: Material resources

For material resources used by the red team, red team members commented on hardware and software issues (1 comment) and the use of open-source software (1 comment).

- Hardware and software issues referred to availability of resources in this area. For example, the red team may not have the hardware and/or software resources to set up a test system and if the customer can provide the red team with the resources for a system such as this, it may affect the performance of the red team.
- Use of open-source software referred to using open-source software because of licensing issues when using other sources, such as Windows. The red team usually modifies open-source software for projects.

3.1.5 Factors that contribute to red team performance: Process criteria of effectiveness

For process criteria of effectiveness (74 total comments), red team members commented on five distinct areas: administrative processes (39 comments), high-level assessment process (12 comments), detailed assessment and planning (18 comments), engage systems (3 comments), and report to customer (2 comments).

The administrative process category consists of 39 comments.

- Documentation (5 comments) included the need to record work-related information in order to communicate individual activities to the other red team members, the importance of documentation so that red team members may cross-reference data, the benefit of a potential web-based system to automate documentation procedures, and the fact that the easier the documentation process, the more people will engage in documenting their work.
- Communication (25 comments) included five sub-categories.
    - Feedback (13 comments).
    - Discussing goals (1 comment).
    - Communicate expectations (2 comments).
    - Communicating with the customer (1 comment).
    - Communication among team members (6 comments).
    - Feedback from management (2 comments).
- Leadership (6 comments) referred to project management in the areas of monitoring the ability of the red team to make adaptive changes to the environment, delegating responsibilities, interfacing with the customer, supporting the red team so that they may carry out their work without having to complete administrative tasks, championing red team processes, and facilitating interested members for red team projects.
- Mentoring (2 comments) referred to mentors cultivating a code of ethics and professionalism in their program.
- Defining processes (1 comment) was described as the formulation and statement of overall practices and goals of red team projects.

High level assessment process (12 comments) consists of five sub-categories

- "Following the information flows" (1 comment) referred to how the red team looks at the flow of information within the organization to look for vulnerabilities.
- Creating a system understanding (2 comments) referred to the importance of generating an accurate and comprehensive understanding of the system under inspection.
- Brainstorming (2 comments) referred to how the red team uses the technique of brainstorming to generate different views of the system.
- Preparation (3 comments) was in regard to the importance of being completely prepared to understand the system under inspection. This includes not only understanding the system's hardware and software, but also the people and processes that work together.
- Project goal formulation (4 comments) referred to how managers need to develop stated goals for the projects, especially in the early project stages.

Detailed assessment and planning (18 comments) consist of four sub-categories.

- Software development (2 comments) refers to the need to document the development of the software used in red team projects.
- Formulating attacks (2 comments) are thoughts regarding the process the red team uses to brainstorm and develop attacks on systems.
- Customer relations (13 comments) includes, the assessment of the customer's security culture, receiving feedback from the customer, the customer's attitude regarding the work of the red team, the customer's interpersonal relationship with the red team, the customer's role with the red team

(i.e. including the customer in the red team), and organizational differences. Organizational differences include organizations that have centralized versus decentralized security functions and private versus public organizations.

- Ideas and thought processes (1 comment) described how ideas are formulated and why decisions are made during exercises. This was related to the fact that this information needs to be recorded.

When the engaging systems (3 comments) the rules of engagement and competing blue team are important.

