HORIZON-MSCA-2023-DN-01
Grant Number 101169473

Funded by the European Union

value-ALIGNed socio-technical systems using large-language models (LLMs)

WP3 - Identification of European Societal Values and User Preferences in the Context of 3 Use Cases

D3.3 User Preference Identification

Contractual Delivery Date	28.02.2026		**Actual Delivery Date**	26.02.2026	
Responsible Beneficiary	TUM		**Contributing Beneficiary**	TUM	
Security	PU		**Nature**	R	
Version	1	**Date**	26.02.2026	**Page Nb.**	27

Authors

Name	Organisation	Email
Gokce Sahin	TUM	gokce.sahin@tum.de
Oana Stanciu	TUM	oana.stanciu@tum.de
Azzurra Ruggeri	TUM	a.ruggeri@tum.de

Document History

Revision	Date	Modification	Contact Person
-	26.02.2026	-	Gokce Sahin

Individuals' Preferences and Cognitive Representations of Artificial Intelligence: A Developmental Perspective

1. Introduction

Over the last few years, Artificial Intelligence (AI) agents (or quasi-agents, Chalmers, 2025) have become increasingly integrated into individuals' everyday lives for a wide range of purposes, including seeking advice, entertainment, learning, mental health support, and having a conversational partner (Li et al., 2023). Furthermore, many AI applications can now support a large variety of tasks. Given the functions that Generative AI (GenAI) tools, especially Large Language Models (LLMs), currently offer (e.g., creating videos, generating an essay, serving as a conversational partner), interactions between humans and AI agents have become more dynamic and reciprocal. These developments fundamentally transform humans' roles from passive consumers of these cultural technologies to active participants (Heeg & Avraamidou, 2024).

Humans, as (ultra)social animals (Tomasello, 2014), learn from *others* (e.g., caregivers, peers, and teachers) through observation (generally followed by imitation; Allen, 2012), immediate interactions (e.g., collaboration; Rogoff, 1991, 2003), and testimony (i.e., second-hand information; Harris & Koenig, 2006). In recent years, people have increasingly interacted with LLMs as cultural tools in daily life (Farrell et al., 2025; Heeg & Avramidou, 2025), often in ways similar to human interlocutors. Consequently, LLM tools influence the dynamics of human learning as *novel* interactive partners. Therefore, it is pivotal that the artificial agents with which people interact are designed to be attuned to human values, needs, and expectations (Jiao et al., 2025; Khamassi et al., 2024).

Childhood and adolescence represent sensitive developmental periods in which early interactions with physical and social environments actively co-construct knowledge (Carpendale & Lewis, 2014; Fernyhough, 2008). As children increasingly engage with AI-embedded technological environments, we need to assess the safety and risks of the tools, taking into account children's developmental vulnerabilities (Heeg & Avramidou, 2025; Kurian, 2025). For instance, children's relatively limited knowledge and epistemic vigilance can render them more vulnerable to the negative aspects of LLM use, such as an increased risk of misinformation (Jiao et al., 2025;

Kurian, 2025). Given that, LLM outputs represent only a narrow subset of cultural diversity, as they are predominantly trained on datasets reflecting Western culture and English-language media sources (but see B. Workshop, 2023), sustained interactions with these systems might result in reproducing biases (e.g., historical prejudices against underrepresented communities) that humanity has not (yet) overcome (Khamassi et al., 2024).

Further, overreliance on these artificial tools could hinder the development of cognitive skills (e.g., critical thinking, memory, planning), as these competencies are exercised less frequently (see Kosmyna et al., 2025, who find less cognitive engagement and deskilling in young adults using AI tools). In addition, LLM outputs might not be sufficiently calibrated to ensure age-appropriate conversation (e.g., long, complex sentences, unsafe content), potentially resulting in comprehension issues and cognitive overload for young end-users. Interacting with AI agents may expose children to psychological and physical harm (Kurian, 2025), underscoring the need for adult monitoring and regulatory frameworks governing children's safe use of AI technologies. Further, as some AI algorithms are designed to mimic a human-like bond (Ben-Zion, 2025), this situation might lead to distorted perceptions and increase a sense of connection with artificial agents (Dignum et al., 2025; Kim et al., 2025).

Attributing mental capacity to LLMs is not a cognitive distortion only found in children; adults also exhibit such predispositions (see Okanda et al., 2021). Nevertheless, given children's developing cognitive abilities, they are likely to be in a more vulnerable position (Xu & Warschauer, 2020). "Animism," the attribution of biological and psychological properties to non-living entities (e.g., objects or abstract concepts; Piaget, 1973), is prevalent in early childhood. Given the rapidly advancing capabilities of GenAI-embedded tools that produce rich contingencies during interactions, agency attributions might be increasingly blurred (i.e., agentic animism; Okanda et al., 2021; see also Goldman & Poulin-Dubois, 2024). Therefore, it is critical to know how children conceptualize LLM tools: what kinds of agents they perceive them to be, which properties they attribute to them (i.e., with respect to their biological, psychological, and cognitive features), how these perceptions shape their understanding of such systems' capacities and limitations (e.g., knowing that AI can make mistakes) (Okanda et al., 2021; Salles et al., 2020).

Knowing the mental representations that children build for LLM tools is necessary to understand children's engagements with these agents (including the purposes of use, preferences in use, and interaction patterns), and to fine-tune the design of LLMs to the cognitive developmental stage of their intended users (Jiao et al., 2025). Ultimately, this would facilitate aligning the affordances of AI-driven learning tools with rigorous protective mechanisms to curb potential harms to sensitive populations (e.g., implementing child-centered safety guardrails; Kurian, 2025). First, it is necessary to define users' perceptions and beliefs about AI to examine how these perceptions and beliefs impact their expectations of AI tools, which can ultimately shape users' preferences in interactive exchanges. The effort to define and integrate human preferences into LLM design could also lead to greater satisfaction and engagement with AI technologies (Zhao et al., 2025), thereby facilitating more efficient learning for users.

Nevertheless, the current research goal is not merely to identify end-users' preferences and directly implement them in the systems, but also to deliberately consider the potential risks associated with certain preferences and avoid them in LLM designs (this will be addressed in collaboration with Doctoral Candidate (DC) 1 in the subsequent phases). We should underline the potential negative consequences of certain preferences when they conflict with human needs and developmental vulnerabilities (e.g., users may dislike LLMs that state their own limitations, but this is an essential feature to avoid misinformation; Li et al., 2023). When such conflicts arise, the creators of LLMs should bear the responsibility of informing users about LLM limitations and encouraging users to actively participate in constructing knowledge, even when it contradicts their preferences. This would contribute to the responsible deployment of AI systems and ethical alignment between humans and artificial agents.

Lastly, it is also important to emphasize the extent to which societal needs are expressed through distorted attributions to AI tools or their misuse. For instance, the overuse of LLMs for socialization (e.g., AI companions) might reflect increasing social isolation and loneliness in modern societies (see Liu et al., 2024). Identification of preferences should also consider such sociological patterns as a matter of scientific responsibility.

2. Current Project

Shedding light on how users conceptualize artificial agents, whether perceiving them as mere informants/learning tools or as agents with *artificial* social capacities, would provide a foundation for further investigations into more nuanced preferences (e.g., preferred assistant persona in mental health contexts). We first aimed to assess how users of different age groups build a basic understanding of AI, their opinions on its benefits and risks, attributions, and their self-perceived competence in using AI. The empirical findings from Study 1 (S1) lay the groundwork for Study 2 (S2), which aims to develop a theoretical framework for identifying end-users' preferences across various AI use contexts. The framework, based on S1 and S2, will be implemented as a preference-specific assessment (e.g., transparency, modality, persona, privacy, utility) across different contexts (e.g., lifestyle, mental health, education, online news consumption) in Study 3 (S3). In Study 4 (S4), we aim to design experimental paradigms to test the consequences of implementing preferences reported by end-users in Studies 1 and 3 (in collaboration with DC1, DC14, and DC17, within the scope of Work Package 3). Additionally, the differences and commonalities in explicit preferences across age groups, as identified in S1 and S3, will inform the design of Study 4 (S4), including the construction of interaction contexts tailored to age-group characteristics using LLMs. Correspondingly, identification of users' mental representations and stated and revealed preferences regarding AI could inform the design of reward models for reinforcement learning from human feedback (RLHF) and system prompts. This, in turn, could support the optimization of safety guardrails calibrated to the developmental needs of different age groups.

