Fieldnotes: Integrating AI Ethic into Philosophy for Children: Account of a Data-Privacy Lesson

Mohammad Reza Vaez Shahrestani

Abstract: This Fieldnotes article documents the design and implementation of a Philosophy for Children (P4C) lesson aimed at introducing children aged 8–12 to the ethical dimensions of Artificial Intelligence (AI), with a particular focus on data privacy. Drawing on the Community of Philosophical Inquiry (CoPI) methodology—specifically the "Plain Vanilla" model associated with Thomas Jackson and disseminated by the IAPC (Yarmel & Gregory, 2023)—the session employed an illustrated narrative stimulus, "Liam's Secret Base," followed by student-generated dialogue and a collaborative "Safe Sharing Scenario" activity. These components were designed to foster critical, creative, and caring thinking (Lipman, 2003) about digital life. This article demonstrates that engaging young learners in ethical inquiry about AI technologies is both timely—given the global call for age-appropriate AI-ethics education (UNICEF, 2021; Dabbagh et al., 2025)—and feasible, since most children already interact with AI-mediated platforms. Classroom reflections showed that the lesson supported higher-order reasoning, metacognitive awareness, and collaborative dialogue in a supportive environment, outcomes consistent with reported benefits of P4C (Adams et al., 2023; Ab Wahab et al., 2022). Rather than prescribing a list of "don'ts," educators used open-ended questions and realistic digital-life scenarios to nurture children's autonomy in ethical decision-making, echoing participatory design approaches from the MIT Media Lab's AI + Ethics curriculum (Wu, 2020; Williams et al., 2022). We conclude that P4C—co-founded by Matthew Lipman and Ann Margaret Sharp (Gregory & Laverty, 2017)—provides an effective pedagogical framework for integrating digital-ethics learning into childhood education, enabling students to become thoughtful, informed, and responsible digital citizens capable of navigating an increasingly data-driven world (Floridi, 2023/2024; Turner, 2019).

Introduction

Artificial intelligence (AI) is an integral component of children's daily environments, including personalized learning platforms, video games, voice assistants, and recommendation systems (UNICEF, 2021). While these technologies promise opportunities for creativity, accessibility, and engagement, they also raise ethical concerns, including data privacy, algorithmic bias, informed consent, and the fairness of automated decisions (Floridi, 2023/2024; Adams et al., 2023). Preparing children to navigate these challenges requires age-appropriate ethics education that complements technical digital literacy skills (Turner, 2019; Dabbagh et al., 2025). Yet many curricula address digital citizenship only superficially, without sustained opportunities to deliberate on normative questions such as: "What counts as private information?", "Who is responsible for protecting it?", or "When is it acceptable to share or restrict data?"

Philosophy for Children (P4C), co-founded by Matthew Lipman and Ann Margaret Sharp at Montclair State College in the late 1960s (Gregory & Laverty, 2017), offers a dialogical and inquiry-based pedagogy that fosters critical, creative, and caring thinking (Lipman et al., 1980; Lipman, 2003). Through the Community of Philosophical Inquiry (CoPI), children engage with stories or dilemmas that serve as conceptual stimuli, then collaboratively formulate questions, exchange reasons,

and reflect on their assumptions (Ab Wahab et al., 2022). The CoPI protocol followed here aligns with the "Plain Vanilla" model (Yarmel & Gregory, 2023), which emphasizes a low-intervention facilitation style, clarity of the guiding question, democratic turn-taking, and reason-giving without premature evaluation.

Integrating AI-ethics themes—particularly data privacy—into P4C aligns with global calls for child-centred AI governance (UNICEF, 2021) and leverages children's familiarity with digital games and social platforms as entry points for ethical reflection. By framing these issues within philosophical dialogue rather than didactic instruction, educators can help children recognize the value-laden nature of technology and make responsible, informed choices in a data-driven society (Floridi, 2023/2024; Williams et al., 2022).

