Domain Knowledge and Self-reflective Skills as Keys to Ethical Design Reflection

Sebastian Wilhelm, S2354896
s.wilhelm@student.utwente.nl
Department of Psychology, Health and Technology, University of Twente
Bachelor Thesis: Human Resources Development
First Supervisor: S.M. Dennerlein, Dr.
Second Supervisor: N. Goossen, Msc.
July 2, 2022

Abstract

Prior research was able to establish frameworks that give guidance on how ethical reflection can be incorporated into the design of modern technologies. Still, little is known about the characteristics of humans that are of importance for ethical design reflection. As previous studies have been able to demonstrate the importance of reflective skills and domain knowledge in ethical reflection, it was hypothesised that people with higher self-reflective skills and domain knowledge are better able than people with lower self-reflective skills and domain knowledge to identify unique and elaborate ethical issues and pose well-structured reflective questions with domain-specific content. The Metaverse hereby served as a case example which was reflected upon. A correlational survey design was employed in which the participants were asked to identify one potential ethical issue in the Metaverse and subsequently asked to formulate respective reflective questions that could be prompted during the design of the Metaverse. The responses were qualitatively assessed and analysed by performing linear regression analyses. All hypotheses had to be rejected except the hypothesis that predicted a significant effect of self-reflective skills on the number of reflective questions. The results of the study mainly implicate that there is a further need to investigate the variables at hand. Still, this study was the first to propose steps for the qualitative assessment of the quantity and quality of ethical design reflection and explored the role of human characteristics in ethical design reflection.

Keywords: Ethical design reflection, self-reflective skills, domain knowledge, Metaverse

“It's not a faith in technology. It's faith in people.”

- Steve Jobs, The Rolling Stone

Nowadays, technology is advancing at a rapid pace. With this advancement, the need for ethically informed decisions is growing too. There are numerous recent instances that underline the susceptibility of modern technologies to ethical issues, such as the Cambridge Analytica case, where data from millions of Facebooks users has been accumulated to influence political elections all over the globe (Venturini & Rogers, 2019). Some further examples are facial analysis software that has shown to hold biases against certain ethnicities, genders and races, or fake news and false information distributed via social media, including Facebook and Instagram, that fuelled riots and violence in India (Khalil, Ahmed, Khattak, & Al-Qirim, 2020; Khan, Alkawaz, & Zangana, 2019). As can be seen, social media is a vulnerable technology when it comes to ethical issues.

In order to avoid unethical consequences like those in the examples above, recent research is concerned with design ethics. In general, design ethics can be defined as ethics that are concerned with the morality, the choices, and the behaviour of designers (Buchanan, 2005). For ethics to be applicable to designers, normative ethics are concerned with capturing the ethical principles that can guide designers' behaviour. Ethical principles, which can be characterised as normative judgements that are meant to guide, justify, and evaluate human behaviour, are then utilised to establish frameworks that can be applied to incorporate design ethics in the workplace to allow for reflection and the avoidance of unethical consequences (Saint Mary’s College, n.d.).

One such framework is the Socio-Technical Reflection of Ethical Principles (SREP) framework (Dennerlein et al., 2020). Important ethical principles that underlie this framework, derived from Jobin et al. (2019), are transparency, justice and fairness, privacy, non-maleficence, and responsibility and accountability. In order to apply these ethical principles, the SREP framework gives direction on when and how reflection can be triggered during the design and usage phases of digital technologies. More precisely, the SREP suggests triggering reflection by identifying ethical issues and then prompting reflective questions between socio-technical design cycles and phases. Ethical issues in this context are understood as issues that are expressed by socio-technical systems as the result of socio-technical interactions (STIs), interactions between humans and technology, and reflective questions are seen as prompts that can trigger reflection during the design of technologies (Dennerlein et al., 2020).

In another ethical framework focused on AI, Ashok et al. (2022) emphasise the human role and the ethical responsibility that humans hold during the design of modern technologies. Still, little is known about what specific characteristics in humans are important to comply with ethical responsibility. Since it is the responsibility of humans to engage in reflective practice during STIs, it will be of value for future professional training to get to know what determinants predict whether humans are able to engage effectively in reflection. Consequently, the aim of this study is to explore human characteristics that are predictive of high-quality ethical design reflection.

Reflective Skills

One determinant that has shown to play an important role in ethical decision-making is the reflective skill of the decision-maker. As Grant and Franklin (2015) define it, self-reflection means observing and evaluating one's feelings, thoughts, and behaviours. Additionally, they suppose it can be measured on two dimensions, the need for self-reflection and the engagement in self-reflection. Regarding ethics, strengthening reflective skills is leading to improvements in ethical decision-making and in learning in general (Tong, 2008). Furthermore, in a study design in which students used blogs to reflect on their learning as well as on their ethical decision-making, Riley et al. (2006) showed that a blog, representing a reflective tool, was successful in integrating ethics within the course learning material. Another example stems from economics. In a business context, where unethical consequences are rather common, the improvement of reflective awareness has shown to contribute to more sustainable and ethical decisions (Bożac et al., 2021). To further explore this domain, the current study is investigating whether reflective skills are of importance in the context of ethical design reflection. More specific, connecting it to the research of Dennerlein et al. (2020), it is being explored what role self-reflective skills play in the identification of ethical issues and the formulation of reflective questions in the Metaverse.

Domain Knowledge

Another determinant that has shown to be important in ethical decision-making is domain knowledge. In a study design concerned with the skills of health care data scientists, Baig & Alzahrani (2019) demonstrated that having domain knowledge is crucial for compliance with ethical considerations, patient privacy and consensual information. This connection of domain knowledge to ethics was also linked to Cognitive Load Theory (CLT). According to Chi et al. (1984; as cited in Moreno 2004), CLT predicts that when prior knowledge on a domain is low, the cognitive load is high because no schema is available to the individual. The other way around, if prior knowledge on a domain is high, the cognitive

load is lower, schemas are available to the individual and cognitive resources for reflection, ethical considerations or prompting are available (Moreno, 2004). This shows that domain knowledge might determine the amount of cognitive capacities that are available for ethical design reflection. As domains can be multifaceted within ethical design reflection, as it might comprise of knowledge on ethics, knowledge on technology, etc., it can also be stated that domain knowledge is multidimensional in ethical design reflection.

Adding to this, a lack of proper consideration of domain knowledge in educational programs is highlighted in research. Oliver & McNeil (2021), for instance, pointed out that there is a general problem in undergraduate data science degrees, namely the lack of incorporation of domain knowledge and ethical considerations into the curriculum. At the same time, Oliver & McNeil (2021) emphasise the importance of greater incorporation.

To conclude, as the importance of domain knowledge in the application of ethics is highlighted frequently in research, this study is investigating whether domain knowledge is of equal importance in the identification of ethical issues and the formulation of reflective questions.

Quantity and Quality of the Reflection

Following Dennerlein et al.(2020), ethical design reflection can be defined as the skill to identify ethical issues and to pinpoint reflective questions that can be posed to avoid a certain ethical issue. Therefore, to distinguish good reflection from bad reflection in the context of ethical design reflection and to fill a gap in research, it had to be established what qualities of ethical issues and reflective questions are in the context of ethical design reflection and how the quantity of formulations can be measured.

First, it was investigated how ethical issues can be assessed. Jobin et al. (2019) ranked eleven ethical principles according to their prevalence in documents. These eleven principles are transparency, justice and fairness, non-maleficence, responsibility and accountability, privacy, beneficence, freedom & autonomy, trust, sustainability, dignity, and solidarity. Hence, there is data available on the prevalence of ethical principles, with some ethical principles being covered frequently and some less frequently. The uniqueness of ethical principles underlying ethical issues can therefore be treated as a quality of ethical issues, as coming up with more unique ethical issues is likely to be more difficult than coming up with less unique issues since less popular ethical principles are covered less in the media or in social movements for example.

The quality of reflective questions can be understood in terms of their content and structure. First, regarding the content of ethical design reflection, the research of Razavian et

al. (2016) offers a proposal for the distinction of key areas in ethical design reflection. Hereby, reflection is categorised into four key areas of software design reflection: *Reflect on the Context and Requirements*, reflection on relevant contextual aspects and requirements, *Reflect on Design Problems*, reflection on problems during the design, *Reflect on Design Solutions*, reflection on solutions to problems in the design, and *Reflect on Design Decisions*, reflection that is concerned with design decisions and the reasoning behind it. These categories are deemed appropriate for the reflection on design problems of digital technologies since possible ethical issues are inexplicably linked to software design.

Regarding the structure of reflective questions, Dennerlein et al. (2020) identified important components for ethical design reflection in their SREP framework. Here, socio-technical systems theory informs the structure of ethical design reflection and suggests incorporating four main components: a *Human actor*, a *Technology*, a *STI*, and an *Ethical Principle*. As these components are deemed important to be considered during ethical design reflection, the SREP can be utilised to assess the structure of reflective questions.

The assessment of the quantity of reflection is straightforward and is, in this study, seen as an indicator of the motivation to reflect. As Ullmann et al. (2013) stated, the length of text can already be a good indicator of the quantity of reflection, especially for shorter pieces of text. Hence, an ethical issue can be assessed by looking at how elaborate its formulation is, and the quantity of reflective questions can be assessed by simply counting the number of questions that can be formulated for a single ethical issue.

To summarise, researchers have established well-designed frameworks that can help in avoiding unethical consequences in the design and usage phases of digital technologies (Dennerlein et al., 2020; Ashok et al., 2022). Nonetheless, these frameworks must be applied by human beings, and there is little known by now about what characteristics and determinants predict people's skills in identifying ethical issues and formulating reflective questions. Eventually, this study is the first to propose an operationalisation of ethical design reflection and represents a first attempt to research the importance of human determinants for ethical design reflection. Since research has shown that *reflective skills* and *domain knowledge* are important determinants in ethical decision-making, we hypothesise that high levels of *reflective skills* and high levels of *domain knowledge* in people are predictive of the ability to identify unique and elaborate ethical issues and predictive of the ability to formulate well-structured reflective questions that are domain-specific in their content. Accordingly, the following hypotheses have been established.

