Spring 5-29-2015

Requirements for IT Security Metrics - an Argumentation Theory Based Approach

Emrah Yasasin
University of Regensburg, emrah.yasasin@wiwi.uni-regensburg.de

Guido Schryen
University of Regensburg, guido.schryen@wiwi.uni-regensburg.de

ISBN 978-3-00-050284-2

REQUIREMENTS FOR IT SECURITY METRICS – AN ARGUMENTATION THEORY BASED APPROACH

Complete Research

Yasasin, Emrah, University of Regensburg, 93053 Regensburg, Germany, emrah.yasasin@wiwi.uni-regensburg.de

Schryen, Guido, University of Regensburg, 93053 Regensburg, Germany, guido.schryen@wiwi.uni-regensburg.de

Abstract

The demand for measuring IT security performance is driven by regulatory, financial, and organizational factors. While several best practice metrics have been suggested, we observe a lack of consistent requirements against which IT security metrics can be evaluated. We address this research gap by adopting a methodological approach that is based on argumentation theory and an accompanying literature review. As a result, we derive five key requirements: IT security metrics should be (a) bounded, (b) metrically scaled, (c) reliable, valid and objective, (d) context-specific and (e) computed automatically. We illustrate and discuss the context-specific instantiation of requirements by using the practically used "vulnerability scanning coverage" and "mean-time-to-incident discovery" metrics as examples. Finally we summarize further implications of each requirement.

Keywords: IT Security Metrics, IT Security Metrics Requirements, IT Security Metrics Design, Argumentation Theory.

1 Introduction

Regulatory, financial and organizational factors drive the demand for measuring IT security performance. For instance, there are various regulatory specifications that require firms to measure IT security (Lennon, 2003), such as the EU's Data Protection Directive 95/46/EC, the Directive 2002/58 on Privacy and Electronic Communications, Clinger-Cohen-Act or the Federal Information Security Management Act (FISMA). In the financial area, organizations that measure success and failure of current and past security measures can use metrics to justify and direct security investments. According to Deloitte's 2010 Financial Services Global Security Study, the increasing number of security threats, alongside more regulation, is driving investments in IT security (Deloitte, 2010). From an organizational point of view, metrics evaluate an IT security program's efficiency, and thereby advance accountability to internal stakeholders and improve customer confidence (Chew et al., 2008; Lennon, 2003). Bayuk (2013) describes that security metrics are typically based on the assumptions that there is a secure way to configure a system and that security management has to maintain configuration. At the system level, IT security metrics are tools that ease decision making and accountability through collection, analysis, and reporting of relevant performance data (Savola, 2013). Based on IT security performance goals and objectives, IT security metrics must be quantifiable, easy to survey, and repeatable in the ideal case (Chew et al., 2008; Littlewood et al., 1993). They identify relevant IT security trends over time, track performance of particular systems and help decision makers to direct appropriate security measures accordingly (Jaquith, 2007; Lennon, 2003; Vaughn et al., 2003).

However, current suggestions of IT security metrics are not based on common methodologically grounded requirements. For instance, Bellovin (2006) mentions the hardness of metrics and denounces

security metrics as ”chimeras for foreseeable future” whereas Cybenko and Landwehr (2012) point out the need to establish sound metrics. Almasizadeh and Azgomi (2014) as well as Fenz (2010) also reveal that quantitative security measurement is “a challenging area” and “one of the grand challenges in IT-security”. Jansen (2011) identifies several factors impeding progress of security metrics: lack of good estimators of system security, entrenched reliance on subjective, human, qualitative input, protracted and delusive means commonly used to obtain measurements, and the dearth of understanding and insight into the composition of security mechanisms. Pfleeger and Cunningham (2010) ask “why is security measurement so hard?” and highlight several reasons by comparing security measurements to metrics. However, the question of “why security metrics are so hard to do” is also posed but not solved by Wong (2012) who points out the discussion of the qualitative vs. quantitative nature of a measurement in security metrics and states that “in some cases, quantitative data simply does not exist.” This is a true statement but it does not provide guidance on how a quantitative metric should be defined. In other words: how do we know whether a (quantitative) IT security metric is useful? We observe a certain consensus among researchers that there is a lack of a methodological basis of how IT security metrics can be derived and evaluated thoroughly against predefined requirements, however only guidelines and proposals of attributes for “good metrics” exist but no theoretically and methodologically driven requirements (Bartol et al., 2009; Chew et al., 2008; Hayden, 2010; Jansen, 2011) although “there is a genuine and increasingly urgent need for variable metrics in information security” (Brotby et al., 2013). To sum up, only best practice suggestions for IT security metrics exist which leads us to propose the following research question: *Which requirements should IT security metrics fulfill?* In this article, we show which requirements IT security metrics should fulfill using the theoretical framework of argumentation theory. We adopt Toulmin's (1958, 2003) theoretical methodology by an argumentative inclusion of evidence, warrants and claims. The theory begins with the thought that an argument is a claim based on data in conformance with a warrant (Aberdein, 2005). To be precise, the data consist of certain facts that support the claim wheres the warrant is an inference conclusion according to which the data proves the claim (Verheij, 2005). Our argumentative investigation contributes therefore, to our best of knowledge, a first methodological attempt to show which requirements IT security metrics should fulfill. The paper also illustrates the need for the proposed requirements by evaluating two IT security metrics which are commonly used in practice, and summarizes further implications of the proposed requirements.

The paper is organized as follows: the next section introduces related work. The third section outlines our methodological approach for answering the research question. Subsequently, we derive and propose five general requirements for IT security metrics (Section 4). In Section 5, we apply the proposed requirements to two metrics and give the requirement’s implications for IT security metrics. Finally, we summarize our contributions, describe limitations of the current study and conclude with an outlook on future research.

2 Related Work

Considering the definition of the term “metric” in the context of IT security, there is a degree of uncertainty in the literature. Generally speaking, the terms metric, including security metric, and measurement tend to be used synonymously. To some extent, this vagueness derives from the fact that many definitions do not categorically distinguish the context in which the term is used. To the best of our knowledge, there is no commonly accepted definition of security metrics – a fact which is also mentioned by Masera and Fovino (2010). Brotby et al. (2013) even state that “metrics” in the information security area are “relatively immature”. The term, as used in practice, appears to represent several different notions: metric (in the sense of quantitative standard function based on process and/or product measures), measurement, score, rating, rank, or assessment result. Cain and Couture (2011) and Chakraborty et al. (2012) further observe that metrics are often confused with measurements. To complicate things, confusion may come up when the term is used in different contexts. Table 1 provides an excerpt of commonly used definitions for IT security metrics and shows the heterogeneity of definitions.