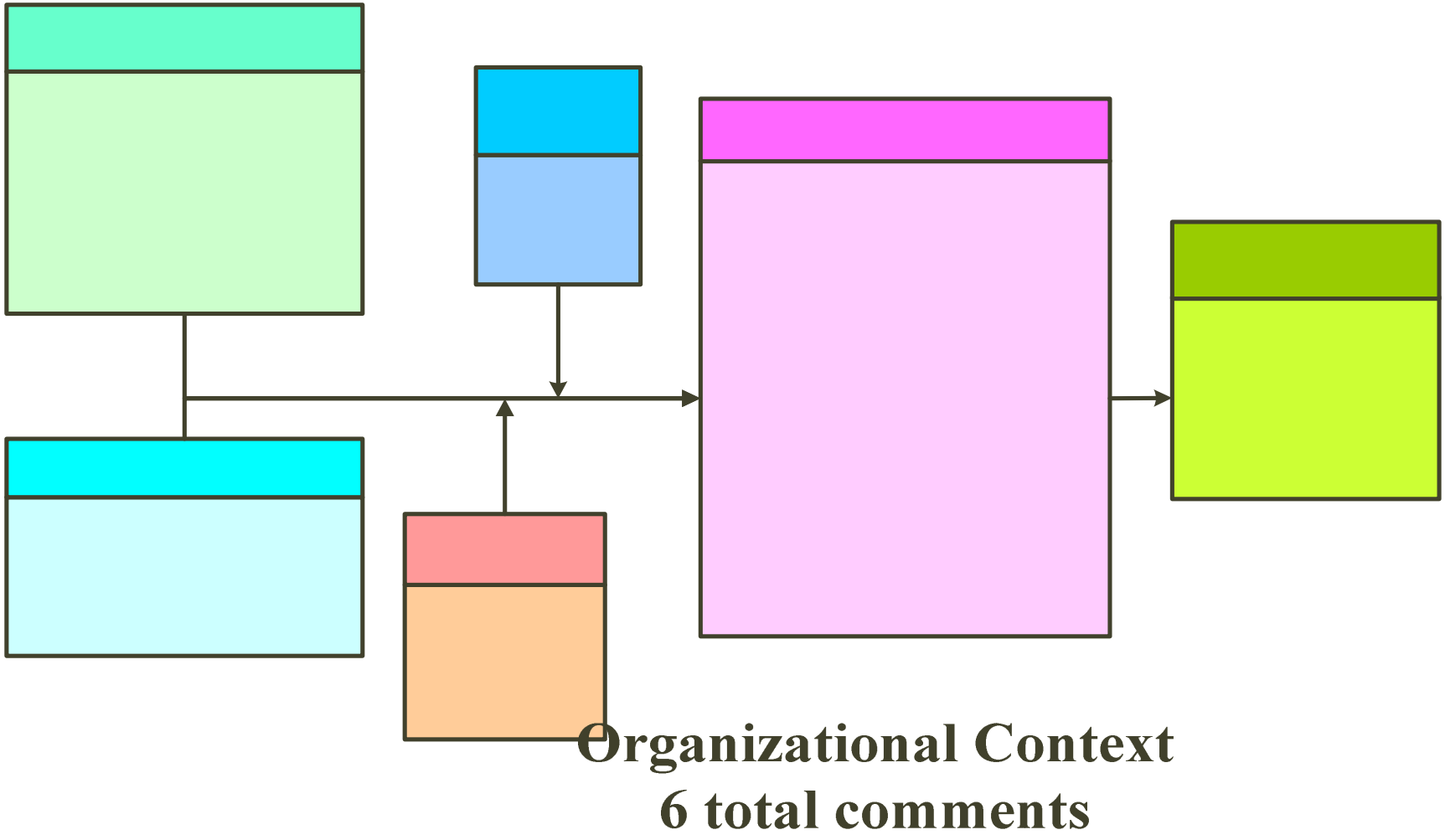
- Rules of engagement (1 comment) are important because when the red team is on the customer's premises, the red team needs to conform to certain boundaries of behavior, which includes interfacing with customer's systems as well as the customers themselves.
- Blue teaming issues (2 comments) include engaging the customer in a blue teaming effort, especially if the security professionals (i.e. the customer) are not aware of the red teaming application to their system. This was relevant because the blue team may change their everyday behavior or method based upon the red team application to their system.

For reporting to the customer, documentation as proof of concepts (2 comments) is important. It is important to record the red team analysis, as well as the output of the red team, so that they may prove their methods and analysis to the customer.

3.2 Factors that hinder red team performance

Findings on factors that hinder red team performance (28 total comments) will be reported in five parts: organizational context (6 comments), team design (4 comments), team synergy (3 comments), material resources (1 comment), and process criteria of effectiveness (14 comments). Refer to figure 7 for a summary of the comments on the factors that hinder red team performance.

