Computers & Security 140 (2024) 103795

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

Ethical principles shaping values-based cybersecurity decision-making

Joseph Fenech [a], Deborah Richards [a,*], Paul Formosa [b]

[a] School of Computing, Macquarie University, Australia
[b] Department of Philosophy, Macquarie University, Australia

ARTICLE INFO

Keywords:
Cybersecurity sensitive decision-making
Ethical principles
Ethics training
Profiling
Human agency

ABSTRACT

The human factor in information systems is a large vulnerability when implementing cybersecurity, and many approaches, including technical and policy driven solutions, seek to mitigate this vulnerability. Decisions to apply technical or policy solutions must consider how an individual's values and moral stance influence their responses to these implementations. Our research aims to evaluate how individuals prioritise different ethical principles when making cybersecurity sensitive decisions and how much perceived choice they have when doing so. Further, we sought to use participants' responses to cybersecurity scenarios to create profiles that describe their values and individual factors including personality. Participants ($n = 193$) in our study responded to five different ethically sensitive cybersecurity scenarios in random order, selecting their action in that scenario and rating and ranking of the ethical principles (i.e., Beneficence, Non-Maleficence, Justice, Autonomy, Explicability) behind that action. Using participants' demographics, personality, values, and cyber hygiene practices, we created profiles using machine learning to predict participants' choices and the principle of most importance to them across scenarios. Further, we found that, generalising, for our participants Autonomy was the most important ethical principle in our scenarios, followed by Justice. Our study also suggests that participants felt they had some agency in their decision making and they were able to weigh up different ethical principles.

1. Introduction

Our increasing online presence results in our greater vulnerability to cybersecurity attacks. Further attention is therefore needed on the decision making that occurs during cybersecurity events. When looking at how cybersecurity systems fail, much of the focus is often directed towards the technical aspects and creating better policies to mitigate that risk, even though cybersecurity attacks continue to target the weakest link in the system, which is usually humans (Anderson, 1993) (Køien, 2019) (Pfleeger et al., 2014). While much research focuses on the issue of cybersecurity professionals and their decision making, we shall instead explore the decision-making processes of untrained individuals to determine how they prioritise different ethical principles when making cybersecurity decisions. Our focus on untrained individuals is important as untrained individuals are commonly victims of cybersecurity events with ethical ramifications, such as ransomware (Hampton and Baig, 2015), and yet there has been little focus on user behaviour in such cases (Anderson, 1993). Further, policies to direct what actions should be taken are unlikely to exist in the case of non-workplace cybersecurity events commonly faced by untrained individuals, which necessitates that they draw on other decision-making frameworks, such as their personal ethical principles. However, these personal ethical principles can conflict with each another in cybersecurity contexts (Formosa et al., 2021). Cybersecurity ethics commonly draws upon a principlist approach, which focuses on specifying and weighing a small group of domain-relevant ethical principles (Christen et al., 2020). To capture ethical reasoning in a cybersecurity context, we utilise the five ethical principles developed in the Principlist framework of Formosa et al. (2021): Beneficence, Non-Maleficence, Justice, Autonomy, and Explicability. Our research aims to evaluate how untrained individuals prioritise these five ethical principles when making cybersecurity sensitive decisions. Focusing on how untrained individuals respond to cybersecurity events assumes those individuals can exercise, and perceive that they can exercise, their agency in such cases. We further aim to verify whether this assumption holds by examining how much perceived choice participants felt they had in our cybersecurity scenarios.

Participants ($n = 193$) in our study responded to five different ethically sensitive cybersecurity scenarios in random order, selecting between two responses (one taking action and the other taking no

[*] Corresponding author.
E-mail address: deborah.richards@mq.edu.au (D. Richards).

https://doi.org/10.1016/j.cose.2024.103795
Received 17 July 2022; Received in revised form 2 February 2024; Accepted 28 February 2024
Available online 2 March 2024
0167-4048/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

action) to that scenario and the rating and ranking for each of the five ethical principles (i.e., Beneficence, Non-Maleficence, Justice, Autonomy, Explicability) behind that choice. Using participants' demographics, personality, values, and cyber hygiene practices, we created profiles using machine learning to predict which single ethical principle is generally most important to an individual based on their responses across the five scenarios. Further, we found that, generalising, for our participants Autonomy was the most important ethical principle in our scenarios, followed by Justice. Our study also suggests that participants felt they had some agency in their decision making and they were able to weigh up different ethical principles. We conclude the paper with consideration of the limitations of our study and areas for future research.

2. Literature review

Several ethical issues arise in cybersecurity (Abomhara and Køien, 2015; Christen et al., 2020; Manjikian, 2018). While professionals in the ICT space are expected to follow a code of ethics, such as the ACM code of ethics that expresses general ethical principles and professional responsibilities (ACM, 2018), no such code exists for ICT end users. We therefore need to turn to ethical principles to deal with such cases. Principlism is the most common approach in applied ethics, and is used extensively in related fields such as bioethics (Beauchamp and Childress, 2001) and artificial intelligence (Floridi et al. (2018)). The widely-used AI4People Framework's five ethical principles of Beneficence, Non-Maleficence, Autonomy, Justice and Explicability has recently been ported into the cybersecurity ethics domain by Formosa et al. (2021) and used in a recent review of the ethical issues raised for cybersecurity by quantum computing (Coates et al., 2023). These principles are summarised in Table 1. Loi and Christen (2020) liken these principles to prima facie duties (Ross, 2002), where stronger prima facie duties can overrule weaker prima facie duties on a case by case basis. This hierarchy can be used to deal with conflicts in a cybersecurity domain between the principles and allow an agent to balance the conflicting principles to find the ethically right resolution.

End users face a range of ethical dilemmas, where different ethical principles compete with one another, relating to ICT use. For example, online hate is harmful to its victims (Non-Maleficence) and can involve targeted discrimination of vulnerable groups (Justice) (Awan, 2014), however its perpetrators may enjoy it (Beneficence), and attempts to limit it can infringe on freedom of expression (Justice) and individual choice (Autonomy) (Ullmann and Tomalin, 2020) and may involve the use of non-transparent surveillance (Explicability). Given that many non-technical ICT users will be subject to cybersecurity threats and events, such as being subject to online hate or suffering a ransomware attack, that have ethical ramifications, our first research question (***RQ1***) is: *Can users identify, apply and order appropriate ethical principles in cybersecurity scenarios involving ethical dilemmas?*

Table 1
Ethical Principles.

Beneficence	"Cybersecurity technologies should be used to benefit humans, promote human well-being, and make our lives better overall."
Non-Maleficence	"Cybersecurity technologies should not be used to intentionally harm humans or to make our lives worse overall."
Autonomy	"Cybersecurity technologies should be used in ways that respect human Autonomy. Humans should be able to make informed decisions for themselves about how that technology is used in their lives."
Justice	"Cybersecurity technologies should be used to promote fairness, equality, and impartiality. It should not be used to unfairly discriminate, undermine solidarity, or prevent equal access."
Explicability	"Cybersecurity technologies should be used in ways that are intelligible, transparent, and comprehensible, and it should also be clear who is accountable and responsible for its use."

SOURCE: All material quoted from Formosa et al. (2021).

However, since these principles can conflict with one another, a related question is how those principles get ordered in cases of conflict. We thus investigate what factors might impact how participants will order the ethical principles explored by RQ1. Firstly, researchers have found a range of human factors that influence individual ethical decision-making in cybersecurity, including gender, age, education, experience (Hoonakker et al., 2009), culture (Kharlamov and Pogrebna, 2019), cyber hygiene (Vishwanath et al., 2020), personality (Gratian et al., 2018), moral foundations (Pfleeger et al., 2014), and human values (Kharlamov and Pogrebna, 2019). Vishwanath et al. (2020) found, using the Cyber Hygiene Inventory (CHI), that cyber hygiene can influence cybersecurity decision making by virtue of an individual being more aware of threats. Cyber hygiene differs from cybersecurity policy in that while both are a set of practices one should take to prevent cyber threats, cyber hygiene is a personal factor while policy is an organisational factor (Maennel et al., 2018). Personality measured as the Big Five of Agreeableness, Conscientiousness, Emotional Stability, Extraversion, and Openness to Experience have been shown to influence privacy decision making. Junglas et al. (2008) found that Conscientiousness and Openness to Experience have a positive influence on whether an individual is concerned about privacy, while high Agreeableness has a negative influence on privacy. Conscientious individuals are noted to be more concerned about potential privacy threats, whereas Openness to Experience typically aligns with having high awareness and being more sensitive to threats. Agreeableness indicates that an individual might be more trusting of others than they should be when it comes to privacy (Junglas et al., 2008). More recently, Gratian et al. (2018) also found that personality influences cybersecurity behaviour and decision making. We capture personality in our study using the Ten-Item-Personality-Indicator (TIPI) (Gosling et al., 2003).

The five constructs (Care, Fairness, Ingroup, Authority, and Purity) captured in the Moral Foundations Questionnaire (MFQ), based on Moral Foundations Theory (MFT) (Graham et al., 2011), have been used to predict the frequency of visiting certain website domains (Kalimeri et al., 2019). Those who had notable MFT scores, such as a low Authority, could be used to predict that the individual would visit progressive news sources. MFT has been linked with building stronger security cultures where each dimension promotes behaviour that improves the cybersecurity within an organisation (Pfleeger et al., 2014). Another measure of morality is a human values-based framework for cybersecurity regulation and governance (Kharlamov and Pogrebna (2019)). This draws on the Schwartz Value Survey (SVS) Schwartz (2007) of 57 values which can be grouped together in various ways. Cultures are separated into cooperative and competitive to identify risk aversion or acceptance of those cultures respectively. Cooperative cultures have low commitment to regulate cybersecurity, while competitive cultures have a high commitment to regulate cybersecurity. To capture this aspect, we use the revised version of the Portrait Values Questionnaire PVQ-RR (Schwartz, 2016). Given the various impacts these measures have been shown to have on cybersecurity behaviours and building on our focus from RQ1, we explore the research subquestion (***RQ2.1***): *"What features predict which ethical principles individuals prioritise in different cybersecurity scenarios involving ethical dilemmas?"*

Secondly, we also investigate human agency, as many cybersecurity events are caused by human actions rather than failures of the technology. Moral agency affects decision making (Bandura, 2006), and the desire to act can be suppressed when the action opposes one's moral values or promoted when the action is supported by one's moral values. The Theory of Planned Behaviour (TPB) (Fishbein and Ajzen, 2011) identifies that actions are driven not only by the intention to undertake the action, but also by one's level of control within that context. Individuals untrained in cybersecurity are likely to have lower perceptions of control due to having lower perceived self-efficacy in the cybersecurity decision making environment. Agency relies on self-efficacy to motivate an individual to act as they must believe their action can impact on the environment or they will lose motivation to act (Gerber and Rogers, 2009). This leads to our last research question (***RQ2.2***): "*Does the level of perceived agency influence individual decision-making in different cybersecurity scenarios involving ethical dilemmas?*"

3. Methodology

3.1. Design

To answer RQ1, we created a set of scenarios that involve ethical dilemmas in cybersecurity contexts. Scenarios are commonly used in research as a method to explore complex issues by creating an environment that a participant can understand without needing a technical background (Ramirez et al., 2015). Scenarios provide a structure that enables the discovery of unobserved features and by laying out all the context of a scenario, ambiguity is reduced (Ramirez et al., 2015). If ambiguity is reduced, then participants have more comparable contextual knowledge, and this can lead to improved reliability when comparing between participants. In the cybersecurity space, the CANVAS project (CANVAS, 2020) uses a scenario-based approach for case studies and their materials helped to direct our decision to implement scenarios. The content of our scenarios was chosen to provide unique situations that could plausibly impact an individual with no particular cybersecurity expertise during their online activities and involved the cybersecurity triad of confidentiality, integrity, and availability (Brey, 2007).

To answer RQ2.2*,* we utilised four inventories to collect features that have been found to influence ethical and/or cybersecurity behaviour. Following from our literature review, we collected data on human values (Schwartz, 2016), moral foundations (Pfleeger et al., 2014), personality indicators (Junglas et al., 2008), and cyber hygiene (Vishwanath et al., 2020) practices.

To answer RQ2.2, we collected data on perceived agency in each scenario by asking the following question, "How much of a choice did you feel you had in the above scenario?", on a five-point scale from none at all to a great deal.

