Assessing Scientific Contributions: A Proposed Framework and Its Application to Cybersecurity

Dan Craigen

“The philosophy of science is about as useful to scientists as ornithology is to birds.”

Attributed to Richard P. Feynman (1918–1988)
Theoretical physicist

Through a synthesis of existing work on evaluating scientific theories and contributions, a framework for assessing scientific contributions is presented. By way of example, the framework is then applied to two contributions to the science of cybersecurity. The science of cybersecurity is slowly emerging. As the science and its theories emerge, it is important to extract the key contributions that characterize actual progress in our understanding of cybersecurity. Researchers and funding agencies will be interested in the assessment framework as a means of assessing scientific contributions to cybersecurity. In a nascent research area such as the science of cybersecurity, this article may contribute to a focused research program to accelerate the growth of the science.

Introduction

Hardly a day goes by without yet another report of significant information security vulnerabilities. Some of the most recent attacks, such as the heartbleed bug (MITRE, 2014a) and the shellshock bug (MITRE, 2014b), have focused on core functionalities. The former vulnerability is in an implementation of an Internet-wide protocol (SSL) and the latter vulnerability is in a widely used UNIX command-line interpreter (bash).

After decades of substantial investment into cybersecurity, it is almost unfathomable that such vulnerabilities continue to expose societies to potentially significant exploitation. In the author’s view, the existence of these vulnerabilities reflects the complexity of the cybersecurity space and suggests that the existing paradigms for identifying, responding to, or mitigating vulnerabilities and their potential exploitation are failing. Given the perceived ad hoc nature of cybersecurity, which is usually exemplified by patching systems in response to identified vulnerabilities, there is an emerging belief that the foundations of cybersecurity need to be revisited with a sound theoretical/scientific perspective.

It is through a sound theoretical/scientific perspective that we can evolve cybersecurity from its current (largely) ad hoc nature, to a foundation that is well-principled and informed by scientifically-based tenets (Schneider, 2012). Such a theoretical foundation then informs a rigorous engineering discipline, which, it is hoped, will positively impact cybersecurity postures.

However, a difficulty facing researchers, funding agencies, government, and industry is how to assess putative contributions to such a theory. In this article, we synthesize a framework for assessing scientific contributions to cybersecurity. The framework was motivated by the author’s involvement with various initiatives in the science of cybersecurity and the need to ascertain whether contributions were truly progressing and contributing to such a nascent science. Particularly, given that development of such a science will be a multi-decade exercise, being able to measure progress and contributions, at least incrementally, would provide important objective input into both research and funding decisions.
Assessing Scientific Contributions
Dan Craigen

First, we introduce the concept of theory and an approach to building theory. We then review the literature on measuring progress in science and assessing theories. The key concepts arising from the literature are then synthesized into a framework for assessing scientific contributions to cybersecurity. Finally, we demonstrate the use of the framework by applying it to two scientific contributions in cybersecurity.

Building Theories

Theory refers to "a well-confirmed type of explanation of nature, made in a way consistent with the scientific method and fulfilling the criteria required by modern science" (Wikipedia, 2014). Weber (2012) notes that "theories provide a representation of someone’s perceptions of how a subset of real-world phenomena should be described” and defines theory as “a particular kind of model that is intended to account for some subset of phenomena in the real world”. However, Weber also offered a slightly different definition of theory in an earlier article: “an account that is intended to explain or predict some phenomena that we perceive in the world” (Weber, 2003).

Weber’s work builds upon an ontology described by Bunge (1977, 1979), which is used to define theory-related concepts. The key assumptions, as described by Weber (2003), can be summarized as follows:

• The world is perceived as a collection of “things” and “properties of things”.

• A state is the values associated with the various properties at a particular time and space.

• Events occur that can result in a change of state.

• Phenomena are defined as states of things or events that occur to things.