2.1. Study 1

Nowadays, artificial agents are a compelling source of (*second-hand*) information (Geng et al., 2025). Correspondingly, these *novel* communication partners begin to play an important role in human cognition. Thus, their alignment with human needs and preferences is critical. To improve their alignment, we first need to understand users' goals and preferences. To address this, Study 1 assesses how children (9-10-year-olds), early adolescents (11-12-year-olds), and adults (older than 18 years old) perceive and use AI tools in their daily activities. We chose to focus on 9-to-12-year-olds to address the limited research on children's interactions with AI tools during earlier developmental stages. Most published research focused on adolescents aged 13 and older, who meet the legal age requirement to use social media platforms and GenAI tools, but we know AI tool use is now common among children as young as 9 (e.g., Flanagan et al., 2023; Girouard‑Hallam et al., 2021; Kosoy et al., 2024). Focusing on this age range also allows us to examine how early engagement with AI influences children's AI literacy and, overall, to investigate its potential impacts on their linguistic and cognitive skills in later developmental stages. Accordingly, we aimed to investigate differences and commonalities across age groups in their understanding, perceptions, concerns, attributions, and the contexts in which they use AI.

Participants from all age groups completed age-adapted versions of the same survey, comprising multiple open-ended and multiple-choice questions about AI. Inter-individual differences within age cohorts were not a primary focus of this first study. However, in subsequent steps of the project, additional individual variables could be measured (e.g., question-asking abilities, executive function skills, socioeconomic status). The first study aims to answer the following three main research questions:

1. How do users understand, perceive, and feel about AI?
2. Are there differences in how users of different ages understand, perceive, and feel about AI?
3. Does the users' general knowledge of AI account for differences (as well as commonalities) across age groups?

2.1.1. AI Survey Development

An initial item pool was generated from a literature review, including measurement tools to assess Human-AI interaction (HAII) across different domains. Items were selected based on their informativeness and alignment with the current research's objectives, as well as their clarity, to ensure that the same item can be used with all age groups with only minimal changes. Considering these criteria, 45 items were selected from previous studies, 32 of which were adapted for use with developmental samples. Lastly, 18 items were generated from scratch to enable a more comprehensive and multidimensional assessment.

Final items (see Table 1) were assigned to five thematic categories:
"Generic understanding of AI and AI use (9 items)*"*, *"Perceived Competence in AI Use* (14 items)*"*, *"Perceived Agency of AI* (9 items)*"*, *"Positive Attitudes towards AI* (25 items)*"*, and

"Ethical Concerns about AI and AI Regulation (6 items)*"*. Perceived Competence in AI Use aimed to explain to what extent participants consider themselves capable of using AI and knowledged about AI technologies in general. Perceived Agency of AI assessed the degree to which participants attributed human-like capacities to AI rather than perceiving it as a mere transactional tool. The Positive Attitudes towards AI subscale measured participants' general positive representations of AI. Lastly, Ethical Concerns about AI and AI Regulation was used to explore individuals' perceived risks associated with AI and their views about AI governance.

The survey contains 63 items in total with multiple formats: open-ended, Likert, and multiple-choice. Likert items record participants' ratings on a 5-point scale (1 = Totally disagree, 5 = Totally agree), with an additional option "0 = I prefer not to respond", which did not statistically affect the results. We averaged Likert items within each subscale.

The survey started with an eligibility check asking respondents: "Do you know what Artificial Intelligence (AI) is?" If the participant said "No", we gave them a few examples of well-known AI tools and repeated the question. If the participant answered "No" a second time, the survey ended, and the participant was excluded from the analysis. In addition to the 63 survey items, there were two attention-check questions. Respondents were expected to answer at least one attention-check question correctly to progress with the survey and be included in the analysis.

Items were created and edited in English. The final items were translated into Italian by a native Italian speaker. The resulting items were checked (and edited, when necessary) by two additional native Italian speakers before data collection.

The survey used with children was identical, except that we simplified instructions and modified some items to better match the cognitive and linguistic abilities of the youngest children (i.e., 9-year-olds), and we excluded age-inappropriate content (e.g., asking about AI use for work).

Table 1. AI Survey

Subscale Name & Brief Explanations	Items
Generic Knowledge and AI Use (9 items) Exploratory investigation of participants' generic knowledge and AI use in daily life – Not averaged –	**Knowledge**
	1K1. Do you know what Artificial Intelligence (AI) is? *(yes/maybe/no)* *(eligibility question)*
	1K2. Have you ever heard about at least one of these AI tools? *(yes/no)* (if 1K1 no) *(eligibility question)* ChatGPT, Amazon Alexa, Siri, Google Assistant, SORA, or DALL-E
	1K3. Can you explain what AI is? *(open-ended)* *(AI as a technology; Heeg & Avraamidou, 2025)*
	1K4*. Can you give me a few examples of AI? *(open-ended)* *(AI as a technology; Heeg & Avraamidou, 2025)*
	Preferences
	1P5. Have you ever used AI in your phone, computer, or an app? *(yes/maybe/no)*
	1P6. Please select the services you have used: *(multiple choices)* Options: Siri, ChatGPT, Alexa, Claude, Copilot, DALL-E, Midjourney, DeepL, Notion AI, Gemini, and MetaAI
	1P7. Services like Siri, ChatGPT, Alexa, Claude, Copilot, DALL-E, Midjourney, DeepL, Notion AI, Gemini, and MetaAI use AI. Have you used any of them? (if 1P5 no) *(yes/maybe/no)*
	1P8. What do you use AI for? (You can choose more than one option.) *(multiple choices)* Options: Ask for advice, search for information, entertainment, creativity, art, problem solving, homework, work

	1P9. Where do you use AI? (You can choose more than one option.) *(multiple choices)* <u>Options:</u> Home, school, work, other
Perceived Competence in AI Use (14 items) Participants who score high on these items believe that they have greater knowledge of AI and higher self-efficacy in using AI to help them achieve their goals.	**PC1*.** I know how AI can help me. *(awareness; Wang et al., 2023)*
	PC2*. I know how to use AI like Siri or ChatGPT. *(cognitive learning, know and understand; Ng et al., 2024)*
	PC3*. I know which AI to use as a function of the task I want to complete. *(evaluation; Wang et al., 2023)*
	PC4*. Using AI improves the quality of my work. *(perceived usefulness of AI; Chai et al., 2020)*
	PC5*. Using AI helps me to finish tasks faster. *(perceived usefulness of AI; Chai et al., 2020)*
	PC6*. I can apply AI to solve problems. *(cognitive learning, apply, evaluate, and create; Ng et al., 2024)*
	PC7*. When AI gives me multiple solutions, I can choose the best one. *(evaluation; Wang et al., 2023)*
	PC8*. It is usually hard for me to learn to use new AI technologies in general. (R) *(usage; Wang et al., 2023)*
	PC9*. I can tell which devices (e.g., phones, toys, websites) use AI and which do not. *(detect AI (Long & Magerko, 2020; Wang et al., 2023); Carolus et al., 2023)*
	PC10*. I can distinguish if I interact with AI or a “real human”. *(detect AI (Long & Magerko, 2020; Wang et al., 2023); Carolus et al., 2023)*
	PC11*. I know the advantages of using AI.