H1: Participants with higher reflective skills will create more elaborate ethical issues than participants with lower reflective skills.

H2: Participants with higher domain knowledge will create more elaborate ethical issues than participants with lower domain knowledge.

H3: Participants with higher reflective skills will create more unique ethical issues than participants with lower reflective skills.

H4: Participants with higher domain knowledge will create more unique ethical issues than participants with lower domain knowledge.

H5: Participants with higher reflective skills will create more reflective questions than participants with lower reflective skills.

H6: Participants with higher domain knowledge will create more reflective questions than participants with lower domain knowledge.

H7: Participants with higher reflective skills will create reflective questions with more domain-specific content and better structure than participants with lower reflective skills.

H8: Participants with higher domain knowledge will create reflective questions with more domain-specific content and better structure than participants with lower domain knowledge.

Methods

In this study, a correlational survey design was employed in combination with a preceding qualitative assessment of the ethical issues and reflective questions created by the participants. The design was chosen to allow for an assessment of the relationship between the independent variables *domain knowledge* and *reflective skills* and the dependent variables *scope of elaboration of the ethical issue, uniqueness of ethical issue, number of reflective questions* and *quality of reflective questions*.

Participants

The required sample size was estimated by using G*Power (Faul et al., 2017). Thus, the calculations resulted in a recommended sample size of 42 participants (CI= 95%, p= .5). For the recruitment of participants, convenience sampling was used. First and foremost, the system "Sona" from the Faculty of Behavioural, Management and Social sciences (BMS) was used for recruitment. The system allows participants from the University of Twente to be rewarded with credits which are required for graduation. In addition, an Instagram story and a WhatsApp post were created to recruit further participants. To be allowed for participation,

two inclusion criteria for participants were applied. Firstly, participants must be enrolled at a university, and secondly, participants must be sufficiently proficient in the English language. The demographic data that was collected included the age, gender, nationality, and academic discipline of the participants. In the end, the responses of 19 male participants and 23 female participants between the age of 17 and 26 were analysed ($N = 42$, $M_{age} = 22.1$, $SD_{age} = 1.90$). Thirty-three of the participants were German, 3 Dutch, and 6 had other nationalities. With regard to the academic disciplines of participants, 27 participants came from the field of Psychology, 3 participants from Communication Science and 12 participants came from other study fields.

Material & procedure

To provide a subject for reflection, the Metaverse will serve as a case example in this study which is reflected upon. "Simply put, Facebook's Metaverse is a tightly interconnected set of digital spaces that lets users escape into a virtual world, and the rules of technology are the only limit" (XR Today, 2021). The goal is, as Meta envisions, that the Metaverse becomes the successor of the mobile internet, with its users being able to immerse themselves using human avatars (Meta, 2021). Consequently, Meta is creating a space in which many STIs take place and therefore, the Metaverse resembles an example of a technology for which ethical design reflection is of importance. Hence, the Metaverse was regarded as a suitable case example in the current study design.

Participants that decided to take part in the study were asked to fill in a survey using the survey tool and distribution software Qualtrics. At the beginning of the survey, participants were presented with informed consent, including the study purpose, possible risk of participation, and the possibility to withdraw at any moment (see Appendix A). Specific information on the variables under investigation was withheld. After consent was given, the participants were asked to fill in demographic data and skip to the next page, to begin with the study.

First, to measure the independent variable domain knowledge, participants were asked to fill in the Metaverse Domain Knowledge Questionnaire (MDKQ) that was created prior to the study. The MDKQ is an 18-item questionnaire consisting of a 6-item frequency matrix that assesses the usage frequency of social media and extended reality (XR) and a 12-item familiarity scale that measures domain knowledge on the Metaverse on three subscales: Familiarity with Meta & the Metaverse (4 items; $\alpha = .679$), Familiarity with Technology (4 items; $\alpha = .525$), and Familiarity with Ethics (4 items; $\alpha = .612$). Items that were used are, for instance: "I am familiar with ethical principles." or "I am familiar with XR (extended reality)

in the Metaverse." (see Appendix B). All Items were scored on a 5-point Likert scale. In this study, we found acceptable reliability of the overall questionnaire (18 items; $\alpha = .74$).

For the assessment of the independent variable reflective skills, the subscale SRIS-SR of an already existing questionnaire, the Self-Reflection and Insight Scale (SRIS), was used (Grant & Franklin, 2015; see Appendix C). The SRIS-SR is comprised of two subscales, the "Engagement in self-reflection" (6 items) and "Need for self-reflection" (6 items). An example of an item that is incorporated to assess engagement in self-reflection is: “I rarely spend time in self-reflection.” An example of an item that is incorporated to assess the need for self-reflection is: “It is important for me to evaluate the things that I do.” For this questionnaire, all items were scored on a 6-point Likert scale. Grant and Franklin (2015) reported a seven-week test-retest reliability of .77 ($p < 0.001$) for the SRIS-SR, and good convergent and discriminant validity was demonstrated.

When the participants were finished filling in the two questionnaires that assessed the independent variables of this study, domain knowledge and reflective skills, participants were presented with basic information on the Metaverse (Appendix D). This was done to ensure that a certain amount of knowledge was universal to all participants. Subsequently, the first reflective task was presented.

At first, a definition of what an ethical issue is was given. Then, the participants were instructed to formulate one ethical issue that they consider important for the Metaverse, and that should be avoided in their opinion. A max of five sentences was allowed for the response. Additionally, all participants were presented with the same two examples before the start (Appendix E).

After the formulation of the ethical issue, in the second reflective task, they were asked to reflect on this issue by phrasing reflective questions that could be posed during the design phase of the Metaverse to avoid the ethical issue identified. The participants had 15 minutes of time to come up with as many reflective questions as possible. The instructions emphasised that the quality of the reflective questions matters. After the task was done, the participants reached the end of the study, got debriefed about the precise purpose of the study and were free to leave the website. The study took most participants 10-15 minutes to complete.

Data Analysis

Piloting the qualitative assessment

In the first iterative step, a deductive approach guided the development of the qualitative assessment and literature was reviewed. The body of literature that was reviewed

focused mainly on the assessment of reflection, ethical decision-making processes, and the exploration of ethical principles. For literature to be suitable for the current study design, theories had to be simple enough to be applied to a single ethical issue and multiple ethical questions, as a vast body of literature is on the assessment of reflection of full written texts.

The second iterative step was the testing of the newly created qualitative assessment. To test the applicability of the qualitative assessment, a pilot study was performed with four participants filling out a prototype of the survey. Afterwards, the participants' responses were analysed using the qualitative assessment. Consequently, the framework was iterated until it was applicable to the kind of data created by the participants and until it became evident that the scheme could be used to effectively differentiate between participants.

Finally, the last iterative step was the researchers' agreement regarding the qualitative assessment's suitability for data analysis.

Operationalisation of the uniqueness and the scope of elaboration of ethical issues

First, the decision of whether an ethical issue is seen as realistic or not determined whether the uniqueness of the ethical issue was analysed. This decision was solely left to the coder. Realism was therefore used as a filter criterion. An ethical issue was considered realistic if the coder believed that there was a possibility that the ethical issue could occur in the Metaverse. If the coder did not have the required knowledge to inform this consideration, he conducted online research to evaluate the criterion of realism. If an ethical issue was coded as realistic, the assessment of the uniqueness followed. However, if an ethical issue was not coded as realistic, the assessment of the uniqueness was not considered at all.

The operationalisation of the uniqueness of ethical issues was based on the study of Jobin et al. (2019). They conducted a content analysis on 84 documents about AI ethical guidelines and ranked them based on the number of documents in which they were included (Appendix F). In the current study, an ethical issue that appears in fewer documents received a better evaluation. The evaluation of the ethical issue was therefore seen in proportion to the number of documents in which the underlying ethical principle appeared. As the least prevalent ethical principle *solidarity* appeared in six documents, and the most prevalent principle *transparency* in 73 documents, a score ranging from six to 73 was assigned. Moreover, the same codes as in the research of Jobin et al. (2019) were used to assign an ethical issue to an ethical principle (Appendix F). One example of an ethical issue could be: "Gender discrimination". One of the included codes for the ethical principle *justice and fairness* is "discrimination". Hence the ethical issue "Gender discrimination" was assigned to

the ethical principle *Justice and fairness*. As this principle appeared in 68 out of 84 documents in the research of Jobin et al. (2019), the ethical issue would yield a score of 68.

The *scope of elaboration* of the ethical issue was simply assessed by counting the number of words in the formulation of the ethical issue. Thus, the score that was assigned was equal to the number of words. For the example above, this would mean that a score of two was assigned for the ethical issue “Gender discrimination”.

Operationalisation of the content, structure and quantity of reflective questions

The assessment of the reflective questions that participants came up with was twofold. First, assessing the *content* of the reflective questions, the coder decided for each reflective question which of the four key areas, *Reflect on the Context and Requirements*, *Reflect on Design Problems, Reflect on Design Solutions*, and *Reflect on Design Decisions*, as identified by Razavian et al. (2016), has been incorporated. Subsequently, a score ranging from one to four was assigned for each key area that has been incorporated in the entirety of reflective questions formulated for the ethical issue. This is done to assess the complete reflection on one ethical issue, as a single question can only cover one key area at a time. If a key area was incorporated in one question, no additional scores were added if the key area was also incorporated in more than one question. In the end, the participant is assigned a score from zero to four, depending on the number of key areas that have been accounted for in the reflective questions.