Weber (2003) takes the view that “the choice and articulation of the phenomena we are seeking to explain or predict via our theories are the two most-critical tasks we undertake as researchers.” A role of a theory is to express “laws” that relate various values of a state. Weber (2003) defines the “account of the phenomena” as “the explanation of the laws that are hypothesized to relate them” and normally uses “constructs,” a property of a thing, and association among constructs (a law).

Weber (2012) then introduces the following parts of a theory:

• Constructs: represent an attribute (the way we perceive a property)

• Associations: for static phenomena, relate construct values; for dynamic phenomena, relate histories of values between constructs

• States: identification of state space that is the object of the theory – the range of legal values

• Events: identification of the events that are the object of the theory – the range of legal state transitions.

Using these terms, Weber (2012) then discusses how to build a theory:

1. Articulate the constructs of a theory.

2. Articulate the laws of interaction (relationships) among the constructs of a theory.

3. Articulate the lawful state space of a theory.

4. Articulate the lawful event space of a theory.

Although the process is presented linearly, it is important to recognize that theory building is iterative. The process starts with good observations and descriptions, and it improves through inductive/deductive cycles, with anomalies resulting in evolution of the theories. In the early stages of understanding phenomena, it may be necessary to use the theories of other disciplines to first articulate our understandings. As we better comprehend our phenomena, new theories or adapted theories may be developed.

In a similar manner, Sjoberg and colleagues (2008) describe the theory-building enterprise as:

1. Defining the constructs of the theory

2. Defining the propositions of the theory

3. Providing explanations to justify the theory

4. Determining the scope of the theory

5. Testing the theory through empirical research
Assessing Scientific Contributions
Dan Craigen

Measuring Progress in Science

For researchers and funding agencies, it is pertinent to ascertain whether we are making scientific progress: are the scientific contributions meaningful? One key input into such considerations was written by the Committee on Assessing Behavioral and Social Science Research on Aging (Feller & Stern, 2007). Though motivated by research into aging, their characterization of progress transcends the discipline to other scientific endeavours. The committee identified two kinds of progress: i) internally defined (i.e., characterized as intellectual progress and contributions to science), and ii) externally defined (i.e., characterized by contributions to society).

For internally defined progress in science, the committee identified five types of progress:

1. Discovery: demonstration of the existence of previously unknown phenomena or relationships among phenomena, or when the discovery that widely shared understandings of phenomena is wrong or incomplete

2. Analysis: development of concepts, typologies, frameworks of understanding, methods, techniques, or data that make it possible to uncover phenomena or test explanations of them

3. Explanation: discovery of regularities in the ways phenomena change over time or evidence that supports, rules out, or leads to qualifications of possible explanations of these regularities

4. Integration: linking theories or explanations across different domains or levels of organization

5. Development: stimulation of additional research in a field or discipline, including research critical of past conclusions, and when it stimulates research outside the original field, including interdisciplinary research and research on previously under-researched questions. It also develops when it attracts new people to work on an important research problem.

For externally defined progress in science, the committee identified four types of progress:

1. Identifying issues: identifying problems that require societal action or showing that a problem is less serious than previously believed

2. Finding solutions: developing ways to address issues or solve problems

3. Informing choices: providing accurate and compelling information, thus promoting better informed choices

4. Educating society: producing fundamental knowledge and developing frameworks of understanding that are useful for making decisions in the private sector, and participating as citizens in public policy decisions. Science can also contribute by educating the next generation of scientists.

Assessing Theories

Prior to discussing our criteria for assessing contributions to science, we note various criteria that are used to assess theories (Berg, 2009; Cramer, 2013; Sjoberg et al., 2008):

- Testibility; refutability
- Precision
- Empirical validity/support
- Explanatory power; predictability; quantifiable
- Parsimony; consilience; simplicity; self-consistent; rational; inductive
- Generality; comprehensiveness
- Utility; heuristic and applied value
- Repeatability

Weber (2012) uses the ontological structure, briefly discussed above, to evaluate a theory from two perspectives: evaluating the components of a theory and evaluating the whole theory. Weber notes that the components of the theory must be described precisely because they essentially define the domain of the theory. From his perspective, a key advantage of precision is that tests can be better designed.