Note that we designed our study around answering these three research questions, rather than testing hypotheses, as there are no previous studies asking for ethical principles to be identified in the context of cybersecurity scenarios or which indicate how individual factors might influence participants' prioritisation of these principles.

3.2. Materials

For our online study we created five scenarios as a contextual baseline for participants. We followed a widely adopted approach for embedding scenarios in surveys known as the Experimental Vignettes Methodology (EVM) (Aguinis and Bradley, 2014). The use of scenario-based instruments, rather than item-based instruments, is common practice in conducting behavioural assessments, as in the case of the Cybersecurity Judgment Questionnaire where real-life scenarios capture and assess correctness of individuals' cybersecurity judgments (Yan et al., 2018). By using descriptive (not just single sentence) real-world scenarios, Yan et al. (2018) argue that the approach draws on the benefits of item-based approaches in being able to efficiently access participants with minimal intrusion while also being able to achieve good ecological validity (Chaytor and Schmitter-Edgecombe, 2003) and incorporate the real world complexity of cybersecurity (Dunn Cavelty, 2014). Real reported cases and cases from the CANVAS project were used as inspiration for building our scenarios.

We used a consistent design for each scenario to make them comparable. Scenarios were designed to: be distinct, plausible and engaging; not rely on technical knowledge and be easy to understand; implicate ethical principles relevant to cybersecurity; consider cybersecurity services involving Confidentiality, Integrity, and Availability (Brey, 2007) equitably; support both sides of the decision with at least two ethical principles supporting each possible choice; contain mutually exclusive and invertible decisions i.e., they act or don't act. All scenarios followed a common structure: background context; describe change; introduce cybersecurity concerns; describe trigger that forces a decision; clarify relevance of each choice; and clarify pros and cons of each choice.

Five ethical scenarios were developed, each exploring a different cybersecurity event: spoofing identities (Misinformation), data breaches (Credentials), suffering a ransomware attack (Ransomware), privacy breaches (Health data), and two factor authentication (2FA). The scenario result subsections below provide a brief description of each scenario together with the specific ACT and DON'T ACT options and the statements provided to participants designed to embed the supporting principles. Table 2 shows the mapping for the five principles where two principles support the ACT option, two different principles support the DON'T ACT option and the remaining principle is irrelevant in the context of the scenario description and action options provided.

Fig. 1 provides a full example of one scenario with participant options. Full scenario descriptions for the remaining four scenarios appear in Appendix A.

3.3. Procedure and data collection

The following procedure was approved by our University's Human Ethics Committee (Approval Number: withheld for double-blind review). We recruited participants from two accessible and "untrained" groups: our university's Psychology participant pool, where students can sign up for studies to receive credit, and students enrolled in an introductory first year unit on cybersecurity who might have been interested in our study. However, as noted in Section 4.1, most participants in this study were recruited via the Psychology pool.

The online procedure is outlined in Fig. 1. Participants first provide informed consent, complete demographic information on their age, gender, cultural identification (based on Australian Bureau of Statistics classifications (ABS, 2019)), area of study, and their perceived ethical knowledge in IT. Next, participants undergo "ethical sensitisation" by reading descriptions and examples of the five ethical values of Beneficence, Non-Maleficence, Justice, Autonomy, and Explicability to establish a common understanding for the participants to use for justifying their decisions.

Participants then receive our five scenarios in a random order. Fig. 1 shows the ransomware scenario and associated questions. As shown, after reading the scenarios, a participant is asked to decide between two options (involving taking action or not taking action) and enter an optional free-text response to justify their decision. The participant next reviews five statements, one for each ethical principle, and indicates the relative importance and ranking of each statement for that scenario, allowing us to cross-validate their choices. Each scenario then asks participants to reflect on whether they felt they had choice when making their decision. After each scenario, participants received the next random scenario until all five scenarios had been completed.

Finally, we collect the profiling data including: Cyber Hygiene Inventory (CHI) (Vishwanath et al., 2020) with 18 items measure on 5-pt Likert scale (never – always) that form five factors; followed by the 10-Item-Personality-Indicator (TIPI) that uses the Big Five factor dimensions of Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience which are scored on a "7-point scale ranging from 1 (disagree strongly) to 7 (agree strongly)" (Gosling et al., 2003); the Moral Foundations Questionnaire (MFQ) with 20 items that comprise five moral factors of Harm, Fairness, Ingroup, Authority, and Purity that each sum four items on a scale of 0 (Not at all relevant / Strongly Disagree) to 5 (Extremely Relevant / Strongly Agree) to get an overall score ranging from 0 to 20; and finally the Portrait Values Questionnaire (PVQ-RR) by classifying 57 items using a 6-point scale provided ranging from 1 (Not like me at all) to 6 (Very much like me) into their corresponding 19 value types (Schwartz et al., 2012).

Table 2
Ethical principles mapped to each scenario and action (ACT/DON'T ACT/irrelevant).

Principle	ACT		DON'T ACT		Irrelevant
Beneficence	Credentials; 2FA; Ransomware		Misinformation		Health
Non-Maleficence	Credentials		Ransomware; Health; Misinf		2FA
Justice	Misinformation	Health	Ransomware	2FA	Credentials
Autonomy	Health	Ransomware	Credentials	2FA	Misinformation
Explicability	Misinformation	2FA	Credentials	Health	Ransomware

3.4. Data analysis

Initial statistical analysis was performed to generate the mean and standard deviation for every instrument and scenario, separating by the decision to act or don't act. To validate whether participants could identify the appropriate principle connected to the decision to act or don't act, for each principle we utilised independent samples t-test (0.05 sig) to determine if there was a statistically significant difference between the two choices. To determine if there was a significant difference between pairs of principles (i.e. Beneficence vs Autonomy, etc.) we performed an ANOVA followed by Tukey's HSD post hoc analysis (0.05 sig), where significant differences indicate that the ethical principle was deliberately chosen and used in the decision-making process. The comments provided by participants were used to demonstrate additional reasoning that participants may have had when making their decision. For statistical analyses we used IBM SPSS statistics package V.27. To learn which participant features might predict the importance of each of the ethical principles, five models were created by taking the average ranking across all five scenarios for each individual and converting this continuous variable into three classes of High, Medium, and Low relative importance by creating equal sized bins of the mean importance rankings. Due to their comprehensibility, we chose to build decision trees and rules. We chose the industry standard C5.0 classification algorithm that has been found to be robust in handling missing data and a wide range of datasets, support boosting methods to improve accuracy and does not require large dataset or processing times (Kuhn and Johnson, 2013). For machine learning we used IBM SPSS Modeler V.18.

4. Results

This section first presents demographic (4.1), inventories (4.2) and scenario (4.3) results. Section 4.4 presents models for the five principles, followed by our agency results (4.5).

4.1. Demographics

Our study was run between March and the first week of June in 2021. Of the total 216 responses recorded, four responses were removed from the data as they did not indicate consent for use for research. An additional 19 partial responses were dropped as the partial responses did not complete the five scenarios and as such were unusable data, leaving a total of 193 valid responses. We had 128 males, 59 females and 3 other. Most were psychology students (68 %), followed by "Other" (20 %) mainly from Science and Health, Computing (6 %), Arts (4 %) and Business (2 %). Computing students were recruited from the cybersecurity unit and the remainder of the participants came from the psychology pool. The age of participants ranged from 18[1]–51 years, with the median age of 20, a mean of 23.09 and SD of 7.85. Table 3 contains the distribution of cultural demographics, where over half (59 %) of participants identified as Oceanian, which includes Australian. Table 4 shows that 117 (61 %) participants rated themselves as having average or above knowledge in IT ethics.

4.2. Inventories

We calculated Cronbach's alpha (α) to measure the reliability of each of the inventories. The means and standard deviations for the Cyber Hygiene Inventory (α =0.92), Ten-Item-Personality-Indicator (α =0.46), Moral Foundations Questionnaire (α= 0.80), and Schwartz Portrait Values Questionnaire (α=0.92) and are found in Tables 5–8 below. All inventories demonstrated good reliability, except for TIPI which is known to have a low α because of its brevity, but is nevertheless deemed suitable for our study due to its demonstrated optimized validity compared to other substantially longer personality instruments that suffer from timing and fatigue issues (Gosling et al., 2003).

4.3. Scenarios

The distribution of participant responses to ACT or DON'T ACT is provided in Table 9. The following subsections present the results for each scenario. For clarity in Tables 10–14, we show in bold where results for the most important principles match with embedded principles. Tukey HSD results for each scenario appear in Appendix B.

4.3.1. Misinformation

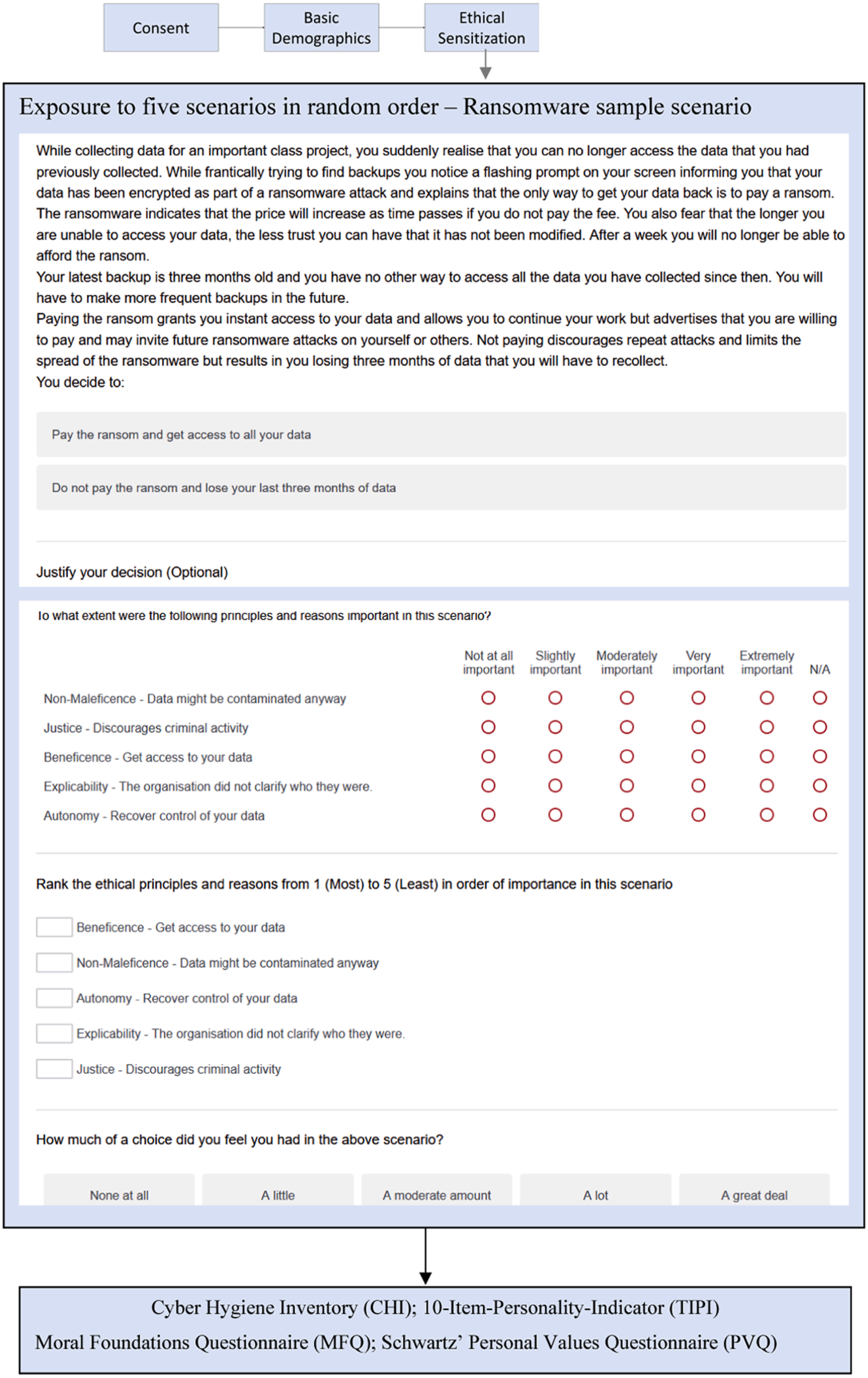
The misinformation scenario involves protesting online, based on misinformation, against a company, causing that company to go bankrupt and its staff to lose their jobs. When the false misinformation is exposed, protesters are asked to take public responsibility for their actions, leading to the choice:

- Publicly acknowledge your involvement and expose your privacy [ACT]
- Do not publicly acknowledge your involvement and protect your privacy [DON'T ACT]

Shown in **bold** in Table 10, consistent with our embedded principles, participants who chose ACT prioritise the importance of Justice and Explicability, while participants who chose DON'T ACT favour the principle of Non-Maleficence. Beneficence (our other embedded principle) is marginally rated higher than Justice and ranks third after Justice for DON'T ACT. T-test results show significant differences between the means for ACT and DON'T ACT options for all principles, except Autonomy, which in this case supported neither action (shown in grey). For Justice and Explicability in the ACT context, and Non-Maleficence in the DON'T ACT context, Tukey HSD Post hoc (found in Appendix B) analysis further confirms consistency between rating and ranking responses and that choices aligned with our design (i.e. our two embedded ACT principles were significantly different to our other principles for that choice and our two embedded DON'T ACT principles were significantly different to our other principles for that choice) and were not due to random chance.