Reference	Definitions and Attributes of an IT Security Metric
Hallberg et al. (2011)	A security metric contains three main parts: a magnitude, a scale and an interpretation. The security values of systems are measured according to a specified magnitude and related to a scale. The interpretation prescribes the meaning of obtained security values.
Jaquith (2007)	Security metrics - a set of key indicators that tell [organizations] how healthy their security operations are, on a stand-alone basis and with respect to peers.
Kaur and Jones (2008)	Security metrics provide a framework for evaluating the security built into products or services available commercially.
Kormos et al. (1999)	A measurable attribute of the result of an SSE-CMM security engineering process (cf. ISO/IEC (2008) for more details) that could serve as evidence of its effectiveness. A security metric may be objective or subjective, and quantitative or qualitative.
Lennon (2003)	IT security metrics provide a practical approach to measuring information security. Evaluating security at the system level, IT security metrics are tools that facilitate decision making and accountability through collection, analysis, and reporting of relevant performance data.
Masera and Fovino (2010)	Security metrics are indicators, and not measurements of security. Security metrics highly depend on the point of reference taken for the measurement, and shouldn't be considered as absolute values with respect to an external scale.
Ouchani and Debbabi (2015)	A security metric is a quantitative measure indicating to which extend the considered entity possesses the attribute of being secure.
Preschern et al. (2014)	Security metrics qualitatively or quantitatively describe the level of security for a system.
Rosenblatt (2008)	A security metric is the standard measurement of computer security.
Rudolph and Schwarz (2012)	A security metric is a security measure with an associated set of rules for the interpretation of the measured data values.
Savola (2007)	A security metric is a quantitative and objective basis for security assurance. It eases in making business and engineering decisions concerning information security.

Table 1. Heterogeneity of definitions for IT security metrics in the literature.

It is not surprising that, besides practitioners, academic researchers also struggle with these terms. Vaughn et al. (2003) note that they are also often confused about what characterizes the measurement or metric, how to interpret it and how to validate it. The authors explicate that measurement simply shows properties like the "extent, dimensions, capacity, size, amount, or some other quantitative characteristic of the software or system" (Vaughn et al., 2003). The usefulness of measures heavily depends on interpretation, except in direct comparison with other measurement results to determine whether one value is more or less desirable than the other. Thus, it is a challenging task to conclude on measures alone (Vaughn et al., 2003; Savola, 2014). We agree with Chew et al. (2008) who argue that information security measures should be used to facilitate decision making and to improve performance and accountability based on information security goals and objectives.

However, as we are not only interested in measuring and quantifying information, we go one step further and, as postulated by Savola (2007), we strive to apply requirements for analysis with an established theory. By doing this, we strictly distinguish between a measurement and a metric and define an IT security metric as follows:

An IT security metric quantifies the security level of an information system and fulfills the following attributes: It is (a) bounded, (b) metrically scaled, (c) reliable, valid and objective, (d) context-specific and (e) computed automatically.

By contrast, a *measurement* is in our opinion the underlying ascertainment and calculation of the IT security metric.

3 Argumentation Theory and Methodology

Our methodology is based on the argumentation theory of Toulmin (1958, 2003) that is well-known and used widely in the IS literature in various contexts (Berente et al., 2011; Gregor, 2006; Schermann et al., 2009; Yetim, 2008). In the model of Toulmin (1958, 2003) three elements play a central role: evidence, warrants and claims. The evidence is adduced in support of a standpoint (claim) and is associated with the claim by means of a - usually implicit - justification (warrant) (Brockriede and Ehninger, 1960; Van Eemeren and Grootendorst, 2004; Simosi, 2003). In principle, the warrant is a general rule that serves to justify the step from the evidence to the claim (Van Eemeren and Grootendorst, 2004), i.e. "what makes the conclusion of an argument a "conclusion" (rather than simply a proposition) is that the reasons for drawing this conclusion on the basis of the premises are (at least partially) spelled out" (Mercier and Sperber, 2011). As we derive our requirements from the literature, our work is an argumentative and narrative discourse. Consequently, we show the line of our arguments and the resulting conclusions precisely as we reach our claims through causal and logical reasoning. Argumentation theory is a method of providing claims by causal and logical justifications (Brockriede and Ehninger, 1960; Van Eemeren and Grootendorst, 2004; Rowland, 2008; Simosi, 2003). We therefore assert that the Toulminian lens is a useful and theoretically neutral concept for thoroughly postulating requirements (Berente et al., 2011) and we use it as our methodological background which is shown in the figure below:

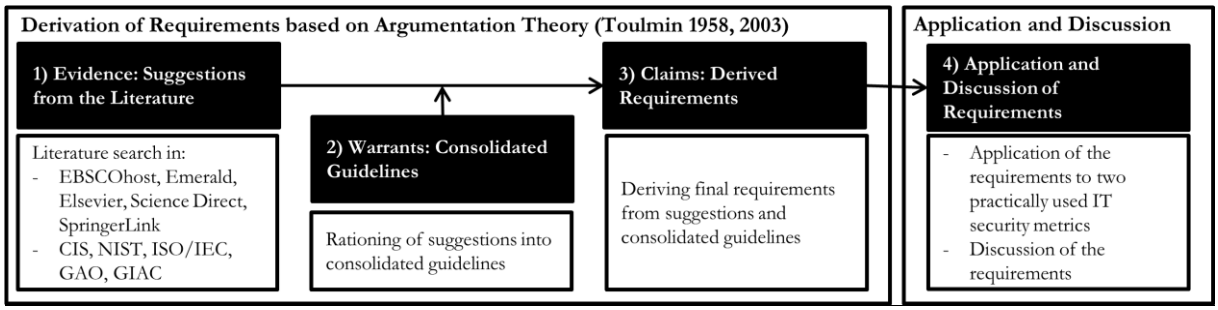

Figure 1. Research methodology.

Our methodology contains the three main components of Toulmin's argumentation theory: evidence, warrants and claims and – in addition – an application and discussion. We address the above-mentioned components in more detail below.

The *evidence*, also known as data or grounds, are statements offered to support the claim (Levy and Ellis, 2006) which is the main proposition of the argument (Berente et al., 2011). The evidence answers the question: "What is your basis?". To systematically identify relevant literature (evidence for deriving requirements), we followed the recommendations of Webster and Watson (2002). We scanned tables of contents, queried journal databases and reviewed selected conference proceedings. We analyzed the articles and extracted suggestions. For the literature search we used the following databases: EBSCOhost, Emerald, Elsevier, Science Direct and SpringerLink, which provide access to a large amount of peer reviewed articles in the major conference proceedings and journals, and are considered important because they cover crucial articles from both, computer science and IS discipline. Beyond these databases, we also focused on research institutes and organizations like CIS, ISO/IEC, NIST, GAO and GIAC as they might provide articles that study requirements for IT security metrics. This is the so-called *additional knowledge* which comprises "experiences and expertise that define the state-of-the-art in the application domain" (Hevner, 2007). We therefore included application-oriented papers (technical and report) from the aforementioned research institutes and organizations. These are necessary to identify the practical aspects and usage of IT security metrics. The use of precise key words or phrases is recommended by Rowley and Slack (2004) which have to be documented accurately (vom Brocke et al., 2009; Webster and Watson, 2002). To transparently report the coverage of our database search, we

provide the key words "IT Security Measurement", "Security Measurement", "Security Metrics", "IT Security Metrics", and "Information Security Measures". We limited the database search to articles published in English and in scholarly journals. The search returned 40 articles from which we removed articles that are not related to our focus on requirements. In particular, we excluded several articles that develop or focus on security metrics but do not address underlying requirements or guiding principles.[1] The remaining articles were read and examined in detail and are listed in Table 2.