Secondly, assessing the *structure* of the reflective questions, the coder decided which of the four main components, a *Human actor*, a *Technology*, a *STI*, and an *Ethical principle*, as identified by Dennerlein et al. (2020), were incorporated per question. Afterwards, a score ranging from one to four, depending on the number of components incorporated, was assigned per question. These scores were then added and divided through the number of questions created. Eventually, the resulting means were rounded to whole figures to allow for exact comparability with the scores obtained from the content analysis and the participant was assigned a score ranging from one to four.

After all, the scores from both assessments were added together, and a score ranging from zero to eight was assigned to the participant to determine the *quality score* as indicated by its content and structure.

To allow for replicability of the assessment of reflective questions, a fictional example is provided. Participant X identified the ethical issue of "Gender discrimination". Subsequently, participant X formulated three questions: "Which policies in the Metaverse can prevent Gender discrimination?", "Can the legal body interfere with what happens in the

Metaverse?" and "If women, for example, are discriminated in the Metaverse, could AI help to detect gender discrimination and act against it?

The assessment of the content of this response would work as follows. For the first question, the key area *Reflect on Context and Requirements*, is assigned, as the question is concerned with the requirement of policies. For the second question, again, the key area *Reflect on Context and Requirements,* is assigned as the question aims at the Metaverse's context within the legal body. Lastly, for the third question, the key area *Reflect on Design Solutions* is assigned since it is reflected upon solutions to the ethical issue. Accordingly, since two out of four key areas were incorporated in the reflective task of participant X, she is assigned a score of two.

Next, the assessment of the structure of the example response is illustrated. In the first question, a *Technology*, "Metaverse", and an *Ethical principle*, "discrimination", are accounted for. Hence, a score of two is assigned to the first question. In the second question, we only find a *Technology*, "Metaverse", and therefore, a score of one is assigned. For the third question, a *Human actor,* "women", a *Technology* "AI", an *Ethical principle*, "discrimination", and a *STI,* as the discrimination against women in the Metaverse is mentioned, are incorporated. Accordingly, a score of four can be assigned to the third question. When adding the scores of the questions together, a score of seven is reached. Dividing this score by three, as three questions have been created, a mean score of 2.33 is calculated. Rounding this number, we arrive at a score of two for the structure of the reflective questions of participant X. Finally, the score for the content and structure are added together, and a quality score of four is assigned to participant X.

The *quantity of reflective questions* was simply operationalised as the number of questions a participant was able to formulate in the second reflective task. Hence, the score assigned was equal to the number of questions formulated. In the example above, participant X came up with three questions. A score of three is then representing the quantity of reflective questions.

Quantitative analysis and hypothesis testing

After data collection had finished, the data was imported from Qualtrics into SPSS for analysis. First, the dataset was checked for missing values, incomplete responses, or responses that signalised comprehension problems. These responses were excluded from the final dataset.

The first hypothesis "Participants with higher reflective skills will create more elaborate ethical issues than participants with lower reflective skills.", was tested by

performing a linear regression analysis with the mean score of the SRIS-SR as the independent and the number of words of the ethical issue as the dependent variable. The second hypothesis “Participants with higher domain knowledge will create more elaborate ethical issues than participants with lower domain knowledge." was tested in the same manner, with domain knowledge being the independent variable.

Next, the third hypothesis “Participants with higher reflective skills will create more unique ethical issues than participants with lower reflective skills." was tested. Again, a linear regression analysis with the independent variable being the mean score of the SRIS and the dependent variable being the uniqueness of the ethical issue was conducted. The fourth hypothesis "Participants with higher domain knowledge will create more unique ethical issues than participants with lower domain knowledge.”, was tested by conducting the same analysis with domain knowledge being the independent variable this time.

In order to test the fifth hypothesis “Participants with higher reflective skills will create more reflective questions than participants with lower reflective skills.”, a linear regression analysis with the mean score of the SRIS being the independent and the number of reflective questions being the dependent variable, was performed. The same analysis, with the independent variable being the mean score of the MDKQ, was performed to test the sixth hypothesis “Participants with high-level domain knowledge will create more reflective questions than participants with low-level domain knowledge”.

After that, the seventh hypothesis “Participants with higher reflective skills will create reflective questions with more domain-specific content and better structure than participants with lower reflective skills.”, was tested by performing a linear regression analysis with the mean score of the SRIS as the independent variable, and the quality score of the reflective questions as the dependent variable. Finally, the eighth hypothesis “Participants with higher domain knowledge will create reflective questions with more domain-specific content and better structure than participants with lower domain knowledge.” was tested using the same analysis, with the independent variable being the mean score of the MDKQ this time**.**

Results

Data was collected from the 14th of April 2022 until the 9th of May 2022. In total, 71 responses were collected. After responses with a completion rate under 100% had been deleted, 44 responses remained. Additionally, two responses have been erased from the dataset because no consent was given. In the end, 42 responses remained on which the following analyses were conducted.

To begin with, the mean scores obtained on the SRIS-SR ranged from 3 to 6 (M = 4.94, SD = 0.63) and the mean scores of the MDKQ from 2 to 4 (M = 2.90, SD = 0.38). The scope of elaboration of the ethical issues ranged from 1 to 60 words (M = 17.62, SD = 17.68) and the scores for uniqueness from 34 to 68 (M = 57.10, SD = 12.64). Participants came up with 0 to 12 reflective questions (M = 4.12, SD = 2.44) and the quality scores of the questions ranged from 0 to 7 (M = 3.38, SD = 1.41).

Qualitative analysis

Uniqueness

The first task required participants to come up with one ethical issue that they deemed important during the design of the Metaverse. At first, the researcher decided whether the ethical issue was realistic. Only then the further analyses were conducted. From 42 responses, 40 were classified as realistic.

Subsequently, the ethical issue was assessed by assigning a score from 6 to 73, representing the uniqueness score of the ethical issue. In this study, the scores for uniqueness ranged from 34 to 68 (M = 26,9, SD = 12,64). The frequencies of the specific principles can be seen in Table 1.

Table 1

Frequencies of ethical principles

	Frequency	Percent
Transparency (73/84)	0	0
Justice and Fairness (68/84)	16	40
Non-maleficence & Responsibility (60/84)	13	32.5
Privacy (47/84)	0	0
Beneficence (41/84)	6	15
Freedom and autonomy (34/84)	5	12.5
Trust (28/84)	0	0
Sustainability (14/84)	0	0
Dignity (13/14)	0	0
Solidarity (6/84)	0	0

Content and structure

Next, to arrive at a single quality score for the reflective questions that participants came up with, the questions were qualitatively assessed regarding their content and their structure. Regarding the content of reflection, the researcher coded what key areas have been accounted for in the totality of questions one participant came up with. The number of key areas accounted for ranged from 0 to 3 (M = 1.79, SD = 0.90).

The key area *Reflect on context and requirements* was incorporated by 18 participants (57.1%) in at least one reflective question. An example of a reflective question that was considered to be about the context and requirements of ethical design reflection in the Metaverse is: “Can the laws be discriminating by any means without veto by a power which can object to laws?”.

Next, the key area *Reflect on design problems* was the most frequent key area, and it was accounted for by 23 participants (54.8%) in at least one question. An example of a question that incorporated this key area is: “How can we prevent people from transferring status and value of property in the metaverse to the outside world which leads to discrimination?”

The key area Design solutions was accounted for by 13 participants (31%) in at least one question and, hence, it was the key area that was accounted for the least. A reflective question that incorporated this key area was, for instance: “Could it have good consequences, if people could live out their fantasies of violence in the metaverse? Meaning that this might decrease the likelihood of violence in the real world.”.

Lastly, the key area Design decisions was accounted for by 22 participants (52.4 %), and an example of a reflective question that incorporated this key area is: "Should we enable the possibility to "play" in the Metaverse with a "fake profile" or under a different name to prevent discrimination?".

Subsequently, the reflective questions’ structure was qualitatively assessed. Hereby, the researcher decided whether a question included a Technology, a Human actor, a Socio-technical interaction (STI), and an Ethical principle. The scores obtained ranged from 0 to 4 (M = 1.62, SD = 0.80).

An example of a question which illustrated poor structure and which gave a score of one to the participant is "Can the power be monopolised?". Here, only the ethical principle, justice and fairness, is incorporated, whereas the other three components are neglected. Furthermore, an example of a question that yielded a score of two is: “How can people be influenced negatively by the metaverse and how do you enlighten about the risks?”. For this

example, the participant has incorporated Human actors, “people”, and the Technology “Metaverse”. Nonetheless, the STI and Ethical Issue are not stated precisely. An instance of a question which was given a score of three is “What technologies could be put in place to avoid it being a base for possible discrimination by other users or even the software itself?”, as here the participant accounted for a Technology, a Human actor, an Ethical issue but no specific STI. At last, an example of a question in which all four components have been accounted for is “Is the Metaverse designed in a way that digital components that the user interacts with could make him/her addicted to the Metaverse?”.