	(understand AI (Ng et al., 2022); Carolus et al., 2023)
	PC12*. I know the disadvantages of using AI. *(understand AI (Ng et al., 2022); Carolus et al., 2023)*
	PC13*. I can tell what AI cannot do. *(understand AI (Ng et al., 2022); Carolus et al., 2023)*
	PC14. I can tell what AI can do.
Agentic Perception of AI (9 items) Participants who score high on these items attribute agent-like capacities to AI-driven tools and perceive them as intentional agents rather than merely technological artifacts.	**AA1*.** The AI can understand how others are feeling. *(mind perception, Kim et al., 2023)*
	AA2*. The AI can remember things (e.g., previous conversations, my preferences). *(mind perception, Kim et al., 2023)*
	AA3*. The AI can tell what is right and wrong. *(mind perception, Kim et al., 2023)*
	AA4. The AI behaves according to moral rules. *(moral agency, Kim et al., 2023)*
	AA5. The AI can only do what humans tell it to do. (R) *(perceived control, AI control, Kim et al., 2023)*
	AA6*. The AI can feel emotions (e.g., fearful, pleased). *(experience, Kim et al., 2023)*
	AA7. AI is intelligent.
	AA8*. The AI can make plans and work toward goals. *(mind perception, Kim et al., 2023)*
	AA9*. I could have an AI as a friend. *(communal relationship; Cheng et al., 2022)*
Positive Attitudes toward AI (25 items)	**PA1*.** I enjoy using AI. *(attitude toward using AI; Chai et al., 2020)*
	PA2*. Sometimes interactions with AI frustrate or frighten me. (R) *(AI self-competency, AI emotion regulation (Carolus et al., 2022); Carolus et al., 2023)*

Participants who score high on these items have positive attributions to AI, believe that AI will overall improve human lives, and people who use AI possess positive traits.	**PA3.** I want to use technologies that rely on AI. *(behavioral; Stein et al., 2024)*
	PA4. I trust AI. *(Sindermann et al., 2021)*
	PA5. I fear AI. (R) *(Sindermann et al., 2021)*
	PA6*. AI will be good for humankind. *(Sindermann et al., 2021)*
	PA7. AI offers solutions to many world's problems. *(cognitive; Stein et al., 2024)*
	PA8. AI will make this world a better place. *(cognitive; Stein et al., 2024)*
	PA9*. AI can be dangerous. (R) *(AI and ethics; Heeg & Avraamidou, 2025)*
	PA.T5*. How do you feel about AI? (You can choose more than one option.) *(multiple choices)* *(affective; Stein et al., 2024)* <u>Options:</u> Anxious, enthusiastic, scared, hopeful, worried, sad, happy, neutral, other: fill in…
	PA10. I find AI tools cool.
	PA11. I find people who are good at using AI cool.
	PA12. I wish I were better at using AI.
	PA13. I wish I knew how to develop AI systems and tools.
	PA14. I think people who know more about AI become richer.
	PA15. I am hopeful about my future in a world where AI is commonly used. *(AI optimism; Chai et al., 2020)*
	PA16. AI has more disadvantages than advantages. (R) *(cognitive; Stein et al., 2024)*

	PA17. I think people who know more about AI have more power.
	PA18*. The AI provides people with useful options and choices. *(perceived control, user control, Kim et al., 2023)*
	PA19. I think people who rely more on AI lose some skills (e.g., critical thinking and creativity). (R)
	PA20*. AI can help us. *(AI as a tool; Heeg & Avraamidou, 2025)*
	PA21*. AI can make mistakes. (R) *(AI and ethics; Heeg & Avraamidou, 2025)*
	PA22*. Learning about AI is interesting. *(affective learning, intrinsic motivation; Ng et al., 2024)*
	PA23. I wish I knew more about AI.
	PA24. I am curious about discovering new AI technologies. *(affective learning, intrinsic motivation; Ng et al., 2024)*
	PA25. I am excited about a future in which everyone will use ChatGPT or similar AI tools for education.
Ethical Concerns of AI Regulation (6 items) Participants who score high on these items believe that AI needs to be regulated to ensure ethical governance and responsible use.	**E1*.** We should create AI that benefits all people. *(AI for social good; Chai et al., 2020)*
	E2*. I think that people should be taught how AI works and what limitations may be expected. *(ethical learning, AI ethics; Ng et al., 2024)*
	E3. I understand how misuse of AI could result in substantial risk to humans. *(ethical learning, AI ethics; Ng et al., 2024)*
	E4. Students should not be allowed to use AI for their homework.
	E5. I think there should be rules about how AI is created and used. *(ethical learning, AI ethics; Ng et al., 2024)*
	E6*. I think that what I tell AI should stay private. *(ethics; Wang et al., 2023)*

Note 1. The presentation of some items is conditional on answers given to previous questions; thus, the flow of the questions could vary across respondents. *For adults survey (English version):* **https://tummgmt.eu.qualtrics.com/jfe/form/SV_bIXC8IjCREC1Spo**
Note 2. Some items may be eliminated based on the Exploratory Factor Analysis.
Note 3. "*" denotes items modified from the original source.
Note 4. "R" refers to reverse-coded items.

3. Empirical work

We completed data collection with two of the three age groups (adults and 11-to-12-year-olds) and report these results here. Data collection with 9-to-10-year-olds is ongoing.

3.1. Methods

3.1.1. Participants

3.1.1.1. Children Sample

A total of 98 10-12-year-old children were recruited from schools in Italy. We excluded 22 participants because they had no knowledge or experience of AI ($N = 3$), did not complete any subscale ($N = 7$), or failed the attention check ($N = 3$).

Additionally, eight participants were excluded from the dataset because they were outside the target age range. The final child sample consists of 77 children aged between 11 and 12 years ($M_{age} = 11.82$, $SD = 0.53$). The sample is gender-balanced (50% female, 49% male, 1% other).

3.1.1.2. Adult Sample

A total of 199 adults were recruited at a large public event (Comic-Con festival) and at a museum in Italy. An additional 10 minors were recruited at the same time and were excluded from the sample. Seven participants were excluded from the adult data because of missing valid consent forms, and an additional 20 respondents failed to complete the attention-check questions or answered them incorrectly. In addition, 2 participants aged 60 or older were excluded from the dataset to avoid unwarranted extrapolation. The final adult sample consists of 171 individuals ($M_{age} = 38.66$, $SD = 9.24$, age range = 18-56 years old; 39% female, 60% male, 0.6% other).

3.1.2. Procedure

Ethical permission was obtained from the Ethical Board of the Technical University of Munich (approval number 2024-70_1-NM-BA). Adult participants gave written informed consent to

participate in the study. Child participants consented verbally to participation, and their parents/legal guardians consented in writing. Participants were informed that their participation in the survey is voluntary and that they can withdraw from the study at any moment.
For adults, data collection was completed using convenience sampling in public events in Italy. Trained research assistants approached adults and asked them if they wanted to participate in the study. The survey was implemented in Qualtrics. Adults completed the survey on their own devices by scanning a QR code or on provided tablets. Children completed the survey in their classrooms in a group setting under the supervision of trained research assistants.

The survey completion time was on average 13.73 minutes for children and 10.63 minutes for adults.

3.2. Planned analysis

The main focus of the quantitative analysis is the Likert items. Means were calculated for each predetermined subscale, and overall correlations between subscales were computed. For each intended subscale, we will compute Cronbach's alpha to assess internal consistency. Here, we report differences in mean ratings between two age groups (11-to-12-year-old children vs. adults), and the continuous effects of age on subscale means within the adult sample. Differences between the child and adult samples in patterns of correlations across subscale means are qualitatively described. Additionally, for each age group, we will also provide a descriptive analysis of responses for each item and the correlations between items.

Once data collection is complete, we will extend the analysis to include an Exploratory Factor Analysis for each age group (if possible, given sample sizes; the two developmental samples, 9–10 and 11–12-year-olds, could potentially be pooled). We will compare the resulting factor structure with our initial proposed categorization of items. Items that generally load poorly or show little meaningful variability may be excluded based on statistical criteria. Analysis at the group level should ideally be followed by a multi-group confirmatory factor analysis to test whether the same factor structure is appropriate across groups. This cross-age comparison would allow us to delineate developmental patterns from middle childhood through early adolescence to adulthood in perceptions and use of AI.

If factor analysis cannot be adequately conducted due to low sample size, we will apply a descriptive analysis, such as principal component analysis, followed by clustering analysis (e.g., k-means), and report whether the resulting clusters map onto the different age groups, suggesting that the age groups are empirically distinguishable.

For the multiple-choice items (e.g., the platforms used), we will present descriptive statistics and visually explore differences across age groups. Open-ended items (i.e., 1K3 and 1K4) will be analyzed using thematic analysis by two independent raters.

The findings will inform refinements to the AI survey's structure and item distribution across subdomains, thereby strengthening the survey's psychometric robustness for future studies.

Study 1 will be submitted for publication, and its results, along with the planned analyses, will be included in the second project deliverable.

3.3. Predictions

We expected certain subscales to correlate, in particular in the adult sample (see pre-registration). Participants with strong ethical concerns about how AI is to be produced and used ("Ethical Concerns about AI") might be likely to exhibit lower positive attitudes toward AI in general ("Positive Attitudes toward AI"). In addition, we predicted a negative association between "Ethical Concerns of AI Regulation" and respondents' awareness and attributions of AI's capacities and limitations as an agent ("Agentic Perception of AI"). The more agentic people perceive AI to be, the more worried they might be about its ethical implications.

