http://dx.doi.org/10.14236/ewic/BCSHCI2023.2

Empowering End-users in Co-Designing AI: An AI Literacy Card-Based Toolkit for Non-Technical Audiences

Freya Smith
Imperial College London
London, UK
smith.freyamk@gmail.com

Malak Sadek
Imperial College London
London, UK
m.sadek21@imperial.ac.uk

Céline Mougenot
Imperial College London
London, UK
c.mougenot@imperial.ac.uk

While there are growing calls to involve end-users in designing AI, the limited understanding of AI's capabilities among the general public hinders their effective participation in AI co-design. This study aims to facilitate the participation of non-technical audiences in AI co-design through the development of a card-based AI literacy toolkit. The toolkit was designed (1) to inform users about AI by presenting 16 competencies extracted from the AI Literacy conceptual framework by Long and Magerko (2020), and (2) to support idea generation and questioning about AI through "What if?" prompts, as done by designers. The effects of the informative and generative toolkit were assessed using a mixed methods approach. First, its impact on AI literacy was evaluated in a design task where non-technical participants (N=50) generated questions about future AI systems. Results showed a statistically significant increase in critical feedback, breadth, and frequency of AI-related keywords in participants' responses after using the toolkit, indicating improved AI literacy. Second, the effects of the toolkit were examined in a co-design workshop involving a group of six participants, half of whom without an AI background. Post-workshop interviews revealed positive effects on collaboration, including the development of a shared vision and common ground within the group. Both studies highlight the potential of AI literacy in enhancing the participation of non-technical audiences in co-designing AI with practitioners.

Keywords: human-centred design; co-design; participatory design; AI literacy; multidisciplinary collaboration.

1. INTRODUCTION

Advancements in AI are transforming our societal landscape, from individual lives to macro-level systems. This creates countless opportunities to optimise and automate society, but not without its costs, and the responsibility of minimising potential threats (Elliott et al., 2021). AIs have already been released which have arguably created more harm than good for people. Racist chatbots and discriminatory recidivism algorithms are both examples of AIs that have been released into society with inbuilt biases (Kusters et al., 2020, van Dijck, 2022). Evidently, it is vital that AIs are released carefully, without using a trial-and-error approach (Floridi, 2019). Demand is therefore growing to apply a Human-Centred Design (HCD) approach to AI (Auernhammer, 2020).

This study aims to support the application of HCD methods in AI, with the overall objective of increasing the AI literacy of end-users who participate in co-design sessions with AI-practitioners. In this paper, end-users refer to people for whom an AI product or system is designed to be used by, and AI practitioners refer to those who develop AI systems.

Here we first present the theoretical background of the study, highlighting the connections between Human-Centred Design, Co-design of AI, and AI Literacy. We also discuss the potential of using card-based tools to enhance the engagement of end-users in co-design processes. We then describe the development of a card-based toolkit, incorporating the AI literacy competencies defined in Long and Magerko's framework (2020). The toolkit serves an informative function by providing AI literacy content and a generative function through the inclusion of "What if?" prompts. Finally, we assess the effects of the informative and generative AI literacy toolkit on non-technical participants' AI knowledge level and their participation in a co-design session.

2. THEORETICAL BACKGROUND

2.1 Participatory Design of AI

AI is by nature an evolving technology, which can lead to them disembodying their anticipated values and causing unintended consequences (Umbrello and van de Poel, 2021). Literature therefore emphasizes the need for an AI's impact to be thoroughly considered before its release so not to threaten society (Floridi, 2019; Umbrello and van de Poel, 2021). This should begin at the start of the design process, before an AI is developed, as the risks of adverse consequences are so high (Floridi, 2019; Umbrello and De Bellis, 2018).

Demand is growing to involve wider stakeholders in the design of AI (Delgado et al., 2022). Participatory design (PD), a Human-Centred Design methodology that includes different stakeholders in the design process, has been described as essential in AI design (Neuhauser et al., 2013). Importantly, involving wider stakeholders considers a range of perspectives from those affected by an AI system. This can help determine and mitigate against any potential risks of the AI to wider society, and minimise the influence of a single designer in decision-making (Delgado et al., 2022, Auernhammer, 2020). It would also allow for AI practitioners to better understand their end-users, leading to the design of a more accessible and appealing product (Auernhammer, 2020, Delgado et al., 2022). Furthermore, research into public attitudes to GDPR and AI has found a willingness from the public to be involved in such consultations (Which?, 2021). A derivative of PD is co-design, which focusses on knowledge sharing between stakeholders and fostering a collective understanding of their different perspectives (Steen, 2013). For simplicity, the term co-design will be used, also referring to PD, throughout this paper. Its potential in AI design will form the basis of this study, with a focus on multidisciplinary groups of AI practitioners and end-users.