Theoretical and Pedagogical Background

The P4C approach is grounded in the idea that children are capable of genuine philosophical thought when provided with structured opportunities for dialogue and reflection (Lipman et al., 1980; Lipman, 2003). Research has consistently shown that P4C enhances not only cognitive outcomes, such as reasoning and problem-solving, but also social-emotional competencies, including empathy, perspective-taking, and collaborative learning (Ab Wahab et al., 2022). Within CoPI, a carefully chosen stimulus sparks conceptual curiosity; participants then generate questions, select a focus through democratic procedures, exchange reasons, and revisit assumptions—practices that cultivate critical, creative, and caring thinking.

Recent scholarship has refined the understanding of how dialogical structure affects the quality of reasoning within the Community of Philosophical Inquiry (CoPI). Yarmel and Gregory (2023) conceptualize the CoPI as a communicative ecosystem in which both cognitive and procedural "noise" can disrupt the flow of collective inquiry. Their analysis shows that excessive facilitator intervention, unclear guiding questions, or unstructured turn-taking can introduce discursive noise that obscures conceptual progress. Drawing on Thomas Jackson's "Plain Vanilla" model, they argue that effective philosophical facilitation involves a process of "noise reduction"—minimizing unnecessary prompts while sustaining clarity, inclusivity, and autonomy in children's reasoning. This framework supports a low-intervention approach in which the facilitator's role is to refine the conditions of dialogue rather than to direct its content, thereby enabling participants to co-construct meaning through focused, self-regulated inquiry (Yarmel & Gregory, 2023).

In digital-ethics education, scholars emphasize moving beyond purely technical instruction to address autonomy, consent, fairness, privacy, and accountability (Adams et al., 2023; Turner, 2019; Petrina, 2020). UNICEF's Policy Guidance on AI for Children (2021) calls for participatory, age-appropriate approaches that respect children's rights and agency. Similarly, the MIT Media Lab's AI + Ethics curriculum demonstrates that narrative prompts and collaborative inquiry can promote ethical awareness and metacognitive skills in middle-school learners (Wu, 2020; Williams et al., 2022).

Lesson Design and Implementation

The lesson was piloted with a mixed-age group of ten children (ages 8–12) in a 60-minute session. The structure followed three interconnected phases: (1) an illustrated narrative stimulus, (2) a Community of Philosophical Inquiry (CoPI) dialogue, and (3) a hands-on ethical-decision-making activity. The lesson was subsequently repeated in two after-school sessions and one primary-school class; reflections synthesize insights from these implementations.

1. Narrative Stimulus: "Liam's Secret Base"

The session opened with a short illustrated story, inspired by Minecraft, to ensure cultural and age relevance. Liam is an avid player who has built a private "secret base" to store his digital treasures and shares the coordinates only with close friends. One day, he discovers unfamiliar players entering the base and using his resources. Upset and confused, he learns that a friend posted the coordinates publicly online. After the storytelling- *the visuals were presented as a short PowerPoint slideshow with basic animations and captions to maintain children's attention and visual engagement*- the facilitator posed reflective questions: "Have you ever played a game where you had to share or protect information?", "How would you feel if someone accessed your information without permission?", "Why do you think protecting personal information matters, even inside a game?"

2. Community of Philosophical Inquiry (CoPI)

Following P4C practice, the facilitator invited children to generate their own questions in response to the story. The questions were recorded on a board and the group voted on which to discuss first. Sample student questions included: "Should people always ask permission before sharing information about someone else?", "Is it fair for game companies to collect players' data?", and "Who is responsible when private information is shared without consent?" The class selected: "Should you always ask before sharing something about another person?" The ensuing dialogue revealed nuanced reasoning about consent, trust, fairness, and responsibility. The facilitator employed open-ended prompts, paraphrasing, and probing questions to deepen inquiry while avoiding prescriptive answers (Lipman, 2003).

