Identifying Core Concepts of Cybersecurity: Results of Two Delphi Processes

Geet Parekh, David DeLatte, Geoffrey L. Herman, *Member, IEEE*, Linda Oliva, Dhananjay Phatak, Travis Scheponik, and Alan T. Sherman, *Senior Member, IEEE*

***Abstract*—This paper presents and analyzes results of two Delphi processes that polled cybersecurity experts to rate cybersecurity topics based on importance, difficulty, and timelessness. These ratings can be used to identify core concepts–cross-cutting ideas that connect knowledge in the discipline. The first Delphi process identified core concepts that should be learned in any first course on cybersecurity. The second identified core concepts that any cybersecurity professional should know upon graduating from college. Despite the rapidly growing demand for cybersecurity professionals, it is not clear what defines foundational cybersecurity knowledge. Initial data from the Delphi processes lay a foundation for defining the core concepts of the field and, consequently, provide a common starting point to accelerate the development of rigorous cybersecurity education practices. These results provide a foundation for developing evidence-based educational cybersecurity assessment tools that will identify and measure effective methods for teaching cybersecurity. The Delphi results can also be used to inform the development of curricula, learning exercises, and other educational materials and policies.**

***Index Terms*—Concept inventory, conceptual learning, cybersecurity, cybersecurity assessment tools (CATS), Delphi process, information assurance, student assessment, assessment tools.**

I. Introduction

CYBERSECURITY is a vital area of growing importance for national competitiveness, yet there is a lack of conceptual clarity and consensus about what it is and how it should be taught [1]. This project conducted two Delphi processes to identify the core concepts of cybersecurity.

The aim of these processes is to use expert ratings of cybersecurity topics to identify "core concepts." Concepts cut across topics, creating a unifying structure of knowledge upon which students build their knowledge [2], [3]. For example, mechanics courses in physics are organized around the concepts of *force* and *energy* to inform context-bound topics such as boxes sliding down inclined planes [3], [4]. Because cybersecurity is a rapidly evolving discipline, the criteria of "timelessness" may help identify which concepts are core and which are relevant because of current technology. Similarly, when identifying core concepts for the purpose of education, it is prudent also to identify topics that are difficult, since those topics may provide the greatest barriers to mastery [5], [6].

These Delphi processes lay a foundation necessary for developing educational cybersecurity assessment tools that will provide rigorous evidence-based infrastructure to advise educators about effective ways to engage, inform, educate, nurture, and retain cybersecurity students, as well as effective ways to structure cybersecurity curricula to prepare professionals for careers in this field [6]. The numerical ratings from these Delphi processes provide a resource for prioritizing concepts and content in developing curricula, learning exercises, and other educational materials and policies.

Cybersecurity lies at the confluence of several disciplines, including computer science, engineering, information systems, networks, cryptology, human factors, and policy [1]. To identify its core concepts, the large number and variety of potential topics motivates a selection process that incorporates multiple expert perspectives and systematically distills the results. This paper presents and analyzes results from a pair of Delphi processes that were carried out in fall 2014 to identify core concepts for cybersecurity.

A Delphi process solicits input from a set of subject matter experts to create consensus about contentious decisions [5], [7]. Topics are refined and prioritized over several rounds, where participants share comments without attribution so that the logic of a contributed remark is most significant [7].

These Delphi processes are part of a larger project: educational Cybersecurity Assessment Tools (CATS - http://www.cisa.umbc.edu/cats/index.html). This larger project aims to create a Cybersecurity Concept Inventory (CCI) and a Cybersecurity Curriculum Assessment (CCA). The CCI is for students completing any first course in cybersecurity; CCA is for students graduating from college about to enter the workforce as cybersecurity professionals. Accordingly, the project completed two separate Delphi processes in parallel, one for CCI, and one for CCA.

Manuscript received May 31, 2016; revised January 16, 2017 and May 12, 2017; accepted May 25, 2017. This work was supported in part by the U.S. Department of Defense under CAE-R Grant H98230-15-1-0294 and Grant H98230-15-1-0273, in part by the National Science Foundation under Grant SFS 1241576, and in part by the NSF under a subcontract of INSuRE under Grant 1344369. *(Corresponding author: Geoffrey L. Herman.)*

G. Parekh was with the Cyber Defense Lab, University of Maryland, Baltimore County, Baltimore, MD 21250 USA (e-mail: geetparekh@gmail.com).

D. DeLatte, D. Phatak, T. Scheponik, and A. T. Sherman are with the Cyber Defense Lab, University of Maryland, Baltimore County, Baltimore, MD, USA (e-mail: dad@umbc.edu; phatak@umbc.edu; tscepnik@umbc.edu; sherman@umbc.edu).

G. L. Herman is with the Department of Computer Science, University of Illinois at Urbana–Champaign, Urbana, IL 61801 USA (e-mail: glherman@illinois.edu).

L. Oliva is with the Education Department, University of Maryland, Baltimore County, Baltimore, MD, USA (e-mail: oliva@umbc.edu).

Digital Object Identifier 10.1109/TE.2017.2715174

The CCI aims to assess an individual's mastery of a set of important concepts drawn from a minimal common core of any first course in cybersecurity, regardless of the department in which the course is taught. The CCA aims to assess how well a college curriculum has prepared an individual for a career in cybersecurity. The purpose of both CCI and CCA is to assess conceptual understanding, but at different depths of technical knowledge. For example, a CCI question might ask about authentication assuming minimal knowledge, whereas a CCA question about authentication might assume knowledge of certain fundamental facts, technologies, and principles.

The CCI and CCA aim to be relevant in a wide range of educational contexts, from professional training through education of future researchers. Using a minimal common core will enable comparisons of the effectiveness of instruction across institutions while respecting the differing curricular priorities and decisions of each institution.

To the authors' knowledge, these are the first Delphi processes for cybersecurity to identify core concepts. Furthermore, while professional certification exams (such as CISSP [8]) exist, their questions are largely informational and hence do not assess mastery of core concepts.

The contribution of this paper is a numerical rating of the importance and difficulty of concepts in cybersecurity that can guide the design of curriculum and assessment tools.