Despite the potential for co-designing AI, most of the existing research is still in the conceptual stage, and therefore lacks empirical evidence (Delgado et al., 2022). Second, friction is created by getting technical and non-technical stakeholders working together (Delgado et al., 2022). This friction is heightened by the fact that the public has a limited knowledge of AI; as an example, only 13% of participants in a study on public attitudes to AI believed they could explain the term 'AI' (CDEI, 2021). This lack of public awareness has caused end-users to develop their own false mental models of AI, and overestimate its capabilities, known as Superhuman Fallacy (Alizadeh et al., 2020). Furthermore, end-users often interact with AI systems without realising, and their limited knowledge of personal data collection in AI systems creates an inability to make informed decisions (Alizadeh et al., 2020). At present, AI companies are making technology easier to use but harder to understand and a lack of transparency in AI products is hindering the ability for it to be understandable by anyone apart from those developing it (Miller et al., 2018). This distinct knowledge gap between AI practitioners and end-users is a concern given that end-users must have some understanding of the capabilities of AI, to be involved in its design (Bratteteig & Verne, 2018).

2.2 AI Literacy and Level of Participation in Designing AI

This study proposes to address this knowledge gap by increasing end-users' AI literacy. AI literacy refers to a set of competencies required for people to be able to critically engage with and interact with AI (Long and Magerko, 2020). Demand is growing to cultivate AI literacy among the public through education in schools and businesses. This necessitates research in facilitating AI literacy in people without technical backgrounds, however progress in this area is only just emerging (Long and Magerko, 2020). As a result, there is a lack of empirical studies on how to implement AI literacy in practice. However, there is a growing body of transferable studies from related areas applicable to this study. Research is emerging on how public perceptions of AI and data practices change as they increase their knowledge of them and measuring changes in AI perception can be an indicator of improved AI literacy. Studies have found that the more well-informed people are about data uses, the more concerned and critical they become towards them (Kennedy et al., 2021). Studies have also shown that the more people know about automated decision-making, the higher their standards for regulation and data rights (Which?, 2021).

These findings are important for this study for two reasons. First, they support the case for why AI literacy is essential for end-users participating in AI co-design sessions, as end-users must give informed opinions and not unwittingly consent to a technology. Second, these studies give tangible indicators of how people respond to AI systems when they have a better knowledge of them.

2.3 AI Literacy Conceptual Framework

Our study builds upon Long and Magerko's AI literacy conceptual framework (2020). This framework has been developed through a scoping review of 150 publications about AI education, with the objective of providing a list of core AI competencies and principles for non-technical audiences. It includes 17 AI competencies needed "to critically evaluate, interact with, communicate with, and use AI technologies", divided into overarching categories, as shown in the table below.

***Table 1:** Long and Magerko's AI Literacy Framework*

Categories	Competencies
What is AI?	**Recognizing AI**: Distinguish between technological artifacts that use and do not use AI **Understanding Intelligence**: Critically analyze and discuss features that make an entity "intelligent", including discussing differences between human, animal, and machine intelligence; **Interdisciplinarity**: Recognize that there are many ways to think about and develop "intelligent" machines. Identify a variety of technologies that use AI, including technology spanning cognitive systems, robotics, and ML; **General vs. Narrow**: Distinguish between general and narrow AI.
What can AI do?	**AI's Strengths & Weaknesses**: Identify problem types that AI excels at and problems that are more challenging for AI. Use this information to determine when it is appropriate to use AI and when to leverage human skills; **Imagine Future AI:** Imagine possible future applications of AI and consider the effects of such applications on the world.
How does AI work?	Representations: Understand what a knowledge representation is and describe some examples of knowledge representations; **Decision-Making:** Recognize and describe examples of how computers reason and make decisions. **Machine Learning Steps**: Understand the steps involved in machine learning and the practices and challenges that each step entails. **Human Role in AI:** Recognize that humans play an important role in programming, choosing models, and fine-tuning AI systems. **Data Literacy:** Understand basic data literacy concepts such as those outlined in; Learning from Data: Recognize that computers often learn from data (including one's own data; **Critically Interpreting Data:** Understand that data cannot be taken at face-value and requires interpretation. Describe how the training examples provided in an initial dataset can affect the results of an algorithm. **Action & Reaction:** Understand that some AI systems have the ability to physically act on the world. This action can be directed by higher-level reasoning (e.g., walking along a planned path) or it can be reactive (e.g., jumping backwards to avoid a sensed obstacle). **Sensors:** Understand what sensors are, recognize that computers perceive the world using sensors, and identify sensors on a variety of devices. Recognize that different sensors support different types of representation and reasoning about the world.
How should AI be used?	**Ethics:** Identify and describe different perspectives on the key ethical issues surrounding AI (i.e., privacy, employment, misinformation, the singularity, ethical decision making, diversity, bias, transparency, accountability).
How Do People Perceive AI?	**Programmability:** Understand that agents are programmable.

