Measures, Metrics, and a Scale for Appraisal of Cyber Threat Intelligence-Informed Decision-Making

Mona Kriesten, Mamello Thinyane and David Ormrod
University of South Australia, Adelaide, Australia

mona.kriesten@mymail.unisa.edu.au
mamello.thinyane@unisa.edu.au
dave.ormrod@unisa.edu.au

Abstract: Cyber threat intelligence (CTI) is information from past, present, and evolving threats which, if correlated and put in context, aims to enhance cybersecurity decision-making at strategic, operational, and tactical levels. Despite the multiple benefits of CTI, such as identifying and profiling threat actors, tuning systems and cybersecurity controls, and providing context to incidents, the field faces challenges that must be overcome for effective implementation of CTI. The bulk of existing research tackling these challenges focuses on the technical aspects of collecting, analysing, using, and sharing CTI. However, one of the main benefits of CTI lies in its intelligence affordances to inform decision-making for key actors in cybersecurity. Unfortunately, there is generally a dearth of research on human factors associated with disseminating and utilising CTI. Further, while some research has been undertaken investigating the quality of CTI, there has not been much research investigating the quality of CTI-informed decision-making. This research is targeted to address this gap within the context of a larger project investigating the effectiveness of gamification in enhancing CTI use for defence against cyberattacks. To measure the benefits of CTI throughout the decision-making process, this research has developed a gamification platform and some of the relevant metrics and measures. Firstly, this paper presents these proposed measures and the derived metrics that can be used to quantify the benefits of using CTI at the individual decision level to measure the overall effectiveness of CTI. Secondly, the paper presents a scale that is developed to provide a yardstick for future CTI performance testing – specifically for CTI gamification solutions and generally for CTI-informed cybersecurity decision-making. The research addresses the need to quantify the impact of CTI on decision-making processes in cybersecurity through the measures, metrics, and a scale to inform the actual assessments.

Keywords: Cyber threat intelligence, Decision-making, Cybersecurity, Gamification, Quality metrics

1. Introduction

Cyber threat intelligence (CTI) is a critical resource that helps organisations gain awareness of the current threat landscape and vulnerabilities and decide on courses of action (COA) to mitigate cyber threats. Further, CTI helps organisations to balance the knowledge asymmetry between attackers and defenders and to make informed decisions. Similar to the military use of intelligence, where it originated, CTI helps to prevent attacks, reduce time to detection, and support decision-making by helping to illuminate the threat landscape from a wide range of information and intelligence sources. Raw data from these sources needs to be analysed, correlated, and contextualised before they become actionable CTI (Ampel et al., 2024; Schlette et al., 2021).

The key contribution of CTI in cybersecurity is to improve situational awareness and enhance decision-making. However, despite this being the primary impact pathway, a lot of current research mainly focuses on the technical aspects of the field instead of shedding light on its application and impact on decision-making (Shin & Lowry, 2020). Research on measuring the quality and effectiveness of CTI similarly focuses on optimising sharing standards and assessing the quality of CTI feeds. Besides initial attempts to define CTI quality dimensions and associated metrics, research on CTI quality, and specifically its impact on decision-making, is still in its infancy and more research needs to be done (Schlette et al., 2021).

This research builds upon the need to assess CTI use, focusing on its impact on decision-making and cybersecurity outcomes; shifting from a solely technical perspective to a socio-technical one that incorporates elements of human cognition and decision-making. Recognising the need for relevant measures and metrics in the domain, this research proposes a set of instruments to measure CTI-informed decision-making. First, the paper introduces foundational literature on decision-making, measures, metrics, and CTI quality. This is followed by the presentation of the measures, metrics, and scale developed in this research to assess CTI use focusing on the correlation between CTI use and the decision outcomes. The paper concludes by discussing the effectiveness and impact of these instruments and makes recommendations for their further utilisation to improve cybersecurity operations and outcomes in organisations.

2. Background and Theoretical Framing

The theoretical foundations for this research are within the domains of decision-making, measures, metrics, and CTI quality. These are brought together to inform the solutions developed in this research to assess the impact of CTI on cybersecurity decision-making.