Review of the free-text reasons for choosing a principle reveals that "*privacy*" was a commonly mentioned issue, followed by references to "*family*". Those who chose not to acknowledge their involvement gave variations of reasons aimed at preventing future harm to themselves and/or their families (e.g. "*As much as it is a bad situation I would not expose myself as it may result in serious harm*" (18, Female)) while still being concerned about the ethical ramifications of the decision. Other

[1] One response indicating 14 years of age is considered an invalid response and wasn't used for age analysis.


Consent
Basic Demographics
Ethical Sensitization
Exposure to five scenarios in random order – Ransomware sample scenario

While collecting data for an important class project, you suddenly realise that you can no longer access the data that you had previously collected. While frantically trying to find backups you notice a flashing prompt on your screen informing you that your data has been encrypted as part of a ransomware attack and explains that the only way to get your data back is to pay a ransom. The ransomware indicates that the price will increase as time passes if you do not pay the fee. You also fear that the longer you are unable to access your data, the less trust you can have that it has not been modified. After a week you will no longer be able to afford the ransom.

Your latest backup is three months old and you have no other way to access all the data you have collected since then. You will have to make more frequent backups in the future.

Paying the ransom grants you instant access to your data and allows you to continue your work but advertises that you are willing to pay and may invite future ransomware attacks on yourself or others. Not paying discourages repeat attacks and limits the spread of the ransomware but results in you losing three months of data that you will have to recollect.

You decide to:

Pay the ransom and get access to all your data

Do not pay the ransom and lose your last three months of data

Justify your decision (Optional)

To what extent were the following principles and reasons important in this scenario?

	Not at all important	Slightly important	Moderately important	Very important	Extremely important	N/A
Non-Maleficence - Data might be contaminated anyway	○	○	○	○	○	○
Justice - Discourages criminal activity	○	○	○	○	○	○
Beneficence - Get access to your data	○	○	○	○	○	○
Explicability - The organisation did not clarify who they were.	○	○	○	○	○	○
Autonomy - Recover control of your data	○	○	○	○	○	○

Rank the ethical principles and reasons from 1 (Most) to 5 (Least) in order of importance in this scenario

□ Beneficence - Get access to your data

□ Non-Maleficence - Data might be contaminated anyway

□ Autonomy - Recover control of your data

□ Explicability - The organisation did not clarify who they were.

□ Justice - Discourages criminal activity

How much of a choice did you feel you had in the above scenario?

None at all　A little　A moderate amount　A lot　A great deal

Cyber Hygiene Inventory (CHI); 10-Item-Personality-Indicator (TIPI)
Moral Foundations Questionnaire (MFQ); Schwartz' Personal Values Questionnaire (PVQ)

Fig. 1. Survey Procedure.

viewpoints saw no value in acknowledging their involvement as "*The damage is done*" (32, Male) and "*It would merely be tokenistic*" (20, Male). This contrasts with those who chose to acknowledge their involvement who focused on guilt ("*living in silence would cause the guilt to eat away at me*" (44, Female)), and taking responsibility ("*I would take responsibility for the injustice suffered by the employees and spend my lifetime making it up to my family members*" (24, Male)).

Table 3
Culture.

Culture	N	%
Oceanian (incl. Australian)	109	59 %
North Western European	6	3 %
South Eastern European	10	5 %
North African & Middle Eastern	13	7 %
South East Asian	23	12 %
North East Asian	7	4 %
South and Central Asian	4	2 %
People of the Americas	2	1 %
Sub Saharan African	2	1 %
Unidentified	10	5 %
Total	186	100 %

Table 4
Knowledge in IT Ethics.

Knowledge in IT Ethics	N	Percent
Terrible	11	6 %
Poor	63	33 %
Average	96	50 %
Good	18	9 %
Excellent	3	2 %
Total	191	100 %

Table 5
Cyber Hygiene Inventory - scale 1 (Never) – 5 (Always).

Construct	Mean	SD
Storage and Device hygiene	2.68	1.17
Transmission hygiene	2.52	1.11
Facebook and social media hygiene	3.47	1.04
Authentication and Credential hygiene	2.90	1.09
Email and Messaging hygiene	3.54	1.13

Table 6
TIPI Personality Results – scale 1(strongly disagree)–7(strongly agree).

Dimension	Mean	SD
Extraversion	3.99	1.53
Agreeableness	4.76	1.02
Conscientiousness	4.87	1.29
Emotional Stability	4.04	1.23
Openness to Experiences	4.93	1.27

Table 7
Moral Foundations Questionnaire – Sum range (0–20).

Construct	Sum	SD
Harm	16.41	2.86
Fairness	16.79	2.53
Ingroup	10.16	3.77
Authority	10.92	3.49
Purity	12.64	3.74
MATH	0.47	0.60
GOOD	4.64	0.64

4.3.2. *Credentials*

In the context of an identified breach of a friend's password, the Credentials scenario asks the participant whether they would:

- Attempt to access your friends' social media accounts with the exposed password credentials to change their password without their permission to try to protect their accounts [ACT]
- Do not attempt to access your friends' social media accounts with the exposed password credentials to change their password and thereby leave their accounts potentially exposed [DON'T ACT]

The results presented in Table 11 show that the ratings and rankings support our design for Beneficence, Non-Maleficence and Autonomy, but only ratings are supported for Explicability. T-tests confirm significant differences between ACT and DON'T for all relevant principles, but not for ranking of Justice which is the irrelevant principle in this scenario. Tukey HSD reveal significant differences between most principles, mostly aligning with our design. Anomalous results, such as the significant difference in the T-test for Justice rating and Tukey HSD exceptions will be discussed later.

Most of the justifications provided for accessing the account involved participants viewing themselves as acting in their friends' best interests ("*If they are a good friend, they will understand you were doing what was best for them*" (19, Female)), and preventing something worse from happening ("*Although I am breaching their trust in logging into their accounts, I am effectively doing something than helps prevent something even worse happening to them*" (20, Male)). On the other side, justifications for why an individual would not access the account included protecting themselves from risk ("*as much as i care for my friend it is not right and i can get caught*" (19, Female)), concerns about legality issues ("*It is illegal without their consent. I'd find a way to contact them*" (19, Female)), and privacy concerns ("*the access of another's accounts is a breach of privacy and individuals shouldn't have the right to login to someone else's account without their consent*" (18, Male)).

4.3.3. *Ransomware*

The Ransomware scenario involves deciding within a time limit whether to:

- Pay the ransom and get access to all your data [ACT]
- Do not pay the ransom and lose your last three months of data [DON'T ACT]

Table 12 confirms that the most important and highest ranked principles, respectively, match with our design. T-tests show a significant difference for the four relevant principles between the ACT and DON'T act responses, in alignment with our design. Tukey HSD results confirm significant differences between the principles that support ACT compared with those that support DON'T ACT for both ratings and rankings, with some significant differences between the irrelevant principle (explicability) with other principles.

Justifications for choosing to pay the ransom were primarily concerned with the cost of losing valuable data ("*How could anyone afford to lose 3 months of data that has been painstakingly collected?!*" (15, Female)). Individuals who chose not to pay the ransom instead did not want to support future attacks ("*By paying the ransom, I would feel I have supported the blackmailers.*" (46, Male)), and they also believed the data would have been maliciously modified ("*The data can not be trusted now as it could have been modified therefore either way I will have to recollect the data*" (19, Female)) or were concerned the data would not be returned after paying ("*paying it won't guarantee you get it back*" (30, Female)).

4.3.4. *Health data*

The Health Data scenario involves the use of data collected from wearable fitness trackers to impact insurance costs, and a decision must be made whether to:

- Ignore the petition to acquire cheaper insurance premiums for 'healthy' customers [DON'T ACT]
- Sign the petition to attempt to prevent the use of this data [ACT]

Table 13 reveals that, according to our design, the principles of Justice and Autonomy drove the decision to ACT, however, contrary to

Table 8
Schwartz PVQ Constructs – scale 1(Not like me at all)−6 (Very much like me).

Construct	SDT	SDA	ST	HE	AC	POR	POD	FAC	SEP	SES
Mean	5.08	4.98	4.39	4.96	4.69	3.03	3.49	4.48	4.87	4.21
SD	0.75	0.71	0.93	0.87	0.93	1.11	1.04	0.91	0.78	1.13
Construct	**TR**	**COR**	**COI**	**HU**	**BED**	**BEC**	**UNC**	**UNN**	**UNT**	
Mean	3.42	4.36	4.46	4.71	5.15	5.19	5.23	4.48	5.09	
SD	1.32	1.10	1.07	0.90	0.72	0.68	0.76	1.02	0.74	

SDT – Self Direction Thought, SDA – Self Direction Action, ST – Stimulation, HE – Hedonism, AC – Achievement, POR – Power Resources, POD – Power Dominance, FAC – Face, SEP – Security Personal, SES – Security Societal, TR – Tradition, COR – Conformity Rules, COI – Conformity Interpersonal, HU – Humility, BED – Benevolence Dependability, BEC – Benevolence Caring, UNC – Universalism Concern, UNN – Universalism Nature, UNT – Universalism Tolerance.

Table 9
Distribution of participant responses.

Distribution	Misinformation	Credentials	Ransomware	Health Data	2FA	Total
ACT	96	51	34	154	113	448
DON'T ACT	97	142	159	39	80	517

Table 10
Misinformation: Ratings - *not at all (1) to extremely (5) important; Rankings- most (1) to least (5) important.*

	ACT				DON'T ACT				T-TEST	
	Rating		Ranking		Rating		Ranking		Rating	Ranking
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	p-value	p-value
Beneficence	2.88	1.12	4.29	1.02	**3.79**	**1.02**	3.22	1.31	<0.001*	<0.001*
Non-Maleficence	3.65	0.97	3.23	1.12	**4.54**	**0.81**	**1.90**	**1.40**	<0.001*	<0.001*
Justice	**4.48**	**0.71**	**1.99**	**1.11**	3.74	0.95	2.86	1.12	<0.001*	<0.001*
Autonomy	3.67	1.15	3.12	1.33	3.49	1.17	3.58	1.41	.277	.415
Explicability	**4.54**	**0.63**	**2.05**	**0.99**	3.49	1.01	3.45	1.16	<0.001*	<0.001*

EMBEDDED DESIGN. *Autonomy* – Decide for yourself whether a product is effective or not (Neither option).
ACT: *Justice* - Rectifying injustices to employees; *Explicability* – Accept accountability and transparency.
DON'T ACT: *Beneficence*– Maintain well-being by avoiding unnecessary stress; *Non-Maleficence* - Avoid harm to your family and protect privacy.
[*] indicates which principles had a significant difference between choosing to act or not act.

Table 11
- Credentials Ratings- not at all (1) to extremely (5) important; Rankings- most (1) to least (5) important.