2.4 Card-based Toolkits

Card-based toolkits, such as IDEO's Method Cards, are a popular medium to facilitate group design sessions (Roy et al., 2019, Välk et al., 2022). They allow for grouping of cards in different ways, viewing multiple cards at once, and they can be written on (Wölfel et al., 2013). The applications of card-based toolkits have progressed from broadly focussing on team collaboration and creative thinking to more recently domain-specific toolkits. A review of 155 card-based design tools (Roy et al., 2019) shows a significant lack of AI domain-specific toolkits, as most of them were created prior to AI's rapid advancements from 2012 (Schuchmann, 2019).

In terms of educating people from non-technical backgrounds on technology, research has found that people are very capable of grasping complex topics including AI when it is presented in an accessible format (Which?, 2021). The 'TechCards' (Ocnarescu et al., 2011) are a successful example of this, albeit without an AI focus. The 'TechCards' are a design toolkit that decompose technological concepts for non-technical audiences to understand. They were created to foster a shared knowledge of technologies within multidisciplinary groups of technical and non-technical stakeholders. Each of the cards features imagery, a technology description, and its characteristics. There is an opportunity to apply similar principles in this study.

A major barrier to collaboration in ideation sessions between non-designers was found to be their deficit of design and communication abilities (Zhang et al., 2022). The application of 'boundary objects', which are abstract or tangible artifacts designed to stimulate ideation and knowledge sharing, shows promise in addressing this shortfall (Zhang et al., 2022). A card-based toolkit will therefore be designed as the boundary object in co-design sessions.

2.5 Summary and Research Question

The literature review has revealed the need to apply Human-Centred Design (HCD) methods to AI at the very start of the design process before an AI is in development. While HCD literature also shows that co-design is an effective way of including end-users in the design process, end-users must have some level of AI literacy in order to participate effectively with AI-practitioners. As of now, there are few practical implementations of AI literacy, or methods of assessing it. There is hence an opportunity to develop an AI literacy toolkit for use in co-design sessions, which specifically targets non-technical audiences. This can be consolidated in the form of a card-based toolkit, which have proven successful as a means to upskill non-technical audiences in technical subjects.

The scope of the project has therefore been distilled into the following research question: “How might we increase the AI literacy of end-users to allow for collaboration and meaningful ideation in co-design sessions between end-users and AI practitioners?”.

3. METHODOLOGICAL APPROACH

The research question is addressed through the design of an AI literacy toolkit, which is evaluated in two stages using a mixed-method approach. As summarised in Table 2, we first quantitatively assessed the toolkit’s effects on AI literacy, then we qualitatively assessed the toolkit’s usage in an AI co-design session.

Table 2: *Summary of the evaluation study*

Study	Focus	Method
1	*AI literacy*	*Between-subject quantitative study.*
2	*AI co-design*	*Inter-subject qualitative study.*

3.1 Design of the card-based toolkit

The toolkit design requirements were condensed from the literature findings into seven specifications:

- The audience is end-users of AI so it should be understandable by those with a non-technical background.
- It will be used in co-design sessions between end-users and AI practitioners.
- It will be used at the start of the design process, so should support early-stage ideation.
- The toolkit should be card-based.
- It will build on Long and Magerko’s AI literacy framework as a structural basis for the cards.
- It will include an informative function and a generative function.
- It should include illustrative examples for each card topic.

This informed the key toolkit attributes regarding audience, thematic and textual content, and visual structure. A selection of the final printed cards can be seen in Figure 1.

Figure 1: *A selection of the printed AI literacy cards.*

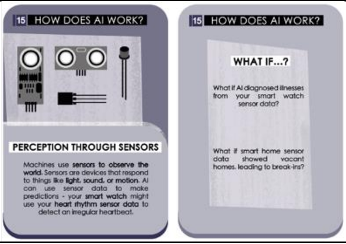
3.1.1 Structure

The toolkit builds upon Long and Magerko’s AI literacy conceptual framework (2020). This framework has been developed through the review of 150 publications in AI Education in a ‘scoping studies’ approach, with the objective of providing a list of core AI competencies and principles for non-technical audiences. It includes 17 AI competencies needed to critically evaluate, interact with, communicate with, and use AI technologies, divided into four overarching categories, as shown in Table 1. The cards were divided into these categories, with the corresponding competencies within them. Competency 17 was omitted, as it covered children’s perceptions of AI, who were not the toolkit’s audience. There is also an additional introductory card, which explains the set. The content is divided across both card faces: The front face acts as an informative side, covering the details of the specific competency, and the reverse acts as a generative side, featuring two divergent “What if?” questions to stimulate further ideas.