Figure 7. Comments on factors that hinder red team performance



Sandia's organization

3.2.1 Factors that hinder red team performance: Organizational context

In the organizational context (6 total comments),

- Job control (1 comment) was cited as an organizational factor that hinders red team performance. One red team member felt they experience less job control working at a government laboratory than in a private sector job. In government agencies, going through the bureaucratic process to accomplish their job may be burdensome.
- Red team members also thought that Sandia's matrix organization may also hinder performance (5 total comments).
    - "Matrixed" red team members felt that being in a consulting role is difficult at times because they have a full time job in their respective organizations (2 comments). Sometimes there are time they cannot make all of the meetings, look over the strategic plans for the attacks and they feel this affects the overall progress of the team.
    - The fluidity in team membership affects performance (3 comments). When people pull out of a project and those positions are replaced by others, there is a learning curve to bring the person up to speed on the project, since those persons were not present in the developmental stages of the systems view and brainstorming sessions. The team may loose a considerable amount of time (i.e. weeks or months). This may not be critical if the red team is just working on operating systems. If the team is working on vulnerabilities, the time element is more critical, since the longer the vulnerabilities are exposed, the long the company is at risk for an exploit.

3.2.2 Factors that hinder red team performance: Team design

In the team design stage (4 total comments), expertise (2 comments) is a key issue for two reasons.

- Firstly, having more than one or two persons in the same expertise area may not create enough diversity to cover all the subject areas that are needed for the project. For example, if the work is mostly like network or application analysis, having more than one network person on the team or more than one programmer on the team can be useful because you bring another perspective to the same problem. If there are more than two or three people in one area of expertise there is not enough work to go around, and in splitting it up each person has too small a view of what they are working on. This results in duplicating efforts and the process under inspection becomes too fragmented to be able to obtain the system view.
- Secondly, the wrong mix of expertise can affect performance. Not having the knowledge areas covered in the group will not allow the team to cover all of the areas needed to understand the system.

Team size (1 comment) is another issue affecting team performance. This is particularly relevant since the Sandia's red team business has grown over the years. In one core member's view, the red team personnel have not grown according to the increase in business. As a result, they have not had the time to spend cultivating a shared vision of red team goals. Team design is also affected by the individualistic nature of computer scientists and system analysts (1 comment). This makes it difficult to organize as a team, because of their singular nature; and as a result individuals do not always think and work in the team context. To address this, more team meetings and teleconferences with the customer with the red team members are conducted to keep the red team operating in the team context.

3.2.3 Factors that hinder red team performance: Team synergy

With team synergy (3 total comments),

- There are gender issues that may arise in the red team (2 comments).
    - This includes men "talking over" women in meeting (1 comment). An interviewee also pointed out that when a female runs a meeting the men respect her leadership and do not interrupt or engage in talking over her.

- There have been instances when men have "side-stepped" females on the team to get information (1 comment). For example, two members with similar expertise (one female, one male) were participating on a project where other red team members would ask the female's male counterpart questions only the female could answer. She would answer the questions and he would report her answer to the group.

Unmotivated team members (1 comment) may also sap team performance levels. Whether the reasons for the team member to be unmotivated is a lack of interest in the project or does not care to work on projects that they feel they cannot perform well at, the overall performance of the team is hindered by their lack of initiative and progress.

3.2.4 Factors that hinder red team performance: Material resources

For the category of material resources (1 total comment), customer's budget constraints (1 comment) may adversely affect red team performance. This includes the amount of money they are willing to spend on equipment. In a lot of cases, the red team would like to get duplicate equipment at the laboratory to practice on prior an engagement. If the customer does not approve it, they have to learn on site, at the time of the engagement.

3.2.5 Factors that hinder red team performance: Process criteria of effectiveness

For the category of process criteria of effectives there were 14 total comments. For the administrative processes, there were no comments made by red team members.

In the high level assessment category (5 total comments) there are three sub-categories.