	ACT				DON'T ACT				T-TEST	
	Rating		Ranking		Rating		Ranking		Rating	Ranking
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	p-value	p-value
Beneficence	**4.65**	**0.52**	**1.61**	**0.85**	3.73	0.94	3.13	1.32	<0.001*	<0.001*
Non-Maleficence	**4.40**	**0.81**	**2.45**	**1.08**	3.94	0.95	3.14	1.29	0.002*	.001*
Justice	4.18	1.01	3.02	1.46	3.66	1.24	3.47	1.41	.008*	0.057
Autonomy	3.37	0.92	3.75	1.13	**4.56**	**0.70**	**2.01**	**1.29**	<0.001*	<0.001*
Explicability	3.28	0.75	4.18	0.79	**3.99**	**0.98**	3.25	1.31	<0.001*	<0.001*

EMBEDDED DESIGN: *Justice* - Your friends were unfairly targeted (neither option).
ACT: *Beneficence* - Help your friend; *Non-Maleficence* - Prevent illegitimate access.
DON'T ACT: *Autonomy* - No permission to access their accounts; *Explicability* - Not acting transparently.

Table 12
Ransomware: Ratings- not at all (1) to extremely (5) important; Rankings- most (1) to least (5) important.

	ACT				DON'T ACT				T-TEST	
	Rating		Ranking		Rating		Ranking		Rating	Ranking
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	p-value	p-value
Beneficence	**4.32**	**0.95**	**2.27**	**1.44**	3.55	1.09	3.40	1.19	<0.001*	<0.001*
Non-Maleficence	3.46	1.09	3.62	1.23	**3.97**	**0.97**	**2.87**	**1.35**	0.007*	.003*
Justice	3.58	1.12	3.32	1.25	**4.29**	**1.05**	**2.30**	**1.49**	.001*	<0.001*
Autonomy	**4.41**	**0.86**	**2.06**	**1.13**	3.60	1.15	3.15	1.36	<0.001*	<0.001*
Explicability	3.44	1.24	3.74	1.16	3.76	1.29	3.29	1.41	0.190	0.086

EMBEDDED DESIGN: *Explicability* - The organisation did not clarify who they were (neither option).
ACT: *Beneficence* - Get access to your data; *Autonomy* - Recover control of your data.
DON'T ACT: *Non-Maleficence* - Data might be contaminated anyway*;* *Justice* - Discourages criminal activity.

Table 13
Health Data: Ratings- not at all (1) to extremely (5) important; Rankings- most (1) to least (5) important.

	ACT				DON'T ACT				T-TEST	
	Rating		Ranking		Rating		Ranking		Rating	Ranking
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	p-value	p-value
Beneficence	3.53	1.11	3.60	1.14	3.89	1.08	2.97	1.46	0.008*	0.004*
Non-Maleficence	2.67	1.22	3.75	1.20	3.41	1.19	3.44	1.35	0.001*	0.160
Justice	**4.13**	**1.07**	**2.39**	**1.10**	3.90	1.09	2.74	1.27	0.228	0.085
Autonomy	**4.56**	**0.84**	**1.63**	**1.14**	3.97	1.03	2.62	1.52	<0.001*	<0.001*
Explicability	2.30	1.09	3.64	1.11	3.65	1.03	3.23	1.39	0.001*	0.053

EMBEDDED DESIGN: *Beneficence* – Being healthy is good for people (neither option).
ACT: *Justice* - People without any data are automatically categorised as 'unhealthy; *Autonomy* – No consent was given for this data usage .
DON'T ACT: *Non-Maleficence* – Putting your name on public petitions open you up to harm; *Explicability* – Openly rewarding healthy behaviours increases transparency.

Table 14
2FA: Ratings- not at all (1) to extremely (5) important; Rankings- most (1) to least (5) important.

	ACT				DON'T ACT				T-TEST	
	Rating		Ranking		Rating		Ranking		Rating	Ranking
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	p-value	p-value
Beneficence	**3.90**	**0.98**	**2.66**	**1.29**	2.97	1.03	3.78	1.33	<0.001*	<0.001*
Non-Maleficence	2.48	1.28	3.60	1.23	3.56	1.10	2.86	1.35	<0.001*	<0.001*
Justice	2.75	1.28	3.65	1.36	3.49	1.10	2.90	1.27	<0.001*	<0.001*
Autonomy	3.23	1.28	2.86	1.25	**4.29**	**0.88**	**1.78**	**1.13**	<0.001*	<0.001*
Explicability	**3.64**	**0.98**	**2.63**	**1.43**	3.14	1.04	3.69	1.01	0.001*	<0.001*

EMBEDDED DESIGN: *Non-Maleficence* – The exclusive content might cause harm (neither option).
ACT: *Beneficence*-Retain access to exclusive content; *Explicability*–Satisfied with justification for 2FA implementation.
DON'T ACT: *Justice* – Protest unfairness as the use of 2FA would exclude those without phones or the technical skills;.
Autonomy – Retain control of who has access to your phone information by not sharing it.

our design, these same principles also drove the decision DON'T ACT. Despite Autonomy rating and ranking highest for both decisions, the T-test identifies highly significant differences in level of importance (ranking 4.13, 3.90; rating 2.39, 2.74, for ACT versus DON'T ACT, respectively), which is consistent with our design. Similarly, Tukey HSD results confirm our design for ACT, but not for DON'T ACT.

Participants who chose to sign the petition considered equal opportunities ("*all people should have the same opportunities regardless of their health*" (21, Male)), privacy concerns ("*I consider this kind of information sensitive and private and it doesn't help people not using the apps*" (23, Male)) and justice concerns ("*It's unfair to use people's confidential data to promote business, even if it does not directly effect you. It causes injustice against the minority*" (18, Female)). Some of those who chose to stay silent could see the potential benefits of the new scheme ("*Taking steps in life to be healthy is very important to me and I believe those who do the same deserve to paid[sic] cheaper premiums*" (26, Female)).

4.3.5. 2FA

The 2FA scenario involves the implementation of a 2-factor authentication system to an existing service and the decision whether to:

- Provide your mobile number and retain access to exclusive content [ACT]
- Do not provide your mobile number and lose access to exclusive content [DON'T ACT]

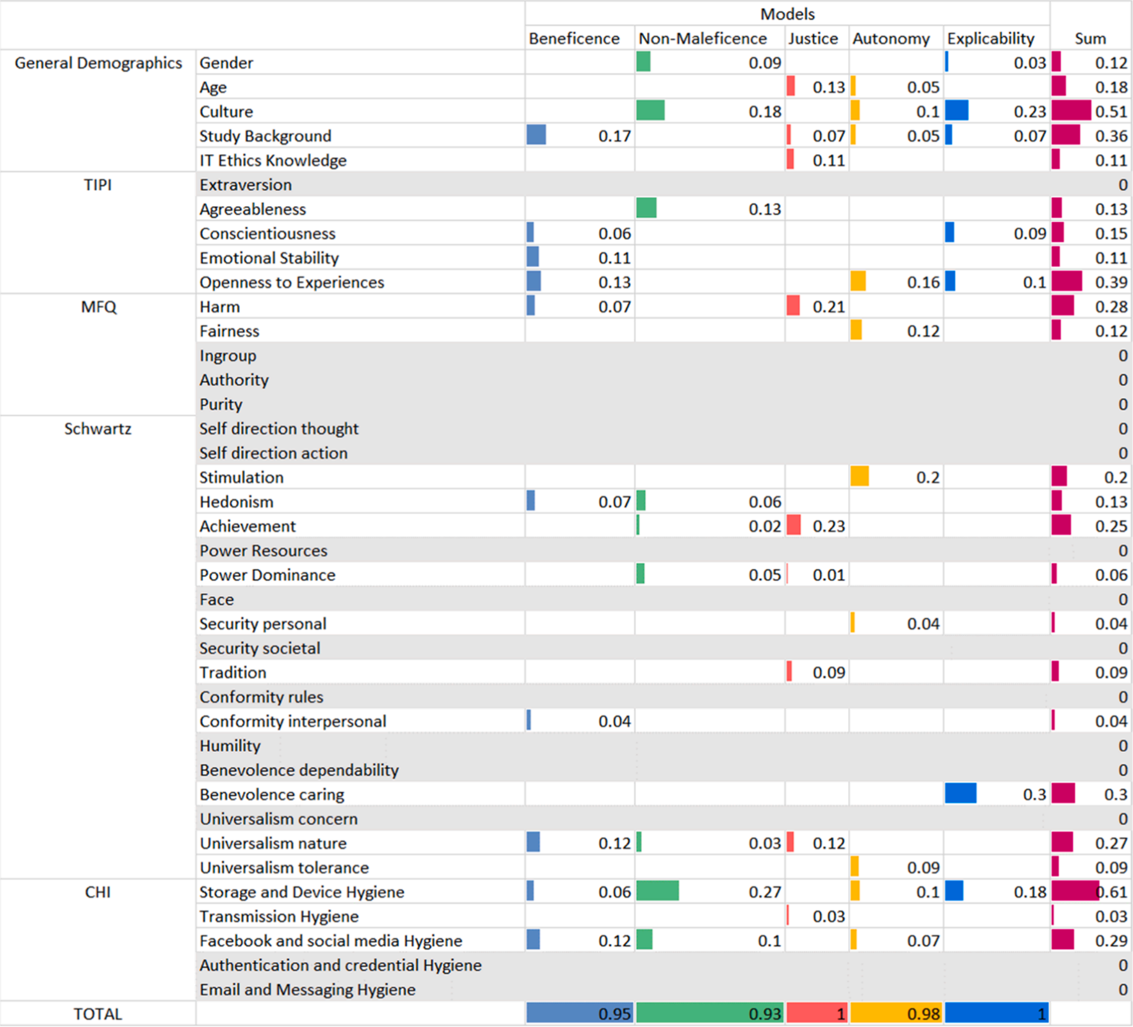
Table 14 shows that both embedded principles (Beneficence and Explicability) for ACT were identified as most important for both ratings and rankings. For DON'T ACT, only one of our embedded principles, Autonomy, was highest for rating and rankings, whereas the irrelevant principle, Non-Maleficence, was the second most important principle instead of Justice. T-tests reveal highly significant differences between all principles, including Non-Maleficence, for both actions. Tukey's HSD confirms that, for ACT, Beneficence was significantly different from all principles except Explicability, which was the other supporting principle for the ACT option and, for DON'T ACT, Autonomy was significantly different from all other principles, while the responses to Non-Maleficence was not significantly different to Justice (with 3 out of 4 p-values 1 or approaching 1).

Participants who chose to provide additional information justified their decision by considering the process to be the "*new normal*" (26, Female), did not consider their phone number information to be "*a serious form of personal information*" (18, Female) and valued that the 2FA service might provide improved security ("*By having to provide the 2FA I would hope that provides me with a little security*" (46, Male)). Others saw the value of the service to be greater than the cost ("*I want the service, so I would 'reluctantly' provide my mobile number - under the assumption / expectation that this information won't be misused by the video-streaming company*" (32, Male)). On the other hand, those who chose not to provide phone number information valued their privacy highly ("*You might have data shared to other companies which breaches your privacy details if you give your number and not having that data shared is more important over keeping the exclusive content*" (18, Female)) and also considered fairness ("*It is unfair to require such sensitive information against an individual's will*" (19, Male)).