Warrants, reflect the principles which suggest that the inference from the evidence to claim is appropriate. Warrants are "statements that justify the inference of the claim from data" (Yetim, 2008) and answer the question: "How did you get there?" (Berente et al., 2011). Toulmin (2003) states that "warrants are general, certifying the soundness of all arguments". To derive our arguments soundly, we therefore determine consolidated guidelines (warrants) from the evidence. According to Hitchcock (2010), warrants can be seen as "inference licenses" or field-specific "standards of reasoning". In our approach, we refer to the second type because the guidelines are not only statements but also causal reasons (Toulmin, 1958; Toulmin, 2003) which we use as an underlying to propose our requirements for IT security metrics. The consolidation into guidelines takes place as follows: in our investigation, the evidence is represented as suggestions from several papers. We account for similarities between different suggestions by summarizing sets of suggestions as "consolidated guidelines"; we named every guideline with an appropriate umbrella term. The results of our guidelines and the reasoning, why the guidelines are suitable, are our warrants. Thus, we present the guidelines and the reasons why they are appropriate. By doing this, we explicate the warrants for our requirements. For example, the Guideline "I: Bounded" is composed of the suggestions 1), 7), 9) and 18):

1) Measures must yield quantifiable information (percentages, averages, and numbers).
7) Expressed as a cardinal number or percentage, not with qualitative labels like "high", "medium", and "low".
9) Simple, precisely definable – so that is clear how the metric can be evaluated.
18) A security metric contains three main parts: a magnitude, a scale and an interpretation.

Based on the evidence and the warrants we derive our requirements (*claim*). The proposition of the claims works as follows: we first state the requirements and give afterwards from which guideline(s) they are derived. In this way we continuously maintain the inclusion from the argumentation theory: suggestions (evidence) ↪ consolidated guidelines (warrants) ↪ derived requirements (claims). To strengthen the claims, we illustrate and discuss the meaning of the requirements.

Furthermore and in accordance with GAO (2009) and Chew et al. (2008), we apply the requirements and discuss their implications, as application and evaluation of the results is an important cornerstone of research in information systems (Venkatesh et al., 2013).

4 Derivation of IT Security Requirements

The following section presents along the inclusions of the argumentation theory, as previously described, the derivation of requirements which an IT security metric should fulfill.

4.1 Evidence: Suggestions from the Literature

The following table shows the suggestions for IT security metrics which have been extracted from the literature:

[1] We give some examples of IT security standards that were not considered in our work: The latest standard ISO/IEC (2013) was excluded because it does not focus on our context like its predecessor ISO/IEC (2005): "Note that information security measurements are outside of the scope of this standard" and thus both outside of our scope. The standard of ISF (2011) deals exhaustively with IT security metrics but does not provide any suggestions/guidelines an IT security metric should fulfill and was therefore not considered either.

Reference	Suggestions for IT Security Metrics
Chew et al. (2008)	1) Measures must yield quantifiable information (percentages, averages, and numbers). 2) Data that supports the measures needs to be readily obtainable. 3) Only repeatable information security processes should be considered for measurement. 4) Measures must be useful for tracking performance and directing resources.
Jaquith (2007)	5) Consistently measured, without subjective criteria. 6) Cheap to gather, preferably in an automated way. 7) Expressed as a cardinal number or percentage, not with qualitative label. 8) Contextually specific – relevant enough to decision makers so that they can take action.
Thangavelu et al. (2010)	9) Simple, precisely definable – so that is clear how the metric can be evaluated. 10) Objective, to the greatest extent possible. 11) Easily obtainable. 12) Valid – the metric should measure what it is intended to measure. 13) Robust – relatively insensitive to insignificant changes in the process or product.
Chapin and Akridge (2005)	14) They should measure organizationally meaningful things. 15) They should be reproducible. 16) They should be objective and unbiased. 17) Over time, they should be able to measure some type of progression toward a goal.
Hallberg et al. (2011)	18) A security metric contains three main parts: a magnitude, a scale and an interpretation.
ISO/IEC (2009)	19) Ease of data collection. 20) Availability of human resources to collect and manage data. 21) Availability of appropriate tools. 22) Number of potentially relevant indicators supported by the base measure. 23) Ease of interpretation. 24) Number of users of developed measurement results. 25) Evidence as to the measure's fitness for purpose or information need. 26) Costs of collecting, managing, and analysing the data.

Table 2. Suggestions for IT security metrics.

4.2 Warrants: Consolidated Guidelines

The result of our warrants (consolidated guidelines) are illustrated below and explained thereafter. To ascertain the objectivity of the requirement development process, the reliability of agreement between the researches was measured with Cohen's Kappa (Cohen, 1960). The table shows for each consolidated guideline the inter-rater reliability statistics with a substantial to (almost) perfect agreement between the researchers (Landis and Koch, 1977).

Consolidated Guidelines	Suggestions	Inter-rater Reliability (Cohen's Kappa Coefficient)
I: Bounded	1), 7), 9), 18)	1.00
II: Quantified	1), 7), 9), 18)	1.00
III: Obtainable metric input data	2), 3), 5), 6), 11), 19), 20), 22), 26)	0.82
IV: Reliable	3), 5), 13), 15)	0.87
V: Valid	4), 12), 14), 17), 25)	0.87
VI: Objective	5), 10), 13), 16)	1.00
VII: Contextually specific	8), 9), 14), 23)	0.60
VIII: Automated	6), 11), 21), 26)	0.62

Table 3. Consolidation of suggestions for IT security metrics into guidelines.

I: Bounded: Suggestion 1) targets at percentages and averages which are both normally ranged. Suggestion 7) refers to cardinal number or percentages and thus directly implies the metric to be bounded. The aim of suggestion 9) may be understood literally in this context as a boundary as well. The clearest demand for a boundary is stated by suggestion 18). Without a magnitude, the characteristic "precisely

definable" of suggestion 9), for instance, is no longer useful when there is no boundary because of the representation of (perfect) insecurity and (perfect) security.

II: Quantified: Note that suggestions 1), 7), 9) and 18) have been just explained. Common to all is that they stand each in their different way for quantitative values.