Weber (2012) evaluates the components of a theory using the following key concepts:

1. Constructs: Should be defined precisely; underlying variables clearly identified
Assessing Scientific Contributions

Dan Craigen

2. **Associations:** Described to various levels of precision. With static phenomena, there is a relationship, but no sign; the sign of association between constructs identified; and a functional relationship is described. With dynamic phenomena, there is a relationship, but no sign or direction; the sign of association between constructs identified but not the direction; the direction of association known (implying causality) or time relationship; and a functional relationship identified.

3. **States:** How clear and precise is the description of the state space?

4. **Events:** How clear and precise are the events?

Weber (2012) evaluates a whole theory using the following key concepts:

1. **Importance:** Does the theory address important phenomena from either a practice or research perspective?

2. **Novelty:** Does it resolve anomalies? Does it change research paradigms?

3. **Parsimony:** Is the theory sufficiently simple?

4. **Level:** Is the theory sufficiently abstract? Weber discusses micro-level and macro-level theories, both of which have associated pros and cons.

5. **Falsifiability:** Can the theory be refuted?

**Assessing Scientific Contributions**

From the above literature review, we synthesize our framework for assessing scientific contributions. There are two aspects to assessing a scientific theory: **Evaluation** and **Contribution.** These two aspects and their components are summarized in Table 1.

**Evaluation** has two constituents: i) **Well-formedness** and ii) **Testing and Analysis.** Broadly speaking, **Evaluation** refers to expectations of how a theory should be expressed and the means through which the scientific and philosophical communities test and analyze theories for acceptance. In large part, evaluation focuses on technical attributes of the theory.

**Contribution** has three constituents: i) **Contribution to Science,** ii) **Contribution to Society,** and iii) **Depth of the Contribution.** The first two constituents align directly with the work by the Committee on Assessing Behavioral and Social Science Research on Aging (Feller & Stern, 2007) in that the **Contribution to Science** aligns to a subset of internally defined progress, while **Contribution to Society** aligns to a modified subset of externally defined progress. In large part, contribution focuses on social attributes of the theory – its role within scientific and societal communities.

**Evaluation:** **Well-formedness**
In the framework illustrated in Table 1, we identify six attributes to determine if a theory is well-formed:

1. **Components:** Evaluation was discussed by (Weber, 2012) as:

   - **Components**
   - **Precision** (Formalism)
   - **Consistency**
   - **Completeness**
   - **Measurability**
   - **Testability**

   **Falsifiability**
   - **Accuracy**
   - **Validity**
   - **Consilience**

   **Discovery**
   - **Analysis**
   - **Explanation**

   **Identifying issues**
   - **Finding solutions**
   - **Making educated choices**

   **Generality**
   - **Comprehensiveness**
   - **Non-obvious results/observations**

   **Novelty**

**Table 1.** Proposed framework for assessing a scientific theory

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Contribution to Science</th>
<th>Contribution to Society</th>
<th>Depth of the Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-formedness</td>
<td>Discovery</td>
<td>Identifying issues</td>
<td>Generality</td>
</tr>
<tr>
<td>Testing and Analysis</td>
<td>Analysis</td>
<td>Finding solutions</td>
<td>Comprehensiveness</td>
</tr>
<tr>
<td></td>
<td>Explanation</td>
<td>Making educated choices</td>
<td>Non-obvious results/observations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Novelty</td>
</tr>
</tbody>
</table>
Assessing Scientific Contributions
Dan Craigen

2012), as summarized earlier in this article. We expect each of these components to be present.

2. Precision (Formalism): Consistent with Weber, we argue that the components of a theory should be described as precisely as possible. Although natural languages are often used in stylized manners to describe concepts “precisely”, the "gold standard" is to describe the components formally using mathematical concepts.

3. Consistency: The expression of the theory should be internally consistent; that is, there are no contradictions.

4. Completeness: In our context, we view completeness from an “expressively complete” perspective in which the theory can describe all of the properties for which it has been developed.