2.1 Foundations of Decision-Making

Decision-making, which is the subject of various disciplines such as cognitive informatics, computer science, management science, economics, sociology, and psychology, is commonly described as a process through which the decision-maker weighs decision options against each other to achieve a desired outcome (Lunenburg, 2010; Uzonwanne, 2016; Wang & Ruhe, 2007). Decision-making is influenced by internal and external factors to the decision maker. Internal factors include personal attributes such as perception, experience, and values, which make decision-making highly subjective. External factors include resources, organisational structures, social systems and macro-environmental factors (Lunenburg, 2010; Svenson, 1979).

Different decision-making styles have been identified at the individual level, such as rational, intuitive, dependent, and avoidant decision-making. The rational style is characterised by the logical evaluation of available alternatives while the intuitive style is largely based on feelings and assumptions. Dependent decision-makers typically search for advice and direction from others while avoidant decision-makers try to avoid any conflict completely in the decision-making process (Scott & Bruce, 1995). Wang & Ruhe (2007) provide another categorisation of decision-making styles into intuitive, empirical, heuristic, and rational. They define intuitive and empirical styles as both relying on human intuitive cognitive psychology without applying rational decision-making models. Examples of these styles are arbitrary decisions based on familiarity, personal preferences or the common sense for intuitive decisions, and experimentation, experience or consultation for empirical decision strategies. The heuristic style refers to probability judgement and can be linked to cognitive psychology and AI (Wang & Ruhe, 2007). Two contrasted styles are rational and intuitive. Rational decision-making is based on facts and information, an analytical process, and a step-by-step procedure to arrive at a decision and is therefore considered an advanced model. In contrast, the intuitive style is identified as representing the gradual adaptations grounded in a profound and intimate understanding of the situation. Intuition is a synthetic psychological function that grasps the entirety of a given situation. It is often linked to having a hunch or a strong sense of knowing what is likely to happen. The intuitive model happens often daily and is seen as basic due to its reliance on predisposition (Uzonwanne, 2016). Intuition involves reliance on judgement, experience, and the use of gut feeling. Intuitive decision-making can support the rational, strategic approach by recognising the value of relationships and the social environment as key factors in decision-making and allowing a more holistic view of an issue by including personal experiences and an increased capability for processing information (Dijksterhuis, 2004; Whitworth et al., 2000).

Apart from the distinction of decision-making styles, researchers have also produced formalisations of the decision-making process including variations depending on the decision-making style, namely, rational and intuitive. These decision-making processes define the steps of problem identification, development and evaluation of alternatives, decision-making, implementation of the decision, and evaluation after the decision (Lancaster & Lancaster, 1982; Lunenburg, 2010). However, decision-making can falter under stress, with rational processes hindered by incomplete information and cognitive biases, leading to disastrous consequences. Examples such as the USS *Vincennes* incident illustrate the importance of balancing decision-making styles and refining processes to minimise errors in high-pressure scenarios, particularly where teams work together (Tingle, 2018).

Relating these theories of decision-making to the cybersecurity domain and CTI-supported decision-making, the rational style corresponds to the decision-making process informed and influenced by CTI due to its information and fact-based approach. The use of CTI allows for a rational and informed decision, provided that the available CTI information is correct and provided at the right time to support a timely decision. However, the intuitive model also applies in cybersecurity because if the decision-maker has great experience in defending against cyber-attacks and therefore has a high level of tacit knowledge, they may tend to decide intuitively based on their gut feeling and experience of dealing with similar situations. Therefore, in certain situations, the formal CTI and the rational decision-making may be combined with or superseded by a deep and intimate understanding of the situation and an intuitive approach. In the event a decision must be made quickly, the intuitive decision may be the only viable option if relevant additional information is not at hand. Thus, both the rational and intuitive decision-making styles must be considered in cybersecurity and this research.

2.2 Measures and Metrics

The term "metrics" is defined as standards of measurement that assess efficiency, performance, progress, or quality (Papazov, 2019). While metrics as quality measures can be subjective and abstract depending on the context, they are composed of concrete, objective attributes that measure data points which are defined as "measures". Thus, the accuracy of a metric depends on the accuracy of the attributed measures. It is also important to put the measures into context, as metrics are meaningless on their own (Black et al., 2008).