- Dismissing other team members' perspectives (1 comment) affects red team performance. Dismissal is especially damaging in brainstorming sessions, where this behavior may shut down other people from talking. To overcome this, the red team will employ the rules of brainstorming.
- Customers may provide inadequate information about their system to the red team (3 comments). This may be a result of poor documentation on the customer's part or a lack of documentation in general. One core member did point out that this does not really affect the quality of their output. In cases where companies do have documentation, about 25% less time is spent on the project.
- When completing the various stages and phases of the project, communicating individual learning to the rest of the team can be problematic (1 comment). The red team assessment process at times becomes lengthy. Members may end up doing work/rework if fellow members are not communicating their knowledge to the rest of the team.

In the detailed assessment and planning phase (9 total comments),

- System misunderstanding may hinder red team performance (1 comment). System misunderstanding was defined if the red team does not adequately comprehend the system under consideration.
- Lack of time to work on certain aspects of the assessment (3 comments) may be a limiting factor. One red team member feels that good success in a project occurs when they are given enough time to really analyze a system and discover its vulnerabilities. Adapting a "tunnel vision" view of a system may also limit performance (1 comment).
- Red team members may operate on preconceived approaches and therefore miss novel attacks.
- Issues concerning customer relations also affect the detailed assessment and planning phases.
    - Firstly, the red team may not always be interacting with the owner of the information system; they may be working with a liaison that interfaces between the red team and owners of the information system (1 comment). The 'once-removed', indirect communication between the red team and the customer liaison is a challenge to for the red team to because of the divisive information they receive.

- Inadequate organizational access (1 comment) on the part of the customer also hinders progress. At times, the customer's upper-management does not convey to the appropriate departments what the red team will be doing and consequently, the red is not given the adequate resources and access to perform a thorough assessment.
- There may also be an element of fear from customer's employees (2 comments). If not appropriately described to the organization's employees, the red team may not be perceived in a helping role, but rather an evaluative group whose job is to prove that the IT department is doing an inadequate job at securing their systems.

3.3 Summary of the factors associated with red team performance

Overall, red team members talked more about factors that contribute to red team effectiveness (162 total comments) than factors that hinder red team performance (56 total comments). In this section, similarities and differences between contributing and hindering are discussed.

In the organizational context category, Sandia's organization had factors that both contributed to (11 comments) and hindered (6 comments) red team performance. Specifically, Sandia's matrix organization supports the red team in its pool of available experts, but the ease in which team members may move from project to project and the limited-time role of consultant may impede performance. The cultural aspects of both Sandia and the red team contribute to performance (5 comments). There were no comments on the culture of these organizations that hinder performance.

In the team design category, the team design factors that contribute to red team performance (53 comments) were greater in number than the team design factors that hinder red team performance (4 comments). Skill expertise was mentioned in both the contributing factors category (12 comments) and the hindering factors category (2 comments). Red team members went on to describe how other team design factors contribute to performance, such as diversity in culture, experience level, types of thinking, personality type, communication style, and maturity and professional development. When describing individual contributing factors (25 comments), attitude (11 comments), which included having a strong initiative and commitment, no ego, a developed interest in their work and a competitive disposition. For individual factors that hinder red team performance, the nature of being "individual" thinkers (1 comment) was mentioned. Overall, is seems that creating diversity in different contexts contributes to performance while considering the individualistic nature of computer scientists may hamper team performance.

In the material resources category, issues surrounding hardware and software (1 comment) and use of open-source software (1 comment) contributed to red team performance. Conversely, customer budget constraints (1 comment) may hinder red team performance. However, these two sets of factors are linked: both are dependent upon the resources given by the customer. Having adequate and appropriate software to complete a project most certainly contributes to effectiveness, but if the customer's places budgetary constraints up on the team, these resources are now unavailable.

There were not strong similarities in the comments regarding team synergy. For the factors of team synergy that contribute to red team effectiveness (18 total comments), more "social" aspects such as traveling together, sharing food as a team, sharing common interests, and having fun together were discussed. For the factors of team synergy that hinder red team effectiveness (3 comments), more interpersonal issues were discussed, such as poor inter-team communication and unmotivated team members.

The category of process criteria of effectiveness yielded 74 total comments in the contributing factor category and 14 total comments in the hindering factor category. There were 39 comments in the contributing factors administrative processes sub-category and zero comments in the hindering factors

sub-category. Within the contributing factors administrative process sub-category, issues such as the different facets of communication and the importance of various types of feedback were emphasized. Communication issues included discussing goals and conveying expectations and information to management.