5. Measurability: It should be possible to objectively measure the theory components, particularly the constructs. Key concepts must be quantifiable and the measurements must be objective.

6. Testability: The theory components should be amenable to scientific experimentation. This attribute is closely related to both the measurable attribute above and the falsifiable attribute described below.

Evaluation: Testing and Analysis
In Table 1, we identify five attributes for the evaluation of testing and analysis:

1. Falsifiability: A key attribute/principle of science – it must be possible to show that the theory is incompatible with possible empirical observations.

2. Accuracy: The empirical observations should be in line with the expectations of the theory.

3. Repeatability: The empirical observations should be reproducible.

4. Consilience: Evidence from independent, unrelated sources can “converge” to strong conclusions.

5. Parsimony: Measures the number of kinds of entities postulated by a theory; theories should be as simple as possible for the phenomena being modelled.

Each of these attributes is testing or analyzing the theory and mostly relate to empirical validation. The first four specifically speak to experiments: Can we fail? Are the experimental results being accurately described or predicted by the theory? Can we repeat the experiment and obtain the same results? Can we obtain the same results by different experimental means? If all of these conditions hold, it then makes sense to ask ourselves whether we have elegance in our theory. Have we truly identified the core relationships and constructs?

Contributions to Science and to Society
The elements Contribution to Science and Contribution to Society are largely those identified by the Committee on Assessing Behavioral and Social Science Research on Aging (Feller & Stern, 2007). Contribution to Society merges their "Informing Choices" and "Education" into Making Educated Choices within the proposed framework. Further, for Contribution to Science, only the first three attributes are included; development and integration can be viewed as attributes of an Evaluation of the Contribution.

As depicted in Table 1, the importance and utility of contributions to science and society are captured in Evaluation of the Contribution:

1. Generality: Is the scientific contribution of specific or general validity?

2. Comprehensiveness: Is the scientific contribution inclusive and of broad scope? Is the scientific contribution inclusive and broadly applicable to societal challenges?

3. Non-obvious results: Are there interesting challenges for scientists to explore? Are there unexpected consequences suggested by the theory when contextualized societally?

4. Novelty: Does the theory provide new insights otherwise not explored by science? Is it normal science or paradigm changing? Does the theory provide new insights otherwise not explored by society?

Measuring Evaluation
Having defined the various evaluation attributes, we posit some potential values for each of the attributes. For simplicity, we define only three values per attribute:

- Well-formedness
  - Components: all components present; some components present; no components
Assessing Scientific Contributions

Dan Craigen

- Precision (Formalism): formal/mathematical; semi-formal; informal
- Consistency: provable consistency; unclear; inconsistent
- Completeness: provable completeness; unclear; incomplete
- Measurability: measurable; unclear; not measurable
- Testability: testable; unclear; not testable
- Testing and analysis
  - Falsifiability: falsifiable; unclear; not-falsifiable
  - Accuracy: accurate; unclear; not-accurate
  - Repeatability: repeatable; unclear; not-repeatable
  - Consilience: consilient; unclear; not-consilient
  - Parsimony: parsimonious; unclear; complex
- Depth of the contributions
  - Generality: general; generalized; specific
  - Comprehensiveness: comprehensive; moderately comprehensive; narrow
  - Non-obvious results/observations: non-obvious; unclear; uninteresting
  - Novelty: paradigm/society shifting; substantive normal progress; not substantive

Applying the Framework

Having defined the framework, we now apply it to two contributions from the science of cybersecurity. These assessments are preliminary, but are intended to illustrate how the framework could be applied.

Phishing in International Waters
At the 2014 Symposium and Bootcamp on the Science of Security (hot-sos.org/2014/), Tembe and colleagues (2014) presented the paper "Phishing in International Waters", in which they reported on a survey of American, Chinese, and Indian Internet users and explored the role of culture in the three nationalities responses to phishing attacks. The authors performed various statistical analyses based on responses to questionnaires and found that there were cross-national differences in agreement regarding the characteristics of phishing, the media of phishing, and the consequences of phishing. Conclusions were drawn in part from the individualistic culture represented by Americans and the collectivist cultures represented by China and India.