In cybersecurity, metrics help organisations assess their cybersecurity posture, verify security controls towards compliance, identify strengths and weaknesses as well as security trends inside and outside of the organisation (Black et al., 2008). While numerous metrics have already been developed for cybersecurity areas such as network security or cyber resilience, the field of cybersecurity metrics is still in its infancy, as it is not always clear which metrics and measures organisations should collect (Enoch et al., 2018; Linkov et al., 2013; Papazov, 2019). In general, quality metrics should enable informed decision-making regardless of the area in which they are applied (Black et al., 2008). With regards to CTI-informed decision-making, it is important to assess not only the quality of the decision-making, which is the subject of this paper but also the quality of the CTI, which is discussed in the next section.

2.3 CTI Quality

The quality of CTI, and the associated feeds, is thought to be an important factor when it comes to an organisation's efficacy in defending against cyber threats. Good quality CTI has been defined as actionable, timely, relevant, accurate, complete and verified (Knerler et al., 2022; Oosthoek & Doerr, 2021). Further, CTI quality has been defined in terms of the data quality dimensions associated with relevance, timeliness and accuracy (Knerler et al., 2022) and attributes such as reputation and concise representation (Oosthoek & Doerr, 2021; Rüedlinger et al., 2024). The appropriate amount of data, representational consistency, objectivity and schema completeness have also been added to the list of quality attributes (Schlette et al., 2021). Griffioen et al (2020) define a similar taxonomy using timeliness, sensitivity, originality, and impact as metrics to measure the quality of CTI feeds. Their research grounds these dimensions on the STIX format which defines domain objects for easier categorisation and chaining of CTI information (OASIS Open, 2017). However, researchers acknowledge that identifying data quality dimensions and metrics is just a first step in defining CTI quality metrics and that further research is required (Schlette et al., 2021). To go beyond this first step, this paper focuses on the valorisation of CTI - how its effectiveness in informing cybersecurity decision-making can be assessed.

3. Measuring the Quality of CTI-Informed Decision-Making

Having noted the dearth of instruments for assessing CTI-informed decision-making, this research sought to develop relevant objective measures and metrics and a perception-based performance scale.

The proposed instruments were developed in the context of a broader research project that has developed a gamification solution to enhance the use of CTI for the cyber resilience of satellite cyber supply chains. This gamification solution provides a suitable platform for experimenting with and investigating the effectiveness of CTI for improved cybersecurity decision-making because it is built on a predetermined scenario, synthesised from two real-world cyber-attacks to represent the flow of a holistic satellite cyber supply chain attack step-by-step. The gamification solution has also defined a decision tree that tracks the decision options, relevant CTI generated using open-source platforms such as MITRE ATT&CK and DEF3ND (The MITRE Cooperation, 2025b, 2025a), and the outcomes of the decisions made. In this way, the platform allows for a granular assessment and monitoring of user's decisions, their use of CTI, and the outcome from every decision point. The detailed discussion of the gamification platform is beyond the scope of this paper.

3.1 Proposed Measures and Metrics

Inspiration for the metrics development is from the machine learning (ML) domain. In ML, the confusion matrix provides the foundation to measure classification errors to evaluate the performance of the applied model (Beauxis-Aussalet & Hardman, 2014). Based on true and false positives, and false and true negatives, the precision, sensitivity, accuracy and further metrics are calculated (Naidu et al., 2023). Like the ML metrics, decision measures must be defined to provide a foundation for CTI use metrics.

The proposed measures are defined at the level of the individual decisions and categorised into three decision values - good, neutral and bad - depending on the impact of the decision on how the attack further unfolds. A good decision improves the cyber-attack scenario, while a bad one worsens it, such as enabling further attacker foothold or progression. Neutral decisions represent a middle ground, with no significant impact on the

situation. The delineation of good, bad and neutral is important from the perspective of representing goal-based decision-making, where each decision should theoretically bring the goal closer to realisation. At each decision point, CTI utilisation is recorded as either affirmative or negative. Based on these, the following four measures were defined:

- Informed Decisions (ID) - representing a good decision value and a positive CTI utilisation;
- Favourable Decisions (FD) - for a good decision value but a negative CTI utilisation;
- Misguided Decisions (MG) - if the decision value is bad but CTI has been used; and
- Poor Decisions (PD) - for a bad decision value and a negative CTI utilisation.

Since neutral decisions do not lead to the worsening of the cyber-attack, they are counted towards ID and FD accordingly – see Table 1.