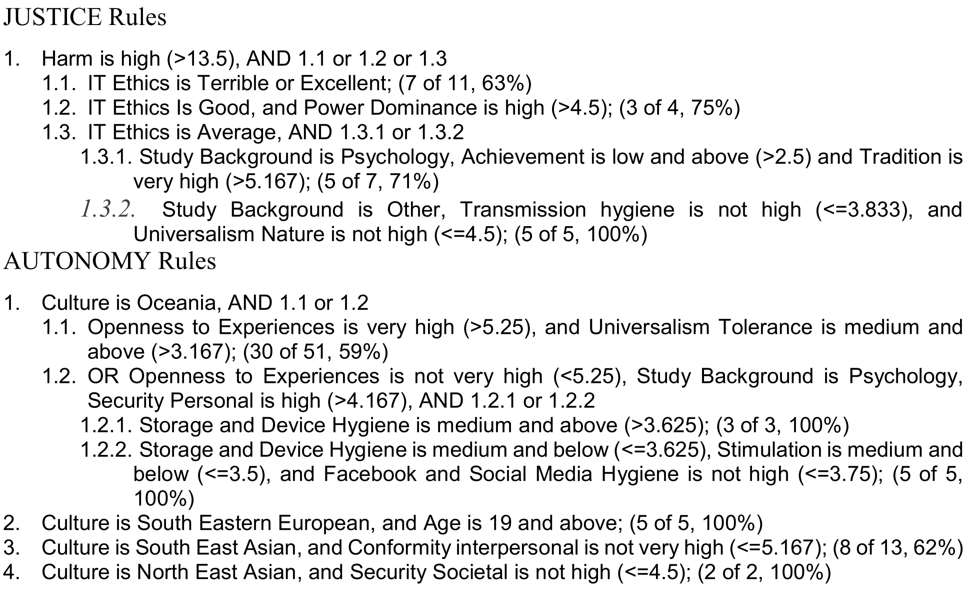
4.4. Principle importance

To explore the predictive power of our various demographic and inventory variables, we employ a machine learning approach (as discussed in Section 3.4) as this helps us to effectively answer RQ 2.1. Table 15 summarises the features that are important in predicting the importance of each ethical principle. The features listed are able to discriminate (i.e. build rules/decision trees) between the low, medium and high classes for each principle, ranging from 93 %—100 % coverage of the cases. We observe that CHI storage and device hygiene (61 %), culture (51 %), openness to experiences (39 %), study background (36 %) and IT Ethics Knowledge (36 %) are highly salient, however, not across all principles. For example, the importance of CHI storage and device hygiene ranges from 6 to 27 % across 4 ethical principles. We see culture is important, though not necessarily the most salient feature, in predicting the importance of Explicability (18 %), Non-Maleficence (18 %) and Autonomy (10 %), but for Justice and Beneficence, the main predictor is Schwartz' achievement (23 %) and study background (17

Table 15
Features important to predicting principle importance.



		Models					
		Beneficence	Non-Maleficence	Justice	Autonomy	Explicability	Sum
General Demographics	Gender		0.09			0.03	0.12
	Age			0.13	0.05		0.18
	Culture		0.18		0.1	0.23	0.51
	Study Background	0.17		0.07	0.05	0.07	0.36
	IT Ethics Knowledge			0.11			0.11
TIPI	Extraversion						0
	Agreeableness		0.13				0.13
	Conscientiousness	0.06				0.09	0.15
	Emotional Stability	0.11					0.11
	Openness to Experiences	0.13			0.16	0.1	0.39
MFQ	Harm	0.07		0.21			0.28
	Fairness				0.12		0.12
	Ingroup						0
	Authority						0
	Purity						0
Schwartz	Self direction thought						0
	Self direction action						0
	Stimulation				0.2		0.2
	Hedonism	0.07	0.06				0.13
	Achievement		0.02	0.23			0.25
	Power Resources						0
	Power Dominance		0.05	0.01			0.06
	Face						0
	Security personal				0.04		0.04
	Security societal						0
	Tradition			0.09			0.09
	Conformity rules						0
	Conformity interpersonal	0.04					0.04
	Humility						0
	Benevolence dependability						0
	Benevolence caring					0.3	0.3
	Universalism concern						0
	Universalism nature	0.12	0.03	0.12			0.27
	Universalism tolerance				0.09		0.09
CHI	Storage and Device Hygiene	0.06	0.27		0.1	0.18	0.61
	Transmission Hygiene			0.03			0.03
	Facebook and social media Hygiene	0.12	0.1		0.07		0.29
	Authentication and credential Hygiene						0
	Email and Messaging Hygiene						0
TOTAL		0.95	0.93	1	0.98	1	



JUSTICE Rules

1. Harm is high (>13.5), AND 1.1 or 1.2 or 1.3
   1.1. IT Ethics is Terrible or Excellent; (7 of 11, 63%)
   1.2. IT Ethics Is Good, and Power Dominance is high (>4.5); (3 of 4, 75%)
   1.3. IT Ethics is Average, AND 1.3.1 or 1.3.2
      1.3.1. Study Background is Psychology, Achievement is low and above (>2.5) and Tradition is very high (>5.167); (5 of 7, 71%)
      1.3.2. Study Background is Other, Transmission hygiene is not high (<=3.833), and Universalism Nature is not high (<=4.5); (5 of 5, 100%)

AUTONOMY Rules

1. Culture is Oceania, AND 1.1 or 1.2
   1.1. Openness to Experiences is very high (>5.25), and Universalism Tolerance is medium and above (>3.167); (30 of 51, 59%)
   1.2. OR Openness to Experiences is not very high (<5.25), Study Background is Psychology, Security Personal is high (>4.167), AND 1.2.1 or 1.2.2
      1.2.1. Storage and Device Hygiene is medium and above (>3.625); (3 of 3, 100%)
      1.2.2. Storage and Device Hygiene is medium and below (<=3.625), Stimulation is medium and below (<=3.5), and Facebook and Social Media Hygiene is not high (<=3.75); (5 of 5, 100%)
2. Culture is South Eastern European, and Age is 19 and above; (5 of 5, 100%)
3. Culture is South East Asian, and Conformity interpersonal is not very high (<=5.167); (8 of 13, 62%)
4. Culture is North East Asian, and Security Societal is not high (<=4.5); (2 of 2, 100%)

Fig. 2. C5.0 rules for Justice and Autonomy. Rule accuracy shown in bracket (7 correct out of 11 cases, 63 % accuracy).

%).

To unpack how these features influenced high rankings for the principle, we analysed the C5.0 rules that predict High across all five scenarios. As might be expected, not all principles were considered of high importance in equal distributions. Only 6 % of our participants considered Explicability to be the most important ethical principle. Similarly, there were low percentages for Beneficence (8 %) and Non-Maleficence (14 %). Given the smaller set of examples, we thus have less confidence in the rules for these principles and only present in Fig. 2 the rules for Justice, which was deemed most important by 23 % of our participants with a 60.7 % (10-fold) cross validation accuracy, and Autonomy representing 34 % of our participants which achieved 49.2 % cross validation accuracy. Given that there are 5 classes to predict, prediction accuracy greater than 20 % is better than chance. Individual rule accuracy is provided in brackets on the leaf nodes. For example, looking at the Autonomy ruleset, rule 1.1 states “IF Culture is Oceanian, Openness to Experience is very high, Universalism Tolerance is medium and above THEN Autonomy is of high importance”. There are 51 participants who are covered by this rule (i.e. fit this description) and 30 of them selected Autonomy as the most important principle. Thus the accuracy of this rule on this dataset is 59 %.

4.5. Agency

For each of our 193 responses there were five scenarios, resulting in 965 cases where an individual’s sense of agency was recorded. Only one category of “None at all” captures individuals who perceive they do not have a choice, while the other four categories on the scale capture varying levels of agency. Table 16 shows that only in 8.9 % of the cases did participants feel that there was no choice in the scenario, indicating that individuals believed they possessed some level of agency.

A T-test was performed between responses that chose ACT and DON’T ACT and their perceived agency in Table 17. There was no significant difference in the perceived agency between responses that picked ACT with those that picked DON’T ACT.

5. Discussion

5.1. RQ1 discussion – identifying and applying ethical principles

In this section we answer our three research questions. In terms of our first research question, RQ1, about whether our participants could identify and order relevant ethical principles in cybersecurity scenarios, we found that individuals could, broadly, identify the ethical principles that were important to their decision. The data in support of this result is that in all five scenarios for the ACT decision, rankings and ratings for both relevant principles were most important. For the DON’T ACT decision, both principles for the Ransomware were most important for rankings and ratings, one of the embedded principles was most important for both rankings and ratings for Misinformation, Credentials and 2FA scenarios, with a second embedded principle rating second highest for Misinformation and Credentials. Only the Health Data scenario did not have highest ratings or rankings for either of the embedded principles for DON’T ACT; instead participants chose the same two principles (Autonomy and Justice) as most important for both actions. Further, T-tests mostly showed a statistically significant difference between the choices of ACT and DON’T ACT for each relevant ethical principle (but not for the irrelevant principle for the Misinformation, Credentials and Ransomware scenarios), indicating that the differences were not due to chance and thus the ethical principle influenced the decision made. Statistical significance in Tukey’s HSD post-hoc analysis further found significant differences between many pairs of principles, often confirming that the supporting ethical principles were significantly different from the non-supporting principles, indicating again that in most cases a relevant ethical principle had influenced the decision.

Table 16
Agency Frequency.

Frequencies	None at all	A little	A moderate amount	A lot	A great deal
Credentials	14 (7.3 %)	65 (33.7 %)	65 (33.7 %)	38 (19.7 %)	11 (5.7 %)
Ransomware	28 (14.5 %)	74 (38.3 %)	59 (30.6 %)	23 (11.9 %)	9 (4.7 %)
Health	9 (4.7 %)	58 (30.1 %)	66 (34.2 %)	44 (22.8 %)	16 (8.3 %)
2FA	21 (10.9 %)	56 (29.0 %)	56 (29.0 %)	44 (22.8 %)	16 (8.3 %)
Misinformation	14 (7.3 %)	51 (26.4 %)	67 (34.7 %)	38 (19.7 %)	23 (11.9 %)
Total	86 (8.9 %)	304 (31.5 %)	313 (32.4 %)	187 (19.4 %)	75 (7.8 %)

Table 17
Perceived Agency significance.

	Picked ACT			Picked DON’T ACT			T-TEST
Scenario	N	Mean	SD	N	Mean	SD	Sig.(2-tailed)
Credentials	51	2.92	0.87	142	2.8	1.066	0.449
Ransomware	34	2.38	0.82	159	2.57	1.07	0.331
Health	154	3.06	1.01	39	2.77	1.09	0.116
2FA	113	2.93	1.16	80	2.83	1.10	0.530
Misinformation	96	3.18	1.13	97	2.88	1.07	0.060

In line with the theory of planned behaviour (also known as the theory of reasoned action) (Fishbein and Ajzen, 2011) that underpinned our question about perceived agency (control beliefs), individuals’ actions are also driven by their behavioural beliefs, such as the acceptability or consequence of a certain behavior, and by their normative beliefs, such as how others would behave or what is expected of them by others. These individual-specific behavioural and normative beliefs seem to have influenced participants’ ethical reasoning for some scenarios and potentially “overridden” the reasoning we provided in the embedded principle. In other words, even though we did not raise a certain ethical issue, the nature of the scenario has triggered certain beliefs and assumptions not articulated in the scenario. As examples, the DON’T ACT decision in the Credentials and Health Data scenarios has triggered a concern for Justice. In the Credentials scenario, the significant T-test result for Justice and the lack of a significant difference between the irrelevant Justice and relevant Non-Maleficence principles in the Tukey HSD results indicate that even though the reason “Justice - Your friends were unfairly targeted” was not raised in the scenario, participants exhibit a normative belief that this would be unfair if this happened to one of their friends and thus may have empathised with the hypothetical friend in the scenario. In the Health Data scenario, for DON’T ACT, participants perhaps chose “Justice - People without any data are automatically categorised as unhealthy” to support not signing the petition because they may have felt that it is fair that people who generate data showing that they are taking care of their health should not pay the same insurance premium as others who may not be taking as much care of their health. This interpretation is consistent with the description in the scenario that discusses who will receive a discount through not acting and the alternative of all individuals paying the higher premium resulting from signing the petition. The selection of Autonomy in this scenario as the most important principle for both actions (in terms of ratings but not rankings for DON’T ACT) is consistent with reasoned action and behaviour concerning health decisions. A study by Cullati et al. (2011) found a general desire for Autonomy in health decision-making. Those who had made several medical decisions in the 6 months prior to their study were more likely to have a strong desire for autonomy; however, this was not the case for those who had frequent contact with medical professionals but had not participated in decision making (Cullati et al., 2011). It thus appears that making medical decisions increases the desire for autonomy. It may be that our participants who chose to act, and thus exercise their autonomy, had been involved in their own medical decision-making in the recent past. However, we did not gather any data from participants to confirm this. Autonomy is one of the most important principles in the health domain as identified by Loi et al. (2019), where they note that Autonomy is usually at odds with other bioethics principles. Informed consent in the health context is complicated by the unpredictable future use of data which challenges Autonomy by removing control of data from the individual (Yaghmaei et al., 2017). This indicates either a flaw in the design of the Health Data scenario or unique differences in that domain needing further investigation. Another interpretation could be related to the alternate ordering of the two choices offered as the Health Data scenario was the only one that offered DON'T ACT before ACT. Further studies would need to be conducted to investigate the plausibility of either interpretation.