II. Background and Related Work

This section briefly explains what a Delphi process is and reviews related work that contributes to the discussion of defining core concepts of cybersecurity. Delphi processes are commonly used to identify the core concepts of disciplines when developing educational assessment tools [5], [6].

A. Delphi Process

Addressing a foundational question such as "what are the core concepts of cybersecurity?" by identifying and rating proposed topics can be challenging and contentious. A process adds rigor and reduces bias if it effectively combines the wisdom of a diverse set of experts, generates ideas for relevant topics, and creates an opportunity to collaboratively assess the topics against metrics such as importance and difficulty [7].

The Delphi process, originally developed by the RAND corporation in the 1950s [7], [9]–[11], seeks to build consensus among a group of subject matter experts through a structured process of (1) topic identification, (2) provisional ratings against one or more metrics, (3) negotiations that articulate the reasons behind the ratings that differ significantly, and (4) iteration until convergence to final ratings is achieved. The leaders of the Delphi process orchestrate an anonymous written communication among the panel of experts; this prevents a few of the experts from having excessive influence, as may occur during round-table discussions or face-to-face debates [7], [12]. The experts are asked to give reasons for their answers and those reasons are shared anonymously with the others. The process emphasizes informed judgment [7] using anonymity to focus negotiation on the merit of comments rather than the reputation of the experts [5].

A Delphi process proceeds in multiple rounds. In each round, experts provide numerical ratings (i.e., 1-to-10) on a criterion of interest (e.g., the importance or difficulty of a topic for a field) [5]. The Delphi leaders compute statistics such as median and interquartile range (the range between the 25^{th} and 75^{th} percentiles) of these ratings after each round, and share these data with the experts to elicit thoughts and induce consensus building [5]. If an expert disagrees with the majority, he or she is given an opportunity to sway the consensus with anonymous, written comments. The Delphi process terminates when consensus is achieved or after a fixed number of rounds [5], [7].

B. Concurrent and Related Work

To the authors' knowledge, there has been no previous Delphi process to identify the core concepts of cybersecurity. A concurrent project at Purdue University, led by Melissa Dark and Jenny Daugherty, is conducting focus groups to identify fundamental topics in cybersecurity, for the purpose of developing educational modules [13].

The NICE Framework [14] established a common lexicon to define the activities of cybersecurity professionals. In 2013, the IEEE/ACM [15] proposed content areas to be included in cybersecurity curricula. While these frameworks provide a list of topics and concepts that could be targeted during a cybersecurity curriculum, they do not provide a numerical rating system that can guide priorities in instruction and assessment. Also, in contrast with these frameworks, the list of concepts generated by this Delphi process is not intended to be exhaustive. There will almost certainly be topics and concepts included in NICE or ACM that will be excluded by the Delphi process. These topics may be critical to particular sub-fields of cybersecurity, but are likely not part of the minimal common core that all cybersecurity students will need to know.

Professional certification exams, such as CISSP [8], define cybersecurity topics, but these exams are largely informational and not conceptual.

In formulating its Cybersecurity National Action Plan (CNAP) [16], the White House identified important areas of cybersecurity. Relatedly, Mozilla [17]–[19] conducted a Delphi process to "identify consensus on areas of cybersecurity policy" towards improving Internet security from a global perspective. Two additional Delphi processes explored other aspects of cybersecurity: Davidson and Hasledalen [17] investigated cyber threats to online education, and Pruitt-Mentle [20] investigated research priorities in cyberawareness. By contrast, the CATS project seeks to create conceptual clarity about the core cybersecurity topics.

Inspiration for the CATS project grew from a 2014 NSF cybersecurity education workshop [1].

III. Methods

In fall 2014, this project carried out two Delphi processes aimed at identifying core cybersecurity topics. Selection of experts, topic identification, and the methods used to collect and assess topic ratings are described below. The CCI and

CCA Delphi processes were conducted electronically, through emails between Delphi leaders and the panel of experts, and through Web forms to collect survey data. The project website [21] lists all questions and instructions posed to the experts.

A. Cybersecurity Experts

Delphi leaders selected experts based on their education, background, and profession. Prospective experts responded to calls for participation announced at conferences (e.g., CISSE), through email solicitations, and by word-of-mouth.

The selected experts constitute a diverse group of men and women from over a dozen U.S. states and from Canada, working as cybersecurity authors, educators, and professionals from industry and government. Each expert holds a Ph.D. in a cybersecurity-related field and either teaches cybersecurity or works as a cybersecurity professional. Twenty-six experts are faculty: 16 at research-focused universities, seven at teaching-focused colleges, and three at community colleges. Five work in industry, and five in government. The project website [21] lists the experts and their affiliations. The Delphi leaders did not accept potential experts who were new to their job or who were graduate students straight from college.

A total of 36 experts participated in the topic generation phase, including 33 for the CCI and 31 for the CCA. A total of 29 experts participated in both processes. For each of the two Delphi processes, approximately 20 experts sustained their support through the entire process of rating topics. This number of experts exceeds the minimum number of 15 recommended for educational Delphi studies [22].

B. Delphi Rounds

Both CCI and CCA began with topic identification followed by three rounds of topic rating and Delphi-leader-mediated sharing of expert comments. Here a "round" consists of a communication from Delphi leaders to the experts and the collection and aggregation of the responses from experts. Within the scope of the Delphi process no direct communication between experts occurred and all shared comments were anonymous. Delphi leaders created Web forms hosted by SurveyMonkey to solicit expert comments and ratings. The University of Maryland, Baltimore County's IRB office approved the research protocols.

The CCI Delphi process took place October 21 through December 18, 2014; the CCA process took place October 21 through December 9.

C. Topic Identification

In the first round responses of the CCI and CCA Delphi processes, experts listed ten cybersecurity topics as important, difficult, and/or timeless. Using principles of grounded theory [23], Delphi leaders grouped similar responses to produce a shorter reconciled list that included all topics mentioned by at least two experts. The results of this reconciled list are member checked by the Delphi experts who can verify whether their opinions were represented in the final list.