The defined measures link back to the decision-making styles of rational and intuitive decision-making. A rational decision-maker will most likely end up with a high amount of ID decisions as CTI will presumably be consumed and ideally lead to a good decision while an intuitive decision-maker, relying on tacit knowledge, experience, and intuition, will be more likely to ignore the option of CTI information when deciding how to react to the posed cyber-attack scenario. However, luck can also lead to FD as an outcome. Misguided decisions might be the result of rational decision-making if CTI was misinterpreted but still consumed. Poor decisions can be linked back again to the intuitive approach but with presumably less knowledge and experience or just based on guesswork.

Table 1: Decision-Level Measures

		CTI Utilisation	
		Affirmative (A)	**Negative (N)**
Decision Value	Good (G)	Informed Decision (ID)	Favourable Decision (FD)
	Neutral	Informed Decision (ID)	Favourable Decision (FD)
	Bad (B)	Misguided Decision (MD)	Poor Decision (PD)

Similar to the confusion matrix, the measures ID, FD, MD, and PD have been defined to provide a foundation for calculating metrics for measuring the decision quality with and without CTI. The following metrics were developed to measure the decision quality, expressing the correlation between CTI use and decision outcomes:

CTI Success Rate (CSR):

The *CTI Success Rate* evaluates how many good (and neutral decisions) were made based on the utilisation of CTI. A high *CTI Success Rate* indicates that CTI utilisation was critical and pivotal for good decision-making and suggests that the CTI was highly actionable. This metric shows how helpful CTI is in the overall decision-making process and answers the question *"How much does CTI use contribute to good decisions?"*.

$$\boldsymbol{CTI\ Success\ Rate} = \frac{\boldsymbol{ID}}{(\boldsymbol{ID} + \boldsymbol{FD})}$$

CTI Impact Score (CIS):

The CTI Impact Score measures how much CTI led to a positive decision outcome compared to misguided decisions. A higher score indicates that CTI positively contributes to good decision-making with a low likelihood of misuse or misunderstanding of CTI. A low score indicates difficulties with the utilisation of CTI for decision-making. It can indicate issues with CTI quality, and the decision-maker's ability to interpret and act on CTI, and therefore hints at opportunities to improve the integration of CTI in the decision-making process. The metric answers the question "How effective is CTI utilisation in decision-making?".

$$\boldsymbol{CTI\ Impact\ Score} = \frac{\boldsymbol{ID}}{(\boldsymbol{ID} + \boldsymbol{MD})}$$

CTI Utilisation Accuracy (CUA):

The *CTI Utilisation Accuracy* is the overall decision accuracy based on whether CTI has been involved in the decision-making process or has been ignored. It shows the overall reliance or non-reliance on CTI. A higher accuracy indicates that CTI effectively supports the decision-making towards a positive outcome and vice versa. The *CTI Utilisation Accuracy* metric answers the question *"How well does CTI support the decision-making process overall?"*.

$$\boldsymbol{CTI\ Utilisation\ Accuracy} = \frac{(\boldsymbol{ID} + \boldsymbol{PD})}{(\boldsymbol{ID} + \boldsymbol{FD} + \boldsymbol{MD} + \boldsymbol{PD})}$$

TI Decision Misalignment Rate (CDMR):

The *CTI Decision Misalignment Rate* represents the misguided decisions that produced an unintended outcome and reflects the degree to which the application of CTI failed within the decision-making process. The metric measures the percentage of scenarios that failed because CTI was either misunderstood or not used in decision-making. The lower the value the better CTI was incorporated into the decision-making process. It analyses the impact of CTI misuse or non-utilisation on the overall decision quality and answers the question *"How poorly did the utilisation or ignorance of CTI influence the decision quality?"*.

$$\boldsymbol{CTI\ Decision\ Misalignment\ Rate} = \frac{(\boldsymbol{FD} + \boldsymbol{MD})}{(\boldsymbol{ID} + \boldsymbol{FD} + \boldsymbol{MD} + \boldsymbol{PD})}$$

To validate and illustrate the utility of these metrics, they have been calculated and interpreted for a selection of participants from the testing of the CTI gamification platform – see Table 2.