III: Obtainable metric input data: Suggestion 2) addresses the obtainability of input data for the measurement. Suggestion 3) mentions "only repeatable information" which is a synonym for obtainability. Suggestion 11) focuses on the obtainability of input data whereas the attributes of suggestions 5), 6), 19), 20), 22) and 26) can be regarded as an aim to obtainability.

IV: Reliable: Our regard of reliability is mainly led by the observation of the quality of a measurement indicating the degree to which the measure is consistent, that is, repeated measurements would yield the same result. Nunnally (1967), for instance, defined reliability as "the extent to which measurements are repeatable and that any random influence which tends to make measurements different from occasion to occasion is a source of measurement error". The definition mainly conforms to suggestion 3) and 5). We agree that these suggestions do not distinctly explain it as reliability but we can conclude from their context that it aims certainly at reliability. This can be explained by the fact that repeatable and consistent measuring leads to reliable values. The same rationale also justifies suggestions 13) and 15).

V: Valid: For the definition of validity, we refer to Shadish et al. (2001) who define it "as the approximate truth of an inference" and denote that validity, regarded as truth of inference, plays an important role in evaluating tests and experiments. This is crucial because true inferences can also be about situations whose occurrence is detrimental. For instance, a true inference is that a computer virus has damaged the system. However, this should not detract from the need of inference truth because as Shadish et al. (2001) state, "it is good if inferences about a test are true, just as it is good for the causal inference made from an experiment to be true." This is a very important point, as the definition indicates the degree to which proposition and theory confirm the findings of test results. Note that validity refers to values that are not only reliable but also true and accurate so that reliability is a precondition for validity but not vice versa. With other words, validity is the extent to which a measurement does what it is supposed to do whereas reliability only guarantees stability of output values for repeated measurements. The definition of suggestion 4), 12), 14), 17) and 25) are quite similar and they directly refer to validity as an attribute of an IT security metric.

VI: Objective: We define objectivity in our context as the state of being even, unbiased and not influenced by any subjectivity or personal judgments. This view is instantly outlined by suggestion 10), 13) and 16). Suggestion 5) calls for an IT security metric without any subjectivity. The suggestions propose only objective facts that can be proved or disproved and furthermore the outcome of the value as objective. We have to distinguish at this point between a) the metric's input data and b) the metric's value itself. The first one is not covered by "III: Obtainable metric input data". It addresses only the availability of input data; objectivity is excluded from this consideration. So, it is necessary to discuss this aspect briefly. There might be metrics with subjective input data but with an objective (computed) value (e.g. by estimating the input data by decision makers.). This means that repeating the computation or determination may lead to the same IT security metric's value as gained before. With this, denote that not only objectivity is given but also reliability, so that we can say that objectivity can also be regarded as a prerequisite of reliability. This conclusion is already mentioned by Rothstein (1989) as early as 1989, in which she states that she "made the argument that an objective measurement is a reliable measurement." Beside the objectivity of the IT security metric's input data, there is a second point to discuss: the IT security metric value's objectivity. We could think of IT security metrics with objective input data. However, this cannot be a guarantee for an objectively determined value. As a result, it would produce a subjective measurement of an objective phenomenon (Rothstein, 1989). Nevertheless, this is not covered by this guideline: this suggests guaranteeing an objective value.

VII: Contextually specific: We distinguish between "V: Valid" and "VI: Contextually specific". Although it might seem paradoxical at first sight to differentiate, it is necessary. The aim of validity is to guarantee measuring the underlying properly (Golafshani, 2003). A contextually specific measurement is a property which guarantees the usefulness for decision makers that is described by suggestion 8) and also indirectly addressed in suggestion 9), 14) and 23). The view between validity and contextual specificity is different depending on which side of the border one resides on. While validity considers an ex ante view (cf. suggestion 12), contextual specificity is an ex post regard. Validity supports decision makers in guaranteeing that the metric is sound for the purpose they want to measure ("truth of inference") (Shadish et al., 2001). In this case we are dealing with an ex ante situation, mainly a view before the measurement of the actual values is carried out. Contextual specificity investigates the extent to which the metrics can actually create value. This implies that decision makers can take action after the measurement and is thus an ex post view. Consider a metric that is valid. While the purpose of that metric is ensured, it is not guaranteed that its output value is useful. Attention should therefore be drawn at this point to meet the need that such a metric's purpose for taking action is given as management decisions require hard facts rather than conjecture. Consequently, there is a fundamental difference between validity and contextual specificity.

VIII: Automated: Automation is considered in suggestions 6), 11), 21) and 26) which propose that the input data should be automatically gathered which enables an automatic computation of IT security metric values. We are aware that automation of a metric may cause additional monetary resources but the main thought behind is a time-saving metric that does not require a manual calculation.

4.3 Claims: Derived Requirements

From the results of the evidence and warrants, we derive the following requirements:

R1: *IT Security metrics are bounded*: IT security metrics should represent a value with an upper and a lower bound. This requirement is derived from Guideline I and II. Guideline I is a summary of several suggestions which aim at boundedness. In our view, Guideline II also suggests that IT security metrics should be in a range. Often, quantitative values like percentages or numbers are accompanied by ranges. In order to represent the extreme values (absolute) insecurity and (absolute) security, we have to postulate a range. Think of an unbounded metric that represents the value for security of an information system and assume that this information system does not have any security precautions. We can assume that the value (of this metric) exhibits the worst value – but which value shall we assign to the metric? We can generalize this question: How can we determine security and insecurity when we do not, as yet, know exactly what they are to be? Therefore we need an upper bound that represents the best and a lower bound which illustrates the worst value. The bounds are left to the authors who develop the metric. We only state that there is a need to cover both extremes whereas we agree that the reality normally lies between these extremes.

R2: *IT Security metrics are scaled metrically*: IT security metrics shall reflect characteristics which can be measured exactly in terms of quantity. Several suggestions (cf. Table 3, Guidelines I and II) address the need for quantification. Thus, to guarantee both the monitoring how the IT security level changes over time and the quantitative evaluation of measures, we require metrics to be metrically scaled. This can either be an interval scale or a ratio scale. In both scales the difference between two metric values is meaningful. Imagine a metric that measures the password strength of a login and let the metric values be percentages for instance. Usually, the password length of an information system and the character set is technically limited. As a side note, we can remark that R1 is fulfilled due to the fact that the range is [0 %;100 %] and ergo bounded. A perfectly insecure password is mapped to the value of 0 % whereas 100 % denotes a perfectly secure password with the given technical limits. So, R2 guarantees that differences between two user's passwords can be compared. Imagine user 1 has password strength of 70 % and user 2 password strength of 90 %. The difference between these two values

equals 20 % and is interpretable. Furthermore, such a metrically scaled value has advantages for firms or decision makers as well but this is mainly covered by R4 and consequently explained there.