3.1.2 Text

The basis for each card’s textual information came from Long and Magerko’s competency descriptions. The exact content was created by combining and interpreting academic sources relevant to the competency area. These same sources also inspired the “What if?” questions. The copy was written to suit the intended audience and was balanced with positives and negatives of AI. Creating the text began with the descriptions of each AI competency, before finding relevant academic papers in the area to elaborate on the competency. Relevant content from the papers was then combined and summarised into simplified English. The example below shows a competency extracted from Long and Magerko's framework (2020) and its application into a card, informative on one side, generative on the other side. The same content production process was applied to the creation of a total of 16 cards.

Table 3 : *Example of a card based on source (Long and Magerko, 2020)*

Source	AI Competency definitions (Long and Magerko, 2020)
Example	*Competency 15 (Sensors): Understand what sensors are, recognize that computers perceive the world using sensors, and identify sensors on a variety of devices. Recognize that different sensors support different types of representation and reasoning about the world.*
Output	AI Literacy Cards: Information side (Definition) + Generation side (Contextualised question)
Example	

3.1.3 Visual Design

The four card categories each have their own colour theme and title at the top. Alongside the colour themes, only varying greyscale tones were used for clarity. The main front face elements are the competency title, text, and image representing the competency. It was noted from the 'TechCards' that the images could restrict participants' imaginations (Ocnarescu et al., 2011). Illustrations were therefore created for each competency instead of using photographic images, which the 'TechCards' used. This was to avoid literal interpretations of the technologies, and behave as abstract stimuli for ideation. Using illustrations also aided the simplification of complex AI topics - the intention of the toolkit. The illustrations were created using the textual information as inspiration. The printed cards are A7 in size and on laminated card, for ease of handling and reading.

3.1.4 Use Case

The toolkit is for early-stage AI co-design sessions between end-users and AI practitioners. It is intended to increase end-users' AI knowledge, and provide a concrete support from which to make informed contributions. The divergent faces are designed to encourage generative thinking, making it appropriate for ideation sessions. The toolkit should be provided to end-users at the start of design sessions to familiarise themselves with the content, but also have them available for reference throughout. A digital version is available; however, the printed version is preferable for in-person sessions for flexible card grouping.

3.2 Evaluation of the Impact of the Toolkit

3.2.1. Study 1: Impact on AI Literacy

Table 4: *Summary of the quantitative study.*

Objective	Approach	*Metrics*
Assessing the effects of the toolkit on end-users' AI knowledge.	*Between-subject quantitative study.*	*Sentiment analysis, breadth and frequency of keywords, scale rating of task ease/difficulty.*

3.2.1.1. Objective and Method

The first study aims to assess the efficacy of the toolkit in achieving the following hypothesis: H1: The toolkit is expected to increase the AI literacy of end-users. To do this, a between-subject quantitative study *(N=50)* was performed.

3.2.1.2. Protocol and Materials

Participants received a problem statement and an AI company's early-stage solution for it. Before receiving these, the test group participants familiarised themselves with an online copy of the toolkit for 10 minutes. The main task took 10 minutes, where participants gave feedback in the form of questions that they wanted to ask the company e.g., for proposal clarity or to raise concerns. To justify each question, a supporting comment was also required. Afterwards, there was a short questionnaire regarding task difficulty and toolkit feedback. Each study was completed individually over video call, with responses typed directly into an online form.

3.2.1.3. Participants

Participants covered an age range of 18-62+, and had a gender split of 26:24 F:M. A snowball sampling technique was used to recruit participants (Parker et al., 2019). The exclusion criterion was that they had no academic or professional knowledge of AI, which was confirmed upon their recruitment. Participants were first briefed on the study, and given opportunity to ask questions. They then gave implicit consent by completing the task using an online form. The study receives the Ethics approval from SETREC under 21IC7361.

3.2.1.4. Data Collection and Analysis

A data collection summary is found in Table 5.

Table 5: *Data collection and analysis summary.*

Data Collection Method	Data Collected	Data Analysis
Online Form	*Written Text*	*Sentiment Score*
	Written Text	*Breadth of Keywords*
	Written Text	*Frequency of Keywords*
	Questionnaire Rating	*Task Difficulty*
	Questionnaire Response	*Toolkit Feedback*

The sentiment score of participants' question and comment pairs was measured, as a key literature finding was that increased knowledge of AI and data uses leads to more negative views of them (Kennedy et al., 2021). It was therefore expected that test group sentiment scores would be lower, indicating more negative responses. To collect the sentiment scores of each participants' responses, every question and comment pair was fed into Google's Natural Language API. This assigned each one a sentiment score between -1 and 1. These scores were then compared between the groups for any significant difference in their means. An independent samples t-test was used for this, as the data was normally distributed and between two independent groups.