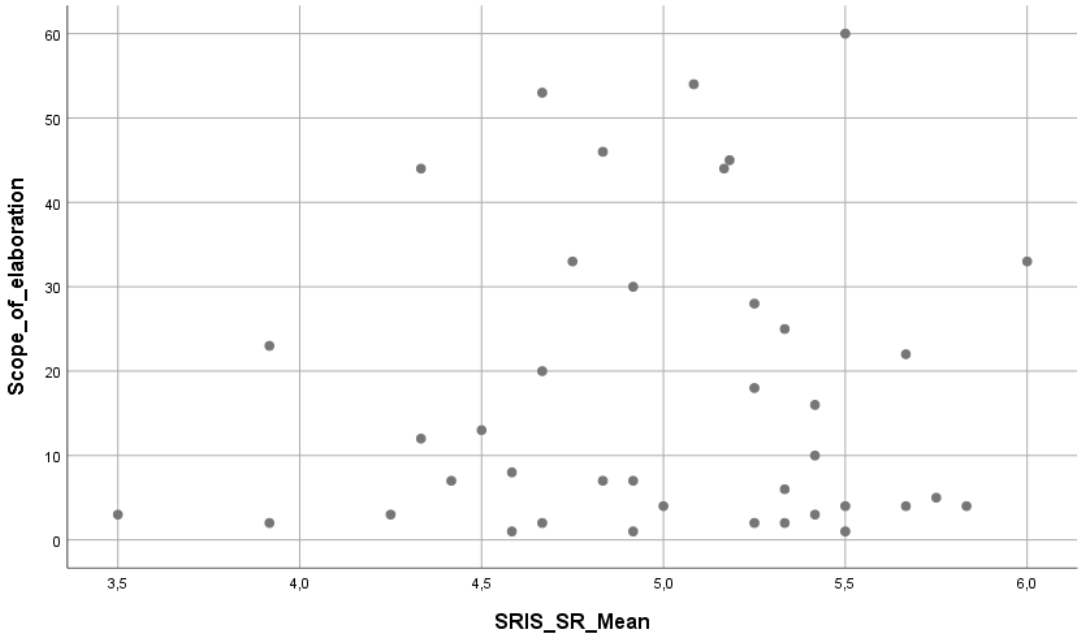
Afterwards, the participants were assigned a score ranging from zero to four, depending on the mean number of components of the SREP incorporated in one question. This score was then added to the score from the first analysis, and a score ranging from zero to eight was assigned. Consequently, this score equals the quality score of a participant for the reflective question he or she created.

Quantitative analysis

In order to test the first hypothesis, a linear aggression analysis with the scope of elaboration of the ethical issue as the dependent variable and the mean score of the SRIS-SR as the independent variable has been performed. A scatterplot describing the relationship between the two variables can be seen in Figure 1. The analysis indicated that the means score of the SRIS-SR did not explain a significant proportion of the variance in the scope of elaboration ($R^2 = .00$, $F (1, 38) = .17$, $p = .685$). The first hypothesis was therefore rejected.

Figure 1

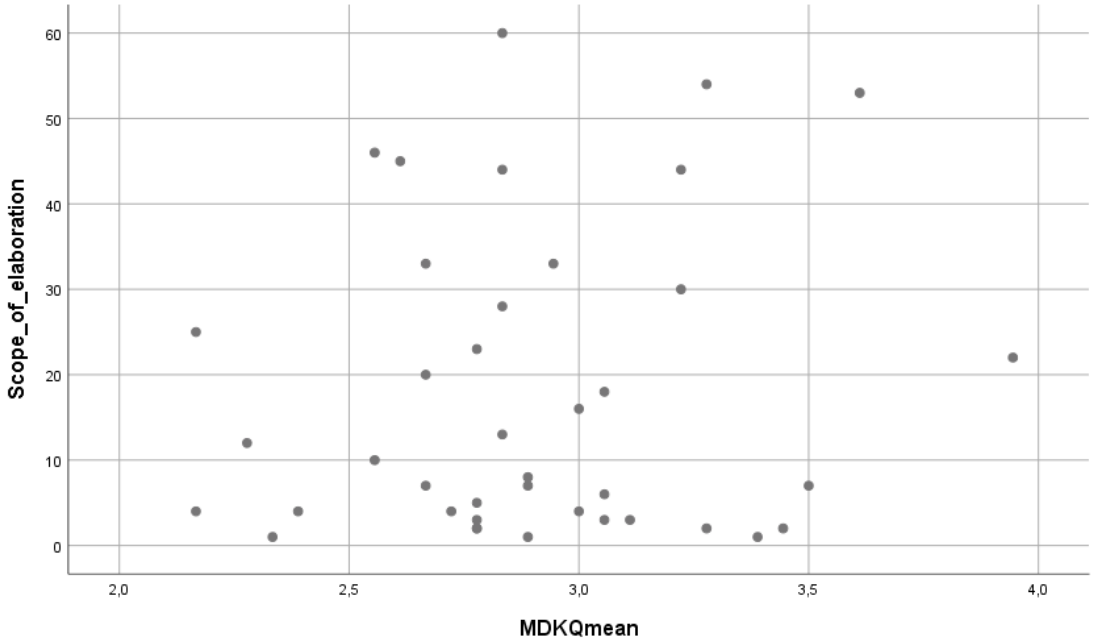
Scatterplot of the scope of elaboration and the SRIS-SR mean



The second hypothesis was tested by performing a linear regression analysis with the *scope of elaboration* of the ethical issue as the dependent and the mean score of the MDKQ as the independent variable. A scatterplot that illustrates the relationship between the two variables can be seen in Figure 2. The analysis revealed that the mean score of the MDKQ did not explain a significant proportion of the variance in the *scope of elaboration* (R^2 = .01, F (1, 38) = .53, p = .470). The second hypothesis was therefore rejected as well.

Figure 2

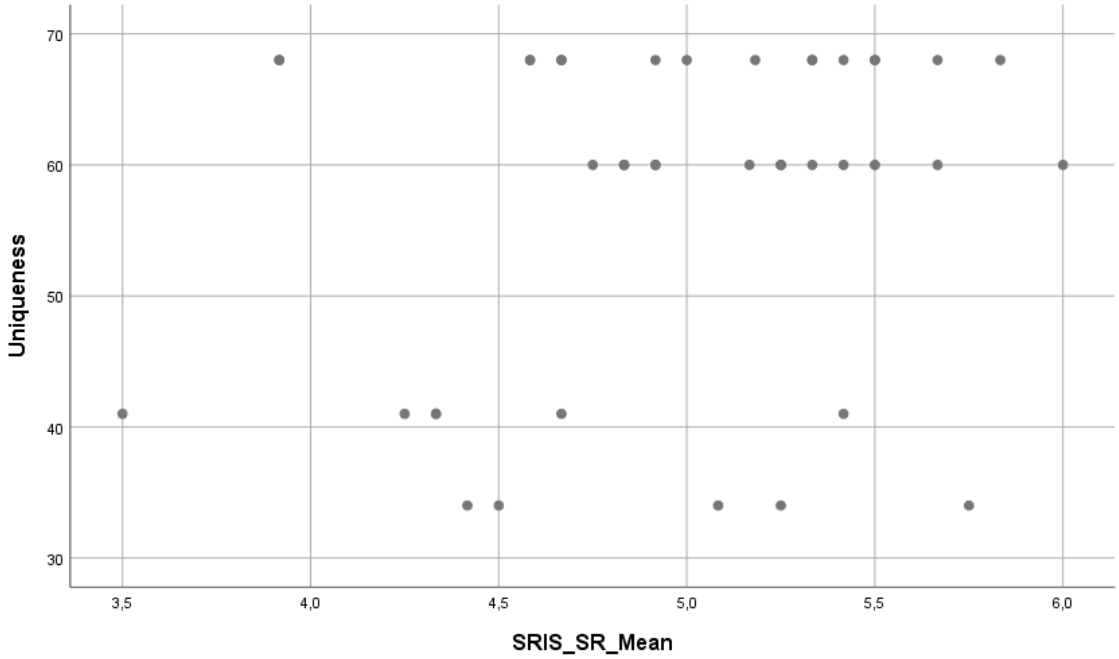
Scatterplot of the scope of elaboration and the MDKQ mean



To test the third hypothesis, a linear regression analysis with the *uniqueness* of the ethical issue as the dependent and the mean score of the SRIS-SR as the independent variable was performed. The scatterplot in figure 3 illustrates the relationship between the two variables. The results suggest that the mean score of the SRIS-SR did not explain a significant proportion of the variance in *uniqueness* (R^2 = .05, F (1, 38) = 2.17, p = .182). Hence, the third hypothesis was rejected.

Figure 3

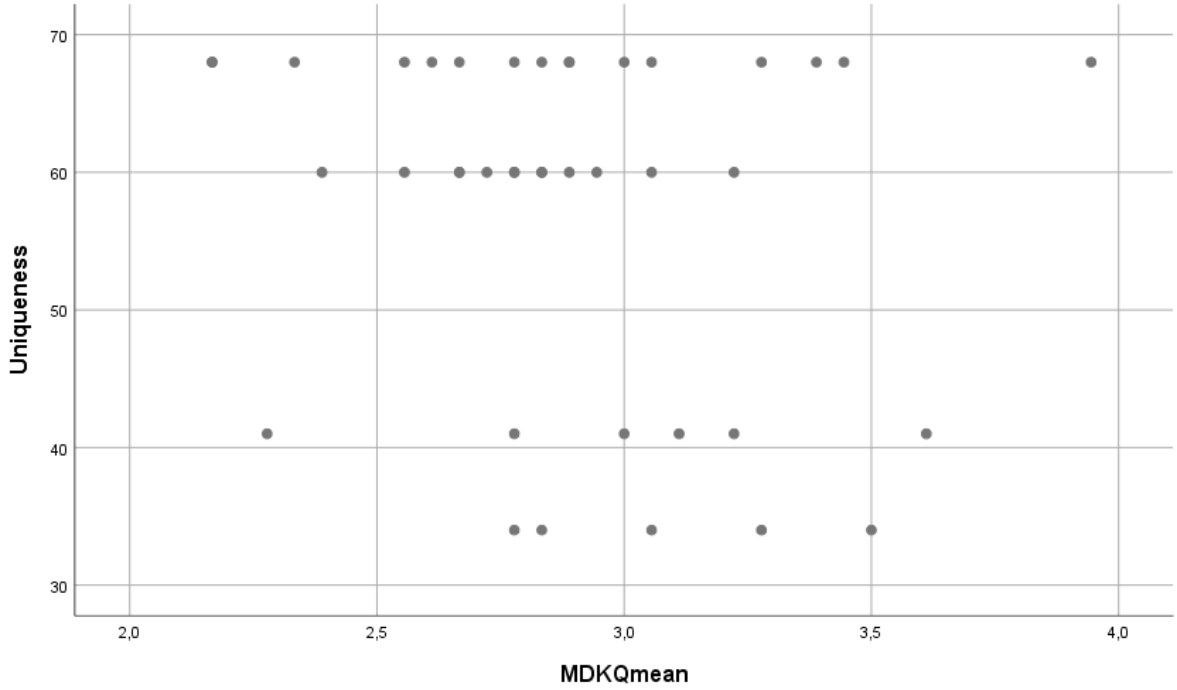
Scatterplot of the uniqueness and the SRIS-SR mean



Next, a linear regression analysis with the *uniqueness* of the ethical issue as the dependent and the mean score of the MDKQ as the independent variable was performed to test the fourth hypothesis. An illustration of the relationship between the two variables can be seen in Figure 4. The analysis indicated that the mean score of the MDKQ did not explain a significant proportion of the variance in the *uniqueness* (R^2 = .04, F (1, 38) = 1.75, p = .194). Accordingly, the fourth hypothesis was rejected.