R3: *IT security metrics satisfy the criteria of quality*: IT security metrics should satisfy the criteria of quality: validity, reliability and objectivity (cf. Table 3, Guidelines IV, V and VI). In other fields, approved and experienced assessment methods and procedures which meet the requirements objectivity, reliability and validity are well-known. We are going to adopt these quality criteria as R3, knowing that this requirement is an umbrella term for these three quality criteria. The need of such quality criteria is explained by Golafshani (2003) who states that "the definitions of reliability and validity in quantitative research reveal two strands: Firstly, with regard to reliability, whether the result is replicable. Secondly, with regards to validity, whether the means of measurement are accurate and whether they are actually measuring what they are intended to measure." Objectivity assures as we have seen before "the extent to which the findings are free from bias" (Ali and Yusof, 2011). The guidelines we observed target to these quality criteria and quantification needs reproducible and objective solutions that are valid. As these quality criteria guarantee sound and replicable values, we can adopt these criteria.

R3a: *IT security metrics are reliable*: IT security metrics should be reliable. Let us revisit our password security example again. The determination of the likelihood is reliable as it will always give the same percentage limit within the technical limits. Consider a password containing 8 characters with accepting English letters and special characters. The extreme cases may be 0 % for "password not set" and 100 %, for instance, when using a password containing 4 letters and 4 special characters. Hence, with 0 % and 100 %, the extreme cases are covered. In other cases, we can define our ranges for various password combinations and map them to percentages. So, we can always get a reliable metric.

R3b: *IT security metrics are valid*: IT security metrics should be valid in the view we described before. In the password security metric example we just explained that the metric's aim is clearly outlined. Of course, we can think of other IT security metrics whose values are percentages but the goal might be different. This metric cannot be used for any other purpose than for the security of a user's password. For instance, such a metric cannot be used to measure the security of a database as it is developed for a dissimilar context. The consequence is that we will not be able to draw true inferences (Shadish et al., 2001) when using this metric in another situation. This example can be generalized so we can say that, probably, a metric used for another purpose than the considered context should be changed substantially. Only then we can conclude that such a metric satisfies the required validity criterion.

R3c: *IT security metrics are objective*: IT security metrics should be objective. Recall our example again. The usage of the English letters and special characters is highly subjective to a user's will. A user is not restricted in selecting the password but the technical limits and the clear mapping of letters and special characters to percentages will guarantee an objective output of the metric. In this example, one can see clearly that subjectivity of the input parameters for the computation of the metric can lead to objectivity. In other words and to sum up, we get an objective measurement of a subjective phenomenon (Rothstein, 1989), i.e. the free determination of combinations. The subjectivity lies in the input data whereas the calculation of the metric value is objective.

R4: *IT security metrics are context-specific*: The design of IT security metrics should be context-specific which is proposed by four suggestions and Guideline VII. As contextual specificity is the validity extent to which the metrics can create value, we illustrate why it is necessary. Recall our example of a percentage password security metric again and let us assume a firm's policy wants to ensure at least 80 % password strength for every user. A contextually specific metric enables decision makers to a) detect how many users do not satisfy this criterion, b) quantify how much improvement is necessary to comply with the policy, c) suggest measures to improve password security (e.g. to use special characters) and d) evaluate whether the measures are effective. So, contextual specificity is required to ensure that IT security metrics are useful for decision-making.

R5: *IT security metrics can be computed automatically*: We adopt Guideline VIII as requirement 5 because automation is important in practice. Gathering input data manually is inefficient and costs valuable time. Furthermore, machine-computable metrics accurately execute the instructions the same way each time and are less error-prone. As a consequence, frequently computed metrics should be automated and designed as to save computational time. That means that not only the formula of the metric should not be too time-consuming but it should also be possible to gather the metric's input source data automatically. However, Jaquith (2007) states that some IT security metrics require time consuming methods and cannot be automated. We agree that automation is not obligatory if the cost of managing the data (ISO/IEC, 2009) and calculating the metric are not too high.

Note that R1 to R4 are stronger propositions compared with R5. We conclude that IT security metrics are useful and appropriate if they fulfill the requirements R1 to R4. R5 guarantees that the metric is also practically usable and economically efficient. Imagine a metric which requires lots of input data and fulfills R1 to R4 but not R5. Such a metric tends to be a theoretical measure being rarely relevant in practice.

5 Application and Discussion of Requirements

We apply our requirements by evaluating two IT security metrics defined by "The Center of Internet Security", a nonprofit organization that provides consensus-oriented products in information security, services, tools, metrics, suggestions, and recommendations (the "CIS Products") as a public service to internet users worldwide (CIS, 2010). We show exemplarily that the Vulnerability Scanning Coverage (VSC) fulfills our requirements while the Mean-Time-to-Incident-Discovery (MTTID) violates R1 and R3. CIS (2010) confirms that the MTTID lacks acceptable goal values and needs further empirical investigation to interprete the computed value clearly (CIS, 2010). We will also discuss how this inaccuracy in the interpretation relates to the violation of R1 and R3. In addition, we discuss the differences between both metrics, the requirements and summarize general implications for decision makers when an IT security metric fulfills these requirements.

5.1 Application 1: Vulnerability Scanning Coverage

The "Vulnerability Scanning Coverage" (VSC) is a metric that quantifies the percentage of an organization's systems that have been checked for known vulnerabilities. This metric measures the extent to which an organization's environment is searched for known vulnerabilities. Organizations can use this metric to evaluate their risk position in terms of concentrations of unknown vulnerability states of systems. It is calculated by dividing the total number of systems scanned by the total number of systems within the metric scope such as the entire firm (CIS, 2010):

$$\mathrm{VSC} = \frac{\sum \mathrm{Scanned\_Systems}}{\sum \mathrm{All\_Systems\_within\_Organization}} \cdot 100\ \%.$$

R1: The VSC metric is bounded. ✓A VSC value of 0 % means that none of the organizations systems were scanned, a value of 100 % means that all systems are checked for known vulnerabilities. Thus, the VSC metric has a range of [0 %;100 %] and is bounded, so that requirement R1 is satisfied.

R2: The VSC metric is metrically scaled. ✓ Due to the fact that the VSC metric is a percentage, it is of a ratio scale and therefore metrically scaled. Given that, requirement R2 is also fulfilled.

R3: The VSC metric is valid, reliable and objective. ✓ We can assume that the data of scanned systems can be characterized binary - 0 if a system is not scanned and 1 if a system is scanned for vulnerabilities. This information can be stored in a database and generally assumed to be collected correctly. So, concerning reliability, the metric always computes the same percentage if carried out repeatedly. Objectivity is also given as the information of scanned or not scanned is binary and consequently not subjective so

that the VSC represents a non-subjective ratio. Due to the definition of the metric, i.e. the fact that it provides information about how much of the organization's environment is checked for known vulnerabilities, the metric's context is clearly specified so that it cannot be used in a dissimilar area. Hence, the metric is valid as well. As the quality criteria are given, we conclude that requirement R3 is satisfied as well.