In the high-level assessment sub-category of process criteria of effectiveness category, there were more comments on the contributing factors (12 total comments) than hindering factors (5 comments). For the contributing factors, more process-related issues were discussed, such as preparation for project, brainstorming the different views of the system, and formulating the project goals. For hindering factors, some individual aspects were troublesome, such as dismissing other member's perspectives in the brainstorming process and not reporting the individual work to the team. In addition, inadequate information of the customer's system may also hinder formulating the system view.

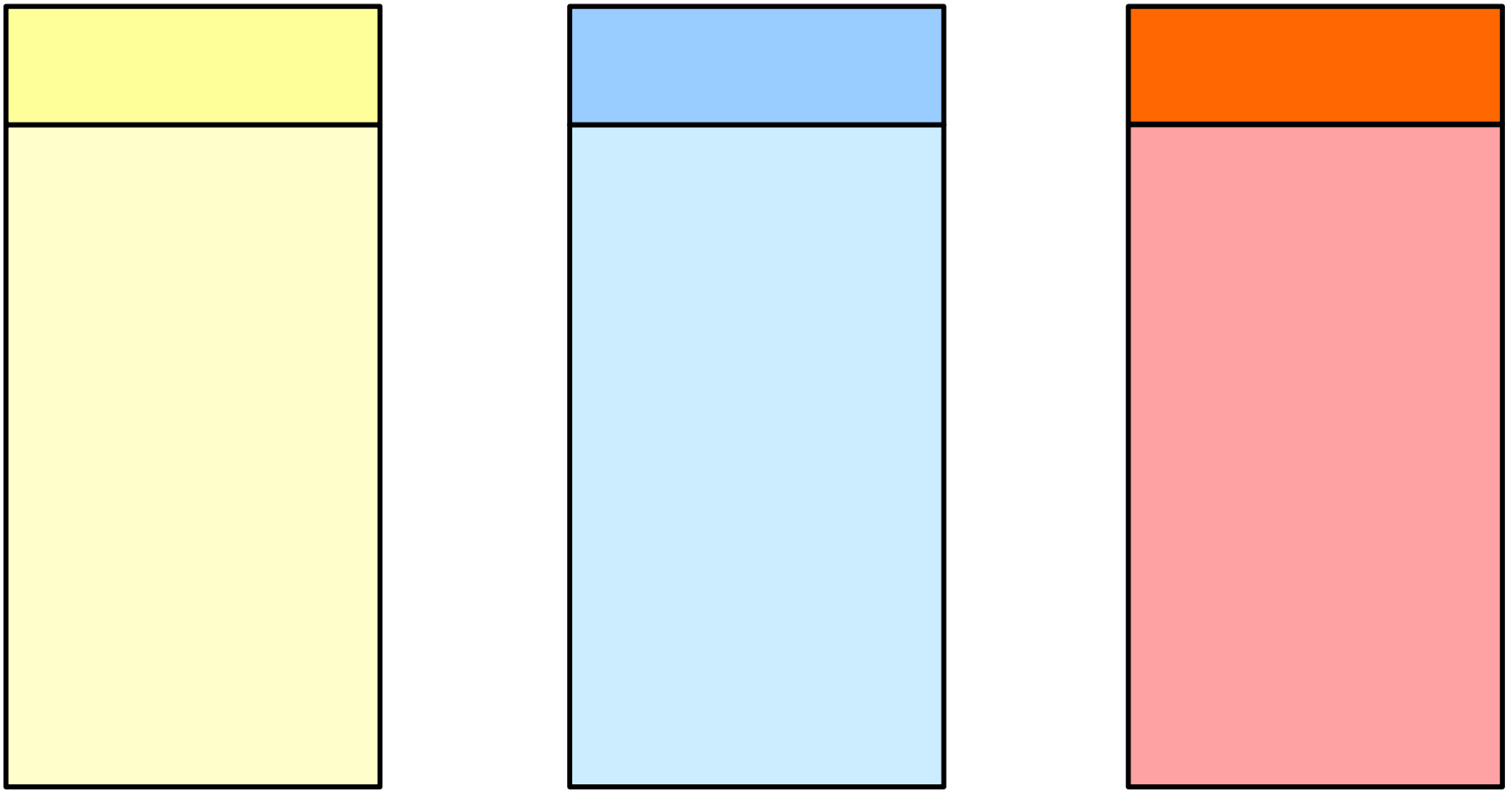
In the detailed assessment and planning sub-category of the process criteria of effectiveness category, factors that contribute to red team performance (18 total comments) again outweighed the factors that hinder red team performance (9 comments). Interacting with the customer, including receiving feedback from the customer, the customer's attitude about a red team assessment, and the interpersonal relationship with the customer was emphasized in as contributing factors of performance (13 comments). For factors that hinder red team performance in this sub-category, there were also issues related to customer relations. Fear of the evaluative assessment by the red team by the customer's employees (2 comments) also red team performance and success on projects. This fear usually results in difficulty in organizational access and limited information about the system under review.