Furthermore, younger age groups will score higher in the "Perceived Competence in using AI" subscale, indicating they have higher self-perceived competence in their interactions with AI than adolescents or adults. Children and early adolescents are predicted to score higher on the "Positive attitudes toward AI" subscale than adult participants. We also predict that the extent of concerns about the need to regulate AI ("Ethical Concerns of AI Regulation") will be greatest for adults, compared to younger samples. On the other hand, younger participants will be more likely to attribute agentic capabilities to AI ("Agentic Perception of AI") than adults.

4. Preliminary Results

On average, children rated themselves as competent in AI use ($M = 3.16$, $SD = 0.57$), but tended to have neutral or slightly negative attitudes towards AI ($M = 2.65$, $SD = 0.59$) and expressed concerns about AI regulation ($M = 3.74$, $SD = 0.65$). Additionally, general attribution to AI tools as an agent was low in the child sample ($M = 2.61$, $SD = 0.61$) (see Figure 1A).

Comparatively, adults rated themselves as more competent in AI use ($M = 3.66$, $SD = 0.53$), but were more positive in their attitudes about AI on average ($M = 3.32$, $SD = 0.54$). They were also fairly concerned about AI regulation ($M = 4.24$, $SD = 0.46$), and did not perceive AI as agentic ($M = 2.7$, $SD = 0.53$), showing a trend similar to the child sample (see Figure 1B).

In both child and adult samples, except for the Ethical Concerns of AI Regulation subscale, all subscales demonstrated acceptable reliability (i.e., $\alpha > .60$, for exploratory studies like the current research; Hair et al., 2021). Cronbach's alpha (α) values for each subscale were as follows: in the child sample, $\alpha = .73$ for Perceived Competence in AI Use, $\alpha = .88$ for Positive Attitudes toward AI, $\alpha = .49$ for Ethical Concerns of AI Regulation, and $\alpha = .68$ for Agentic Perception of AI. In the adult sample, the corresponding values were $\alpha = .84$, $\alpha = .90$, $\alpha = .57$, and $\alpha = .66$, respectively.

Figure 1. *Distribution of subscale means for Likert items*
Note. Each dot represents the subscale mean for one participant. The dotted line corresponds to "Neither agree nor disagree".

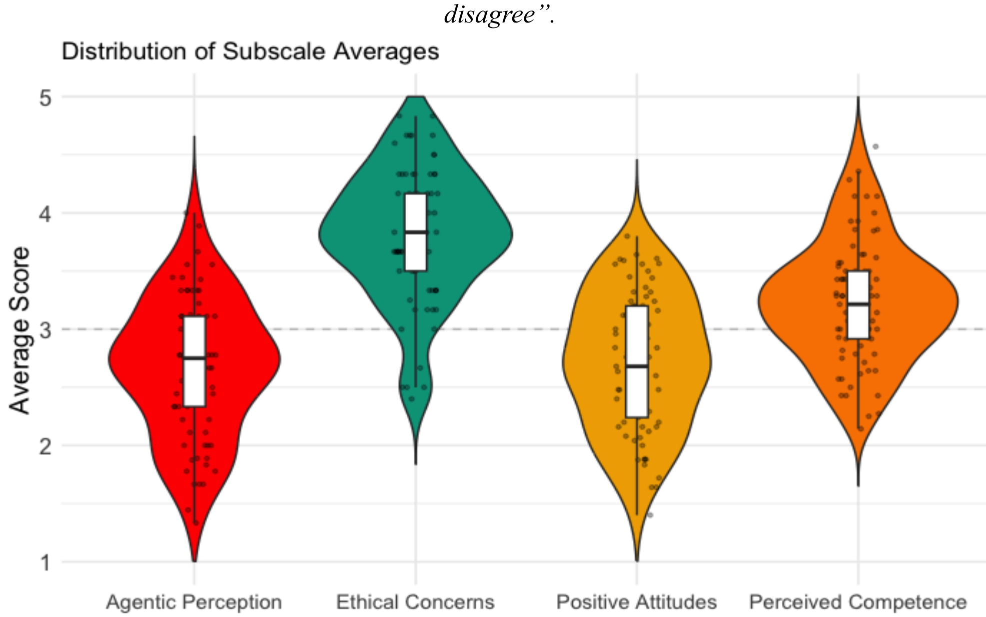


Figure 1A. Child Sample

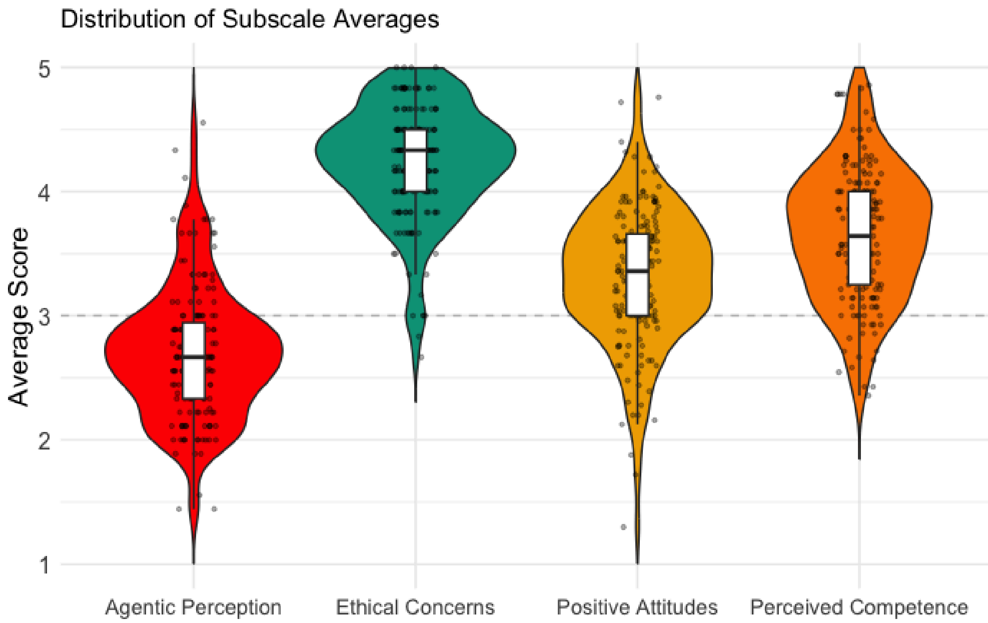


Figure 1B. Adult Sample

4.1. Relationships between Subscales

In the child sample, the results revealed a strong positive relationship between individuals' perception of AI as an agent and their positive attitudes toward AI (r *(75)* = .68, $p < .001$) (see Figure 2A). Additionally, a statistically significant positive association was found between users' self-perceived competence in AI use and their positive attributions toward AI (r *(75)* = .36, p = .001) (see Figure 3A).

These findings were replicated in the adult sample. That is, "Perceived Agency of AI" and "Positive Attitudes toward AI" subscales revealed a significant positive correlation (r *(169)* = .45, $p < .001$) (see Figure 2B). Perceived Competence in AI Use and Positive Attitudes toward AI subdomains also displayed a significant positive association (r *(168)* = .60, $p < .001$) (see Figure 3B).

Overall, the results suggest that individuals who attribute human-like capacities to AI-embedded technologies also hold more favorable evaluations of AI. Furthermore, the results indicate that users who feel more competent in AI are likely to exhibit more positive representations of AI.

Figure 2. *Relationship between AP & PA Subscales*
Note. Dots represent individual participants, and the blue line is a loess fit.