Figure 4

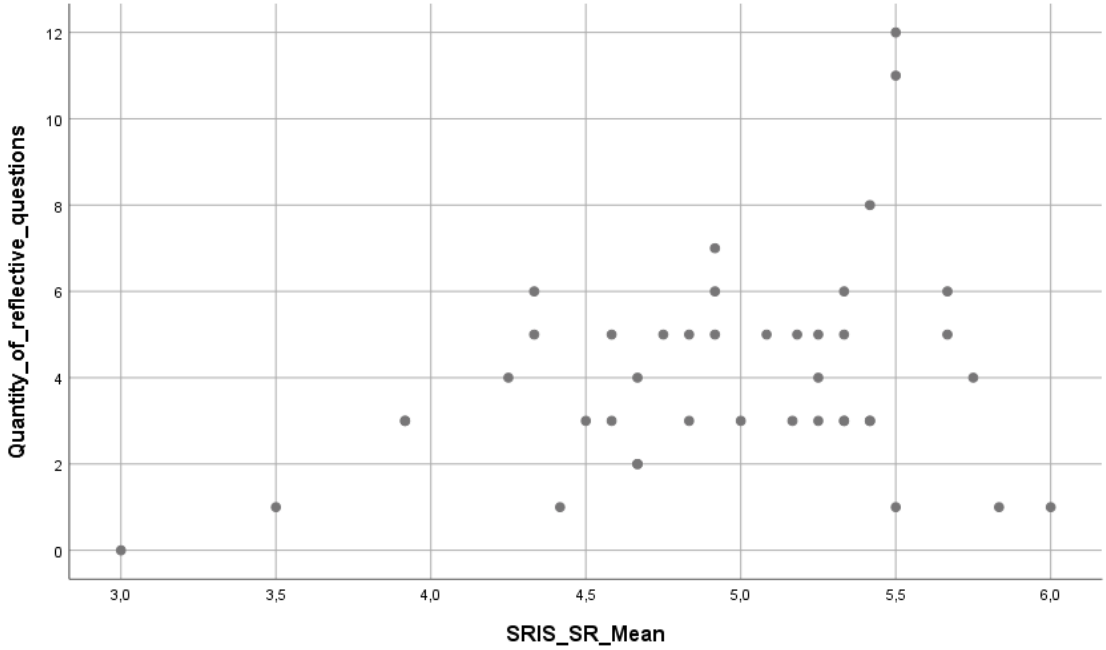
Scatterplot of the uniqueness and the MDKQ mean



Subsequently, the *quantity of reflective questions* created by the participants was in the focus of analysis. The fifth hypothesis was tested by performing a linear regression analysis with the *quantity of reflective questions* as the dependent and the mean score of the SRIS-SR as the independent variable. The relationship between the two variables is illustrated in a scatterplot in Figure 5. The analysis indicated that the regression model explains a significant proportion of the variance in the number of questions that were created by the participants (R^2 = .09, F (1, 40) = 4.2, p = .047). The mean score of the SRIS significantly predicted the number of questions that were created (B = 1.19, t (40) = 2.05, p = .047). Due to the significant coefficient and the slight linear relationship that can be obtained in Figure 5, it can be concluded that the assumption of linearity is met.

Figure 5

Scatterplot of the quantity of reflective questions and SRIS-SR mean



Next, the assumption of independence was tested, and a Durbin-Watson value of 2,04 suggested that the residuals are not correlated as values near two indicate independence of the residuals. Hence, the assumption of independence is met (see Table 2).

Table 2

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,31[a]	,10	,07	2,35	2,04

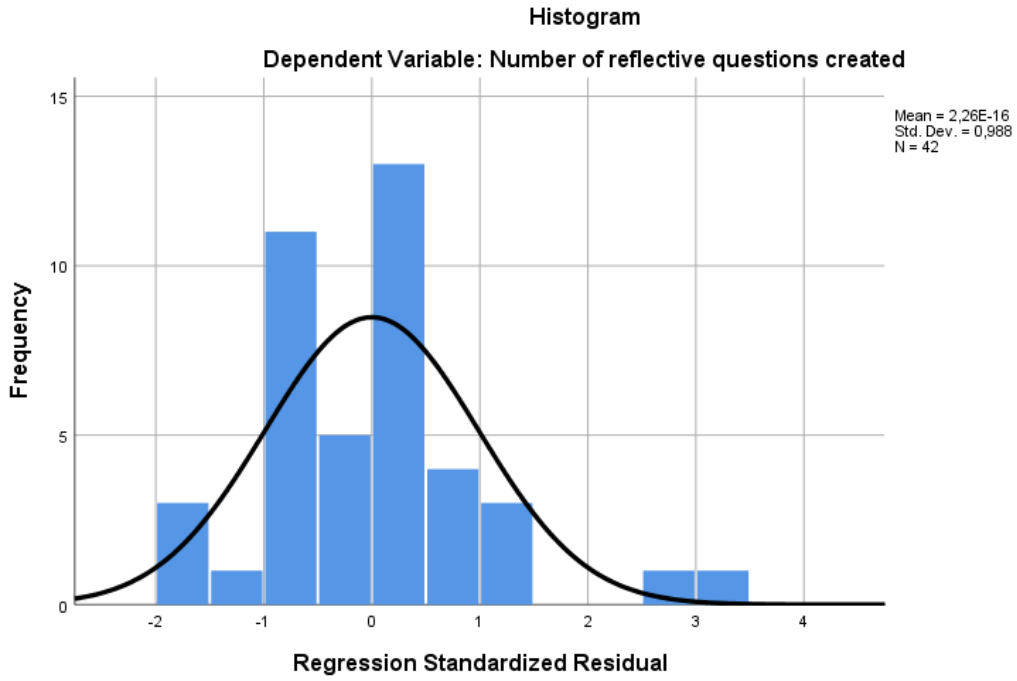
a. Predictors: (Constant), RSmean

b. Dependent Variable: NOQ Questions Number of reflective questions created

After that, the assumption of normality was tested, and the histogram shows that it cannot be met completely as the residuals are not normally distributed but are slightly skewed to the left (see Figure 6).

Figure 6

Histogram of the distribution of residuals (quantity of reflective questions)

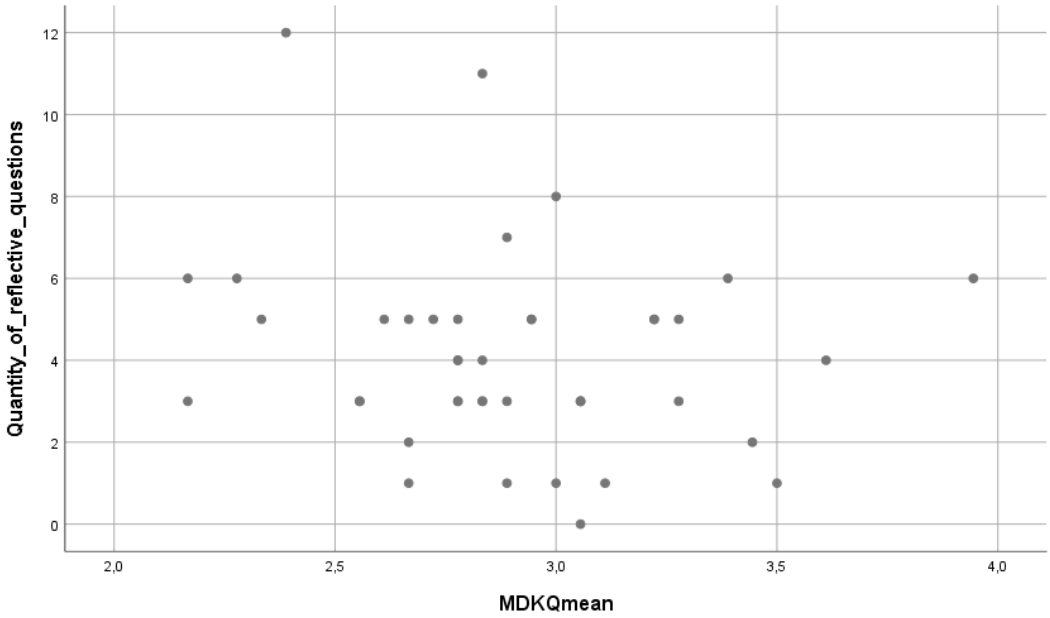


The last assumption tested, the assumption of homoscedasticity, is violated to some extent as well. As can be seen in Figure 2, the plot is slightly funnel-shaped, and therefore the assumption of homoscedasticity cannot be met completely.

Next, for the sixth hypothesis, a linear regression analysis with the *quantity of reflective questions* formulated by the participants as the dependent and the mean score of the MDKQ as the independent variable was performed. The relationship between the two variables can be seen in Figure 7. The analysis indicated that the mean score of the MDKQ did not explain a significant proportion of the variance in the number of questions (R^2 = .03, F (1, 40) = 1.48, p = .23). The sixth hypothesis was therefore rejected.