Responses from the first rounds of CCI and CCA were unexpectedly similar, although adversarial thinking was a prevalent theme among CCI responses. To ensure that CCI was headed in a distinct direction from CCA, a second topic identification round was performed for CCI only. Delphi leaders asked participants to provide topics focused on adversarial thinking, which the Delphi leaders and the experts felt constitutes a vital core of cybersecurity. The restarted CCI process produced 30 topics. The next CCI round included a supplemental question asking the experts to propose additional topics. Eight new topics were added, resulting in a list of 38 that would be rated in subsequent CCI rounds (see Table I). Meanwhile, CCA Round 1 produced a total of fifty-three topics (see Table II).

Adversarial Thinking: involves reasoning about actions and goals in a context in which bad actors might be attempting to defeat those goals and carry out their own nefarious actions. Such reasoning requires an understanding of the goal requirements, and of the bad actors and their objectives, resources, access, capabilities, knowledge, motivations, and risk tolerance. In the CCI Delphi process, experts identified what they considered to be the important topics of adversarial thinking.

To help the CCI experts identify specific important topics of adversarial thinking, and because the first CCI round produced responses similar to those in CCA, the Delphi experts specifically encouraged the experts to identify important tasks.

As was described earlier, concepts are the abstract, underlying structure that connects knowledge within a domain [3]. Conceptual knowledge is often tacit, meaning that experts use conceptual knowledge but are unaware that they are using it [4]. Consequently, it is common that when asked to identify core concepts, experts will describe topics (a broad term that includes concepts, skills, and applications) that have those core concepts embedded within them [5]. Consequently, throughout this paper, although the project seeks to identify core concepts, the experts are described as rating topics rather than concepts. Identifying core concepts is an interpretive step based on the expert topic ratings [4].

D. First Round of Topic Rating

Experts rated CCI and CCA topics according to three distinct metrics: (1) Importance, (2) Difficulty, and (3) Timelessness using a 1-10 Likert-type scale [24]. Once the deadline for the round passed, the Delphi leaders compiled summary statistics for each topic. Delphi data can be considered as ordinal data [25]. Consequently, to analyze the data, the team applied nonparametric statistics [26], including medians and interquartile ranges (as opposed to sample means and standard deviations) because there are no strong reasons to believe that the absolute score numbers (as opposed to their relative values) had strong meanings in the experts' minds.

These descriptive statistics and data visualization provided the Delphi leaders with information about the level of consensus (i.e., deviation from the median [27]). Consensus is high for a topic when experts tend to give the same numerical rating. The interquartile range is an appropriate estimator of

TABLE I
FINAL LIST OF RECONCILED CCI TOPICS SORTED BY MEDIAN IMPORTANCE (I) AND THEN BY MEDIAN DIFFICULTY (D) AFTER THE THIRD TOPIC RATING ROUND

	Topic	I	D		Topic	I	D
1	Identify vulnerabilities and failures	9	8	20	Technology vs Policy	7	7
2	Identify attacks against CIA triad and authentication	9	8	21	Assess the risk of acting and of not acting	7	7
3	Devise a defense	9	7	22	Given a policy, devise a way to evade it	7	7
4	Identify the security goals	9	6	23	Assess the difficulty of various attacks	7	7
5	Identify potential targets and attackers	9	5	24	Rank a set of possible corrective actions	7	7
6	Devise an attack	8	8	25	Assess the risks for two different types of users	7	7
7	Given a breach, explain how to recover from it	8	8	26	Rank a set of vulnerabilities	7	7
8	Explain why a failure happened	8	7	27	Devise attacks that exploit the role of actors and information outside of the system	7	7
9	Identify risky behaviors	8	7	28	Identify and classify vulnerabilities by categories	7	6
10	Identify vulnerabilities based on usability issues	8	7	29	Identify a vulnerability	6	9
11	Identify which assumptions of a system are most likely to be exploitable	8	7	30	Identify a vulnerability in software	6	8
12	Given two security solutions, compare their pros and cons	8	7	31	Explain how to exploit a software vulnerability	6	8
13	Devise a social engineering attack	8	5	32	Solve a puzzle requiring “out-of-the-box” thinking	6	8
14	Identify new vulnerabilities caused by a change	7	8	33	Explain how to exploit traffic analysis	6	7
15	Identify vulnerabilities based on gaps between theory and practice	7	8	34	Identify ways to influence people	6	5
16	List assumptions that a system makes implicitly	7	8	35	Identify possible phishing emails from a set of samples	6	4
17	Devise a security plan	7	7	36	Devise an attack that analysts can't identify	5	10
18	Identify vulnerabilities caused by a faulty functionality or incorrect assumption	7	7	37	Given a multi-party protocol, identify vulnerabilities based on people cheating	5	8
19	Rank the relative risks of certain possible actions	7	7	38	Given a malware example, characterize its behavior	5	8

consensus. Reduction of interquartile range from one round to the next indicates convergence toward consensus.

Delphi leaders provided the following guidance to improve consistency in the way experts used the numerical ratings.

For Importance:

10 – Absolutely essential; leaving this topic out would be egregious, and topic is appropriate for the target.

7 – Foundational but some perspectives may not find this topic as important. Alternatively, topic is important but may be too advanced for the target.

4 – Could be important to some perspectives, but not generally foundational. Alternatively, topic is important but likely too advanced for the target.

1 – Topic is too trivial or too advanced for the target.

For Difficulty:

10 – Few, if any, students will have mastered this topic after the target course or curriculum.

7 – The best students will understand this topic, but many students will struggle to understand it.

4 – Weak students will struggle to understand this topic, but most students will understand it.

1 – Prerequisite knowledge that students should already have mastered.

For Timelessness:

10 – Foundational and highly relevant across essentially all technologies throughout the foreseeable future.

7 – Quite important through the near future, but somewhat dependent on current technology.

4 – Somewhat important today but unlikely to offer a long-lasting principle, and dependent on current technology.

1 – Will likely soon become irrelevant; highly dependent on current technology.

E. Second Round of Topic Rating

For the second round, the summary statistics for each topic were provided to the experts as data to inform their subsequent rating and to promote consensus [5]. If an expert chose to rate a topic outside the interquartile range, they were asked to provide a written justification for their deviation from the consensus. These comments enabled dissenting experts to sway the majority. Once the deadline for the round passed, the Delphi leaders compiled summary statistics (median and interquartile range) and written comments for each topic.