The statistical analyses included multivariate analysis of covariance and logistic regression analysis. According to the paper, a logistic regression was used to compare nationality with phishing and the characteristics of the risk profile. Further, the authors reported that a multivariate analysis of covariance was used to compare nationality with characteristics of phishing, types of media, and the consequences of phishing. Notably, neither age nor education had any influence on the likelihood of being phished.

Table 2 summarizes our analysis of "Phishing in International Waters" using our framework for assessing scientific contributions.

Selective Interleaving Functions
McLean (2014) presented one of the keynote presentations at the Science of Security conference (HOTSoS, 2014). His presentation, "The Science of Security: Perspectives and Prospects", provided two case studies: one on access control models and the second on information flow models. Here, we assess the scientific contribution of the second case study using our proposed framework. In this second case, McLean examined the evolution of information-flow models and how our understanding in this area has improved over time and has resulted in a compelling framework that could be used to explain information flow models. Table 3 summarizes our analysis of portion of his paper on "Selective Interleaving Functions" and his related earlier paper (McLean, 1994).

Contribution

In this article, we have presented a framework for assessing scientific contributions to cybersecurity and then applied the framework to two contributions to the Science of Cybersecurity. Our assessment framework consists of two parts: Evaluation and Contribution. Through these two parts, we have synthesized and structured a number of approaches cited in the literature for assessing scientific contributions. Prior work, such as that of Weber and the Committee on Assessing Behavioral and Social Science Research on Aging has focused on one part solely (either evaluation or contribu-
Assessing Scientific Contributions
Dan Craigen

Table 2. Assessing the scientific contribution of "Phishing in International Waters" (Tembe et al., 2014) using the proposed framework

<table>
<thead>
<tr>
<th>Well-formedness</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Components</td>
<td>Some components present</td>
</tr>
<tr>
<td>Precision (Formalism)</td>
<td>Semi-formal</td>
</tr>
<tr>
<td>Consistency</td>
<td>Unclear</td>
</tr>
<tr>
<td>Completeness</td>
<td>Incomplete</td>
</tr>
<tr>
<td>Measurability</td>
<td>Measurable</td>
</tr>
<tr>
<td>Testability</td>
<td>Testable</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Testing and Analysis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Falsifiability</td>
<td>Falsifiable</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Unclear</td>
</tr>
<tr>
<td>Repeatability</td>
<td>Unclear</td>
</tr>
<tr>
<td>Consilience</td>
<td>Unclear</td>
</tr>
<tr>
<td>Parsimony</td>
<td>Unclear</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contributions to Science and Society</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>To Science</td>
<td>Explanation</td>
</tr>
<tr>
<td>To Society</td>
<td>Making educated choices</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Depth of the Contributions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Generality</td>
<td>Specific</td>
</tr>
<tr>
<td>Comprehensiveness</td>
<td>Narrow</td>
</tr>
<tr>
<td>Non-obvious results/observations</td>
<td>Unclear</td>
</tr>
<tr>
<td>Novelty</td>
<td>Not substantive</td>
</tr>
</tbody>
</table>
Assessing Scientific Contributions
Dan Craigen

Table 3. Assessing the scientific contribution of "Selective Interleaving Functions" (McLean, 2014) using the proposed framework

<table>
<thead>
<tr>
<th>Well-formedness</th>
<th>Components</th>
<th>All components present</th>
<th>Not evaluated here</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision (Formalism)</td>
<td>Formal/mathematical</td>
<td>Not evaluated here</td>
<td></td>
</tr>
<tr>
<td>Consistency</td>
<td>Consistent</td>
<td>See McLean (1994) for mathematical presentation</td>
<td></td>
</tr>
<tr>
<td>Completeness</td>
<td>Unclear</td>
<td>Not evaluated here</td>
<td></td>
</tr>
<tr>
<td>Measurability</td>
<td>Measurable</td>
<td>Systems could be implemented to demonstrate whether, for example, composition claims hold.</td>
<td></td>
</tr>
<tr>
<td>Testability</td>
<td>Testable</td>
<td>Systems could be implemented to demonstrate whether, for example, composition claims hold.</td>
<td></td>
</tr>
</tbody>
</table>