Figure 7

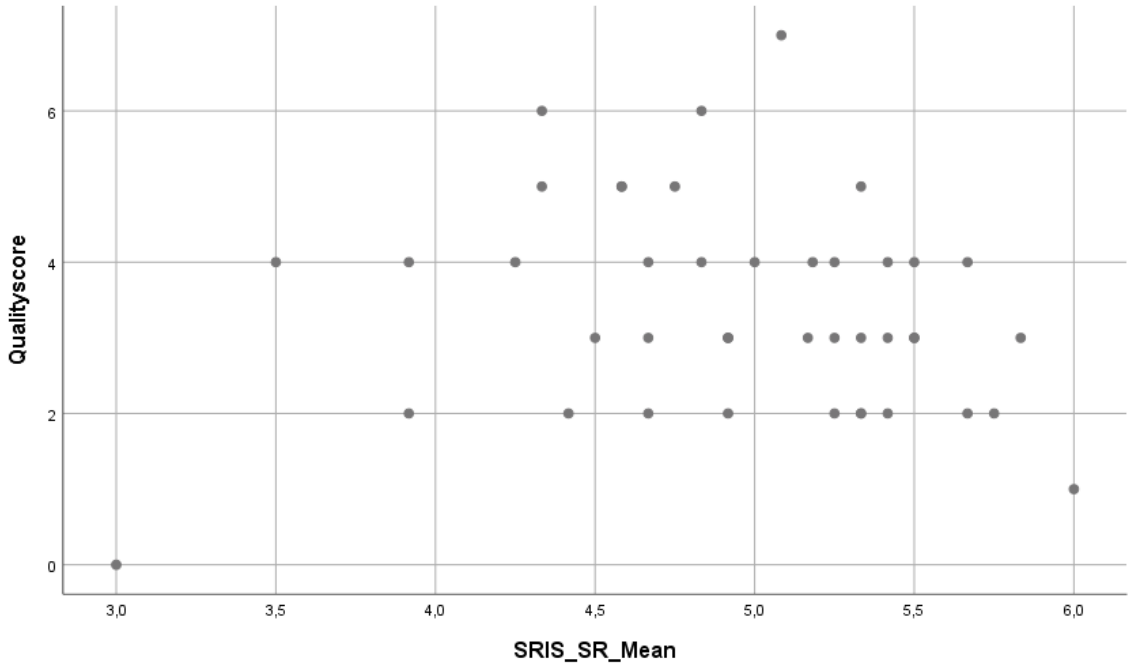
Scatterplot of the quantity of reflection and MDKQ mean



The seventh hypothesis was tested by running a linear regression analysis with the *quality score* of the reflective questions created by the participants as the dependent and the mean score of the SRIS as the independent variable. Figure 8 illustrates the relationship between the two variables. The analysis suggests that the mean score of the SRIS did not account for a significant proportion of the variance of the quality scores (R^2 = .00, F (1, 40) = .20, p = .66). Consequently, the seventh hypothesis was rejected.

Figure 8

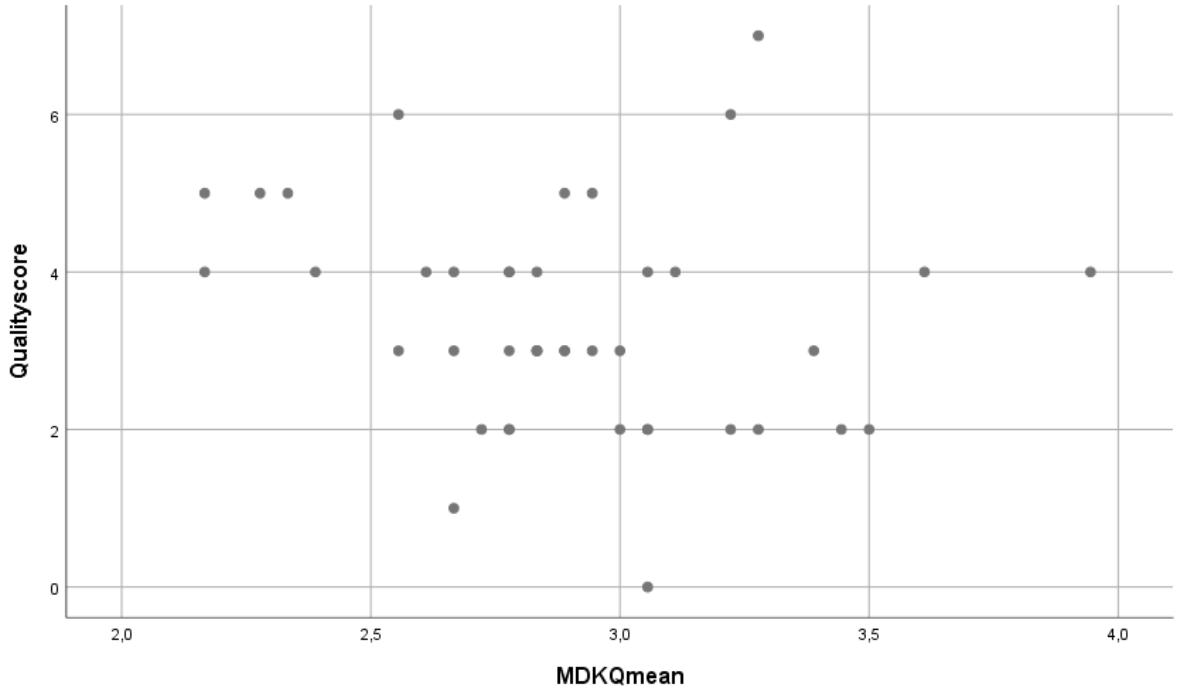
Scatterplot of the quality score and SRIS-SR mean



Lastly, the eighth hypothesis was tested by performing a linear regression analysis with the quality score of the questions created by the participants as the dependent and the mean score of the MDKQ as the independent variable. In Figure 9, the relationship between the two variables can be seen. The results suggest that the mean score of the MDKQ did not account for a significant amount of variance in the quality scores (R^2 = .03, F (1, 40) = 1.48, p = .23). Hence, the eighth hypothesis was rejected.

Figure 9

Scatterplot of the quality score and MDKQ mean



Discussion

The main purpose of the current study was to investigate the role that self-reflectional skill and domain knowledge play in ethical design reflection, with the Metaverse serving as a case example. Additionally, this study explored the operationalisation of ethical design reflection and offered a study design that can be used to assess participants' reflectional skills with regard to the identification of ethical issues and the formulation of respective reflective questions.

With regard to self-reflectional skills, the current study was not able to demonstrate the importance of reflective skills in ethical reflection that prior research was able to demonstrate (Boŝac et al., 2021; Riley et al., 2006; Tong, 2008). Nonetheless, some indication was found that self-reflective skills positively predicted the number of reflective questions participants formulated to trigger an ethical issue. Nonetheless, against the expectations, self-reflective skills did not predict the scope of elaboration of the ethical issue,

the uniqueness of ethical issues, or the quality of reflective questions as indicated by their content and structure.

Furthermore, against the expectations, the level of prior domain knowledge on the Metaverse that participants held was neither predictive of the scope of elaboration and uniqueness of ethical issues formulated by participants nor of the quality of reflective questions as indicated by their content and structure. Hence, this study was not able to prove the importance of prior knowledge for reflection in ethical design reflection that prior research has shown (Baig & Alzahrani, 2019; Chi et al., 1984; as cited in Moreno, 2004).

Self-reflective skills

In the available literature on the value of reflective skills, researchers were able to prove the positive effects of reflective skills on ethical reflection (Tong, 2008). Bozac et al., in addition, investigated the role of reflective awareness and did find a significant predictive value in ethical decision-making. Further, Riley et al. (2006) showed that using a blog as a reflective tool was successful in integrating ethical learning material. All of these previous studies demonstrate the surplus value that reflection can have in ethics. Nonetheless, in the previous studies, the object of reflection was external and not the reflector herself, as it would be in self-reflection. It might be argued that self-reflective skills are not capturing the right form of reflection that is important in ethical decision-making. It must also be noticed that in this study, the quality of reflection was captured by the uniqueness of ethical issues and by the domain-specificity of the content and structure of reflective questions. It is very likely that these dimensions of quality are not exhaustive at capturing the concept of qualitative reflection. Jones (1991), for example, emphasises that moral intensity of ethical issues is deemed important when assessing the characteristics of ethical issues. Hence, moral intensity might be another dimension that was not incorporated in the current study design.

Moreover, this study was partly able to find an effect of self-reflective skills on the quantity of reflection. Self-reflective skills were not predictive of the scope of elaboration of ethical issues that participants identified but significantly predicted the number of reflective questions that participants formulated. As was stated in the operationalisation of this study, it was hypothesised that the quantity of reflection reflects the level of motivation to reflect that a participant holds. Hence, we might cautiously interpret the significant finding as that the level of self-reflective skills in this study was predictive of the motivation to reflect regarding reflective questions but not regarding the identification of ethical issues. It must also be noted that due to the instructions given, participants were limited to a maximum of five sentences in their formulation of an ethical issue, and this might have influenced the results.

In addition, in this study, most of the participants scored highly on the SRIS-SR (*M* = 4.94), possibly distorting the results and possibly explaining why this study was not able to fully demonstrate the importance of self-reflective skills in ethical design reflection. It might be that the scale did not differentiate enough between the participants, or the characteristics of the sample accounted for the high scores. As most of the participants were psychology students from the University of Twente, and self-reflective skills are taught extensively in this program (University of Twente, 2022), the sample was not random regarding reflective skills. It would be of interest to see whether a more random sample would have yielded similar results.