R4: The VSC metric is context-specific. ✓ The context of the VSC metric is clearly described, so it is useful for decision makers. They can take the metric as input to implement security measures, e.g. to scan more systems in their organization. Thus, requirement R4 is fulfilled.

R5: The VSC metric can be computed automatically. ✓ Requirement 5 is satisfied as the stored data inputs for the computation of the metric can be read directly from a database which has a relatively short runtime.

5.2 Application 2: Mean-Time-To-Incident-Discovery

The "Mean-Time-To-Incident-Discovery" (MTTID) measures the efficiency of detecting incidents in terms of the average elapsed time between the initial occurrence of an incident and its subsequent discovery. For each record, the metric is calculated by subtracting the date of occurrence from the date of discovery. These periods of time are then averaged across a scope of incidents, for example by time (CIS, 2010):

$$\text{MTTID} = \frac{\sum(\text{Date\_of\_Discovery - Date\_of\_Occurence})}{\sum\text{Incidents}}.$$

R1: The MTTID metric is not bounded. ↯ The value "0 hours" indicates a hypothetical instant detection time which is the lower bound of the metric (CIS, 2010). However, the upper bound of the metric lacks clarity. Imagine the case that the date of discovery of vulnerability was much further in the future than the date of occurrence. This leads to a growing numerator with no upper bound. In the most extreme theoretical case the limit value of the numerator is infinite and thus the value of MTTID limits to infinite. That means there is no upper bound and it is difficult to interpret what the metrics value does mean which is also stated by CIS (2010): "Because of the lack of experiential data from the field, no consensus on the range of acceptable goal values for MTTIDs exist." So, requirement R1 is not fulfilled.

R2: The MTTID metric is metrically scaled. ✓ The definition of the unit is time per incident which can be measured on a weekly, monthly, quarterly or annual basis per definition. So far, requirement R2 is fulfilled as time per incident is of metrical scale.

R3: The MTTID metric is not valid but reliable and objective. ↯ It does not contravene reliability and objectivity. The parameters "date of discovery", "date of occurrence" and the number of incidents are objective; as a result, a recomputation yields the same value for the metric. Even though the CIS states that the metric "measures the effectiveness of the organization in detecting security incidents" (CIS, 2010), they admit that the MTTID lacks goal values so that the metrics validity is not given. A valid IT security metric's computed value must be defined and interpreted clearly (Golafshani, 2003) which, in fact, is not the case for MTTID.

R4: The MTTID metric is context-specific. ✓ The context of the VSC metric is clearly described, so that utility for decision makers is given. They can use the metric to evaluate possible security measures (e.g. make efforts to reduce the mean time to discover incidents). Thus, requirement R4 is fulfilled.

R5: The MTTID metric cannot be computed automatically. ↯ We recognize that the input source "date of discovery" and the number of incidents can be gathered and computed automatically. However, we cannot think of an example detecting "date of occurrence" automatically without contradicting the date of discovery. Therefore, requirement R5 is not given.

5.3 Discussion

Our comparison of both metrics underlines the necessity of requirements for IT security metrics. Taking a closer look at MTTID, we see that the violation of R1 and R3 has severe consequences for the usefulness of the metric. Concerning the unboundedness of MTTID (violation of R1), it is sufficient that only one incident discovery lies farther in the future than the occurrence and the value becomes biased due to the fact that there is no upper bound. If there was an upper bound, it would map this outlier to a value in a bounded range which is less affected by irregularities and outliers. Besides, MTTID's violation of validity (R3) leads also to a non-valid metric whose output value needs further empirical investigation. If validity is compromised, the obtained/computed results of the metric cannot be explained and interpreted soundly. In other words, how can the difference between two output values of the MTTID metric be understood? Thus, validity is crucial for the exact explanation (Golafshani, 2003) of the metric's value. A minor shortcomming of the MTTID is that it can only be computed semi-automatically (violation of R5), however this does not have serious impacts on the metric's value.

In contrast, the VSC fulfills all requirements. VSC has a defined metrical range between [0;1] with a clear scope and delivers the very same results when the metric is determined repeatedly. Besides, it can be computed automatically. As VSC is bounded, every value can be assigned unambiguously to an action that was taken and vice versa. For instance, when more systems are scanned in a firm, the metric's value approaches 1 but cannot exceed it which leads us to the next important aspect: in contrast to MTTID, validity is clearly given and indicates the degree to which the systems of a firm have been scanned. From the line of reasoning which lead to the postulated requirements and from the analysis of two specific metrics, we conclude that our derived requirements ensure that the metric and its value are appropriate and meaningful. If one of the requirements is violated, interpretation and computational efficiency are compromised. In our example, validity was not given in MTTID which leads to non-interpretable values. Further, the metric was unbounded so that it is difficult to interprete the value of the metric. There might be other side effects as well if one of the requirements is violated: for instance, if VSC did not rigidly count the amount of scanned systems but rather a subjective estimation of scanned systems done by experts, the metric's value would be difficult to reproduce. Thus, evaluation in such highly subjective measures must be carefully thought out and its results must be interpreted with caution.

This discussion can be generalized for all requirements. A metric fulfilling the requirements is not only useful but also supports decision makers as it allows them to initiate appropriate measures to be taken. We would like to highlight at this point that it is not necessary to propose usefulness as a separate requirement because it is implied by the proposed requirements. We summarize these in the implications, in particular for decision makers, below:

Derived Requirement	Implications
R1: IT security metrics are bounded.	▪ Mapping of (perfect) insecurity and (perfect) security to concrete quantities. → By normalizing the metric's value in a range, the (numerical) values of the metric can be transformed into a manageable dimension.
R2: IT security metrics are scaled metrically.	▪ The difference between the metric's values is meaningful and interpretable. ▪ Decision makers can determine the improvement of the security level. → They can therefore implement measures to improve IT security.
R3: IT security metrics satisfy the criteria of quality: validity, reliability and objectivity.	▪ The metric is objective and its computation is replicable under similar circumstances. ▪ The metric measures what it is designed for. ▪ Considering the metric's value, decision makers can make informed decisions to increase the security level because the results are consistent, stable, dependable, and are therefore reliable as well as objective.