Tukey HSD analysis identifies which pairs of principles were deemed significantly different for each scenario and sheds light on how the pairs may be overlapping in some scenarios. For example, in the Credential scenario, Beneficence and Non-Maleficence are only significantly different for ACT rankings. This highlights the close connection between these two principles where in many contexts, such as helping or protecting a friend, they can be seen as two sides of the same coin. In fact, both of these principles were valid for the ACT context and participants would have had to decide whether they favoured avoiding harm over delivering benefit in their response.

5.2. RQ2.1 discussion – predicting the priorities of different ethical principles

Moving on to our second research question, RQ2.1, concerning what features could predict which ethical principle is most important for an individual in cybersecurity scenarios, we identified several features that could predict the importance of each ethical principle, with Study Background (17 %), CHI storage and device hygiene (27 %), Schwartz' achievement (23 %), openness to experience (16 %) and culture (23 %) being the most salient features for Beneficence, Non-Maleficence, Justice, Autonomy and Explicability, respectively. No features were predictors for all five principles. CHI storage and device hygiene was the only feature that was a predictor for 4 principles; all but Justice which was the only principle where IT Ethics Knowledge was important. This may suggest that, in general, identifying this cyber hygiene behaviour for an individual may help to predict their ethical behaviour, and providing remedial training in this behaviour, if found to be inadequate, may be a low-hanging fruitful approach to sensitising untrained individuals to make ethical decisions. Alternatively, increasing knowledge of IT Ethics may increase awareness of justice concerns in cybersecurity decision making. Our findings, thus, suggest which features might be most useful to measure sensitivity to and/or provide tailored training about specific ethical principles. Testing of these conjectures would require new studies to be conducted.

Collectively, demographic features were strong predictors, including Culture, Age, Study Background, and IT Ethics Knowledge. Culture was a predictor for Non-Maleficence, Explicability, and Autonomy. This fits with previous research that shows that culture influences beliefs and behaviour (Hofstede, 2003). Age was found to predict prioritisation of the principles of Autonomy and Justice. This might be explained by Sheldon et al. (2006) who suggest that individuals value Autonomy more as they age. A study by Brienza and Bobocel (2017) found that different types of justice were important to different aged employees, where informational and interpersonal justice influenced older workers and distributive and procedural justice had a greater influence on younger workers. While Study Background was the main predictor for Beneficence (17 %), this feature is probably specific to our dataset that involved students and may not be a feature relevant to describing other populations of untrained individuals. Thus, more generalizable useful features for Beneficence may be Schwartz' universalism nature and CHI Facebook and Social Media Hygiene (12 %), then emotional stability and openness to experience (11 %).

CHI Facebook and Social Media Hygiene factor considers whether individuals manage their privacy settings, ensuring location data does not leak, and verify the authenticity of correspondents and communication exchanges (Vishwanath et al., 2020). A low score in this factor indicates that the individual values the benefits (Beneficence) of using the service above the potential security risks (Non-Maleficence). Schwartz' Universalism Nature concerns the preservation of nature (Schwartz, 2016) and a non-low score indicates that an individual would not want to cause harm to nature. High Hedonism involves enjoyment (Schwartz, 2016) which is linked to Non-Maleficence where avoiding harm allows for enjoyment. High Facebook and Social Media Hygiene scores indicate that the individual holds themselves accountable for their interactions on social media (Vishwanath et al., 2020) and would value preventing harm.

High Tradition has links to Justice by considering traditional practices to be important (Schwartz, 2016), and this importance links to Justice as these traditions should not be the basis for discrimination. A high MFQ Harm score indicates that participants were concerned with empathy (Graham et al., 2011), and this links to Justice by identifying that people should be treated fairly and with kindness.

For Autonomy, Security Personal involves the security of one's self (Schwartz, 2017), which has a clear link to Autonomy where individuals valuing Autonomy consider acting on their own will to protect their security to be important. Moderate Universalism Tolerance involves understanding and accepting others (Schwartz, 2017), which links to Autonomy with respecting other's will. Conformity Interpersonal seeks to avoid harming others where possible (Schwartz, 2017) and a not very high score could indicate that individuals are willing to act on their own will even if it may upset others. Regarding Explicability, Conscientiousness relates to reliability (Goldberg, 1993) where Explicability is involved in a system being reliable.

To understand how these different features may influence cyber ethical decision making we need to uncover the rules or decision-trees. However, the five rulesets uncovered to predict whether an individual ranks a principle as highly important may not be reliable or generalisable to other populations due to small data sample sizes. Only two of the five principles (Justice and Autonomy) have produced rules with acceptable accuracy percentages and cover a reasonable number of observations. Looking at the ruleset presented for Autonomy, Culture was an important feature. This could be related to the Universalism vs Particularism dimension where individuals follow the rules and processes or they adapt to the changing environment and accommodate themselves to their situation (Hampden-Turner et al., 2020). Regarding the influence of human values (PVQ-RR), Achievement, which involves being successful (Schwartz, 2017), was an important feature for Justice which could be linked by Justice rewarding those who are successful. Stimulation is another important feature for Autonomy where those seeking new experiences (Schwartz, 2017) would value being able to act on their own will to accomplish different goals. Looking at the Justice ruleset, we see that Harm, a moral stance construct from MFQ, is linked to being concerned with being compassionate and caring for the vulnerable (Graham et al., 2011). This aligns with the notion of Justice which involves supporting vulnerable groups and avoiding prejudice.

A further conclusion we draw from creating these models is that certain demographic features and elements of inventories, such as the TIPI, CHI, and PVQ-RR, all show promising predictive power. This suggests there is value in eliciting this data from individuals to predict their possible decisions. It may also suggest appropriate training so that employees and managers are aware of the impact of personal values and culture on decision making.

5.3. RQ2.2 discussion – the impact of perceived agency

We draw on the quantitative results and the justifications provided in comments to assist in answering RQ2.2, concerning how much agency participants felt they had in cybersecurity scenarios. Firstly, our analysis showed that 83.3 % of all the data on sense of agency fell within “a little to a lot of agency”, with the majority (64.9 %) of participants feeling they had little to moderate level of agency. The lack of significant differences in sense of agency between those picking ACT versus DON’T ACT helps to validate our instrument, since we intended that perceived agency would not influence whether an individual chooses to act or not to act. It is unclear from the data whether a lack of agency might result in not acting, or conversely result in randomly choosing either option; this requires future investigation.

We found that 109 of 193 responses (56 %) selected the same perceived agency across three or more scenarios. Since the standard deviation between scenarios for individual sense of agency tends towards below 1, it indicates that individuals perceived they had similar levels of agency across all five scenarios; i.e., an individual would rate their perceived agency consistently across all five scenarios and perhaps suggesting that the scenarios were not biased in terms of providing viable choices. Our data shows that a lack of agency did not influence the decisions made where the level of perceived agency was not significant in any scenario. As a final point, in the C5.0 models we created to predict the importance of principles, we included participants’ responses to perceived agency. None of the C5.0 models included perceived agency and thus we conclude it did not impact on their decision.

6. Limitations, implications and future work

Due to skewed distribution (i.e., uneven selection of principles across scenarios), models for some principles had very few members in some classes, resulting in reduced generalisability for the generated models. To remedy the inadequate numbers for some classes, more data collection is required to potentially increase the numbers in both classes. Another limitation that influences generalisability is the participant pool was primarily comprised of Psychology students, with smaller representation from other untrained student groups. Further studies need to draw participants from a wider pool, including larger age groups and non-university students. While cultural diversity was present in our study, consistent with the highly multicultural nature of Australia, this younger age group is potentially more likely to identify as Oceanian (Australian) than their migrant parents or grandparents, and thus distribution across ethnicities is likely to be different with older populations. We treated all participants as “untrained” and so appropriate for our study, however it may be that participants we recruited via the first-year cybersecurity unit had both more interest and more knowledge about these issues. Due to the small number of participants from the cybersecurity unit, it was not possible to determine if there were any significant differences between them and the participants recruited via the psychology pool. A future study could compare how these different cohorts respond to our study.

Quantitative methods were used as the primary means of data analysis, and it is a limitation of our study that we only used one free-text question to elicit the perceived reasons behind why a decision was taken. Qualitative coding of this open-text data was done by one researcher. While triangulation of coders would have given us more robust qualitative results (Nancy Carter et al., 2014), given that we only used qualitative data here to provide illustrative quotes of some of the main issues raised in the responses, rather than report a detailed theme frequency table, this lack of robustness does not impact the results reported here.

Another limitation is that only the Health Data scenario provided the option to act and not act in a different order to the other scenarios. To reduce the impact of ordering effects (Krosnick and Alwin, 1987), the order should be randomised for all scenarios in future work. Different interpretations of our scenarios due to prior knowledge and beliefs may have impacted some scenarios and some changes to wording might be needed to minimise these differences. For example, the Health Data scenario may have been interpreted as a case of data protection and impacted by changing the context of the scenario from an ethical dilemma to a legal context.

The use of a scenario-based study is another limitation due to the specific scenarios used, cybersecurity ethical decision-making data is only collected on the specific situations described in our study. Future work in this area should build a larger suite of scenarios to improve the generalisability of the results and increase applicability to other cyber-ethical contexts. Survey-based research also has well-known limitations (Coughlan et al., 2009), and the use of other methods, such as observational studies of how people deal with ethical challenges in real-world cybersecurity scenarios, would be a highly beneficial supplement to the results presented here. Observational data would be helpful in confirming whether the way that our participants said they would choose to act in our hypothetical scenarios matches what we would observe them to actually do in real-world cybersecurity scenarios with all their complexities and pressures. To partly address this concern, our current project involves the use of a video game to train cybersecurity students and professionals to be aware of and apply the five cybersecurity ethical principles and captures player’s actions and reasoning during the game (Ryan et al., 2022). While the use of a video game can make cybersecurity scenarios more realistic and complex, the consequences of the in-game choices are still merely fictional and the pressures the game creates are artificial, which consequently does not completely remove the importance of future observational studies to supplement our work.

As part of our profile building, we found that MFQ is not a strong predictor for cybersecurity sensitive situations and can be omitted from future research to reduce effort and complexity. While we found the TIPI to be a useful predictor, given the low Cronbach’s alpha (α) value for this measure, a more reliable but longer version of this measure should be considered for use, although this must be weighed against increasing fatigue concerns (Gosling et al., 2003).

Our work has implications for addressing the gap in training in cyberethics. A report published by Datto (2020) found that lack of cybersecurity training is the third largest origin of ransomware infections caused by 26 % of the users. Moreover, ethical fading has been the major cause of human error that leads to 93 % of cybersecurity breaches (Tenbrunsel and Bazerman, 2011). Pólkowski (2015) has identified an unmet need for educational institutions to provide computer ethics training to their students, academics and other employees. Our study has potential implications for practitioners including cybersecurity students, cybersecurity professionals and managers and their education and training. For cybersecurity professionals and managers, the responses to our scenarios confirm that personality and other individual factors can influence the cybersecurity decision-making of computer users and raise their awareness of the role that ethical priorities and value judgements make in user’s responses to policies, technologies and processes they may design and implement. For the non-technical end-user, our results indicate that exposure to scenarios with embedded ethical principles can elicit responses which demonstrate understanding and appropriate application of the five ethical principles for specific contexts. Thus, our principle-based scenarios have the potential to raise end user awareness of the ethical implications of their cybersecurity actions, thus protecting themselves and/or their organisations from harm. Our findings show that there was a connection between the importance assigned by the participant to specific ethical principles and the scenario choices they made, as well as their receptivity to a range of ethical considerations. Future work could explore how training might be developed based on this. Future work could also explore how our predictive models might be used to create tailored personal and targeted interventions to help different users make more ethical decisions in cybersecurity contexts. For example, if we know certain types of new employees are more likely to ethically fail in certain types of scenarios or fail to prioritise certain types of ethical principles, they can be provided with targeted training interventions to help address this, and this could have much better impacts than non-targeted interventions. However, given the well-known judgment-action gap between what people say they will do and what they will actually do (Stephens, 2018), exploring whether any interventions actually improve real-world moral behaviour, and not just impact self-report measures about what participants would hypothetically do, will be of crucial importance.