Domain Knowledge

As we couldn't find any effect of domain knowledge on neither the quantity nor the quality of ethical design reflection, the importance of domain knowledge in this regard can be questioned. The finding is against the expectation as being acquainted with the topic of reflection is logically sound of importance and has shown to be of importance in research many times (Baig & Alzahrani, 2019; Chi et al., 1984; as cited in Moreno 2004). Hence, the question remains why this predictive value could not be found in the current study.

One line of argument could be that domain knowledge on the Metaverse is yet difficult to grasp since it is a social technology that is being developed and not yet useable for everyone. Accordingly, it is difficult to pinpoint the exact level of domain knowledge people hold. Further, it must be noticed that the MDKQ is a first attempt to capture domain knowledge on the Metaverse, and it has not been validated yet. Although the MDKQ has shown acceptable reliability, it cannot be securely concluded that domain knowledge on the Metaverse was accurately assessed in this study.

Additionally, it must be noticed again that the sample in this study was rather homogenous as 27 out of 42 analysed responses were from Psychology students from the University of Twente. Usually, psychologists are not well acquired with the field of domain knowledge asked for in this study, except for some ethical domains. The homogeneity of the sample possibly might therefore offer another reason for questioning whether domain knowledge was grasped precisely in this study.

Strengths and limitations

A major strength of the current explorative study design is the contribution of this explorative study in the realm of ethical design reflection to the scientific and professional community. This study set a base for the exploration of human characteristics and their importance in ethical design reflection. As prompting reflective questions is an important tool

to incorporate ethics into technical design, professionals must be assessed in their ability to ethically reflect, and the current research made a first step toward enabling named assessment.

Further, this study design was the first to develop steps of qualitative assessment that can be followed to assess ethical issues regarding their uniqueness and reflective questions with regard to their content and structure. The steps of assessment are easy to replicate, and they enable effective discrimination between participants. Based on reviewed literature in the field of ethical design reflection, the steps of qualitative assessment are the first that can be utilised to assess single words to small sentences and hereby, they differ from previous assessment tools that are mostly designed to assess larger pieces of text.

Another strength of this study design was the development of the MDKQ. Due to the fact that no questionnaire was available to measure domain knowledge for ethical design reflection on the Metaverse, this questionnaire was newly created and has shown to be reliable ($\alpha = .74$). Even though further trials and analyses are needed to establish the psychometric properties of this questionnaire, it is the first of its kind and is likely to be of value for future research.

A limitation of the current study design is the fact that self-reflective skills and domain knowledge have only been measured via self-assessment questionnaires. The utility and accuracy of self-assessment questionnaires for the measurement of psychological concepts are questioned frequently in research. Hailikari (2009), for instance, states that self-assessment questionnaires are time and cost-efficient but do not offer an accurate assessment of most psychological concepts. Moreover, the scores on the MDKQ and the SRIS-SR, as mentioned above, might be influenced by the characteristics of the sample. After all, it can be concluded that self-reflective skills and domain knowledge might have been measured in a way that does not represent the actual level of these variables in the population.

Another limiting factor of the current study design is the effect size of the needed sample size that was calculated. The G*Power analysis was conducted with an effect size of 0.5, which is seen as large (Faul et al., 2017). A smaller effect size would have yielded a higher required sample size as the expected differences between the groups are smaller. As Brydges (2019) highlights in his research, small effects might be overlooked when the sample size is not large enough. Hence, it can be questioned whether the same results would have been found if a smaller effect size had been applied and a larger sample had been recruited.

Another limitation is demonstrated by the slight violations of assumptions in the linear regression analysis. Even though a significant effect of self-reflective skills on the

number of reflective questions was found, this result is to be seen cautiously. The assumption of linearity and the assumption of homoscedasticity have been slightly violated, and hence, the significance of the result can be questioned.

A final limitation can be found in the fact that this study focused on the case example of the Metaverse. Since there is not much known about the Metaverse yet, as it is still being developed, the levels of domain knowledge in the population might be generally low, and high levels of domain knowledge were captured less often. As the highest score obtained on the MDKQ was four in this study, it might be argued that in this study, higher levels of domain knowledge still did not represent an actual high amount of domain knowledge on the Metaverse.

Future recommendations

Based on the findings of the current study, it can be concluded that there is more research needed to determine how the quality of ethical issues and reflective questions can be captured effectively. In future research, it would be of interest to see how the relationship between self-reflective skills and the motivation to reflect is operating in specific and how this knowledge can be used to inform professional training. Hereby, studies on the professional use of self-reflective skills and their effect on the motivation to reflect are regarded as beneficial for future professional training. As self-assessment questionnaires have often been found not to be accurate, a manipulation of the variable could result in a better ability to differentiate between participants. Thus, it would be of interest to investigate whether an intervention that aims to improve reflective skills would lead to better ethical design reflection.

Another recommendation for future research is related to the quality of reflection. In this study, no effect of the dependent variables on the uniqueness of ethical issues was found and hereby, more research is needed to firstly establish whether uniqueness can be seen as a quality of ethical issues and, secondly, what other dimensions of quality could capture the characteristics of ethical issues. Additionally, this study was the first to explore how reflective questions can be assessed. Accordingly, it should be further researched what qualities of reflective questions are and whether domain-specific content and structure can indeed be seen as a quality of reflective questions.

Moreover, this study set a starting point for the steps of qualitatively assessing ethical design reflection. Thus, more research would be needed to reveal how accurate the steps of the qualitative assessment are and how well they can help to distinguish between humans with regard to their reflective skills. If a point of saturation can be reached and a reliable

framework can be established, this would facilitate research on the determinants that are predictive of the quality of reflection. Adding to this, it would be of interest to see how high the intercoder reliability is and if this research would yield the same results if another coder would apply the framework or when another case example is given to participants. Hence, a replication of this study is recommended.

Conclusion

Even when recognising the limitations of the current study design, this study contributes to the exploration of human characteristics in ethical design reflection and in the qualitative assessment of named reflection. It was hypothesised that reflective skills and domain knowledge positively predict the scope of elaboration, the uniqueness of ethical issues, and the quantity and domain specificity of the content and structure of reflective questions. Despite the rejection of most hypotheses, the results were able to give valuable input and recommendations for future research and raise important questions about how to operationalise ethical design reflection. Further, the development of steps for the qualitative assessment of ethical design reflection and the development of the MDKQ, a questionnaire created to assess domain knowledge on ethics in the Metaverse, are important contributions that can be utilised in the future to create a pathway for ethics into the design of modern technologies.

References

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for Artificial Intelligence and digital technologies. *International Journal of Information Management, 62*. DOI: 10.1016/j.ijinfomgt.2021.102433

Baig, M. A., & Alzahrani, S.J. (2019). Revisiting the skills of a healthcare data scientist as a field Expert. *Studies in Health and Informatics. 262*, 43-46. DOI: 10.3233/SHTI190012

Božac, M. G., Kostelić, K., Paulišić, M., & Smith, C. G. (2021). Business Ethics Decision Making: Examining Partial Reflective Awareness. *Sustainability, 13*, 2635. DOI: 10.3390/su13052635

Brydges, C. R. (2019). Effect Size Guidelines, Sample Size Calculations, and Statistical Power in Gerontology. *Innovation in Aging, 3*(4), 1-8. DOI: 10.1093/geroni/igz036

Buchanan, R. (2017, October 27). *Design Ethics.* Encyclopedia.com. http://www.encyclopedia.com/science/encyclopedias-almanacs-transcripts-and maps/design-ethics.

Dennerlein, S., Wolf-Brenner, C., Gutounig, R., Schweiger, S., & Pammer-Schindler, V. (2020). Guiding Socio-Technical Reflection of Ethical Principles in TEL Software Development: The SREP Framework. *Addressing Global Challenges and Quality Education*, 386-391. DOI: 10.1007/978-3-030-57717-9_32

Dochy, F.J.R.C., Segers, M., & Buehl, M.M. (1999b). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research, 69*(2), 145–186. DOI: 10.3102/00346543069002145

Goodell, J. (2011, January 17). *Steve Jobs in 1994: The Rolling Stone Interview.* Rolling Stone. https://www.rollingstone.com/culture/culture-news/steve-jobs-in-1994-the-rolling-stoneinterview-231132/

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191

Grant, A. M., Franklin, J. (2002). The Self-Reflection and Insight Scale: A New Measure of Private Self-Consciousness. *Social Behaviour and Personality An International Journal, 30*(8), 821836. DOI: 10.2224/sbp.2002.30.8.821

Hailikari, T. (2019). *Assessing University Students' Prior Knowledge: Implications for Theory and Practice.*[Doctoral dissertation, University of Helsinki]

https://core.ac.uk/display/14915392?utm_source=pdf&utm_medium=banner&utm_campaig=pdf-decoration-v1

Jobin, A., Ienca, M., & Effy, V. (2019). The global landscape of AI ethics guidelines. *Nature MachineIntelligence, 1*, 389-399. DOI: 10.1038/s42256-019-0088-2

Jones, T. M. (1991). Ethical Decision Making by Individuals in Organisations: An Issue-Contingent Model. *Academy of Management Review, 16*(2), 366–395. DOI: 10.5465/amr.1991.4278958

Khalil, A., Ahmed, S. G., Khattak, A. M., & Al-Qirim, N. (2020). Investigating bias in facial analysis systems: A systematic review. *IEEE Access, 8*, 130751–130761. DOI: 10.1109/Access.2020.3006051

Khan, S. A., Alkawaz, M. H., & Zangana, H. M. (2019). The use and abuse of social media for spreading fake news. 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS).

Meta. (2021, October 28). *The Metaverse and How We'll Build It Together -- Connect 2021* [Video]. YouTube. https://www.youtube.com/watch?v=0294iXEPO4Y

Mezirow, J. (1990). Fostering critical reflection in adulthood. San Francisco: Joey Bass

Moreno, R. (2004). Decreasing Cognitive Load for Novice Students: Effects of Explanatory versus Corrective Feedback in Discovery-Based Multimedia. *Instructional Science, 32*(1-2), 99-113. DOI: 10.1023/B:TRUC.0000021811.66966.1d

Razavian, M., Tang, A., Capilla, R., & Lago, P. (2016). Reflective Approach for Software Design Decision Making. *2016 Qualitative Reasoning about Software Architectures (QRASA)*, 19-26. DOI: 10.1109/QRASA.2016.8.