Transcript Excerpt (Plain Vanilla CoPI)

Facilitator: Why do you think Liam was upset?
Student A: Because it was a secret and someone told without asking.
Facilitator: What does it mean to keep something secret? Is every secret private?
Student B: Maybe not. Some secrets are for surprises, like birthdays.
Facilitator: So is privacy always about hiding, or sometimes about choosing?
Student C: Choosing! You get to decide who knows things about you.
Facilitator: Interesting. Does that mean privacy is also about control of information?
Student D: Yeah, like being the boss of your information.
Facilitator: If privacy is control, what makes sharing sometimes okay and sometimes not?
Student E: It depends who asks and why they need it.
Facilitator: What criteria could we use to decide?
Student F: If it's sensitive; if I trust them; and what could happen after.

Analytic Note on Facilitation

This exchange illustrates the Plain Vanilla sequence in which facilitator prompts shift discourse from anecdotal sharing to conceptual exploration. By reframing "secret" as an analyzable concept and linking privacy to choice and control, the facilitator elicited criteria (sensitivity, trust, purpose, consequences) that support principled judgment rather than rote rules (Yarmel & Gregory, 2023).

3. Hands-On Activity: "Safe Sharing Scenario"

To reinforce dialogue with experiential learning, students participated in a role-play activity called "Safe Sharing Scenario." The facilitator presented a realistic online-gaming situation: "You're playing a popular online game. A new friend asks for your real name and home address to send you a gift. Another player asks for your game password, promising to help you unlock a rare item. What would you do in each case?" Children broke into small groups to analyze each situation: identify what kind of information is being requested, decide if it is safe to share and explain why or why not, and predict possible consequences of sharing. Groups then role-played the interactions—some acting as the 'player,' others as the 'requester,' and a third subgroup as 'observers' offering feedback. Finally, each group drafted a "Safe Sharing Checklist" featuring rules such as "Always ask a trusted adult before sharing personal information" and "Never share passwords—even with friends." A closing debrief helped children link these practices to wider issues of consent, trust, and data privacy in their digital lives.

Reflections on Learning Outcomes

Engagement and Higher-Order Thinking. Throughout the session, children displayed high engagement with both the story and the discussion. Their contributions moved beyond opinions toward reason-giving, perspective-taking, and questioning assumptions—key indicators of higher-order cognitive skills. Several students explicitly linked the fictional scenario to real-life digital experiences, such as sharing photos or passwords with friends.

Ethical Reasoning and Metacognitive Awareness. The CoPI dialogue enabled children to identify ethical dimensions—consent, trust, fairness, responsibility—while reflecting on their own reasoning processes. This aligns with research showing that P4C fosters ethical sensitivity and metacognitive growth (Lipman, 2003; Ab Wahab et al., 2022). Across repeated sessions, similar patterns emerged: initial storytelling elicited anecdotal responses, but the facilitator's probing questions (e.g., "Is privacy the same as secrecy?") consistently moved dialogue toward conceptual generalization.

Collaborative Dialogue and Caring Thinking. Students listened actively to peers, paraphrased each other's points, and disagreed respectfully with reasons. This collaborative turn-taking reflected "caring thinking," demonstrating empathy and recognition of others' viewpoints—skills important for responsible digital citizenship.

Transfer to Everyday Digital Life. During the closing discussion, several children reported intentions to check privacy settings, seek adult help before sharing, and avoid password-sharing.

Follow-up conversations one week later indicated recall of the Safe Sharing Checklist, suggesting initial transfer of ethical learning to practical decision-making.

Practical Recommendations for Educators

- Leverage familiar digital contexts: Use games, stories, or platforms already known to children—such as Minecraft—to make abstract ethical concepts more relatable and concrete (Adams et al., 2023).
- Prioritize child-led inquiry: Encourage students to generate and prioritize their own questions. This preserves the democratic spirit of CoPI and supports autonomy in ethical reasoning (Lipman, 2003).
- Balance dialogue with experiential activities: Combine philosophical discussion with hands-on tasks—like the Safe Sharing Scenario—to bridge conceptual reflection and real-life decision-making (Turner, 2019).
- Connect ethical inquiry to family and community: Provide take-home Safe-Sharing Checklists or reflection prompts to involve parents and caregivers, reinforcing learning in broader digital lives (UNICEF, 2021).
- Adopt inclusive facilitation strategies: Create a supportive, non-judgmental space that respects diverse experiences with technology and allows children from varied backgrounds to feel heard (Ab Wahab et al., 2022).