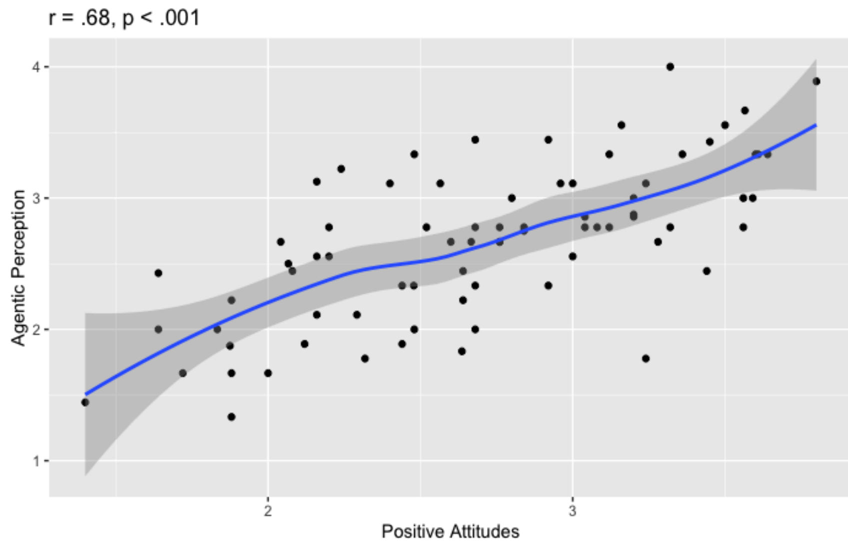


Figure 2A. Child Sample

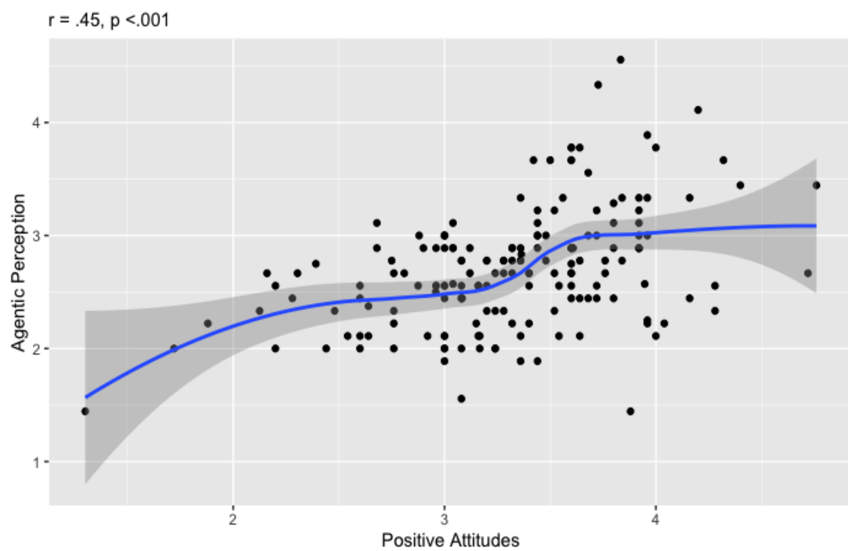


Figure 2B. Adult Sample

Figure 3. *Relationship between PA & PC Subscales*
Note. Dots represent individual participants, and the blue line is a loess fit.

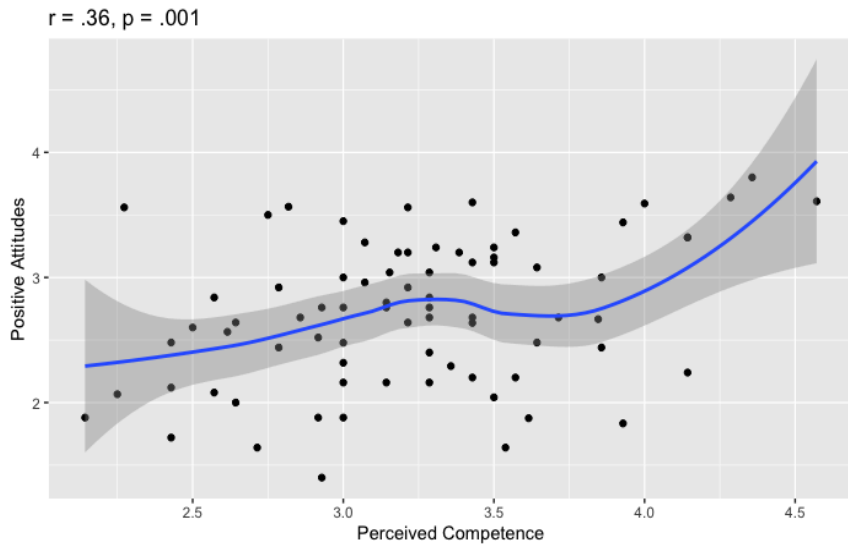


Figure 3A. Child Sample

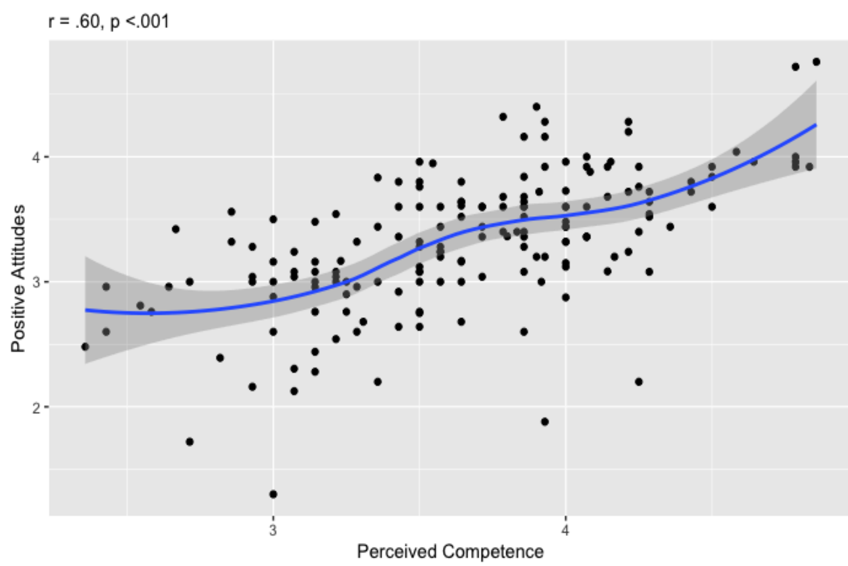


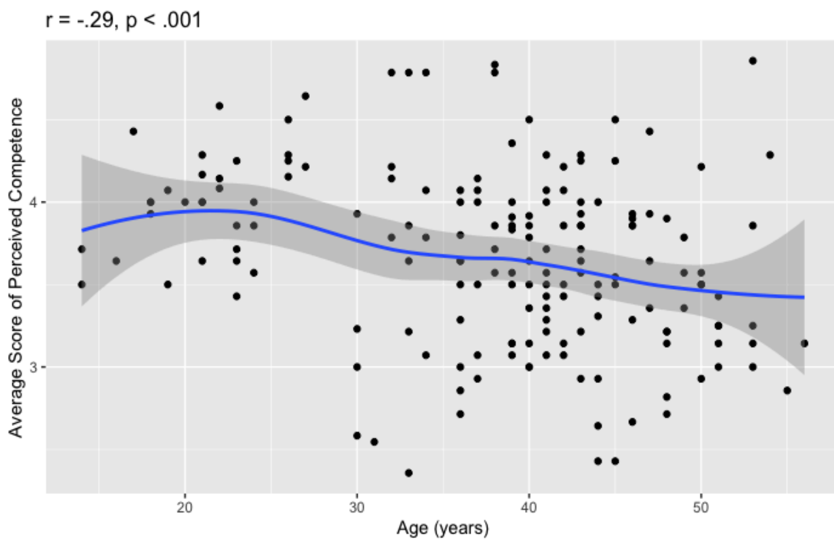
Figure 3B. Adult Sample

On the other hand, Ethical Concerns of AI Regulation and Agentic Perception of AI subdomains on the survey did not show statistically significant relationships in the children (*r (75)* = -.06, *p* = .63) nor the adult sample (*r (169)* = -.09, *p* = .23). Similarly, no significant correlations were found between Ethical Concerns of AI Regulation and Positive Attitudes toward AI subscales, both in the child sample (*r (75)* = -.10, *p* = .38) and adult sample (*r (169)* = .05, *p* = .48).

4.2. Continuous Effect of Adult Age on Subscales

The only significant age effect was observed in the Perceived Competence in AI Use subscale (r (168) = -.29, $p < .001$) in the adult sample. Participants perceived self-competence in interactions with AI tools, and their knowledge of AI decreased with age, as predicted (see Figure 4).

Figure 4. *Adult Sample - PC Average Scores as a Function of Age (years)*
Note. Each dot represents a participant, and the blue line is a loess fit alongside the 95% confidence band (shaded).



4.3. Patterns of AI Use

4.3.1. The Platforms Preferred in AI Use

The majority of children (66.2%) reported using the voice assistant Alexa. This was followed by ChatGPT (62.3%), Siri (44.2%), Meta AI (13%), and Gemini (9.1%). Smaller portions of the sample stated using Copilot (5.2%), Sora (3.9%), Notion AI (2.6%), and Claude (1.3%). In addition, 13% of participants indicated using other AI services, such as Google Assistant (see Figure 5A).