When the red team engages a system, the rules of engagement (1 comment) and having a competing blue team (3 comments) contributes to the performance of the team. There were no comments on engaging systems in the hindering factors category. For reporting to the customer, documenting red team work for proof of concepts (2 comments) was important for project success. There were no comments on reporting to the customer in the hindering factors category.

4. Trade off analysis between red teaming and simulated security methods

In this phase of the study, our aim is to compare and contrast red team performance to modeling/simulation methods. Red team members reported on three main areas: (1) weaknesses of simulation methods (16 comments); (2) strengths of simulation methods (6 comments); and (3) strengths of red teaming (12 comments). Refer to Figure 8 for a summary of comments in these three areas.

Figure 8. Comments on tradeoffs between red team performance and simulated security methods



4.1 Weaknesses of simulation methods

Red team members commented on the various weaknesses associated with using simulation methods for security assessment (16 total comments).

- In this area, red team members felt that humans are superior at connecting data into useful information (7 comments). To a certain extent, there is some value to script attacks because it may save time and labor. However, a human still "needs to be in the loop" because of the unplanned oddities that come up in attacks. A simulated package may not acknowledge and respond accordingly. Further, there are false positives and a need for interpretation that are associated with some of these tools. For example, the half of the output of a Nessus scan may false positives. A human needs to recognize and interpret the findings of scans before useful information can be extrapolated from the output of such scans. Lastly, modeling and simulation tend to be either too abstract to handle the ambiguities of security. One person described security has having two levels of resolution that one needs to keep track of at the same time. There is a very fine resolution that will describe a specific machine that is not protected by a firewall and a higher level that says the system architecture some of systems are sharing trust with places that they shouldn't be sharing places with. The red team interprets at these two levels of detail at the same time and the current simulation methods that the work with do not. These methods usually end up being one or the other. A script is not dynamic, like a human with domain experience, a red team member interpret at a deeper level. A previously unrelated piece of information may how make sense, a memory association, and simulation methods cannot make these connects.
- A lack of human cognitive processing capabilities (2 comments).
- Simulated methods were described as a tool to open doors, but a human with an "extra cognitive level" recognizes the situations that aren't normal.
- Simulation methods also lack social engineering capabilities (2 comments). Since humans are interacting with the systems, maintaining patches, and the like, the methods used to exploit the system need to take into account that fact. An automated or simulation method would be more appropriate for areas of a system that are automated.
- Simulated methods cannot address the complexity and intricacies of systems (2 comments). Specifically, automated systems are not artificial intelligence. The automated system will do what they are programmed to do, where as humans can be innovative and "think on their feet".
- Simulated methods do not address the unpredictability of computer systems (1 comment). The red team may script attacks and it is only successful if all of the conditions required for the way in which they scripted the attack are met. Unexpected responses to an attack in the system may foil an attack. With a live team, a live engagement, the team can respond to a host of unanticipated and unexpected responses.
- A simulated tool does not address the ever-changing nature of computer and information security (1 comment). The cyber world is constantly changing and automated and simulated methods do not keep up with these changes.
- Simulation methods cannot assess a system's "nightmare consequences", or, perform a targeted assessment of the system's key vulnerabilities (1 comment). The current methods lack the empirical evidence to prove these.