F. Third Round of Topic Rating

For the third round, the summary statistics and dissenting comments for each topic were provided to the experts as data to inform their subsequent rating. If an expert chose to rate a topic outside the interquartile range, they were asked to provide a written justification for their deviation from the consensus.

TABLE II
FINAL LIST OF RECONCILED CCA TOPICS SORTED BY MEDIAN IMPORTANCE (I) AND THEN BY MEDIAN DIFFICULTY (D) AFTER THE THIRD TOPIC RATING ROUND

	Topic	I	D		Topic	I	D
1	Privacy	10	7	28	Well-known attacks, such as man-in-the-middle	8	6
2	Ethics	10	5	29	Apply symmetric and asymmetric encryption	8	6
3	Authentication	10	4	30	Operational security	8	6
4	Integrity	10	4	31	Legal aspects	8	6
5	Confidentiality	10	3	32	Economic aspects of cybersecurity	8	6
6	Secure coding	9	8	33	Countermeasures	8	5
7	Assess vulnerabilities	9	7	34	Collaboration skills	8	5
8	Analyze threats	9	7	35	Design secure protocols	7	9
9	Manage risks	9	7	36	Malware analysis	7	8
10	Operating system security	9	7	37	Perform security assessments	7	7
11	Assured operations	9	6	38	Select and apply appropriate cryptographic primitives	7	7
12	Trust, including rooting trust in hardware	9	6	39	Wireless security	7	7
13	Communication skills	9	6	40	Penetration testing	7	7
14	Ability and desire to keep up-to-date	9	6	41	Virtualization and cloud security	7	7
15	Social engineering	9	5	42	Scripting languages, systems programming, low-level programming	7	7
16	Insider threat	9	5	43	Incident analysis	7	6
17	Access control	9	5	44	Design & analyze secure web applications	7	6
18	Forensics	8	8	45	Response & recovery	7	6
19	Design & analyze secure networks	8	8	46	Formulate and evaluate security policies	7	6
20	Adversarial modeling	8	7	47	International aspects of cybersecurity	7	6
21	Attention to detail	8	7	48	Secure development lifecycle	7	5
22	Manage keys	8	7	49	Auditing	7	5
23	Cyberphysical systems	8	7	50	Ability to identify and apply best practices	7	5
24	Software vulnerability analysis	8	7	51	Ability to identify and use modern tools	7	4
25	Usable security	8	7	52	Applications of homomorphic encryption and private information retrieval	5	9
26	Balance competing objectives	8	7	53	Zero-knowledge protocols	4	8
27	Healthy skepticism and paranoia	8	6				

G. Additional Survey Items

In the topic identification round and first topic rating round, the Delphi leaders asked the experts to answer additional directed questions on a four-point Likert scale (strongly disagree, disagree, agree, strongly agree). These questions were asked to clarify the structure and content of the proposed CCI and CCA. Two of these questions address the inclusion or exclusion of topics in the Delphi processes and are described for completeness.

For CCI (first topic rating round), experts were asked whether they thought a well-crafted concept inventory based on adversarial thinking would be predictive of a student's performance in other cybersecurity courses or in the profession. This survey question was used to verify the deeper focus on adversarial thinking for the CCI.

For both CCI and CCA, multiple experts identified ethics and communication skills as important topics. Since these topics are not exclusive to cybersecurity, the experts were polled during the first topic rating round about whether ethics and communications skills are best addressed in separate assessment tools. This question was used to verify the exclusion of these skills from the CCI Delphi process and to inform decisions about whether to include them in the CCA or other subsequent assessment tools.

IV. RESULTS

This section presents the results of the CCI and CCA Delphi processes, summarized in Tables I and II (lists of reconciled topics sorted by importance) and the associated Figs. 1 and 2 (two-dimensional scatter plots of reconciled topics by importance and difficulty). This section also summarizes responses of the experts to selected additional survey questions. The project website and Parekh [28] present more complete data through additional tables and figures.

A. CCI Results

Table I lists the final reconciled CCI topics sorted by the final round median importance rating. Correspondingly, Fig. 1 is a two-dimensional scatter plot of Table I's topics, plotting each topic's median difficulty vs. its median importance.

The authors decided not to include timelessness in the figure because it is highly correlated with importance (the non-parametric Spearman Rank Correlation Test gives a high correlation of 0.67). In contrast, importance and difficulty are somewhat anti-correlated, with value -0.29. Of all the final round individual topic ratings for importance and timelessness

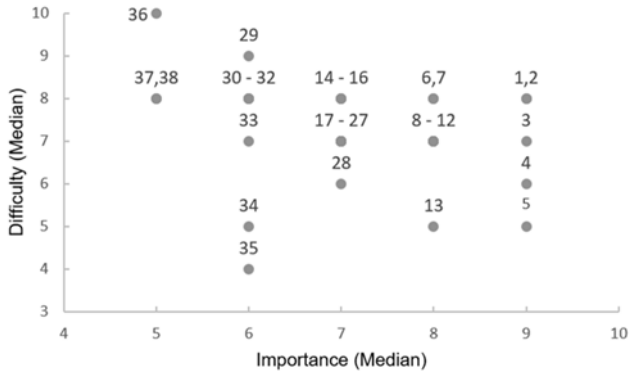

Fig. 1. Final CCI topic ratings: importance vs. difficulty. Five topics were rated nine for importance. Topic numbers refer to those in Table I.

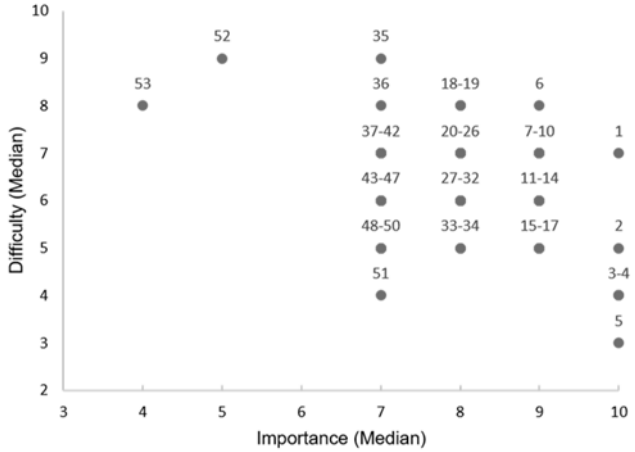

Fig. 2. Final CCA topic ratings: importance vs. difficulty. Five topics were rated 10 for importance. Topic numbers refer to those in Table II.