7. Conclusion

Our study explored how people weigh relevant ethical principles when making cybersecurity decisions. We found that ethical principles did influence decision making in cybersecurity sensitive situations. We were successfully able to design scenarios that allowed participants to exercise their ethical reasoning and, in general, correctly identify the embedded ethical principles. We created profiles that could describe an individual's prioritisation of ethical principles by building suitable models for the Justice and Autonomy principles. We also showed that individuals mostly feel they have a moderate level of agency in cybersecurity sensitive situations and that levels of perceived agency did not impact their decision to act or not act. We found that personality (TIPI), cyber hygiene (CHI), and guiding values (PVQ-RR), along with demographic features including Age, Gender, Study Background, Culture, and IT Ethics Knowledge were strong predictors. This study provides important data to help us to better understand the role of ethical principles in cybersecurity decision making.

CRediT authorship contribution statement

Joseph Fenech: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft. **Deborah Richards:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft. **Paul Formosa:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Joseph Fenech reports financial support was provided by Macquarie University. Deborah Richards reports financial support was provided by Australian Research Council. Paul Formosa reports financial support was provided by Australian Research Council. Deborah Richards reports a relationship with Macquarie University that includes: employment. Paul Formosa reports a relationship with Macquarie University that includes: employment. Nothing additional to declare.

Data availability

Data will be made available on request.

Author Contribution

All authors contributed to the conceptualization, writing, and reviewing of this article. The authors are listed in order of the degree of contribution.

Appendix A

Misinformation Scenario

Table B1,Table B2,Table B3,Table B4,Table B5

You are part of an online garden maintenance forum that explores innovations in automating the care of home gardens. A well-respected reviewer of garden systems posts on your forum that a particularly popular gardening tool has inherent problems that can cause toxins to build up in the garden, which could be a risk to the health of the gardens and even people.

Acting on this information you join in large protests outside the company, harassing their employees. This causes a large backlash that results in the company going bankrupt and most of their staff losing their jobs. A few days after the company collapses, the well-respected reviewer whose post started the backlash indicates that they did not make the post. They were on vacation at the time and had their credentials stolen. The gardening tool in question is in reality safe and extremely effective.

You were a part of the protests that caused stress and harm to the company and its employees and you are asked by other protesters to take public responsibility for acting on unvalidated information.

Publicly acknowledging your involvement involves accepting accountability for your actions and the harms they caused to others and attempts to rectify the injustice suffered by employees, but it could cause harm by exposing your privacy and opening you and your family up to reprisals. If you stay silent about your involvement, you maintain your privacy and prevent any potential reprisals and help to minimise additional stress for all involved.

You decide to:

- Publicly acknowledge your involvement and expose your privacy (1)
- Do not publicly acknowledge your involvement and protect your privacy (2)

Credentials Scenario

An online shopping service that sells turtle themed products stores customer details, including passwords, in plain text in their database. Their database suffers a breach and exposes thousands of customer details and passwords which are posted on message boards around the web.

These exposed password credentials are immediately blocked on the shopping service, but there is a possibility that users were reusing their credentials on different services. You learn of the breach a day after it has occurred, and you search through the exposed details and recognise some of your friends' details there.

You are concerned that your friends might have reused the same credentials across multiple distinct services, including their important social media accounts, but you are unable to contact them all day to alert them to the danger.

You could try to access their social media accounts with the exposed password credentials in order to prevent illegitimate access by malicious third parties by changing their password for them, but you would have to do this without their permission. Leaving their accounts untouched avoids these issues, but leaves their important accounts potentially exposed to malicious third parties.

You decide to:

- Attempt to access your friends' social media accounts with the exposed password credentials to change their password without their permission to try to protect their accounts. (1)
- Do not attempt to access your friends' social media accounts with the exposed password credentials to change their password and thereby leave their accounts potentially exposed. (2)

Ransomware Scenario

While collecting data for an important class project, you suddenly realise that you can no longer access the data that you had previously collected. While frantically trying to find backups you notice a flashing prompt on your screen informing you that your data has been encrypted as part of a ransomware attack and explains that the only way to get your data back is to pay a ransom.

The ransomware indicates that the price will increase as time passes if you do not pay the fee. You also fear that the longer you are unable to access your data, the less trust you can have that it has not been modified. After a week you will no longer be able to afford the ransom.

Your latest backup is three months old and you have no other way to access all the data you have collected since then. You will have to make more frequent backups in the future.

Paying the ransom grants you instant access to your data and allows you to continue your work but advertises that you are willing to pay and may invite future ransomware attacks on yourself or others. Not paying discourages repeat attacks and limits the spread of the ransomware but results in you losing three months of data that you will have to recollect.

You decide to:

- Pay the ransom and get access to all your data (1)
- Do not pay the ransom and lose your last three months of data (2)

Health Data Scenario

You have been a long-time user of a personal fitness tracker which monitors certain metrics of health and stores that data on their organisation's servers. The organisation has decided without specific consent to sell this data to insurance companies who want to use it to tailor insurance premiums to customers based on some of these metrics.

To combat rising prices, insurance companies wish to categorise these positive metrics as 'healthy' customers who will pay less and negative or missing metrics as 'unhealthy' customers who will pay more. Your colleagues bring up that these metrics require research to justify their usage and ask whether you would sign an online petition opposing the use of this data.

You fall in the 'healthy' category as your fitness tracker collects metrics that correlate to their model of a 'healthy' individual and as a result you will pay reduced premiums.

Ignoring the petition allows 'healthy' individuals to pay cheaper insurance premiums but penalises those without any metric data. Signing the petition provides social pressure to prevent the unauthorised use of personal information but results in all customers paying slightly increased premiums rather than just those considered 'unhealthy'.

You decide to:

- Ignore the petition to acquire cheaper insurance premiums for 'healthy' customers (1)
- Sign the petition to attempt to prevent the use of this data (2)

2FA Scenario

You are a customer of a popular video streaming service that provides exclusive access to certain content. Your video streaming service sends an email informing its users that in 3 weeks' time all users will have to nominate a mobile phone number to use with 2 factor authentication (2FA).

The 2FA will use a combination of existing passwords and 1-time use codes delivered via SMS to improve the security of its authentication. The video streaming service provider updates their terms of use requiring all users to utilise 2FA or their accounts will be inaccessible for usage.

Since you regularly consume the exclusive content provided by this service you face a choice.

Providing a phone number will maintain access and improve the security of the service, but will also require you to share sensitive personal information (i.e. your mobile phone number) that raises privacy considerations. Not providing your mobile phone number will avoid these privacy concerns, but result in you losing access to the exclusive content from the service provider.

You decide to:

- Provide your mobile number and retain access to exclusive content (1)
- Do not provide your mobile number and lose access to exclusive content (2)

APPENDIX B

Tukey HSD results

Table B1
Tukey HSD of Misinformation Scenario ethical principle pairs.

Misinformation scenario pairs		ACT				DON'T ACT			
		Rating		Ranking		Rating		Ranking	
		Mean diff	p-value	Mean diff	p-value	Mean diff	p-value	Mean diff	p-value
BEN	N-M	−0.771*	<0.001	1.063*	<0.001*	−0.750*	<0.001*	1.320*	<0.001*
BEN	JUS	-l.604*	<0.001*	2.302*	<0.001*	0.045	0.998	0.361	0.29
BEN	AUT	−0.799*	<0.001*	0.875-	<0.001*	0.298	0.245	−0.361	0.29
BEN	EXP	−1.667*	<0.001*	2.240*	<0.001*	0.298	0.241	−0.237	.701
N-M	JUS	−0.833*	<0.001*	1.240*	<0.001*	0.795*	<0.001*	−0.959*	<0.001*
N*M	AUT	−0.028	1	−0.187	0.773	1.047*	<0.001*	−1.6SO*	<0.001*
N-M	EXP	−0.896*	<0.001*	1.177*	<0.001*	1.047*	<0.001*	−1.557*	<0.001*
JUS	AUT	0.805*	<0.001*	−1.427*	<0.001*	0.253	0.416	−0.722*	.001*
JUS	EXP	−0.062	0.991	−0.063	0.995	0.252	0.411	−0.598*	.011*
AUT	EXP	−0.868*	<0.001*	1.365*	<0.001*	0	1	0.124	.963

Table B2
Tukey HSD of credential scenario ethical principle pairs.

Credentials scenario pairs		ACT				DON'T ACT			
		Rating		Ranking		Rating		Ranking	
		Mean diff	p-value	Mean diff	p-value	Mean diff	p-value	Mean diff	p-value
BEN	N-M	0.247	.554	.843*	.001*	−0.208	.386	−0.007	1
BEN	JUS	.471*	.033*	−1.412*	<0.001*	0.067	.979	−0.331	.219
BEN	AUT	1.275*	<0.001*	−2.137*	<0.001*	−0.836*	<0.001*	1.120*	<0.001*
BEN	EXP	1.373*	<0.001*	−2.569*	<0.001*	−0.264	.16	−0.113	.953
N-M	JUS	0.224	.647	−0.569	.067	0.274	.131	−0.324	.239
N*M	AUT	1.027*	<0.001*	-l.294*	<0.001*	−0.628*	<0.001*	1.127*	<0.001*
N-M	EXP	1.125*	<0.001*	−1.725*	<0.001*	−0.057	.989	−0.106	.962
JUS	AUT	.804*	<0.001*	−0.725*	.008*	−0.902*	<0.001*	1.451*	<0.001*
JUS	EXP	.902*	<0.001*	−1.157*	<0.001*	−0.331*	.039*	0.218	.635
AUT	EXP	0.098	.974	0.431	.269	.571*	<0.001*	l.232*	<0.001*

Table B3
Tukey HSD of ransomware scenario ethical principle pairs.

ransomware scenario pairs		Act				don't act			
		Rating		Ranking		Rating		Ranking	
		Mean diff	p-value	Mean diff	p-value	Mean diff	p-value	Mean diff	p-value
BEN	N-M	.869*	.008*	−1.353*	<0.001*	−0.414*	.009*	.535*	.004*
BEN	JUS	.748*	.035*	−1.059*	.005*	−0.731*	<0.001*	1.107*	<0.001*
BEN	AUT	−0.088	.997	0.206	.96	−0.042	.997	0.258	.442
BEN	EXP	.882*	.007*	−1.472*	<0.001*	−0.206	.479	0.113	.947
N-M	JUS	−0.121	.99	0.294	.867	−0.316	.086	_572*	.002*
N*M	AUT	−0.957*	.003*	1.559*	<0.001*	.372*	.026*	−0.277	.368
N-M	EXP	0.013	1	−0.118	.995	0.209	.462	−0.421*	.047*
JUS	AUT	−0.836*	.013*	1.265*	<0.001*	.689*	<0.001*	−0.849*	<0.001*
JUS	EXP	0.135	.985	−0.412	.653	_525*	<0.001*	−0.994*	<0.001*
AUT	EXP	.971*	.002*	l.676*	<0.001*	0.164	.695	0.145	.878

Table B4
Tukey HSD of health data scenario ethical principle pairs.

Health data scenario pairs		ACT				DON'T ACT			
		Rating		Ranking		Rating		Ranking	
		Mean diff	p-value	Mean diff	p-value	Mean diff	p-value	Mean diff	p-value
BEN	N-M	.684*	<0.001*	−0.149	.779	0.486	.304	−0.462	.592
BEN	JUS	−0.777*	<0.001*	1.208*	<0.001*	−0.003	1	0.231	.95
BEN	AUT	-l.205*	<0.001*	1.968*	<0.001*	−0.082	.998	0.359	.789
BEN	EXP	.360*	.03*	−0.039	.998	0.243	.87	−0.256	.928
N-M	JUS	1.461*	<0.001*	1.357*	<0.001*	−0.489	.292	0.692	.19
N*M	AUT	−1.890*	<0.001*	2.117*	<0.001*	−0.568	.159	0.821	.076
N-M	EXP	−0.324	.068	0.11	.914	−0.243	.87	0.205	.967
JUS	AUT	−0.429*	.004*	.760*	<0.001*	−0.079	.998	0.128	.994
JUS	EXP	1.136.	<0.001*	−1.247*	<0.001*	0.246	.862	−0.487	.54
AUT	EXP	1.565*	<0.001*	2.006*	<0.001*	0.325	.692	0.615	.299

Table B.5
Tukey HSD of 2FA scenario ethical principle pairs.