The breadth and frequency of keywords used in participants' responses were also chosen to assess H1 because of the expectation that better AI knowledge would lead to broader AI-related vocabulary and more AI-related ideas to contribute. To extract keywords, all participants' text responses were grouped together and "stop words" and duplicate words were removed. This created a list of every word used in the assessment. From this, AI-related words were manually identified and compiled into their own list of keywords.

To compare the breadth of keywords used between the groups, the number of unique keywords used by each participant was calculated. To do this, the aforementioned keyword extraction process was followed, but without initially aggregating the responses, so to create a word list for each participant. The keywords in each list were counted, which is the number of unique keywords used by each participant. To analyse any significant difference between the groups, a Mann-Whitney test was used because the data was not normally distributed and concerning the means of independent groups.

Next, to analyse any difference in keyword frequency between groups, the number of keywords used per participant was calculated. This was achieved with a similar process to the breadth of keywords, with the difference that duplicate words were not removed from participants' word lists. This ensured repeat words were recognised, to obtain a total number of keywords used for each participant. To test for statistical difference between the groups, an independent samples t-test was used, as the data was normally distributed and concerning two independent groups.

Task difficulty was measured as it was assumed test group participants would record lower difficulty ratings due to their increased AI knowledge. A Mann-Whitney test was used to ascertain any significant difference in the difficulty scores as the data was not normally distributed and concerning the means of independent groups.

The test group also provided toolkit feedback to understand how the cards were received and in case any iterating was required before the final co-design session, e.g., should participants all struggle to understand a particular card.

3.2.2. Study 2: Impact on Co-design

Table 6: *Summary of the qualitative evaluation.*

Objective	Approach	Metrics
Assessing the toolkit in the context of a co-design session.	*Inter-subject qualitative study.*	*Collaboration of multidisciplinary teams of end-users and AI practitioners.*

3.2.2.1. Objective and Method

The second evaluation comprised a qualitative, inter-subject assessment of the toolkit in the intended use context: a co-design session between end-users and AI practitioners *(N=6)*. The aim was to assess the toolkit's effects on group collaboration, and gain insight into how the cards were utilised and received during a co-design workshop.

3.2.2.2. Protocol and Materials

The study involved a 50-minute in-person co-design workshop and individual 10-minute post-workshop interviews. The main components of the workshop included an initial icebreaker to make participants comfortable (Mepieza, 2023), two ideation tasks to design an AI recruitment product, the printed toolkit intervention, and a final group task listing the harms and benefits of the designed product.

Figure 2: *Experimental set up of study 2.*

Within two days of the workshop, individual interviews were conducted online and transcribed using MS Teams. The interviews were semi-structured and broadly covered the topics of collaboration and the toolkit.

3.2.2.3. Participants

Six university students were selected to represent AI practitioners and end-users: half had a background in AI, and half did not. The inclusion criterion for the AI practitioners was that they had experience developing AI software, and end-users had no technical experience of AI. Participants were asked about their AI experience upon recruitment. Participant ages ranged between 21-23, with an even gender split, and a facilitator was present to provide direction. To adhere to ethics procedures, participants completed a consent form, after being versed on the activities of the session and invited to ask questions. The study receives the Ethics approval from SETREC under 21IC7361.

3.2.2.4. Data Collection and Analysis

The data collected from the study comprised transcriptions of the post-workshop interviews. These were coded and analysed using thematic analysis to group the data into overarching themes and sub-themes (Braun et Clarke, 2008). An inductive, bottom-up approach was used for this, so themes were driven by the data and not limited to the analyst's preconceptions (Braun et Clarke, 2008). The stages of the thematic analysis included data extraction in the form of participant quotes, coding those quotes, grouping these into sub-themes, and finally grouping the sub-themes into overarching themes.

4. RESULTS

We present here the results of the studies, which cover the toolkit's effects on AI literacy and collaboration.

4.1 Impact of AI Literacy

The results were that the 25 test group participants compared to the 25 control group participants showed significantly increased negative sentiment ($t(559) = 2.78$, $p = 0.006$), breadth of keywords ($z = 3.34$, $p < .001$), and keyword frequency ($t(48) = 4.10$, $p < .001$) in their responses. The results also showed there was no significant difference in task difficulty between the groups ($z = -0.36$, $p = 0.72$).

Table 7: *Complete results from the quantitative study.*

Sentiment Analysis						
Independent Samples t-test						
H	*p*	*Confidence Interval*		*t*	*df*	*sd*
		Lower	*Upper*			
1	*0.006*	*0.0132*	*0.0764*	*2.78*	*559*	*0.19*

Keyword Breadth			
Mann-Whitney Test			
H	*p*	*z*	*ranksum*
1	*< .001*	*3.34*	*809.50*

Keyword Frequency						
Independent Samples t-test						
H	*p*	*Confidence Interval*		*t*	*df*	*sd*
		Lower	*Upper*			
1	*< .001*	*3.88*	*11.32*	*4.10*	*48*	*6.55*

Task Difficulty Rating			
Mann-Whitney Test			
H	*p*	*z*	*ranksum*
0	*0.72*	*-0.36*	*618.50*

The most occurring keywords across all responses were "AI", used 75 times in the control group compared to 106 times in the test group, "Data" (12 and 51 times respectively), "Design" (10 and 26 times respectively), and "Bias" (13 and 32 times respectively). Also worth noting are "Model", "Dataset", and "Train", used 11, 9, and 9 times respectively in the test group, but were absent from control group responses.