Table 2: Participant decision data and metrics

Participant Number	ID	FD	MD	PD	CSR	CIS	CUA	CDMR
46	5	5	2	1	0.5	0.7143	0.4615	0.5385
52	4	4	2	3	0.5	0.6667	0.5385	0.4615
53	10	0	2	4	1	0.8333	0.8750	0.125
55	1	4	0	6	0.2	1	0.6364	0.3636
56	6	10	0	5	0.375	1	0.5238	0.4762

The table displays five different participants from the experiment, with data representing first the cumulated numbers for each measure and then the calculated metrics in order. While Participant 53 shows a rather ideal version of CTI utilisation with high scores of CSR, CIS and CUA, and a low score for CDMR, participants 55 and 56 are interesting considering the low value for CSR but a perfect score for CIS. Both participants sparsely used CTI but when they did it turned out to be beneficial hence the high CIS value. However, many decisions were made without the use of CTI with positive results still achieved, which led to a low CSR value but relatively high CUA as FDs were still high.

Medium scores for all four metrics can be found for participants 46 and 52. Both participants show an even distribution between the successful use of CTI and favourable decisions without CTI. A similar picture emerges for MD and PD.

3.2 CTI Performance Scale

The measures and metrics defined above provide an objective assessment of CTI use for decision-making. Along with these, a scale is proposed that complements the metrics to provide an evaluation of CTI use reflecting the user experience. The proposed scale focuses on the user perception of how the provided CTI has informed their decision-making along several attributes. Inspiration for this scale is from the System Usability Scale (SUS) developed by Brooke (1996) which measures the perceived usability of a system by rating the level of agreement of statements that target usability characteristics (Drew et al., 2018). While the usability scale focuses on effectiveness, satisfaction and efficiency, the proposed CTI performance scale (CPS) applies CTI quality dimensions of relevance, accuracy, appropriate amount of data, concise representation, and completeness along with usability attributes related to effectiveness and efficiency (Brooke, 1996; Schlette et al., 2021). The user can rate their subjective experience from "S*trongly Disagree*" to "*Strongly Agree*" on a 5-point Likert scale. The statements in the scale are presented in Table 3.

Table 3: CTI performance scale

Statements	Strongly Disagree				Strongly Agree
1. The provided CTI was relevant to the scenario.	1	2	3	4	5
2. The provided CTI was confusing.	1	2	3	4	5
3. I found the provided CTI helpful and well-presented.	1	2	3	4	5
4. The provided CTI format was difficult to understand.	1	2	3	4	5
5. I found the CTI precise and informative.	1	2	3	4	5
6. I was overwhelmed by the CTI.	1	2	3	4	5
7. Using CTI facilitated my decision-making.	1	2	3	4	5
8. The provided CTI complicated my decision-making.	1	2	3	4	5
9. The CTI provided a comprehensive picture.	1	2	3	4	5
10. The provided CTI was lacking detail and depth.	1	2	3	4	5

The CTI quality dimensions were mapped to the CPS statements to assess how well each attribute is represented. Impact levels from 1 to 3 were used, with 1 indicating low impact and 3 showing a strong connection. Two researchers independently completed the mapping, discussed and compared their ratings for attribute coverage, and finally aligned their ratings through an inter-rater consensus approach, with the final mapping as shown in Table 4.

Table 4: CTI attribute distribution mapping

Attribute	Relevant	Accurate	Appropriate	Concise	Complete	Effective	Efficient
1	***	*				*	
2			**	**		*	*
3	**	**	***	*		***	*
4			***	*		**	*
5	*	***		**		*	**
6				*		**	***
7	*		**			***	*
8			*			***	**
9	*	*			***		
10			*		***	*	

To calculate the final score and balance out positively and negatively worded experiences, the ranking of the odd-numbered rankings gets reduced by 1 while the even-numbered items get deducted from a value of 5. Like the SUS, the points are rescaled to match a range with points from 0 to 100. The rescaling factor is calculated based on the number of statements and the highest score possible after deduction took place which leads to a rescaling factor of 100 / (10 x 4) = 2.5. Instead of subtracting 1 for positive items and the item value of 5 for negative items, a neutral value can be calculated for neutral answers, which results in a 3 on a scale of 1 to 5. The neutral score is calculated by the number of items times the neutral score minus the subtraction values for positive and negative items which would result in the equation (10 x 3) – 10 = 20 for the CPS. Therefore, the formula to calculate the CPS score is as follows:

$$CPS\ score = 2.5 \times \left(20 + \sum(CPS1 + CPS3 + CPS5 + CPS7 + CPS9) - \sum(CPS2 + CPS4 + CPS6 + CPS8 + CPS10)\right)$$

A high score on the CPS scale indicates a high level of satisfaction and positive experience with CTI-informed decision-making, while a low score indicates dissatisfaction and a negative experience. The final score gives an overall indicator however, the statements and their rating can be used individually to trace back specific issues, e.g., accuracy, effectiveness, or relevance.