Testing and Analysis

<table>
<thead>
<tr>
<th>Falsifiability</th>
<th>Falsifiable</th>
<th>Systems could be implemented to demonstrate whether, for example, composition claims hold.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>Unclear</td>
<td>Papers cited within are theoretical contributions. While testable and measurable, it is unclear whether any experiments have been performed.</td>
</tr>
<tr>
<td>Repeatability</td>
<td>Unclear</td>
<td>Papers cited within are theoretical contributions. While testable and measurable it is unclear whether any experiments have been performed.</td>
</tr>
<tr>
<td>Consilience</td>
<td>Unclear</td>
<td>Papers cited within are theoretical contributions. While testable and measurable it is unclear whether any experiments have been performed.</td>
</tr>
<tr>
<td>Parsimony</td>
<td>Parsimonious</td>
<td>Selective interleaving functions are effective in unifying information flow models.</td>
</tr>
</tbody>
</table>

Contributions to Science and Society

| To Science | Discovery | McLean argued that selective interleaving functions provided a common framework for an otherwise incomparable collection of information-flow security models. Further benefits arose because they explained why certain types of compositions were harder on security than others. |
| To Society | Finding solutions | Selective interleaving functions provided an overall characterization of information-flow models and explained difficulties in composability, hence providing information on how systems could be developed with composition in mind. |

Evaluation of the Contributions

| Generality | General | Subsumed prior existing work and a miscellaneous collection of information-flow security models. Provided a common framework. |
| Comprehensiveness | Comprehensive | Within the context of information-flow security models, this approach covers “possibilistic” security properties. |
| Non-obvious results/observations | Non-obvious | While setting a general framework, it appears this framework also can be used to explore other composition properties. |
| Novelty | Substantive normal progress | Built upon prior work on understanding safety/liveness, composition theories, etc. |
Assessing Scientific Contributions

Dan Craigen

Weber provides a significant assessment of an Information Systems paper that can usefully inform how to proceed with theory evaluations. We expand upon Weber’s evaluation by discussing both well-formedness and testing/analyzing criteria a theory more comprehensively.

Particularly, given that development of a Science of Cybersecurity will be a multi-decade exercise, being able to measure progress and contributions, at least incrementally, will provide important objective input into both research and funding decisions and is expected to contribute to a focused research program and accelerate the growth of the science.

Conclusion

The assessment framework presented in this article is preliminary. Specifically, whether the values for each criterion are sensible and whether there should be additional criteria is open for refinement. Weber (2003, 2012) uses an ontological framework to motivate his analysis; future work should build upon these ontological considerations.

Moreover, this type of work can be used to assess “scientific progress”. For example, the science of cybersecurity is in its early stages, and it would be beneficial to measure the progress made in the field. Assessing contributions provides potentially rational inputs into the determination of scientific progress and thereby potentially contribute to a focused research program to accelerate the growth of the science.

References


About the Author

Dan Craigen is a Science Advisor at the Communications Security Establishment in Canada. Previously, he was President of ORA Canada, a company that focused on High Assurance/Formal Methods and distributed its technology to over 60 countries. His research interests include formal methods, the science of cybersecurity, and technology transfer. He was the chair of two NATO research task groups pertaining to validation, verification, and certification of embedded systems and high-assurance technologies. He received his BScH and MSc degrees in Mathematics from Carleton University in Ottawa, Canada.


Keywords: cybersecurity, science of cybersecurity, scientific progress, scientific contributions, societal contributions, assessing science

www.timreview.ca