(taken over all topics and all experts), 80% are within one point and 54% are identical.

In Topic 2, the phrase "CIA triad" refers to Confidentiality, Integrity (broadly interpreted to include authentication), and Availability [29].

B. CCA Results

Table II lists the final reconciled CCA topics from Round 4 sorted by median importance. Correspondingly, Fig. 2 is a two-dimensional scatter plot of these topics, plotting each topic's median difficulty vs. its median importance.

As for CCI, the authors decided not to include timelessness in Fig. 2 or Table II because it is highly correlated with importance (Spearman Rank Correlation of 0.86). Of all the final round individual topic ratings for importance and timelessness (taken over all topics and all experts), 87% are within one point and 58% are identical.

C. Additional Survey Items

For CCI (Round 2), twelve out of fifteen experts (80%) supported the focus on adversarial thinking for the CCI. For the CCI and CCA (Round 2), ten out of 15 experts (67%) in each group supported excluding ethics and communication

Fig. 3. Convergence of top thirteen importance scores (topics with ratings of 9 or 8) from the CCI over the three topic rating rounds. Small and decreasing interquartile ranges show increasing consensus. Topic numbers refer to those in Table I.

skills from the reconciled lists of topics and from subsequent assessment tools.

V. Discussion

This section analyzes the CCI and CCA results, reflects on the Delphi consensuses, discusses implications, summarizes lessons learned, and outlines the project's future work.

A. Analysis of CCI Results

Because the project seeks to identify the core concepts of cybersecurity, Fig. 1 and Table I are sorted by median importance. Five CCI topics have median importance ratings of 9; these are the most important topics identified by the Delphi experts. Furthermore, Fig. 1 shows no clear clustering of topics by a combination of importance and difficulty. By contrast, some other Delphi studies sort topics by a metric that combines importance and difficulty, such as Euclidean distance [5].

Fig. 3 visualizes expert convergence of importance scores for the CCI topics in the top two score groups for importance. Topics added after the initial topic identification round (e.g., 11 and 12) were scored only twice. As quantified by interquartile ranges, this figure illustrates the small interquartile ranges in both of the last two rating rounds as well as an overall trend toward consensus: for each topic, the final interquartile range is no larger (and typically smaller) than the initial range [30]. Thus, the Delphi process is producing its intended result; 37 out of 38 topics have a final interquartile range less than two; one has an interquartile range of three. An interquartile range of two or smaller is also generally deemed as evidence of consensus [30]. Showing all thirty-eight topics makes an unwieldy figure, and it is natural to focus on the topics in the highest score groups.

The topics rated more highly on importance tend to be the relatively more abstract topics that encompass many of the more specific topics. For example, CCI Topic 1 (identify vulnerabilities and failures) generalizes Topics 10, 14, 15, 18, 30, 36, 37, which deal with particular types of vulnerabilities, and Topic 26 is to rank vulnerabilities. In particular, each of the top five topics supports the CIA triad (see Section IV-A). The abstract nature of these topics points to a potential underlying conceptual structure that comprises concepts such as

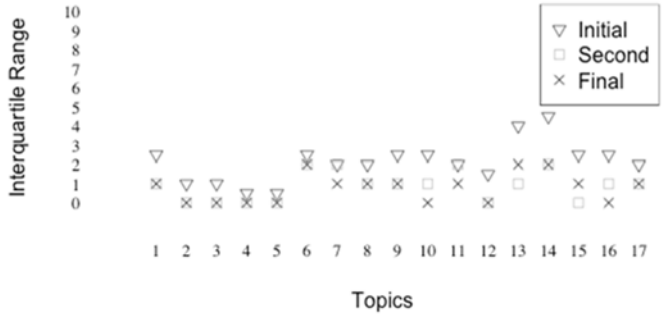

Fig. 4. Convergence of top seventeen importance scores (topics with final rating of 10 or 9) from the CCA over the three topic rating rounds. Small and decreasing interquartile ranges show increasing consensus. Topic numbers refer to those in Table II.

vulnerabilities, the CIA triad, and authentication. For example, a student who can think rigorously about the CIA triad would likely be able to identify vulnerabilities and potential attackers. These core concepts can be used to limit the scope of the CCI [5].

As noted in Section III-C, the framing of the CCI topics in terms of tasks is largely a consequence of instructions from the Delphi leaders. This perspective can be viewed as identifying important security properties of cyber systems, whereas the perspective that emerged from the CCA results can be seen more as identifying important cyber domains of security issues.

B. Analysis of CCA Results

Five topics in CCA have median importance ratings of 10. These compose important cross-cutting concepts of cybersecurity. For the same reasons stated for CCI, the CCA topics are again sorted only by importance. Two topics (zero-knowledge protocols and homomorphic encryption) stand out as low-importance, high-difficulty outliers (see Section V-C).

Fig. 4 visualizes expert convergence of importance scores for the CCA topics in the top two score groups for importance. As measured by interquartile ranges, this figure illustrates a trend toward consensus and strong consensus in the final two rating rounds. As for CCI, for each topic, the final interquartile range is no larger (and typically smaller) than the initial range. Of the 53 CCA topics, 52 have a final interquartile range of two or less.

It is notable that the experts considered privacy and ethics to be among the five most important topics, with privacy also receiving high ratings on difficulty. In cybersecurity, practitioners must exercise wisdom and responsibility to use their knowledge and skills appropriately. Privacy, in particular, involves many complex technical, ethical, legal, cultural, social, and national security issues; it is related to but different from confidentiality. For example, under what circumstances, if any, should Apple be compelled to help the FBI retrieve information from the iPhone of a criminal suspect? Nevertheless, for the CATS project, the authors, in agreement with the majority of the experts in our study, feel that ethics is a topic best assessed through a separate and different instrument (see Section IV-C).