4. Discussion

CTI aims to enhance situational awareness and decision-making before, during, and after a cyber-attack (Schlette et al., 2021). The presented research proposes a method to measure and analyse the influence of CTI on decision-making quality. To this end, measures and metrics are proposed to provide an objective foundation to quantify the influence of CTI on decision-making. Using the concept of the confusion matrix, the measures and metrics target the correlation between CTI utilisation and decision-making focusing on input and outcome while acknowledging the presence of further influences on CTI and decision-making which exceed the targeted context, for which future research is intended. A few of these influences are described below.

The interpretation of the metrics is guided by the specific questions each metric is designed to address. Ideally, CSR, CIS, and CUA should be as high as possible, while a lower CDMR indicates better outcomes. While this generally holds in different situations, alternative analyses and interpretations are possible. For instance, considering the intuitive style of decision-making, over time, security analysts may increasingly rely on their experience and tacit knowledge from previous relevant cybersecurity incidents rather than relying solely on CTI. In this sense, the metrics can be used to objectively analyse and characterise the decision-making profiles of individual analysts.

From an organisational perspective, a low aggregate CSR value could allude to the risk appetite or organisational culture where analysts are encouraged to make intuitive decisions. Since the CSR value reflects the overall reliance on CTI rather than the quality of decisions, a low score may be acceptable in this context. Conversely, for organisations that emphasise informed and rational decision-making, and that are perhaps more risk-averse, a high CSR value might be preferable. On the other hand, CIS and CUA scores represent the effective and accurate use of CTI and should consistently remain high to ensure good CTI application. Similarly, a low CDMR value indicates a low error rate and should always remain low.

The metrics provide information on the helpfulness of the CTI but also show whether there are problems with its integration in the decision-making processes. However, due to the subjective character and individuality of human decision-making, the objective, measured metrics have the limitation of not being able to capture the individual's perceptions and experiences. As such a scale has also been suggested to complement the objective view of the measures and metrics. The scale would target the user perception and experience of the CTI-informed decision-making. The scale statements were mapped to CTI quality attributes to illustrate and confirm the broad coverage of each CTI attribute within the scale. The mapping represents the consensus of two researchers. Further testing of the CPS could finetune the statement-attribute-distribution further and improve the statement phrasing and validation by involving more participants. The nuance between the different attributes could also be further refined. For example, the difference between effectiveness and efficiency in this context must be carefully differentiated to avoid misunderstanding. While effectiveness represents the overall applicability and usefulness of CTI, efficiency can best be understood as the impact of CTI utilisation towards making a 'good' decision quicker or better.

The metrics and scales that are proposed in this research are intended to inform assessments of CTI use for decision-making within the cybersecurity domain in general. They have been operationalised and tested through the CTI gamification platform developed in this research. However, one of the challenges of operationalising these metrics broadly is associated with the difficulty of evaluating the quality of decisions as good, neutral or bad. Determining a 'good' decision is nearly impossible in real life, as even the most obvious choices can lead to failure due to imperfect information. A key challenge in this research is the timeliness and context of information evolving into CTI, affecting its usability and effectiveness. A single data point may hold little relevance alone but can gain significance when combined with others. Context adds value to CTI and must be measured to assess its impact on decision-making. Factors like data volume, speed, variety, and reliability influence interpretation, digestion, and decision-making timeframes. Different datasets have varying relevance durations and audience (Oosthoek & Doerr 2021; Schlette et al. 2021). The contextual issue makes research challenging but highlights the importance of generalisable concepts that provide metrics and measures to support deeper and more comprehensive research in the future.