Adults used a wider variety of IAI services. The most commonly used AI tool was Siri (78.4%), followed by ChatGPT (50.9%), Alexa (33.9%), Claude (28.1%), and Copilot (21.6%), respectively. Fewer adults reported using DALL-E (12.9%), Midjourney (9.4%), DeepL (7.60%), NotionAI (7%), Sora (7%), and Gemini (6.4%) (see Figure 5B).

Figure 5. *AI Services Used*
Note. One participant could choose multiple options.

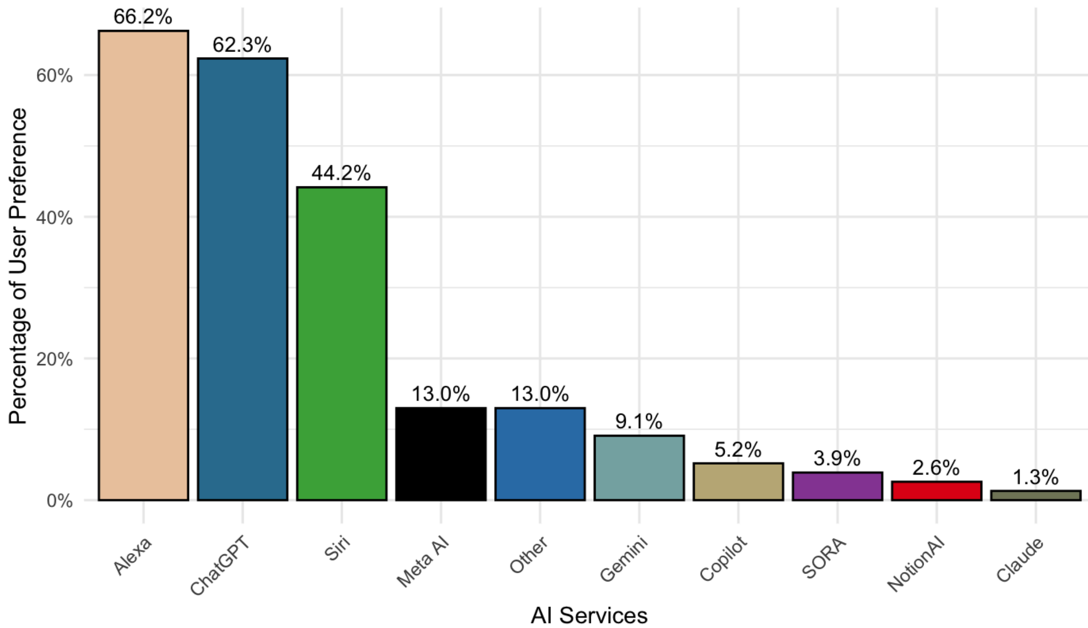


Figure 5A. Child Sample

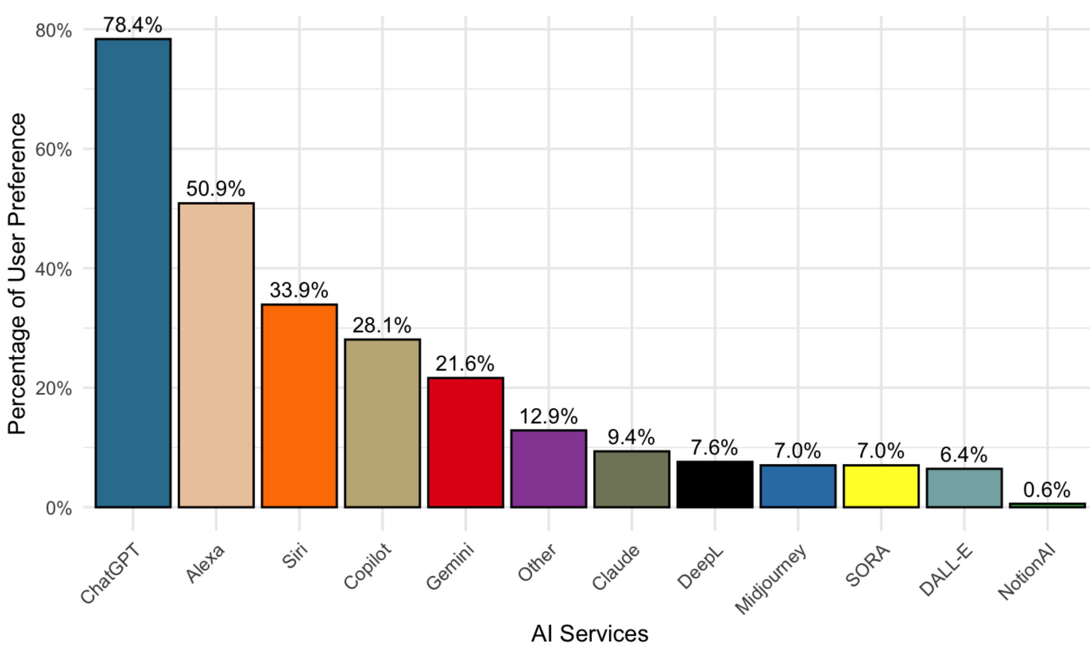


Figure 5B. Adult Sample

4.3.2. Purposes of AI Use

Children primarily used AI to search for information (67.5%). The second most frequently reported use was seeking advice (44.2%), followed by creativity (24.7%), entertainment (16.9%), and problem solving (15.6%). Smaller proportions of participants reported using AI for homework (9.1%) and for creating art (7.8%). An additional 16.9% of the sample indicated AI usage in reasons not included in the provided list (e.g., curiosity about AI) (see Figure 6A).

In the adult sample, the most common context for AI use was likewise information search (77.2%). Other frequently reported reasons for AI usage among adults were: asking for advice on AI tools (44.4%), getting help with work (42.1%), problem solving (37.4%), and using AI for creative purposes (33.3%). Entertainment was mentioned relatively less frequently (17.5%). Smaller proportions of the sample reported using AI for homework (8.2%), artistic purposes (5.3%), and other unlisted contexts (4.1%) (see Figure 6B).

Figure 6. *Reported Purpose of AI Use*
Note. One participant could choose multiple options.