R3a: IT security metrics are reliable.	▪ Decision makers can verify the metric by recomputing and they can rely on the fact that repeated measurements would give the same result.
R3b: IT security metrics are valid.	▪ Decision makers know that the purpose of the metric is clear and specific.
R3c: IT security metrics are objective.	▪ The resulting objectivity leads to more confidence in decisions and a better use of security. ▪ For decision makers, the security level will become unbiased so that they can implement measures with confidence.
R4: IT security metrics are context-specific.	▪ Decision makers can take actions based on the metric's value. ▪ Decision makers can improve the security level by identifying specific deficits.
R5: IT security metrics can be computed automatically.	▪ From an economic perspective, automization tends to be more efficient as algorithmizing the metric enables an automatic computation. ▪ Integration of the metric into tools is feasible. ▪ Manually calculating of the metric's value becomes obsolete.

Table 4. Summary of requirements and implications.

6 Conclusion

We have derived and shown in this research using the argumentation theory of Toulmin (1958, 2003) what requirements should be fulfilled by IT security metrics. Our contribution is threefold: First, to the best of our knowledge, this is the first systematic and methodological derivation of requirements for IT security metrics. Second, we use an established theory in the IS literature as an underlying: Based on argumentation theory, we explore not only requirements (*claims*) but also the rationale behind each derivation (*warrants*) from the current academic and practical literature (*evidence*) of the proposed requirements. Third, we apply our requirements to IT security metrics which are used widely in practice and evaluate them against our proposed requirements. We showed that the MTTID metric needs deeper empirical evaluation in order to use it in practice properly because of its lack of acceptable goal values. Acceptable goal values are covered by bounds in R1 which, in addition to the other requirements, is fully satisfied by the VSC metric requirement so that it has no obstructions for practical usage. We further demonstrated that our proposed requirements ensure an objective, valid, reliable and clearly interpretable metric. Finally, we summarized managerial implications that follow from conformance to the requirements.

Despite the insights in the nature of IT security metrics and the contribution of a first scientific attempt to establish requirements on IT security metrics, we want to make some limitations transparent as well to enable future research. First, we analyzed and discussed two IT security metrics. In the future, we therefore plan to evaluate and discuss a broader set of IT security metrics based on our requirements. Second, as the metrics are drawn from CIS (2010), our choice is solely based on one source. Future research may use, in addition to the suggested metrics of CIS (2010), other sources which may provide additional IT security metrics in other contexts as well. Third, the identification and recognition of further implications in practice may also guide the next steps.

We hope that our work will be a starting point for an academic discourse on the nature and definitions of IT security metrics.

Acknowledgment. The research leading to these results was supported by the "Regionale Wettbewerbsfähigkeit und Beschäftigung", Bayern, 2007-2013 (EFRE) as part of the SECBIT project (http://www.secbit.de) and by the "Bavarian State of Ministry, Education, Science and the Arts" as part of the FORSEC research association (https://www.bayforsec.de).

References

Aberdein, A. (2005). "The Uses of Argument in Mathematics." *Argumentation* 19(3), pp.287–301.

Ali, A.M. and H. Yusof (2011). "Quality in Qualitative Studies: The Case of Validity, Reliability and Generalizability." *Issues in Social and Environmental Accounting* 5(1), pp.25–64.

Almasizadeh, J. and M.A. Azgomi (2014). "Mean Privacy: A Metric for Security of Computer Systems." *Computer Communications* 52, pp.47–59.

Bartol, N. et al. (2009). *Measuring Cyber Security and Information Assurance (State-of-the-Art-Report (SOAR))* K. J. Knapp, ed., IGI Global. URL: https://buildsecurityin.us-cert.gov/sites/default/files/MeasuringCybersecurityIA.PDF.

Bayuk, J.L. (2013). "Security as a Theoretical Attribute Construct." *Computers & Security* 37, pp.155–175.

Bellovin, S.M. (2006). "On the Brittleness of Software and the Infeasibility of Security Metrics." *IEEE Security & Privacy Magazine* 4(4), pp.96–96.

Berente, N. et al. (2011). "Arguing the Value of Virtual Worlds: Patterns of Discursive Sensemaking of an Innovative Technology." *MIS Quarterly* 35(3), pp.685–709.

Vom Brocke, J. et al. (2009). "Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process." In: *Proceedings of the 17th European Conference on Information Systems (ECIS 2009)*. pp. 2206–2217.

Brockriede, W. and D. Ehninger (1960). "Toulmin on Argument: An Interpretation and Application." *Quarterly Journal of Speech* 46(1), pp.44–53.

Brotby, W.K., G. Hinson and M.E. Kabay (2013). *Pragmatic Security Metrics Applying Metametrics to Information Security*, Boca Raton, Florida: CRC Press.

Cain, C.I. and E. Couture (2011). "Establishing a Security Metrics Program." *GIAC Enterprises*, pp.1–27.

Chakraborty, A., A. Sengupta and C. Mazumdar (2012). "A Formal Approach to Information Security Metrics." In: *International Conference on Emerging Applications of Information Technology*. pp. 439–442.

Chapin, D.A. and S. Akridge (2005). "How Can Security Be Measured?" *Information Systems Control Journal* 2, pp.43–47.

Chew, E. et al. (2008). "NIST 800-55 Revision 1: Performance Measurement Guide for Information Security." *National Institute of Standards and Technology (NIST)*. URL: http://csrc.nist.gov/publications/nistpubs/800-55-Rev1/SP800-55-rev1.pdf.

CIS (2010). "The CIS Security Metrics." *The Center for Internet Security*, pp.1–166. URL: https://benchmarks.cisecurity.org/tools2/metrics/CIS_Security_Metrics_v1.1.0.pdf.

Cohen, J. (1960). "A Coefficient of Agreement for Nominal Scales." *Educational and Psychological Measurement* XX(1), pp.37–46.

Cybenko, G. and C.E. Landwehr (2012). "Security Analytics and Measurements." *IEEE Security & Privacy* 10(3), pp.5–8.

Deloitte (2010). "2010 Financial Services Global Security Study: The faceless threat." *Deloitte Touche Tohmatsu*, pp.1–37. URL: http://www.deloitte.com/assets/Dcom-Argentina/Local Assets/Documents/consultoria/ERS/arg_cons_gfsi-ss-2010_13092010.pdf.

Van Eemeren, F.H. and R. Grootendorst (2004). *A Systematic Theory of Argumentation: The Pragma-Dialectical Approach*, New York: Cambridge University Press.

Fenz, S. (2010). "Ontology-based generation of IT-security metrics." In: *Proceedings of the 2010 ACM Symposium on Applied Computing*. pp. 1833–1839.

GAO (2009). "Concerted Effort Needed to Improve Federal Performance Measures." *United States Government Accountability Office*. URL: http://www.gao.gov/assets/300/295160.pdf.

Golafshani, N. (2003). "Understanding Reliability and Validity in Qualitative Research." *The Qualitative Report* 8(4), pp.597–606.

Gregor, S. (2006). "The Nature of Theory in Information Systems." *MIS Quarterly* 30(3), pp.611–642.

Hallberg, J. et al. (2011). "Controlled Information Security." *FOI, Swedish Defence Research Agency*, pp.1–42. URL: http://www.foi.se/ReportFiles/foir_3187.pdf.