As with CCI, the topics rated more highly on importance tend to be relatively more abstract and conceptual (including the CIA triad), encompassing many of the more specific topics. Topics rated lower on importance tend to include more concrete abilities, including more technology-specific ones such as using modern tools. As with the CCI, the most highly rated topics reflect abstract, cross-cutting concepts.

C. Reflections on the Delphi Consensuses

As shown by the CCA Delphi results, cybersecurity is a broad multidisciplinary field encompassing many diverse issues, skills, and topics. Some experts strongly reflected their orientation, advocating for particular topics such as secure programming, cloud security, forensics, or legal aspects. The Delphi process provided a method for harnessing these diverse opinions to distill core concepts from the candidate topics.

Given the controversial nature of identifying core topics for courses and curriculum, it is vital to understand what this consensus implies for future research, curriculum design, and instruction, and also importantly, to understand what this consensus does not imply. These results indicate which topics are important, irrespective of a particular sub-discipline's focus or stakeholder values. Based on this assertion, readers are invited to consider the validity of these findings based on whether they agree or disagree with the assertion that the topics rated as most important are indeed core topics. If these topics are indeed core, then these topics should compose the minimum of what cybersecurity educators should expect students to learn.

In contrast, these ratings do not provide information about the relative importance of topics for specific sub-disciplines or contexts. In other words, it is expected that individuals and entire sub-disciplines would rate specific topics more highly than they are rated in these consensus ratings. Therefore, the lack of inclusion of what a reader may consider to be an important topic does not threaten the validity of the findings. Educators are encouraged to add these other topics to the core to meet the specific learning objectives for their contexts. While these ratings provide insights into the core concepts of a curriculum or course, they are not meant to be a final arbiter dictating an exhaustive list of what topics should be covered in a curriculum or course. Professional judgment and responsiveness to student and stakeholder needs are critical to applying these findings correctly.

For example, even one of the authors feels strongly that the CCI and CCA topics ought to include the principle of "limited capacity"—the idea of restricting the capability of computational devices (to finite state machines) to enable formal analysis of specified security properties, avoiding the curse of undecidability (determining whether a system meets its security specifications is typically undecidable). If a topic, such as "limited capacity," does not appear in the lists, it simply means that the topic cannot be considered core. A potential rating for the topic should not be inferred.

Further, for a topic to be included in the list for rating, at least two experts needed to recommend the topic in their list of top ten topics for cybersecurity. If over 30 experts from diverse sectors and institutions failed to mention a topic twice during

the topic identification phase, it is unlikely that that topic is core, even though that topic may be important or even critical for particular sub-disciplines of cybersecurity. If a person feels that a topic should have been included but was not, it does not mean that the person's opinion is invalid; rather, it means that the person's opinion reflects contextual priorities that may not be equally important for other contexts. For example, no expert mentioned the topic of "limited capacity," but limited capacity may be a vital concept in some curricula and courses.

As for some intriguing topics, such as zero-knowledge protocols and homomorphic encryption, the experts considered them esoteric, unimportant, and highly difficult. Experts can often misjudge the generalized difficulty of topics, so the difficulty ratings of the topics should be taken with less certainty than the importance ratings [31], [32]. The authors conjecture that in time these topics will eventually be seen as very important and not overly difficult.

Given the range of experts' experiences in academia, industry, and government, and given that we maintained sufficient sample sizes throughout the process, the authors assert that the results are reasonable, confirming strongly-held opinions by many that adversarial thinking and the CIA triad are vital cross-cutting concepts in cybersecurity. The ratings for these core concepts are likely to be stable even with a different panel of experts. However, the stability of ratings would likely go down for topics rated with lower importance. Although there is scientific evidence that Delphi processes tend to be stable [33], only additional studies can confirm the authors' stability hypothesis.

D. Implications

Results from the Delphi processes lay a foundation for improving cybersecurity teaching and learning by helping educators design better assessment tools, learning materials, and curricula. The design of these tools should begin by identifying the core concepts of the discipline before selecting and structuring the presentation of topics around those core concepts [3]. Textbooks should not be a hodgepodge of specialized topics. Indeed, most course design methods suggest that course design begin by identifying a small number of core concepts or big ideas that organize the knowledge in a course and connecting all information and tasks in the course to those core concepts [34], [35]. Professional examinations should go beyond information questions such as, "How many bits long is a DES key?" The Delphi results identify these important and timeless concepts on which learning strategies and materials should flow. Cybersecurity courses should focus more on these concepts and less on particular technologies that may soon become obsolete.

In particular, the five CCI topics rated most important comprise meaningful conceptual activities that ought to be included in any first course in cybersecurity, regardless of whether the course is in a computer science, information systems, or business department. These courses, however, do not have to be identical and should not be. For example, a cybersecurity course in a business school might elucidate core cybersecurity concepts from a business perspective. Unfortunately, many courses and textbooks slight the broad complex topics of privacy and ethics, which are among the five most important topics identified by the Delphi experts.

Although learning activities should focus on important concepts, it is also vital to understand these concepts through concrete, practical, hands-on tasks, which requires choices of particular context and technology. These choices of context and technology are not very important *per se*, but their use in stimulating, facilitating, and measuring learning is very important. Trying to learn abstract adversarial thinking without concrete context and technology might be doomed to failure, as might be trying to master detailed technologies without a guiding conceptual framework of adversarial thinking.

The authors hope these results will be helpful to others; the results, however, are not intended to restrict creative educators from pursuing their unique perspectives.

E. Limitations of the Study

Limitations of the Delphi process include the total number, selection, and attrition of the experts. For CCI, out of the pool of 33 experts, 21, 15, 22, 18, and 18 experts responded to each round of the process, respectively, for a mean of 18.8 experts (57.0%) responding to each round. For CCA, out of the pool of 31 experts, 20, 15, 22, and 17 experts responded per round, respectively, for a mean of 18.5 experts (59.7%) responding per round. These numbers are consistent with similar studies [5] and are above the minimum recommended panel size of 15 [22].

A large number of the experts are from universities (24 (72.7%) for CCI, 21 (67.7%) for CCA). Especially for the purpose of identifying core concepts including those that ought to be learned in any first course on cybersecurity, it is appropriate to have a high representation from university educators. While industry experts are less likely to be familiar with the breakdown of courses in an undergraduate curriculum, it would be interesting to see if a different selection of experts (e.g., with higher representation from industry and government) would produce different results.