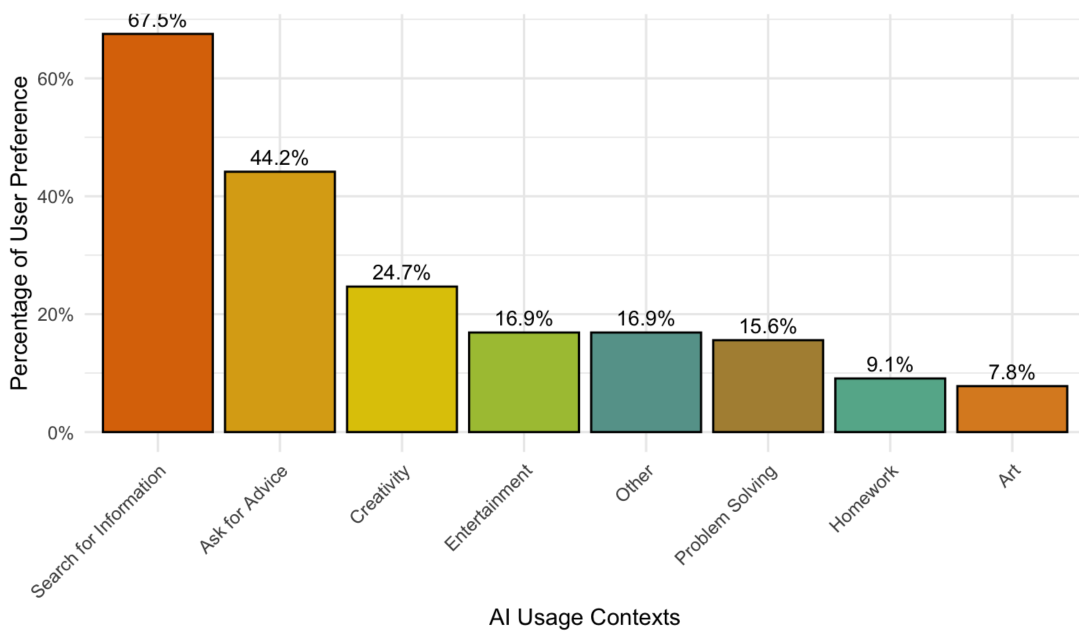


Figure 6A. Child Sample

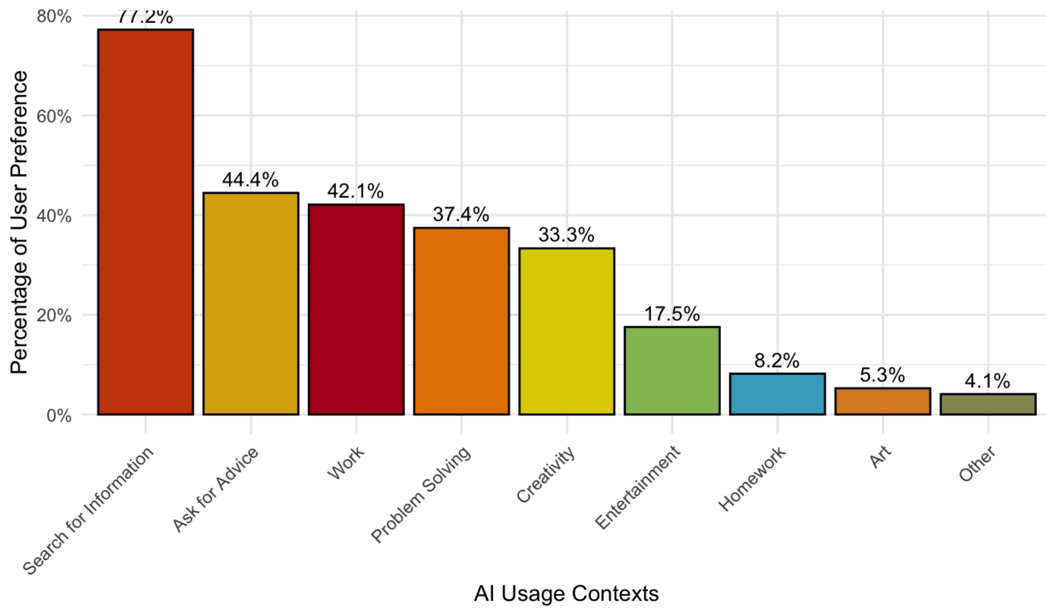


Figure 6B. Adult Sample

4.3.3. Emotional Responses to AI

Among children, happiness was the most frequently reported feeling in response to AI (31.2%). Feelings of amazement (26%) and worry (26%) about AI were the second most frequently reported emotions in the sample. Smaller percentages of the sample reported feeling scared (15.6%), anxious (14.3%), enthusiastic (13%), hopeful (13%), and neutral (13%) about AI. Additionally, 13% of the sample reported experiencing different emotions that were not listed as options. Sadness was the least frequently stated affect (3.9%) among child participants (see Figure 7A).

In the adult sample, participants most often reported feeling worried about AI (42.7%), even though a considerable proportion (36.3%) remained hopeful. Other commonly reported emotions were: amazement (33.3%), enthusiasm (26.9%), fear (19.3%), and anxiety (15.2%) about AI. In contrast to the child sample, adult participants reported feeling significantly less happy about AI (8.8%). Finally, smaller percentages of the sample reported other unlisted feelings (8.2%) and sadness (5.8%) about AI (see Figure 7B).

Figure 7. *Reported Affect toward AI*
Note. One participant could choose multiple options.

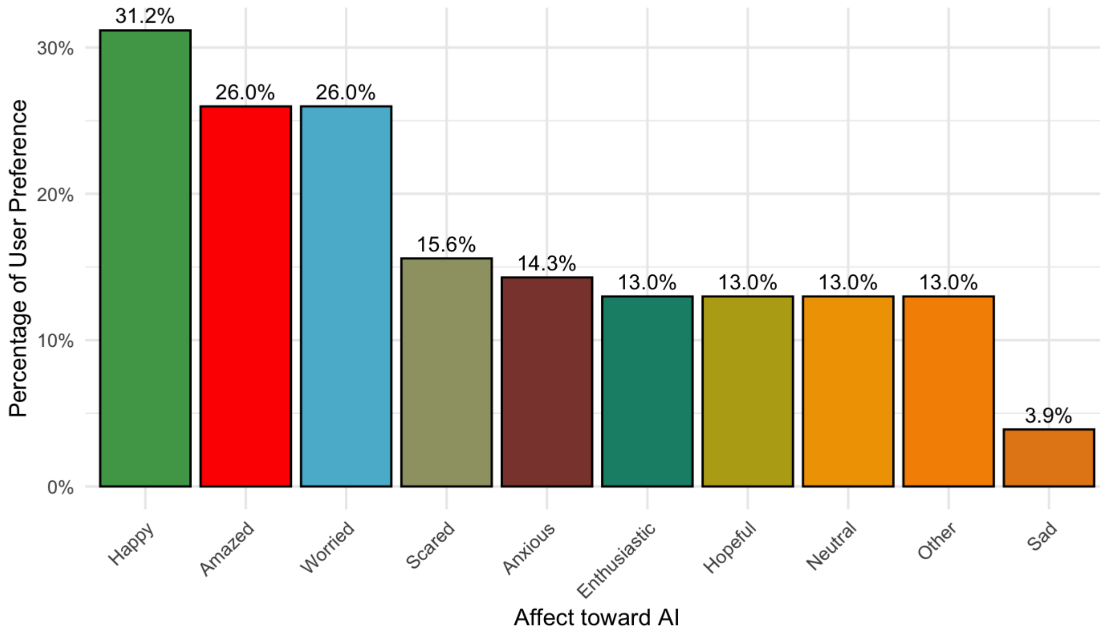


Figure 7A. Child Sample

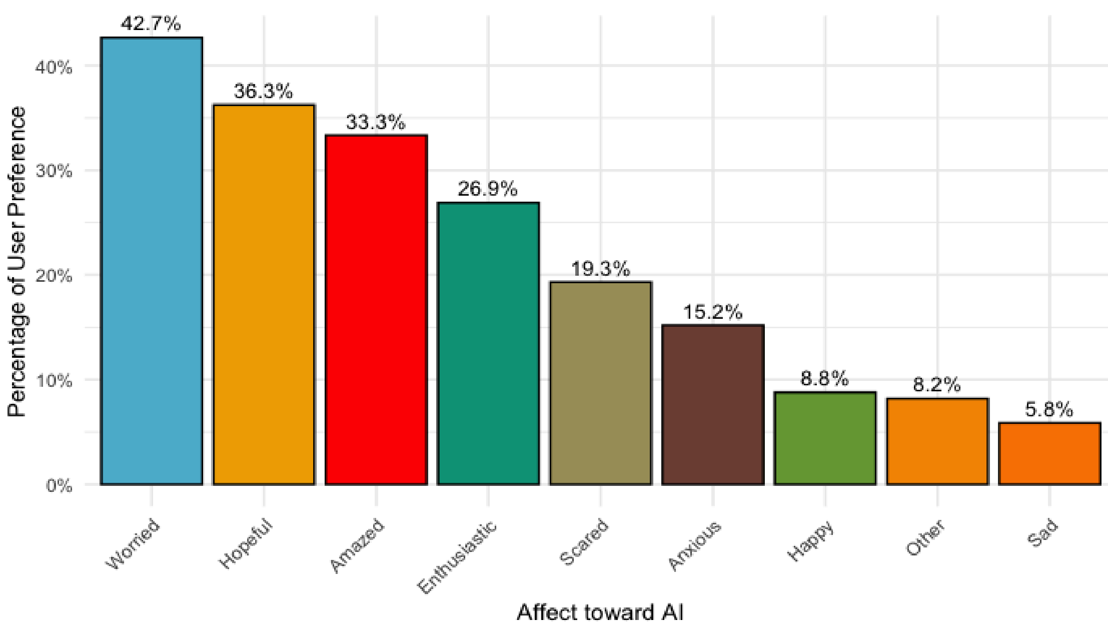


Figure 7B. Adult Sample

References

Allen, J. W. P. (2012). Imitation situations: Learning to use others as a resource for further activity. [Doctoral dissertation, Lehigh University].

Ben-Zion, Z. (2025). Why we need mandatory safeguards for emotionally responsive AI. *Nature*, *643*(8070), 9. https://doi.org/10.1038/d41586-025-02031-w

Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS-Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change-and meta-competencies. *Computers in Human Behavior: Artificial Humans*, *1*(2), 100014.

Carpendale, J. & Lewis, C. (2014). The development of children's understanding of social interaction. In S. Robson, & S. F. Quinn (Eds.), *The Routledge international handbook of young children's thinking and understanding* (pp. 123-132). Routledge.