Enhanced CTI quality should improve decision-making, where the information is up-to-date, relevant and not misinformed or deceptive (Oosthoek & Doerr, 2021). Despite these initial criteria, other considerations are necessary. Too much information, presented incorrectly, or delayed, can impact decision-making. The timeliness of CTI and the impact time has on efforts to collect, disseminate, process and use CTI should also be considered. Decisions must meet timeliness criteria, which may happen on a different scale to CTI processes. Understanding

how these factors interact and their impact on efforts to measure CTI is a component of our next research phase. This research did not account for factors like deception or cultural biases that affect CTI and decision-making. Measuring decision success or the impact of specific information is challenging, hindering CTI metric development in the industry. Overemphasis on one metric can lead to counterproductive behaviours, so balanced and broad metrics are preferred. This study proposes a set of metrics that can be refined to improve CTI's role in decision-making.

On the other hand, the scale is immediately and broadly operationalisable; it can be a useful tool for feedback between security analysts and CTI teams and can help improve the delivery and content of CTI. Answers to each statement can point towards issues of the individual CTI attribute. While the scale has not undergone extensive testing, it is anticipated that its increasing adoption and utilisation will contribute to its refinement and adaptation for specific contexts. The complementarity of measures and metrics and the scale offer a more comprehensive and nuanced picture, helping to identify whether issues in decision-making stem from analyst struggles or CTI-related problems.

5. Conclusion

The paper presented a set of instruments for evaluating CTI performance, one providing an objective evaluation using measures and metrics and the other providing a subjective perception-based scale helping to create a holistic picture of quantitative performance and feedback-providing insights into the applied CTI. A broader framing has been provided explaining the implications of decision-making theory, quality metrics and the challenges faced in measuring CTI quality. Limitations and constraints include the limited testing of the metrics and the scale. However, the utility of the metrics has been demonstrated through their operationalisation in the CTI gamification platform developed in this research. Further experiments are necessary to validate the overall validity, ideally in a real-world setting to enable their generalisability and finetuning. Further activities in this project will focus on refining CTI quality requirements and further using the metrics to analyse the influence of CTI on creating cyber-resilient environments.

6. Acknowledgement

This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

References

Ampel, B. M., Samtani, S., Zhu, H., Chen, H., & Nunamaker, J. F. (2024). Improving Threat Mitigation Through a Cybersecurity Risk Management Framework: A Computational Design Science Approach. *Journal of Management Information Systems*, *41*(1), 236–265. https://doi.org/10.1080/07421222.2023.2301178

Beauxis-Aussalet, E., & Hardman, L. (2014). *Visualization of Confusion Matrix for Non-Expert Users.*

Black, P. E., Scarfone, K., & Souppaya, M. (2008). Cyber Security Metrics and Measures. In J. G. Voeller, *Wiley Handbook of Science and Technology for Homeland Security* (1st ed., pp. 1–15). Wiley. https://doi.org/10.1002/9780470087923.hhs440

Brooke, J. (1996). SUS: A quick and dirty usability scale. *Usability Evaluation in Industry*.

Dijksterhuis, A. (2004). Think Different: The Merits of Unconscious Thought in Preference Development and Decision Making. *Journal of Personality and Social Psychology*, *87*(5), 586–598. https://doi.org/10.1037/0022-3514.87.5.586

Drew, M. R., Falcone, B., & Baccus, W. L. (2018). What Does the System Usability Scale (SUS) Measure?: Validation Using Think Aloud Verbalization and Behavioral Metrics. In A. Marcus & W. Wang (Eds.), *Design, User Experience, and Usability: Theory and Practice* (Vol. 10918, pp. 356–366). Springer International Publishing. https://doi.org/10.1007/978-3-319-91797-9_25

Enoch, S. Y., Ge, M., Hong, J. B., Alzaid, H., & Kim, D. S. (2018). A systematic evaluation of cybersecurity metrics for dynamic networks. *Computer Networks*, *144*, 216–229. https://doi.org/10.1016/j.comnet.2018.07.028

Griffioen, H., Booij, T., & Doerr, C. (2020). Quality Evaluation of Cyber Threat Intelligence Feeds. In M. Conti, J. Zhou, E. Casalicchio, & A. Spognardi (Eds.), *Applied Cryptography and Network Security* (Vol. 12147, pp. 277–296). Springer International Publishing. https://doi.org/10.1007/978-3-030-57878-7_14

Knerler, K., Parker, I., & Zimmermann, C. (2022). *11 Strategies of a World-Class Cybersecurity Operations Center*. The MITRE Corporation.