Despite some rigorous aspects of the Delphi process and the authors' analysis of it, care should be exercised not to infer unduly high quantitative authoritative weight and specificity to the findings.

F. Lessons Learned

The use of a convenient online questionnaire allowed the experts to answer asynchronously, and ensured that each expert received the same instructions and question wordings. As expected, the Delphi processes took approximately eight weeks to complete. For the Delphi leaders, as expected, the hardest and least well-defined task was topic reconciliation. Interactions among the experts seemed to contribute towards convergence, though the authors had hoped for even more interactions.

G. Future Work

CATs is a four-year project, now in its second year; the project has interviewed twenty-six students to understand how

they reason about cybersecurity concepts, and to uncover misconceptions and problematic reasoning [36]. The team devised interview prompts inspired by the five CCI topics rated most important by the experts. After the team finishes analyzing these interviews, they will create assessment questions. The Delphi experts, and more experts to be recruited throughout the development process, will be asked to review whether drafts of the assessment tool adequately assess the core concepts and achieve the assessment goals. After obtaining expert consensus on the quality of the assessment tools, the assessment tools will be psychometrically analyzed.

Future work could use the list of core concepts to develop concept maps that show how the concepts are interrelated. Future work could also develop and refine strategies for teaching these concepts.

VI. Conclusion

In fall 2014, the project carried out two Delphi processes to identify core concepts of cybersecurity. Tables I and II and associated Figs. 1 and 2 summarize the results. The findings provide a foundation for developing evidence-based, cybersecurity educational assessment tools that will identify and measure effective methods for teaching cybersecurity. They can also help prioritize the development of curricula, learning exercises, other educational materials, and policies involving cybersecurity. Importantly, the results point toward a more promising way of teaching and learning cybersecurity by focusing on the important and timeless concepts identified by the Delphi experts rather than simply trying to cover a hodgepodge of idiosyncratic detailed topics.

The results of the Delphi processes, especially the CCA process, identified a range of specialized topics, reflecting the broad, multi-faceted aspects of cybersecurity. This range of facets can make prioritizing content in cybersecurity education difficult and make it difficult for students to discern how topics connect. The five topics rated most important by the Delphi experts in CCI and in CCA stand out as important and timeless concepts that can create priorities in instruction and help students organize their learning.

In addition, these results help clarify, distill, and articulate what is cybersecurity, which this project sees (as supported by the Delphi processes) as *the management of information and trust in an adversarial cyber world*.

Acknowledgment

The authors thank the experts who participated in the Delphi processes. They also thank W. Byrd, R. Dodge, M. Loui, M. McNeary, and J. Thompson for helpful comments.

References

[1] S. Buck and D. Burley, "Cybersecurity education workshop: Final, report," George Washington Univ. at Arlington Center, Arlington, MA, USA, Tech. Rep., 2014.

[2] P. L. Smith and T. J. Ragan, *Instructional Design*. New York, NY, USA: Wiley, 1999.

[3] S. A. Ambrose, M. W. Bridges, M. DiPietro, M. C. Lovett, and M. K. Norman, *How Learning Works: Seven Research-Based Principles for Smart Teaching*. San Francisco, CA, USA: Jossey-Bass, 2010.

[4] G. L. Herman and M. C. Loui, "Identifying the core conceptual framework of digital logic," in *Proc. Amer. Soc. Eng. Educ. Annu. Conf. Expo.*, San Antonio, TX, USA, 2012, pp. 1–25.

[5] K. Goldman *et al.*, "Setting the scope of concept inventories for introductory computing subjects," *ACM Trans. Comput. Educ.*, vol. 10, no. 2, pp. 1–29, 2010.

[6] R. A. Streveler *et al.*, "Rigorous method for concept inventory development: Using the 'assessment triangle' to develop and test the thermal and transport science concept inventory (TTCI)," *Int. J. Eng. Educ.*, vol. 27, no. 5, pp. 968–984, 2011.

[7] B. Brown, *Delphi Process: A Methodology Used for the Elicitation of Opinions of Experts*. Santa Monica, CA, USA: RAND Corporat., 1968.

[8] *CISSP*. Accessed on May 31, 2016. [Online]. Available: https://www.isc2.org/CISSP/Default.aspx

[9] N. Dalkey and O. Helmer, "An experimental application of the Delphi method to the use of experts," *Manag. Sci.*, vol. 9, no. 3, pp. 458–467, 1963.

[10] O. Helmer, *Analysis of the Future: The Delphi Method*. Santa Monica, CA, USA: RAND, 1967.

[11] M. Scheibe, M. Skutsch, and J. Schofer, "Experiments in Delphi methodology," in *The Delphi Method: Techniques and Applications*, H. A. Linstone and M. Turoff, Eds. Reading, MA, USA: Addison-Wesley, 1975, pp. 262–287.

[12] J. Pill, "The Delphi method: Substance, context, a critique and an annotated bibliography," *Socio Econ. Plan. Sci.*, vol. 5, no. 1, pp. 57–71, 1971.

[13] M. Dark, *Personal Communication*. A. Sherman, 2015.

[14] NICE. (2016). *NICE Framework*. Accessed on May 31, 2016. [Online]. Available: http://csrc.nist.gov/nice/framework

[15] M. Sahami *et al.*, "Computer science curricula 2013 curriculum guidelines for undergraduate degree programs in computer science," Joint Task Force Comput. Curricula Assoc. Comput. Mach., New York, NY, USA, and IEEE Comput. Soc., Washington, DC, USA, Tech. Rep., Dec. 2013.