From the feedback, participants found the cards useful in generating responses to the design task, with 92% of participants utilising them. Those who did not use the cards had a personal interest in AI, and so already had some AI knowledge. Participants described the cards as helping them "shape [their] feelings" with "critical thinking" towards the design challenge. The feedback validated the toolkit's usability for non-technical audiences, as it was described as "friendly, non-scary and non-technical", as well as being "easy to read and understand". The imagery on the cards was found to be "enhancing" and helpful in "understanding what the text meant". Importantly, the cards were also successful in making participants "well informed about the negative as well as positive sides of AI". One suggestion was to create a card to explain "writing prompts to AI such as ChatGPT".

4.2 Impact on Co-design

This section presents the qualitative thematic analysis findings, which evaluate the AI literacy toolkit in a co-design session. The thematic analysis derived three overarching themes: improved collaboration, broader AI awareness, and participant perceptions. These themes were constructed from 8 sub-themes that were identified from the coded transcriptions. Participant comments are referenced A(1:3) and N(1:3) to represent those with and without AI backgrounds, respectively.

4.2.1. Improved Collaboration

Reduced Discomfort in Group Interactions. Participants used the toolkit to avoid awkward group interactions by keeping discussions flowing, leading to improved participation. One member recalled "there was a point when the conversation fell, and I was going to run out of ideas, and it was a bit awkward. So, I started having a look through the cards and tried to think if anything was related to the concept." (N3). They also reported: "When I was stuck, [the cards] did help me participate better."

Facilitated Shared Language and Mental Models. A key finding is the toolkit's impact on establishing a shared mental model within the group: a component of collaboration (Yen et al, 2003). This was apparent from comments surrounding the shared understanding the cards facilitated, by putting "everyone on the same playing field" so everyone "knew what was going on" (A2). The cards were also mentioned as opening the non-AI practitioners' "eyes in terms of some of the technical challenges", which hence "enabled better cross-disciplinary collaboration" (A3). One participant without an AI background, explained the cards helped them understand terminology used by the AI practitioners (N1), and following the toolkit's introduction, another mentioned the group was "agreeing better" and "having a more shared vision" (N3).

4.2.2. A Broader AI Awareness

Increased Critical Thinking. A component of critical thinking is deductive reasoning (Lai, 2011), where a general conclusion is then applied to a specific context. This was demonstrated by participants drawing conclusions about their product, having read new information on the cards. One participant noted the discussions prior to the toolkit had not involved thinking about the impact of using an AI. Having read the cards, they "came up with way more negatives and concerns about using AI than benefits" (A3).

Supported Idea Generation. The cards were described as being used to shape wider group discussion topics (A3), as well as generating individual "prompts or sources of inspiration to come up with more things" (N1). One participant recalled specific points they contributed surrounding data that were "sparked with one of the cards" (N2).

Supported Divergent Thinking. The group grasped the application of the "What if?" card questions to stimulate divergent thinking. One distinct example of this was a participant who recalled: "That's what came up in the discussion after the cards: What if the AI gets it wrong? What if I can only learn from patterns in the data?..." (A1).

4.2.3. Participant Perceptions

The Value of Considering Participants' Opinions in AI Design. Participants were cautious of AI and believed having their opinions included in its design would be "valuable" and beneficial "to make sure it's responsibly done" (N3). Two end-user participants wished to be involved in further co-design sessions throughout the design process, so the risks they identified were mitigated against before release, and they "had enough chance to give input" (N3).

The Noticeable Difference in Stakeholder Perspectives. All participants noticed differences in task approach depending on stakeholder background. One member described the differences as "complementary" rather than "contradictory" (N2). The AI practitioners recognised their usual approach solely focusses on "making [a technology] work", and considering wider perspectives is "not something [they have] done a lot before" (A2). Two end-user participants mentioned the value of contributing their unique perspectives in the co-design session, and ability to "offer things that [the AI practitioners] hadn't thought of" (N3, N2).

Concerns about AI. The final sub-theme was constructed from codes expressing concern over AI technologies. One participant found it "unsettling" to have the "big responsibility" of designing an AI product that could "influence millions of people" (A3). Two AI practitioners were concerned by large companies releasing AI. One believed their "sole goal is to make money" (A3), and another said they would only consent to using the product they designed if it is released by an NGO or charity (A1).