Hayden, L. (2010). *IT Security Metrics: A Practical Framework for Measuring Security and Protecting Data*, McGraw-Hill Professional.

Hevner, A.R. (2007). "A Three Cycle View of Design Science Research." *Scandinavian Journal of Information Systems* 19(2), pp.87–92.

Hitchcock, D. (2010). "Obituary: Stephen Edelston Toulmin." *Argumentation* 24(3), pp.399–401.

ISF (2011). "The 2011 Standard of Good Practice for Information Security." *Information Security Forum*.

ISO/IEC (2008). "ISO/IEC 21827:2008 - Information technology -- Security techniques -- Systems Security Engineering -- Capability Maturity Model® (SSE-CMM®)." *International Organization for Standardization*.

ISO/IEC (2005). "ISO/IEC 27002:2005 - Information technology -- Security techniques -- Code of practice for information security management." *International Organization for Standardization*.

ISO/IEC (2013). "ISO/IEC 27002:2013 - Information technology -- Security techniques -- Code of practice for information security management." *International Organization for Standardization*.

ISO/IEC (2009). "ISO/IEC 27004:2009 - Information technology -- Security techniques -- Information security management -- Measurement." *International Organization for Standardization*.

Jansen, W. (2011). "Research Directions in Security Metrics." *Journal of Information System Security Volume* 7(1), pp.3–22.

Jaquith, A. (2007). *Security Metrics: Replacing Fear, Uncertainty and Doubt*, Addison-Wesley Professional.

Kaur, M. and A. Jones (2008). "Security Metrics - A Critical Analysis of Current Methods." In: *Proceedings of the 9th Australian Information Warfare and Security Conference*. pp. 41–47.

Kormos, C. et al. (1999). "Using Security Metrics to Assess Risk Management Capabilitie." In: *National Information Systems Security Conference*.

Landis, J.R. and G.G. Koch (1977). "The Measurement of Observer Agreement for Categorical Data." *Biometrics* 33(1), pp.159–174.

Lennon, E.B. (2003). "IT Security Metrics." *National Institute of Standards and Technology*, pp.1–3. URL: http://csrc.nist.gov/publications/nistbul/bulletin08-03.pdf.

Levy, Y. and T.J. Ellis (2006). "A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research." *Informing Science Journal* 9, pp.181–212.

Littlewood, B. et al. (1993). "Towards Operational Measures of Computer Security." *Journal of Computer Security* 2(2-3), pp.211–230.

Masera, M. and I.N. Fovino (2010). "Security Metrics for Cyber Security Assessment and Testing." *Joint Research Centre of the European Commission* D42, pp.1–26.

Mercier, H. and D. Sperber (2011). "Why Do Humans Reason? Arguments for an Argumentative Theory." *The Behavioral and Brain Sciences* 34(2), pp.57–74; Discussion 74–111.

Nunnally, J.C. (1967). *Psychometric Theory*, New York. URL: McGraw-Hill.

Ouchani, S. and M. Debbabi (2015). "Specification, Verification, and Quantification of Security in Model-based Systems." *Computing*, pp.1–21. URL: http://link.springer.com/10.1007/s00607-015-0445-x.

Pfleeger, S.L. and R.K. Cunningham (2010). "Why Measuring Security Is Hard." *IEEE Security and Privacy* 8(4), pp.46–54.

Preschern, C. et al. (2014). "Quantitative Security Estimation based on Safety Architecture Design Patterns." *Lecture Notes on Software Engineering* 2(4), pp.307–313.

Rosenblatt, J. (2008). "Security Metrics: A Solution in Search of a Problem." *EDUCAUSE Quarterly* 31(3), pp.8–11.

Rothstein, J.M. (1989). "On Defining Subjective and Objective Measurements." *Journal of the American Physical Therapy Association* 69(7), pp.577–579.

Rowland, R.C. (2008). "Purpose, Argument Fields, and Theoretical Justification." *Argumentation* 22(2), pp.235–250.
Rowley, J. and F. Slack (2004). "Conducting a Literature Review." *Management Research News* 27(6), pp.31–39.
Rudolph, M. and R. Schwarz (2012). "A Critical Survey of Security Indicator Approaches." In: *Proceedings of the 2012 7th International Conference on Availability, Reliability and Security (ARES)*. pp. 291–300.
Savola, R. (2007). "Towards a Security Metrics Taxonomy for the Information and Communication Technology Industry." In: *International Conference on Software Engineering Advances (ICSEA 2007)*. pp. 28–30.
Savola, R.M. (2013). "Quality of Security Metrics and Measurements." *Computers & Security* 37, pp.78–90.
Savola, R.M. (2014). "Towards Measurement of Security Effectiveness Enabling Factors in Software Intensive Systems." *Lecture Notes on Software Engineering* 2(1), pp.104–109.
Schermann, M. et al. (2009). "Justifying Design Decisions with Theory-Based Design Principles." In: *Proceedings of the 17th European Conference on Information Systems (ECIS 2009)*. pp. 1065–1076.
Shadish, W.R., T.D. Cook and D.T. Campbell (2001). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* 2nd ed., Boston, New York: Houghton Mifflin.
Simosi, M. (2003). "Using Toulmin's Framework for the Analysis of Everyday Argumentation: Some Methodological Considerations." *Argumentation* 17(2), pp.185–202.
Thangavelu, S.R.K., A. Sumithra and K. Alagarsamy (2010). "The Applicability of Existing Metrics for Software Security." *International Journal of Computer Applications* 8(2), pp.29–33.
Toulmin, S. (1958). *The Uses of Argument*, Cambridge: Cambridge University Press.
Toulmin, S. (2003). *The Uses of Argument*, Cambridge: Cambridge University Press.
Vaughn, R.B., R. Henning and A. Siraj (2003). "Information Assurance Measures and Metrics - State of Practice and Proposed Taxonomy." In: *Proceedings of the 36th Hawaii International Conference on System Sciences*. IEEE Computer Society Press.
Venkatesh, V., S.A. Brown and H. Bala (2013). "Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems." *MIS Quarterly* 37(1), pp.21–54.
Verheij, B. (2005). "Evaluating Arguments Based on Toulmin's Scheme." *Argumentation* 19(3), pp.347–371.
Webster, J. and R.T. Watson (2002). "Analyzing the Past to Prepare for the Future: Writing a Literature Review." *MIS Quarterly* 26(2), pp.xiii–xxiii.
Wong, C. (2012). *Security Metrics: A Beginner's Guide*, New York: McGraw-Hill Professional.
Yetim, F. (2008). "A Framework for Organizing Justifications for Strategic Use in Adaptive Interaction Contexts." In: *Proceedings of the 16th European Conference on Information Systems (ECIS 2008)*. pp. 815–825.