Riley, D., Ngambeki, I., & Claris, L. (2006). *The Ethics Blog: Students Making Connections Among Ethics, Thermodynamics, And Life.* Paper presented at 2006 Annual Conference Exposition, Chicago, Illinois. DOI: 10.18260/1-2—152

Saint Mary's College. (n.d.). *Basic Ethical Principles.* https://www.stmarys-ca.edu/institutional-review-board/basic-ethical-principles

Schön D. A., The Reflective Practitioner: How Professionals Think In Action, 1st ed. New York, NY, USA: Basic Books, 1984.

Tong, R. (2008). Practice Precedes Theory: Doing Bioethics "Naturally" Is There an Ethicist in the House?: On the Cutting Edge of Bioethics. *J Med Humanist, 29*, 133-135. DOI: 10.1007/s10912-008-9056-2

Oliver, J.C., & McNeil, T. (2021). Undergraduate data science degrees emphasise computer science and statistics but fall short in ethics training and domain-specific context. *PeerJ Computer Science, 7*, 1-15. DOI: 10.7717/peerj-cs.441

Ullmann, T. D., Wild, F., & Scott, P. (2013). Reflection - quantifying a rare good. In M. Kravcik, B. R. Krogstie, A. Moore, V. Pammer, L. Pannese, M. Prilla, et al. (Eds.), *Proceedings of the 3rd workshop on awareness and reflection in technology-enhanced learning* (pp. 29–40). Presented at the 8th European Conference on Technology Enhanced Learning: Scaling up learning for sustained impact, Paphos, Cyprus.

University of Twente. (2022). *Bachelor Psychology.* https://www.utwente.nl/en/education/bachelor/programmes/psychology/

Valencia, S.W., Stallman, A.C., Commeyras, M., Pearson, P.D. & Hartman, D.K. (1991). Four measures of topical knowledge: A study of construct validity. *Reading Research Quarterly, 26*, 204–233. DOI: 10.2307/747761

Venturini, T., & Rogers, R. (2019). "API-based research" or how can digital sociology and journalism studies learn from the Cambridge Analytica data breach. *Digital Journalism, 7*(4), 532–540.

XR Today Team. (2021, November 24). *What is Facebook's Metaverse?.* XR Today. https://www.xrtoday.com/virtual-reality/what-is-facebooks-metaverse/

Appendix A

Informed consent

General information

You are being invited to participate in a research study titled 'Keys to an Ethical Metaverse'. This study is being done by Sebastian Wilhelm from the Faculty of Behavioural, Management and Social Sciences at the University of Twente.

The purpose of this research study is to investigate the human role in ethical decision making in the design phase of digital technologies. Hereby, the Metaverse will serve as an instance of a modern digital technology that requires thorough human attention to ethical issues. In particular, it is being assessed what determinants of human characteristics are important in ethical decision making. The study will take you approximately 25 minutes to complete. Optimally, the results of this research will be used in the end to inform the training of professionals who want to design ethically responsible digital technologies.

Your participation in this study is entirely voluntary and you can withdraw at any time. You are free to omit any question.

We believe there are no known risks associated with this research study; however, as with any online related activity the risk of a breach is always possible. To the best of our ability your answers in this study will remain confidential. We will minimize any risks by anonymizing all answers. It is possible that the data will be published, however data that enables one to identify you will be removed. The data will be stored for 10 years in a secure data storage from the University of Twente to which only authorized staff will have access to.

If you ever have any questions after this session has ended you can email me: s.wilhelm@student.utwente.nl. My supervisor, Dr. Sebastian Dennerlein can be reached via s.dennerlein@utwente.nl. This study has been ethically approved. For questions about the ethical approval and your rights you can reach ethicscommittee-bms@utwente.nl.

I agree to participate in this study.

- ○ Yes
- ○ No

Appendix B

Metaverse Domain Knowledge Questionnaire (MDKQ)

Please indicate your level of familiarity for the following statements.

	Not at all familiar	Slightly familiar	Somewhat familiar	Moderately familiar	Extremely familiar
I am familiar with XR (extended reality).	○	○	○	○	○
I am familiar with ethical issues.	○	○	○	○	○
I am familiar with ethics in technology.	○	○	○	○	○
I am familiar with the metaverse.	○	○	○	○	○
I am familiar with ethical principles.	○	○	○	○	○
I am familiar with XR (extended reality) in the Metaverse.	○	○	○	○	○
I am familiar with ethics in general.	○	○	○	○	○
I am familiar with Meta.	○	○	○	○	○
I am familiar with software development.	○	○	○	○	○
I am familiar with technology in general.	○	○	○	○	○
I am familiar with Ethics by Design.	○	○	○	○	○
I am familiar with Meta's products (Facebook, Instagram, Whatsapp, etc.)	○	○	○	○	○

How often do you use...?

	Never	Rarely	Occasionally	A moderate amount	A great deal
Social media	○	○	○	○	○
Facebook	○	○	○	○	○
Instagram	○	○	○	○	○
Whatsapp	○	○	○	○	○
Virtual reality (VR)	○	○	○	○	○
Augmented reality (AR)	○	○	○	○	○

Appendix C

Self-reflection and Insight Scale (SRIS; Grant & Franklin, 2015)

FACTOR LOADINGS FOR THE SELF-REFLECTION AND INSIGHT SCALE FROM STUDY ONE AND STUDY THREE

	Study 1 Factor Analysis		Study 3 Factor Analysis	
	Factor Loadings		Factor Loadings	
	1	2	1	2
Item	α = .91	α = .87	α = .71	α = .82
Engagement in self-reflection				
I don't often think about my thoughts (R)	.68	-.01	.32	-.07
I rarely spend time in self-reflection (R)	.78	-.02	.61	-.12
I frequently examine my feelings	.86	-.07	.85	-.09
I don't really think about why I behave in the way that I do (R)	.72	.10	.57	-.02
I frequently take time to reflect on my thoughts	.72	.01	.77	.04
I often think about the way I feel about things	.72	-.08	.72	-.02
Need for self-reflection				
I am not really interested in analyzing my behaviour (R)	.71	.02	.63	-.05
It is important for me to evaluate the things that I do	.75	.00	.76	-.01
I am very interested in examining what I think about	.77	.01	.70	-.03
It is important to me to try to understand what my feelings mean	.79	-.04	.78	-.14
I have a definite need to understand the way that my mind works	.73	-.03	.72	-.17
It is important to me to be able to understand how my thoughts arise	.72	-.02	.80	-.14
Insight				
I am usually aware of my thoughts	-.13	.67	-.43	-.23
I'm often confused about the way that I really feel about things (R)	-.06	.79	-.18	.80
I usually have a very clear idea about why I've behaved in a certain way	.21	.66	.27	.60
I'm often aware that I'm having a feeling, but I often don't quite know what it is (R)	-.01	.66	-.13	.76
My behavior often puzzles me (R)	-.16	.78	-.17	.76
Thinking about my thoughts makes me more confused (R)	.05	.65	-.03	.73
Often I find it difficult to make sense of the way I feel about things (R)	-.06	.80	-.12	.87
I usually know why I feel the way I do	.07	.78	-.27	.63
Factor Intercorrelations				
Factor 1	1.00	-.03	1.00	-.31**

Appendix D

Basic information on the metaverse

Basic information on the Metaverse

Before the first task begins, I would like to provide you with a basic understanding of the Metaverse. Please read the following text.

"In the broadest terms, the metaverse is understood as a graphically rich virtual space, with some degree of verisimilitude, where people can work, play, shop, socialize — in short, do the things humans like to do together in real life (or, perhaps more to the point, on the internet). Metaverse proponents often focus on the concept of “presence” as a defining factor: feeling like you’re really there, and feeling like other people are really there with you, too." - Oli Welsh, 2022

"Many experts look at the metaverse as a 3D model of the internet. Basically, a place parallel to the physical world, where you spend your digital life. A place where you and other people have an avatar, and you interact with them through their avatars." - Shamani Joshi, 2022

"The metaverse can be defined as a simulated digital environment that uses augmented reality (AR), virtual reality (VR), and blockchain, along with concepts from social media, to create spaces for rich user interaction mimicking the real world." - XR Today Team, 2022

Appendix E

Examples of ethical issues as presented to the participants

Example 1: Gender discrimination.
Example 2: Violations of data privacy in digital spaces are ethically problematic. Third parties could have access to data of users in the Metaverse. Users could have access to data from fellow users or take pictures of them in extended reality (XR), etc.

Appendix F

Ethical principles and their prevalence (Jobin et al., 2019)

Ethical principle	Number of documents	Included codes
Transparency	73/84	Transparency, explainability, explicability, understandability, interpretability, communication, disclosure, showing
Justice and fairness	68/84	Justice, fairness, consistency, inclusion, equality, equity, (non-) bias, (non-)discrimination, diversity, plurality, accessibility, reversibility, remedy, redress, challenge, access and distribution
Non-maleficence	60/84	Non-maleficence, security, safety, harm, protection, precaution, prevention, integrity (bodily or mental), non-subversion
Responsibility	60/84	Responsibility, accountability, liability, acting with integrity
Privacy	47/84	Privacy, personal or private information
Beneficence	41/84	Benefits, beneficence, well-being, peace, social good, common good
Freedom and autonomy	34/84	Freedom, autonomy, consent, choice, self-determination, liberty, empowerment
Trust	28/84	Trust
Sustainability	14/84	Sustainability, environment (nature), energy, resources (energy)
Dignity	13/84	Dignity
Solidarity	6/84	Solidarity, social security, cohesion