Conclusion

This Fieldnotes study demonstrates that integrating AI-ethics education—specifically data-privacy topics—into P4C is both pedagogically feasible and developmentally appropriate for upper-primary learners (children aged 8–12).

By employing a story-driven stimulus, student-generated philosophical dialogue, and a hands-on role-play activity, the lesson enabled children to recognize ethical dimensions of everyday digital practices, reason about consent, trust, fairness, and responsibility, and begin transferring ethical insights to their real-life online behaviours. Rather than treating digital ethics as a list of rules to memorize, the approach encouraged critical, creative, and caring thinking (Lipman, 2003) in line with participatory principles advocated by UNICEF (2021) and demonstrated in initiatives like the MIT Media Lab's AI + Ethics curriculum (Wu, 2020; Williams et al., 2022). We conclude that P4C offers an effective and scalable framework for embedding ethical literacy into childhood digital-citizenship programs, empowering young learners to become thoughtful, reflective, and responsible participants in a rapidly data-driven society (Floridi, 2023/2024; Adams et al., 2023; Dabbagh et al., 2025). Future research should explore long-term impacts on children's reasoning, behavioural choices online, and their evolving understanding of AI-mediated environments.

References

Ab Wahab, M. K., Zulkifli, H., & Abdul Razak, K. (2022). Impact of Philosophy for Children and its challenges: A systematic review. Children, 9(11), 1671. https://doi.org/10.3390/children9111671

Adams, C., Pente, P., Lemermeyer, G., & Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. Computers & Education: Artificial Intelligence, 4, 100131. https://doi.org/10.1016/j.caeai.2023.100131

Dabbagh, H., Earp, B. D., Porsdam Mann, S., Plozza, M., Salloch, S., & Savulescu, J. (2024). AI ethics should be mandatory for schoolchildren. AI and Ethics, 5(1), 87–92. https://doi.org/10.1007/s43681-024-00462-1

Floridi, L. (2023). *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities.* Oxford University Press. https://doi.org/10.1093/oso/9780198883098.001.0001

Gregory, M., & Laverty, M. (Eds.). (2017). In community of inquiry with Ann Margaret Sharp: Childhood, philosophy and education. Routledge.

Lipman, M. (2003). Thinking in education (2nd ed.). Cambridge University Press.

Lipman, M., Sharp, A. M., & Oscanyan, F. S. (1980). Philosophy in the classroom (2nd ed.). Temple University Press.

MIT Media Lab, Personal Robots Group. (2020, April 13). Family-friendly data privacy + AI activities. https://www.media.mit.edu/

Petrina, S. (2020). Philosophy of technology for children and youth. In P. J. Williams & D. Barlex (Eds.), Pedagogy for technology education in secondary schools: Contemporary issues in technology education (pp. 311–323). Springer.

Turner, K. H. (Ed.). (2019). The ethics of digital literacy: Developing knowledge and skills across grade levels. Rowman & Littlefield.

UNICEF. (2021, November). *Policy guidance on AI for children: Recommendations for building AI policies and systems that uphold child rights* (Version 2.0). Office of Global Insight and Policy, United Nations Children's Fund. https://www.unicef.org/innocenti/media/1341/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf https://www.unicef.org/innocenti/reports/policy-guidance-ai-children

Williams, R., Park, H. W., & Breazeal, C. (2022). AI + ethics curricula for middle-school youth. International Journal of Artificial Intelligence in Education, 32(4), 1035–1063. https://doi.org/10.1007/s40593-021-00262-0

Yarmel, A., & Gregory, M. (2023). Noise reduction in the community of philosophical inquiry. Journal of Philosophy of Education, 57(4), 705–724. https://doi.org/10.1111/1467-9752.12766

Address Correspondences to:
Mohammad Reza Vaez Shahrestani
Director of Yara: The Iranian Center for the Advancement of Philosophy for Children and Youth (ICAPCY)
Researcher in Philosophy, University of Bonn, Germany
https://rooyeshedigar.ir/director
mrvaez@uni-bonn.de