2FA scenario pairs		ACT				DON'T ACT			
		Rating		Ranking		Rating		Ranking	
		Mean diff	p-value	Mean diff	p-value	Mean diff	p-value	Mean diff	p-value
BEN	N-M	1.421*	<0.001*	−1.336*	<0.001*	−0.590*	.004*	.913*	<0.001*
BEN	JUS	1.153*	<0.001*	−1.381*	<0.001*	.−0.513*	.018*	.875*	<0.001*
BEN	AUT	.673*	<0.001*	−0.593*	.006*	−1.317*	<0.001*	2.000*	<0.001*
BEN	EXP	0.265	0.431	−0.363	0.23	−0.167	0.852	0.087	.991
N-M	JUS	−0.268	0.426	−0.044	0.999	0.077	0.99	−0.038	1
N*M	AUT	−0.748	<0.001*	.743*	<0.001*	−0.727*	<0.001*	1.088*	<0.001*
N-M	EXP	−1.155	<0.001*	.973*	<0.001*	0.423	0.081	−0.825*	<0.001*
JUS	AUT	−0.480*	.013*	.788-	<0.001*	−0.804*	<0.001*	1.125*	<0.001*
JUS	EXP	−0.887*	<0.001*	1.013*	<0.001*	0.346	0.226	−0.788*	.001*
AUT	EXP	−0.407	0.069	0.23	0.68	1.150*	<0.001*	−1.912*	<0.001*

Joseph Fenech, School of Computing, Macquarie University

Joseph Fenech was awarded a Masters of Research (MRes) in 2021 for a thesis on Ethical principles shaping cybersecurity decision-making motivated by an interest in human agency in cybersecurity decision-making. Prior to this he completed a Master of IT in Networking where he undertook a research project focused on Authentication protocols on Wireless Sensor Networks. He also completed a BCompSci at University of New South Wales. He aims to commence a PhD building on the MRes project in the near future.

Deborah Richards, School of Computing, Macquarie University

Deborah Richards is a Professor in the School of Computing at Macquarie University. Following 20 years in the IT industry during which she completed a BBus (Comp and MIS) and MAppSc (InfoStudies), she completed a PhD in artificial intelligence on the reuse of knowledge at the University of New South Wales and joined academia in 1999. While she continues to work on solutions to assist ethical decision-making and knowledge acquisition, for the past decade, her focus has been on intelligent virtual agents, virtual worlds and serious games for education, health learning and well-being to challenge attitudes and empower users to make good choices.

Paul Formosa, Department of Philosophy, Macquarie University

Paul Formosa is an Associate Professor in the Department of Philosophy at Macquarie University, and the Director of the Centre for Agency, Values and Ethics. Paul has published widely in topics in moral and political philosophy with a focus on Kantian ethics, the nature of evil, and the ethical issues raised by videogames, technology, cybersecurity and AI. His-work has been published with *Oxford* and *Cambridge University* Presses and in journals such as *Ethics and Information Technology, Games and Culture, European Journal of Philosophy*, and *Ethical Theory and Moral Practice*.

References

Abomhara, M., Køien, G.M., 2015. Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks. J. Cyber Secur. Mobility 65–88.

ABS. Australian standard classification of cultural and ethnic groups (ASCCEG); 2019. Available from: https://www.abs.gov.au/statistics/classifications/australian-standard-classification-cultural-and-ethnic-groups-ascceg/2019. [Accessed 25/10/2021 2021].

ACM. ACM code of ethics and professional conduct; 2018. Available from: https://www.acm.org/code-of-ethics. [Accessed 24/4/2021 2021].

Aguinis, H., Bradley, K.J., 2014. Best practice recommendations for designing and implementing experimental vignette methodology studies. Organ Res. Methods 17 (4), 351–371.

Anderson, R., 1993. Why cryptosystems fail. In: Proceedings of the 1st ACM Conference on Computer and Communications Security, pp. 215–227.

Awan, I., 2014. Islamophobia and Twitter: a typology of online hate against Muslims on social media. Policy Internet 6 (2), 133–150.

Bandura, A., 2006. Toward a psychology of human agency. Perspect. Psychol. Sci. 1 (2), 164–180.

Beauchamp, T.L., Childress, J.F., 2001. Principles of Biomedical Ethics. Oxford University Press, USA.

Brey, P., 2007. Ethical aspects of information security and privacy. Security, Privacy, Trust Modern Data Manage. 21–36.

Brienza, J.P., Bobocel, D.R., 2017. Employee age alters the effects of justice on emotional exhaustion and organizational deviance. Front. Psychol. 8, 479.

CANVAS. canvas_reference-curriculum_q-and-a_case-studies; 2020. Available from: https://canvas-project.eu/assets/results/canvas_reference-curriculum_q-and-a_case-studies.pdf. [Accessed 20/11/2021 2021].

Chaytor, N., Schmitter-Edgecombe, M., 2003. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol. Rev. 13, 181–197.

Christen, M., Gordijn, B., Loi, M., 2020. The Ethics of Cybersecurity. Springer Nature.

Coates, R., Baruwal Chhetri, M., Liu, D., Pieprzyk, J., Richelle, R., Kang, W., Kwashie, S., Wu, T., Nepal, S., 2023. Risks of quantum computing to cybersecurity: a responsible innovation approach (ed). In: Pulications, CR (Ed.), CSIRO. Research Pulications Repository.

Coughlan, M., Cronin, P., Ryan, F., 2009. Survey research: process and limitations. Int. J. Ther. Rehabil. 16 (1), 9–15.

Cullati, S., Courvoisier, D.S., Charvet-Bérard, A.I., Perneger, T.V., 2011. Desire for autonomy in health care decisions: a general population survey. Patient Ed. Counseling 83 (1), 134–138.

Datto. Datto's global state of the channel ransomware report. 2020.

Dunn Cavelty, M., 2014. Breaking the cyber-security dilemma: aligning security needs and removing vulnerabilities. Sci. Eng. Ethics 20, 701–715.

Fishbein, M., Ajzen, I., 2011. Predicting and Changing behavior: The reasoned Action Approach. Taylor & Francis.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., 2018. AI4people-an ethical framework for a good ai society: opportunities, risks, principles, and recommendations. Minds. Mach. (Dordr) 28 (4), 689–707.

Formosa, P., Wilson, M., Richards, D., 2021. A principlist framework for cybersecurity ethics. Comput. Secur. 109, 102382.

Gerber, A.S., Rogers, T., 2009. Descriptive social norms and motivation to vote: everybody's voting and so should you. J. Polit. 71 (1), 178–191.

Goldberg, L.R., 1993. The structure of phenotypic personality traits. Am. psychol. 48 (1), 26.

Gosling, S.D., Rentfrow, P.J., Swann Jr, W.B, 2003. A very brief measure of the Big-Five personality domains. J. Res, Pers. 37 (6), 504–528.

Graham, J., Nosek, B.A., Haidt, J., Iyer, R., Koleva, S., Ditto, P.H., 2011. Mapping the moral domain. J. Pers. Soc. Psychol. 101 (2), 366–385.

Gratian, M., Bandi, S., Cukier, M., Dykstra, J., Ginther, A., 2018. Correlating human traits and cyber security behavior intentions. Comput. Secur. 73, 345–358.

Hampden-Turner, C., Trompenaars, F., Hampden-Turner, C., 2020. Riding the Waves of culture: Understanding diversity in Global Business. Hachette UK.

Hampton, N., Baig, Z.A., 2015. Ransomware: emergence of the cyber-extortion menace. In: 13th Australian Information Security Management Conference. SRI Security Research Institute, Edith Cowan University, pp. 47–56.

Hofstede, G., 2003. Culture's consequences: Comparing values, behaviors, Institutions and Organizations Across Nations. Sage publications.

Hoonakker, P., Bornoe, N., Carayon, P., 2009. Password authentication from a human factors perspective: results of a survey among end-users. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Los Angeles, CA. SAGE Publications Sage CA, pp. 459–463.

Junglas, I.A., Johnson, N.A., Spitzmüller, C., 2008. Personality traits and concern for privacy: an empirical study in the context of location-based services. Eur. J. Inf. Syst. 17 (4), 387–402.

Kalimeri, K., Beiró, M.G., Delfino, M., Raleigh, R., Cattuto, C., 2019. Predicting demographics, moral foundations, and human values from digital behaviours. Comput. Human Behav. 92, 428–445.

Kharlamov, A., Pogrebna, G., 2019. Using human values-based approach to understand cross-cultural commitment toward regulation and governance of cybersecurity. Regul. Gov. n/a(n/a).

Køien, G.M., 2019. Why Cryptosystems Fail Revisited. Wirel. Pers. Commun. 106 (1), 85–117.

Krosnick, J.A., Alwin, D.F., 1987. An evaluation of a cognitive theory of response-order effects in survey measurement. Public Opin. Q. 51 (2), 201–219.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer.

Loi, M., Christen, M., 2020. Ethical frameworks for cybersecurity. Eds.. In: Christen, M, Gordijn, B, Loi, M (Eds.), The Ethics of Cybersecurity. Springer International Publishing, Cham, pp. 73–95.

Loi, M., Christen, M., Kleine, N., Weber, K., 2019. Cybersecurity in health–disentangling value tensions. J. Inf., Commun. Ethics Soc.

Maennel, K., Mäses, S., Maennel, O, 2018. Cyber Hygiene: The Big Picture. Nordic Conference On Secure IT Systems. Springer, pp. 291–305.

Manjikian, M., 2018. Cybersecurity ethics: an Introduction. Routledge, New York.

Nancy Carter, R., Bryant-Lukosius, D., Alba DiCenso, R, 2014. The Use of Triangulation in Qualitative research. Oncology nursing Forum. Oncology Nursing Society, p. 545.

Pfleeger, S.L., Sasse, M.A., Furnham, A., 2014. From weakest link to security hero: transforming staff security behavior. J. Homeland Secur. Emergency Manage. 11 (4), 489–510.

Pólkowski, Z, 2015. Ethical Issues in the Use and implementation of ICT. Sankalpa: J. Manage. Res. 2–5 ed R Khajuria, R Banerjee i K Sinha, 4th International Conference on "Business Ethic for Good Corporate Governance & Sustainability", Gujarat Technological University, Ahmedabad.

Ramirez, R., Mukherjee, M., Vezzoli, S., Kramer, A.M., 2015. Scenarios as a scholarly methodology to produce "interesting research. Futures. 71, 70–87.

Ross, D., 2002. The Right and the Good. Oxford University Press.

Schwartz S. Coding and analyzing PVQ-RR data (instructions for the revised Portrait Values Questionnaire)2016.

Schwartz, S.H., 2007. Basic human values: theory, measurement, and applications. Revue française de sociologie 47 (4), 929.

Schwartz, S.H., 2017. The Refined Theory of Basic Values (eds). In: Roccas, S, Sagiv, L (Eds.), Values and Behavior: Taking a Cross Cultural Perspective. Springer International Publishing, Cham, pp. 51–72.

Schwartz, S.H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., 2012. Refining the theory of basic individual values. J. Pers. Soc. Psychol. 103 (4), 663.

Sheldon, K.M., Houser-Marko, L., Kasser, T., 2006. Does autonomy increase with age? Comparing the goal motivations of college students and their parents. J. Res. Pers. 40 (2), 168–178.

Stephens, J.M., 2018. Bridging the divide: the role of motivation and self-regulation in explaining the judgment-action gap related to academic dishonesty. Front. Psychol. 9, 246.

Tenbrunsel, A.E., Bazerman, M.H., 2011. Blind Spots: Why We Fail to Do What's Right and What to Do about It. Princeton University Press.

Ullmann, S., Tomalin, M., 2020. Quarantining online hate speech: technical and ethical perspectives. Ethics Infor. Technol. 22 (1), 69–80.

Vishwanath, A., Neo, L.S., Goh, P., Lee, S., Khader, M., Ong, G., 2020. Cyber hygiene: the concept, its measure, and its initial tests. Decis. Support Syst. 128, 113160.

Yaghmaei E., van de Poel I., Christen M., Gordijn B., Kleine N., Loi M., et al. Canvas white paper 1–cybersecurity and ethics. Available at SSRN 3091909 2017.

Yan, Z., Robertson, T., Yan, R., Park, S.Y., Bordoff, S., Chen, Q., 2018. Finding the weakest links in the weakest link: how well do undergraduate students make cybersecurity judgment? Comput. Human. Behav. 84, 375–382.