4.2 Strengths of simulation methods

Red team members described the strengths of simulation methods (6 total comments).

- Simulation methods are good at testing against a set of known vulnerabilities (3 comments). This is an ample method for capturing the low-hanging fruit and assessing the basic vulnerabilities of the system.
- Simulation methods are also useful for compiling a database of red team findings (2 comments). The most beneficial of this application would be to compile a database of basic vulnerabilities, so that the red team could be more effective at eliminating "low-hanging fruit".

- Simulation methods are most useful in simple environments, such as single-purpose information systems, such as an isolated section of an information system that has few machines and systems interconnected to it (1 comment). With this limited interconnectivity, a limited number of hosts communicating within that information system generate very predictable traffic within that system.

4.3 Strengths of red teaming

Red team members commented on the specific strengths of red team performance (12 total comments).

- Red team tactics are anticipatory and "ahead of the curve" in the hacking world (3 comments). Their methods are at least current but certainly even anticipatory or advanced in terms of developing attacks that have not been seen in current hacking efforts.
- Red teams are a creative source for developing attacks (3 comments). Simulated tools will allow the red team to cover a lot of known territory, but, creativity is needed to discover new strategies. Further, the red team's creativity and innovation allows them to understand the truly important issues for the customer.
- Red teams are able to respond dynamically to changes in system states (2 comments). A system is constantly changing, in terms of number of users, user access levels, adding and removing functions and applications, and modifying firewalls. The ability to respond to those changing states is a red team strength.
- The red team produces a specified analysis of a system (2 comments). Red teaming is a process in which there is a targeted research, analysis, and solution formulation to a system. Red team members feel that a comprehensive red team analysis needs to not only identify the specific vulnerabilities of a system (i.e. number of open ports) but link this information into the broader system picture. How does this information tie into their security policy and implementation of security policy? Answering questions such as these allows the red team to provide an analysis that is tailored specifically to the customer's system.
- Writing tools to aid in targeting vulnerabilities are strengths of the red team (1 comment). Red team members rewrite and write programs to aid in attacks. These programs are tailored to specifically target a system for vulnerabilities after the red team has identified the unique characteristics of the system under assessment.
- The red team is able to find patterns in large amounts of data (1 comment). A simulation program's strength is to exhaust the realm of possibilities for vulnerabilities, while a human's strength is their heuristic nature; finding patterns and associations within large amounts of data.

5. Conclusion

This research studied red team performance in several ways. First, we identified and defined some preliminary measures of red team performance. Findings included measures that were categorized into four different perspectives: individual team member, the team as a whole, management, and customer. These perspectives were further stratified into process or outcome measurement dimensions: diagnostic (i.e. process measures), evaluative (i.e. outcome measures), and descriptive (i.e. process measures). An emphasis of was placed the team and customer perspectives as well as dimensions of process measurement (i.e. both the descriptive and diagnostic types of measurement). A fifth category of comments included some of the difficulties in measuring red team performance. Second, we identified factors that contribute to and hinder red team performance. In our model of red team performance, five categories of organizational factors were described: (1) organizational context; (2) team design; (3) team synergy; (4) process criteria of effectiveness; and (5) material resources. In both sets of factors, an emphasis was placed on the categories of process criteria of effectiveness and team design. Third, we identified some elements in a trade-off analysis comparing red team performance to automated/simulated methods. There were three categories in this analysis: (1) strengths in red teaming; (2) strengths in automated/simulation methods; and (3) weaknesses in automated/simulation methods.