[16] The White House. (2016). *Cybersecurity National Action Plan*. Accessed on May 31, 2016. [Online]. Available: https://www.whitehouse.gov/the-press-office/2016/02/09/fact-sheet-cybersecurity-national-action-plan

[17] S. Dave. (2014). *Ensuring a MORE SECURE Internet: Reset the Net and the Cyber Security Delphi. Mozilla's Official Blog on Open Internet Policy Initiatives and Developments.* Accessed on May 31, 2016. [Online]. Available: https://the-net-and-the-cyber-security-delphi

[18] Mozilla. (2014). *Netpolicy/Cybersecurity Delphi*. Accessed on May 31, 2016. [Online]. Available: https://wiki.mozilla.org/Netpolicy/Cybersecurity_Delphi

[19] Mozilla. (2015). *Mozilla Cybersecurity Delphi 1.0: Towards a User-Centric Policy Framework Available*. Accessed on May 31, 2016. [Online]. Available: http://blog.mozilla.org/netpolicy/files/2015/07/Mozilla-Cybersecurity-1.0.pdf

[20] D. Pruitt-Mentle, "A Delphi study of research priorities in cyberawareness," in *Proc. 15th Colloq. Inf. Syst. Security Educ.*, Fairborn, OH, USA, 2011, pp. 1–3.

[21] *Cybersecurity Assessment Tools*. Accessed on May 1, 2017. [Online]. Available: Available: http://www.cisa.umbc.edu/cats

[22] M. J. Clayton, "Delphi: A technique to harness expert opinion for critical decision-making tasks in education," *Educ. Psychol.*, vol. 17, no. 4, pp. 373–386, 1997.

[23] B. Glaser and A. Strauss, *The Discovery of Grounded Theory; Strategies for Qualitative Research*. New Brunswick, Canada: Aldine, 1967.

[24] R. Likert, "A technique for the measurement of attitudes," *Archives Psychol.*, vol. 140, pp. 5–55, Jun. 1932.

[25] J. Carifio and J. P. Rocco, "Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes," *J. Soc. Sci.*, vol. 3, no. 3, pp. 106–116, 2007.

[26] A. Agresti, *Analysis of Ordinal Categorical Data*, 2nd ed. Hoboken, NJ, USA: Wiley, 2010.

[27] F. Hasson, S. Keeney, and H. McKenna, "Research guidelines for the Delphi survey technique," *J. Adv. Nursing*, vol. 32, no. 4, pp. 1008–1015, 2000.

[28] G. Parekh, "Identifying significant, difficult and timeless concepts for cybersecurity assessment tools: Results and analysis of two Delphi processes," M.S. thesis, Dept. Comput. Sci., Univ. Maryland at Baltimore County, Baltimore, MD, USA, 2015.

[29] B. Schneier, *Applied Cryptography, Algorithms, and Source Code in C*, 2nd ed. New York, NY, USA: Wiley, 1996.

[30] I. R. Diamond *et al.*, "Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies," *J. Clin. Epidemiol.*, vol. 67, no. 4, pp. 401–409, 2014.

[31] R. Clark, D. Feldon, J. V. Merrienboer, K. Yates, and S. Early, "Cognitive task analysis," in *Handbook of Research on Educational Communications and Technology*, vol. 3. New York, NY, USA: Erlbaum, 2008, pp. 1801–1856.
[32] G. L. Herman, C. Zilles, and M. C. Loui, "A psychometric evaluation of the digital logic concept inventory," *Comput. Sci. Educ.*, vol. 24, no. 4, pp. 277–303, 2014.
[33] R. B. Akins, H. Tolson, and B. R. Cole, "Stability of response characteristics of a Delphi panel: Application of bootstrap data expansion," *Biomed. Central Med. Res. Methodol.*, vol. 5, no. 37, pp. 1–12, 2005.
[34] E. J. Hansen, *Idea-Based Learning: A Course Design Process to Promote Conceptual Understanding*. Sterling, VA, USA: Stylus, 2011.
[35] R. A. Streveler, K. A. Smith, and M. Pilotte, "Aligning Course content, assessment, and delivery: Creating a context for outcome-based education," in *Outcome-Based Science, Technology, Engineering, and Mathematics Education: Innovative Practices*. Hershey, PA, USA: IGI Global, 2012, pp. 1–26.
[36] T. Scheponik *et al.*, "How students reason about cybersecurity concepts," presented at the IEEE Frontiers in Education Conf., Erie, PA, USA, 2016, pp. 1–5.

Geet Parekh received the B.S. degree in information technology from Gujarat University, India, in 2010 and the M.S. degree in computer science from the University of Maryland, Baltimore County in 2015. From 2010 to 2013, he was with Indian Multinational, Tech Mahindra, on a network security product that discovered devices and connections. In 2015, he joined Deutsche Bank, Cary, NC, USA, as a Software Development Engineer and is currently working there on the bank's global net-banking portal for its corporate customers.

David DeLatte is a Research Associate Professor of computer science with University of Maryland, Baltimore County, affiliated with the Cyber Defense Laboratory and Mathematics Department

Geoffrey L. Herman is a Teaching Assistant Professor of computer science with the University of Illinois at Urbana-Champaign. His research focuses on studying how students learn engineering and computer science concepts and assessing students' knowledge of these concepts.

Linda Oliva received the Doctorate degree from Boston University. She is an Associate Chair of the Education Department, University of Maryland, Baltimore County. She teaches both graduate and undergraduate courses in the secondary teacher education program. She has over 30 years teaching experience at the elementary, middle, high schools, and post-secondary levels. Her research interests include teacher induction and retention, educational technologies, teacher dispositions, development of domain specific expertise, cybersecurity education, and program evaluation.

Dhananjay Phatak is an Associate Professor with the Department of Computer Science and Electrical Engineering Department, University of Maryland, Baltimore County.

Travis Scheponik is currently pursuing the Ph.D. degree in computer science with the CSEE Department, University of Maryland, Baltimore County, under the supervision of A. T. Sherman. He is a Software Engineer with Analysis, Computing and Engineering Solutions. His main research interest is cybersecurity. He has carried out research in cybersecurity education and open source software security.

Alan T. Sherman received the Ph.D. degree in computer science from MIT in 1987 under the supervision of R. L. Rivest. He is a Professor of computer science with the CSEE Department, University of Maryland, Baltimore County (UMBC), and the Director of UMBC's Center for Information Security and Assurance. His main research interest is high-integrity voting systems. He has carried out research in election systems, algorithm design, cryptanalysis, theoretical foundations for cryptography, applications of cryptography, cloud forensics, and cybersecurity education. He is also a Private Consultant performing security analyses.