5. DISCUSSION

The findings of this paper can be divided into the two toolkit evaluation stages. In the first stage, the toolkit's effect on end-users' AI literacy was discerned. Participants' task responses, using the toolkit, showed an increased negative sentiment (-25.46%), breadth (+53.94%) and frequency (+80.17%) of AI-related vocabulary compared to the control group. These findings suggest the toolkit increased the AI literacy of participants without AI backgrounds. The increased keyword results are particularly interesting as they suggest the toolkit helped participants develop and contribute more relevant opinions to the AI design proposal. This was also evident by test group participants using specialised keywords that were absent from the control group's responses, including "Model", "Dataset", and "Train". This therefore suggests the toolkit can better equip participants to contribute to AI co-design sessions.

However, participants using the toolkit did not find the task significantly easier than those without the toolkit, and instead gave similar difficulty ratings. This does not align with the hypothesis, as it was expected that the toolkit would reduce the challenge difficulty, given participants would be better equipped to contribute. An explanation could be that participants who upskilled their AI literacy before completing the task had the added difficulty of learning and applying new information, whereas the control group only utilised their existing knowledge in the challenge.

In the second evaluation stage, the toolkit was assessed in an AI co-design session. Thematic analysis showed the toolkit successfully supported multidisciplinary collaboration. This was evident as after exposure to the cards, participants reported an improved group shared mental model – an indicator of team collaboration (Yen et al, 2003). Participants also reported better understanding of technical language used in discussions. This is important as a limitation of co-design is the lack of shared language within multidisciplinary groups (Delgado et al., 2022). The cards were also used as a conversational aid to keep discussions flowing and reduce awkward interactions. In these respects, the toolkit successfully improved collaboration in multidisciplinary groups of technical and non-technical members.

The thematic analysis also validated the initial study regarding the toolkit's ability to increase AI awareness. This was evident from the result that the toolkit increased critical and divergent thinking, and supported idea generation, all within the domain of AI. This study also validated literature findings of a willingness from end-users to be involved in the AI design process.

5.1 Limitations and Further Suggestions

The study was limited by the co-design workshop's selection of participants. The toolkit was intended to be used in a commercial co-design setting between end-users and professional AI developers. Whereas for this study, the context was simplified due to time constraints by using students with and without a background in AI. In the intended use case, the interactions and participant dynamics would be more realistic and hence different insights might emerge. This is an area that would benefit from further testing and exploration.

A second limitation was performing a within-group assessment in the qualitative evaluation. With greater time and resources, the scalability of the findings could be better validated through executing multiple workshops. This is because a single workshop could not account for personality differences, which affect collaboration. It would therefore be worth investigating a quantitative study using the toolkit in multiple co-design workshops.

A toolkit limitation is that as AI rapidly progresses, so will the AI knowledge required of the public. Therefore, the cards may need regular iterations to stay relevant. This was noted by one participant who suggested a ChatGPT card could improve the set. As it was not released when Long and Magerko's AI Literacy framework was created, it can therefore be assumed that numerous new cards will need adding over time. One way to allow this, could be an online open-source version of the toolkit, which can easily be added to as technology develops.

6. CONCLUSION

This paper details the design, development, and evaluation of an AI Literacy toolkit to support non-technical audiences in AI co-design sessions. An initial literature review ascertained the opportunity, format, and specifications for the toolkit. The effects of the toolkit on end-users' AI literacy were then quantitatively assessed, as well as qualitatively assessing the toolkit's effects on collaboration in a multidisciplinary AI co-design workshop.

The results of the study were that the AI literacy cards significantly increased the negative sentiment, breadth and frequency of AI-related keywords used in participants' responses to a design challenge. These support the hypothesis that the cards increased participants' AI literacy. However, the cards were not found to reduce the difficulty of evaluating an AI product proposal for participants. In the qualitative assessment, the cards were found to support multidisciplinary collaboration by establishing a shared language and mental model within the group, providing topics for discussion, and as prompts to keep conversations flowing.

The AI literacy toolkit has filled an unmet need for an AI domain-specific card-based toolkit. Crucially, it has furthered existing research in AI literacy by creating a toolkit for non-technical audiences to utilise, and establishing metrics to assess AI literacy. This is important as there are currently very few studies which involve practical implementations of AI literacy, or empirical studies of assessing it. The versatility and breadth of topics covered in the AI literacy toolkit gives it the potential for application well beyond the context of AI co-design sessions. It was designed for non-technical audiences so would be a suitable tool for public information, use in education, or even upskilling company employees. With a growing need for national AI literacy, the toolkit has many possible applications.

7. REFERENCES

Alizadeh, F., Esau, M., Stevens, G. and Cassens, L., (2020). eXplainable AI: take one step back, move two steps forward. Mensch und Computer 2020-Workshopband.