5.1 Study limitations

Limitations of this study include the fact that descriptions of performance were based solely upon red team members' perceptions. IDART's red team represents one type of red team. There are other red teams in existence and it would be interesting to assess if the same or different facets and dimensions of performance would be identified in other groups. This includes extending this preliminary work by interviewing other red team members and leaders at Sandia, as well as other red teams. Further, teams evolve over time (Morgan, Glickman, Woodard, Blaiwes, & Salas, 1986) and the length of time that they have worked together can have a significant effect on group processes (Foushee, Lauber, Baetge, & Acomb, 1986). Performance assessment suggests the need to collect examples of complex performance over longer periods of time to gain a truer picture of team performance. Lastly, this research studied team performance in the contexts of interviews and some observation. Future research in this area may include other contexts that would allow for an understanding of performance that cannot be measured in interviews. Some alternate contexts could include an assessment of team behavior and process in a live engagement or brainstorming session.

5.2 Recommendations

The purpose of this research was to describe red team performance to improve red team performance. Unfortunately, evaluating team performance is no easy or straight-forward task. It almost always requires complex and time-intensive research methods to adequately assess *some* aspects of performance. In this preliminary research, we have described some dimensions and factors of red team performance. More clearly defined and comprehensive team performance measures and factors are needed to improve red team performance. In order formulate an effective team improvement intervention, not only is a developed understanding of the team's current performance state important, but so is an understanding of the team's current and future goals. There are many types of interventions and tools for team performance improvement. Some of these include team training, introducing a new feedback mechanism for team members, or a database that houses team strategies or case findings.

Recommendations for future red team performance work include both extending this preliminary work to build a more comprehensive view of red team performance and identifying key goals for improvement intervention. We need to further develop and define the measures and factors identified in this study. This may include more observation and interviews of not only previously interviewed red team members, but members that are new or were not interviewed in this study which would also include management. It is important to establish a baseline set of measures and factors so that performance may be monitored over time. The red team members who were interviewed emphasized process measures and process factors,

and that may be an area of further investigation in designing upcoming red team performance research. Some of areas, such as issues of feedback or team design, could be some specific areas to target for development.

We have recently submitted two NSF proposals, one to the Information Technology Research (ITR) program and one to the CyberTrust program, to further our study of the performance of red teams. In these proposals, we plan to advance our development of red team performance measures, as well as the factors associated with red team performance. We plan to examine red team performance with an experimental design that will test various team configurations with associated outcomes. We also plan to develop tools to enhance red team performance. This includes including human error taxonomies and case-based reasoner, which is a framework for categorizing and merging data from multiple sources including the many human, organizational, and team factors involved in cybersecurity problems. We will continue our collaboration with Sandia on both of these proposals. We are collaborating with Daniel Schwartz and Sara Stoeklin, professors of Computer Science at Florida State University on both proposals. On the CyberTrust proposal, we hope to collaborate with Sami Saydjari and Dave Farrell's red teams at the Cyber Defense Agency, LLC.

We have submitted our findings on measures of red team performance to the 2004 Human Factors and Ergonomics Society Conference, which will convene this September. We are continuing our publications efforts with journal papers on the factors contributing to and hindering red team performance and on the tradeoffs between red teams and automated security methods.

References

Foushee, H. C., Lauber, J., Baetge, M., & Acomb, D. (1986). *Crew factors in flight operations: III. The operational significance of exposure to short-haul air transport operations* (NASA Technical Memorandum 88322). Sunnyvale, CA: National Aeronautics and Space Administration-Ames Research Center.

Hackman, J. R. (1987). The design of work teams. In J. Lorsch (Ed.), *Handbook of Organizational Behavior* (pp. 315-342). Englewood Cliffs, NJ: Prentice Hall.

Morgan, B. B., Glickman, A. S., Woodward, E. A., Blaiwes, A. S., & Salas, E. (1986). *Measurement of team behaviors in a Navy environment* (Technical Report No. NTSC TR-86-014). Orlando, FL: Naval Training Systems Center.

Paris, C. R., Salas, E., & Cannon-Bowers, J. A. (2000). Teamwork in multi-person systems: A review and analysis. *Ergonomics, 43*(8), 1052-1075.

Appendix 1

Interview Questions

1. Various factors affect red team performance and red team performance can be evaluated on different dimensions. What are the various criteria for evaluating team performance?

2. What are the factors that contribute to red team performance/effectiveness?

3. What are the factors that hinder red team performance/effectiveness?

4. What are the trade-offs of red teams to modeling/simulation techniques?