Chai, C. S., Wang, X., & Xu, C. (2020). An Extended Theory of Planned Behavior for the Modelling of Chinese Secondary School Students' Intention to Learn Artificial Intelligence. *Mathematics, 8*(11), 2089. https://doi.org/10.3390/math8112089

Chalmers, D. J. (2025, November 20). What we talk to when we talk to language models. PhilPapers. https://philarchive.org/rec/CHAWWT-8

Cheng, X., Zhang, X., Cohen, J., & Mou, J. (2022). Human vs. AI: Understanding the impact of anthropomorphism on consumer response to chatbots from the perspective of trust and relationship norms. *Information Processing & Management*, *59*(3), 102940.

Dignum, V., Ericson, P., & Tucker, J. (2025, September 17). *AI chatbots are not therapists: Reducing harm requires regulation*. Tech Policy Press. https://www.techpolicy.press/ai-chatbots-are-not-therapists-reducing-harm-requires-regulation/

Farrell, H., Gopnik, A., Shalizi, C., & Evans, J. (2025). Large AI models are cultural and social technologies. *Science*, *387*(6739), 1153-1156.

Fernyhough, C. (2008). Getting Vygotskian about theory of mind: Mediation, dialogue, and the development of social understanding. *Developmental Review, 28*(2), 225–262. https://doi.org/10.1016/j.dr.2007.03.001

Ferrarello, L., Fiadeiro, R., Mazzon, R., & Cavallaro, A. (2022). Reframing the narrative of privacy through system-thinking design. *DRS2022: Bilbao*, *2022*, 1-18.

Flanagan, T., Wong, G., & Kushnir, T. (2023). The minds of machines: Children's beliefs about the experiences, thoughts, and morals of familiar interactive technologies. *Developmental Psychology, 59*(6), 1017–1031. https://doi.org/10.1037/dev0001524

Geng, Z., Zeng, B., & Huang, J. (2025). Living in a digital ecology: Children's selective trust in technological informants. *Journal of Applied Developmental Psychology*, *101*, 101872.

Girouard‐Hallam, L. N., Streble, H. M., & Danovitch, J. H. (2021). Children's mental, social, and moral attributions toward a familiar digital voice assistant. *Human Behavior and Emerging Technologies, 3*(5), 1118–1131. https://doi.org/10.1002/hbe2.321

Goldman, E. J., & Poulin-Dubois, D. (2024). Children's anthropomorphism of inanimate agents. *Wiley Interdisciplinary Reviews. Cognitive Science, 15*(4), e1676. https://doi.org/10.1002/wcs.1676

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer international publishing.

Harris, P. L. & Koenig, M. A. (2006), Trust in testimony: How children learn about science and religion. *Child Development, 77*, 505-524. https://doi.org/10.1111/j.1467-8624.2006.00886.x.

Heeg, D.M. & Avraamidou, L. (2025). Young children's understanding of AI. *Education Information Technologies* 30, 10207–10230. https://doi.org/10.1007/s10639-024-13169-x

Jiao, J., Afroogh, S., Chen, K., Murali, A., Atkinson, D., & Dhurandhar, A. (2025). LLMs and Childhood Safety: Identifying Risks and Proposing a Protection Framework for Safe Child-LLM Interaction. *arXiv preprint arXiv:2502.11242*.

Khamassi, M., Nahon, M. & Chatila, R. (2024). Strong and weak alignment of large language models with human values. *Scientific Reports 14*, 19399. https://doi.org/10.1038/s41598-024-70031-3.

Kim, P., Xie, Y., & Yang, S. (2025). "I am here for you": How relational conversational AI appeals to adolescents, especially those who are socially and emotionally vulnerable. *arXiv preprint arXiv:2512.15117*.

Kim, T., Molina, M.D., Rheu, M., Zhan, E.S., & Peng, W. (2023). One AI Does Not Fit All: A Cluster Analysis of the Laypeople's Perception of AI Roles. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*.

Kocielnik, R., Amershi, S., & Bennett, P. N. (2019, May). Will you accept an imperfect AI? exploring designs for adjusting end-user expectations of ai systems. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1-14).

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A.V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing tasks [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2506.08872.

Kosoy, E., Jeong, S., Sinha, A., Gopnik, A., & Kraljic, T. (2024). Children's mental models of generative visual and text based AI models. *arXiv preprint arXiv:2405.13081*.

Kurian, N. (2025). 'No, Alexa, no!': designing child-safe AI and protecting children from the risks of the 'empathy gap' in large language models. *Learning, Media and Technology*, *50*(4), 621–634. https://doi.org/10.1080/17439884.2024.2367052

Li, J., Sun, S., Yuan, W., Fan, R. Z., Zhao, H., & Liu, P. (2023). Generative judge for evaluating alignment. *arXiv preprint arXiv:2310.05470*.

Liu, A. R., Pataranutaporn, P., & Maes, P. (2024). Chatbot companionship: a mixed-methods study of companion chatbot usage patterns and their relationship to loneliness in active users. *arXiv preprint arXiv:2410.21596*.

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and validation of the AI Literacy Questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology, 55*(3), 1082–1104. https://doi.org/10.1111/bjet.13411

Okanda, M., Taniguchi, K., Wang, Y., & Itakura, S. (2021). Preschoolers' and adults' animism tendencies toward a humanoid robot. *Computers in Human Behavior*, *118*, 106688.

Piaget, J. (1973). *The child's conception of the world. Transl. by Joan and Andrew Tomlinson*. Paladin.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.

Rogoff, B. (2003). *The cultural nature of human development*. Oxford University Press.

Salles, A., Evers, K., & Farisco, M. (2020). Anthropomorphism in AI. *AJOB Neuroscience*, *11*(2), 88–95. https://doi.org/10.1080/21507740.2020.1740350

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H.S., Li, M., Sariyska, R., Stavrou, M., Becker, B., & Montag, C. (2020). Assessing the Attitude Towards Artificial Intelligence: Introduction of a Short Measure in German, Chinese, and English Language. *KI - Künstliche Intelligenz, 35*, 109–118.

Stein, J. P., Messingschlager, T., Gnambs, T., Hutmacher, F., & Appel, M. (2024). Attitudes towards AI: measurement and associations with personality. *Scientific reports*, *14*(1), 2909.

Tomasello, M. (2014). The ultra‐social animal. *European Journal of Social Psychology, 44*(3), 187–194. https://doi.org/10.1002/ejsp.2015

Wang, B., Rau, P. L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology*, *42*(9), 1324-1337.

Workshop, B., Scao, T. L., Fan, A., Akiki, C., Pavlick, E., Ilić, S., ... & Bari, M. S. (2022). Bloom: A 176b-parameter open-access multilingual language model. *arXiv preprint arXiv:2211.05100*.

Xu, Y., & Warschauer, M. (2020). What are you talking to?: Understanding children's perceptions of conversational agents. In *Proceedings of the (CHI) Conference on Human Factors in Computing Systems* (pp. 1-13).

Zhao, S., Hong, M., Liu, Y., Hazarika, D., & Lin, K. (2025). Do LLMs recognize your preferences? evaluating personalized preference following in LLMs. *arXiv preprint arXiv:2502.09597*.

DC2.WP3 - Tentative PhD Progress

TASK NAME	2025			2026				2027				2028
	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1
Study 1 - A Generic Overview of Individuals' Cognitive Landscape about Artificial Intelligence												
Literature Review												
Development of the AI Survey												
Data Collection (9-10, 11-12, and >18 yos)												
Analyses												
Deliverable I												
Writing the Manuscript of Study 1												
Study 2 - Stated Preferences Theoretical Study												
Literature Review												
Collaboration with DC1 for Defining Value Alignment of Preferences												
Writing an Opinion Paper about User Preferences-S2												
Mapping Preferences (e.g., factors and contexts, age-based)												
Conference/Sympossium Submissions												
Study 3 - Stated Preferences Empirical Study												
Transition from S1 & S2 to S3, Construction of Theoretical Framework for Survey												
SRCD 2027 - Submission *https://www.srcd.org/event/srcd-2027-biennial-meeting/call-submissions*												
Development of the Measurement Tool for Assessing User Preferencess												
Data collection												
Analyses												

Writing the Manuscript of S3												
Study 4 - Revealed Preferences												
Design of the Experimental Paradigm												
Secondment at EPFL - Collaboration with DC14 & DC17												
Data Collection												
Analyses												
Conference/Sympossium Submissions												
Writing the Manuscript of S4												
Writing the Dissertation												
Deliverable II												

Note. *Lighter shades represent tasks that are ongoing during the indicated period but are not as a main objective of the related research phase.*