Auernhammer, J. (2020) Human-centered AI: The role of Human-centered Design Research in the development of AI, in Boess, S., Cheung, M. and Cain, R. (eds.), *Synergy - DRS International Conference 2020*, 11-14 August.

Braun, V. & Clarke, V. (2006) 'Using thematic analysis in psychology', Qualitative Research in Psychology, 3:2, pp. 77–101.

Bratteteig, T., & Verne, G. (2018). Does AI make PD obsolete? exploring challenges from artificial intelligence to participatory design. In Proceedings of the 15th Participatory Design Conference - Volume 2 (PDC '18). ACM, New York, NY, USA, Article 8, 1–5.

CDEI (2021) Public attitudes to data and AI: Tracker survey . rep. Centre for Data Ethics and Innovation.

Delgado, F., Barocas, S. and Levy, K. (2022) 'An uncommon task: Participatory design in Legal Ai', Proceedings of the ACM on Human-Computer Interaction, 6(CSCW1), pp. 1–23.

Elliott, K., Price, R., Shaw, P. *et al.* (2021) Towards an Equitable Digital Society: Artificial Intelligence (AI) and Corporate Digital Responsibility (CDR). *Soc* 58, 179–188.

Floridi, L. (2019) Establishing the rules for building trustworthy AI. Nat Mach Intell 1, 261–262.

IDEO Method Cards 51 Ways to Inspire Design / IDEO. London] EC1R 5DF, UK: IDEO, 2010. Print.

Kennedy, H. et al. (2021) Living with Data Survey. rep. Living with Data.

Kusters R, et al. (2020). Interdisciplinary Research in Artificial Intelligence: Challenges and Opportunities. Front Big Data. 2020 Nov 23;3:577974.

Lai, E. (2011). Critical thinking: A literature review. Pearson's Res Rep. 6. 40-41.

Long, D. & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). ACM, New York, NY, USA, 1–16.

Mepieza, R. (2023). The Power of Ice Breaker Activity: Examining the Impact of Icebreakers on Student Participation and Engagement in the Classroom . *European Journal of Learning on History and Social Sciences*. 1, 1 (Apr. 2023), 22–36.

Neuhauser, L. et al. (2013). Using design science and artificial intelligence to improve health communication: Chronology MD case example. Patient Education and Counseling, 92(2), 211-217.

Ocnarescu, I. et al. (2011). Improvement of the industrial design process by the creation and usage of intermediate representations of technology, "TechCards". Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces (DPPI '11). ACM, New York, NY, USA, Article 50, 1–8.

Parker, C., Geddes, A. and Scott, S. (2020) 'Snowball sampling', SAGE Research Methods Foundations.

Roy, R. & Warren, J. (2019). Card-based design tools: a review and analysis of 155 card decks for designers and designing. Design Studies, 63 pp. 125–154.

Schuchmann, S. (2019). Analyzing the Prospect of an Approaching AI Winter.

Steen, M. (2013) Co-Design as a Process of Joint Inquiry and Imagination. *Design Issues* 2013; 29 (2): 16–28.

Umbrello, S., & Angelo, B. (2018). A Value-Sensitive Design Approach to Intelligent Agents.

Umbrello, S., & Van de Poel, I. (2021) Mapping value sensitive design onto AI for social good principles *AI Ethics* **1**, 283–296 (2021).

Välk, S., Chen, Y., Dieckmann, E. and Mougenot, C. (2023) Supporting collaborative biodesign ideation with contextualised knowledge from bioscience, CoDesign, 19:2, 142-161.

van Dijck, G. (2022) 'Predicting recidivism risk meets AI act', European Journal on Criminal Policy and Research, 28(3), pp. 407–423.

Which? (2021) The Consumer Voice: Automated Decision Making and Cookie Consents proposed by "Data: A new direction". Policy Research Paper). https://www.which.co.uk/policy-and-insight/article/the-consumer-voice-automated-decision-making-and-cookie-consents-proposed-by-data-a-new-direction-a6lPz2D3bzsC [accessed June 2023]

Wölfel, C. and Merritt, T. (2013) Method card design dimensions: A survey of card-based Design Tools, Human-Computer Interaction–INTERACT 2013, pp. 479–486.

Yen, John, Xiaocong Fan, Shuang Sun, Rui Wang, Cong Chen, Kaivan Kamali and Richard A. Volz. (2023) "Implementing Shared Mental Models for Collaborative Teamwork.

Zhang, R., Tan, Z., Su, Y. (2022). Non-designers in Interdisciplinary Design: Facilitating Communication and Ideation Using Boundary Objects. In: Bruyns, G., Wei, H. (eds) [ ] With Design: Reinventing Design Modes. IASDR 2021. Springer, Singapore.