Lancaster, W., & Lancaster, J. (1982). Rational Decision Making: Managing Uncertainty. *JONA: The Journal of Nursing Administration*, *12*(9), 23–28. https://doi.org/10.1097/00005110-198209000-00007

Linkov, I., Eisenberg, D. A., Plourde, K., Seager, T. P., Allen, J., & Kott, A. (2013). Resilience metrics for cyber systems. *Environment Systems and Decisions*, *33*(4), 471–476. https://doi.org/10.1007/s10669-013-9485-y

Lunenburg, F. C. (2010). The decision making process. *National Forum of Educational Administration & Supervision Journal*, *27*(4).

Naidu, G., Zuva, T., & Sibanda, E. M. (2023). A Review of Evaluation Metrics in Machine Learning Algorithms. In R. Silhavy & P. Silhavy (Eds.), *Artificial Intelligence Application in Networks and Systems* (Vol. 724, pp. 15–25). Springer International Publishing. https://doi.org/10.1007/978-3-031-35314-7_2

OASIS Open. (2017, 2024). *Introduction to STIX*. https://oasis-open.github.io/cti-documentation/stix/intro.html

Oosthoek, K., & Doerr, C. (2021). Cyber Threat Intelligence: A Product Without a Process? *International Journal of Intelligence and CounterIntelligence*, *34*(2), 300–315. https://doi.org/10.1080/08850607.2020.1780062

Papazov, Y. V. (2019). Cybersecurity Metrics. *NATO Science and Technology Organization (STO)*, 1–18.

Rüedlinger, A., Klauser, R., Lamprakis, P., Happe, M., Tellenbach, B., Veyisoglu, O., & Trammell, A. (2024). FeedMeter: Evaluating the Quality of Community-Driven Threat Intelligence: *Proceedings of the 10th International Conference on Information Systems Security and Privacy*, 54–66. https://doi.org/10.5220/0012357600003648

Schlette, D., Böhm, F., Caselli, M., & Pernul, G. (2021). Measuring and visualizing cyber threat intelligence quality. *International Journal of Information Security*, *20*(1), 21–38. https://doi.org/10.1007/s10207-020-00490-y

Scott, S. G., & Bruce, R. A. (1995). Decision-Making Style: The Development and Assessment of a New Measure. *Educational and Psychological Measurement*, *55*(5), 818–831. https://doi.org/10.1177/0013164495055005017

Shin, B., & Lowry, P. B. (2020). A review and theoretical explanation of the 'Cyberthreat-Intelligence (CTI) capability' that needs to be fostered in information security practitioners and how this can be accomplished. *Computers & Security*, *92*, 101761. https://doi.org/10.1016/j.cose.2020.101761

Svenson, O. (1979). Process descriptions of decision making. *Organizational Behavior and Human Performance*, *23*(1), 86–112. https://doi.org/10.1016/0030-5073(79)90048-5

The MITRE Cooperation. (2025a). MITRE ATT&CK. *ATT&CK Matrix for Enterprise*. https://attack.mitre.org/

The MITRE Cooperation. (2025b). MITRE D3FEND. *D3FEND - A Knowledge Graph of Cybersecurity Countermeasures*. https://d3fend.mitre.org/

Tingle, A. (2018, July). *The Human-Machine Team Failed Vincennes*. U.S. Naval Institute. https://www.usni.org/magazines/proceedings/2018/july/human-machine-team-failed-vincennes

Uzonwanne, F. C. (2016). Rational Model of Decision Making. In A. Farazmand (Ed.), *Global Encyclopedia of Public Administration, Public Policy, and Governance* (pp. 1–6). Springer International Publishing. https://doi.org/10.1007/978-3-319-31816-5_2474-1

Wang, Y., & Ruhe, G. (2007). The Cognitive Process of Decision Making: *International Journal of Cognitive Informatics and Natural Intelligence*, *1*(2), 73–85. https://doi.org/10.4018/jcini.2007040105

Whitworth, B., Van De Walle, B., & Turoff, M. (2000). Beyond Rational Decision Making. *Group Decision and Negotiation*.