Advancing Human-Robot Collaboration: The Impact of Flexible Input Mechanisms

Helen Beierling, Kira Loos, Robin Helmert and Anna-Lisa Vollmer

Abstract—**This position paper calls for an increase in the number and flexibility of input mechanisms in user-robot interactions, highlighting their potential to enhance learning algorithms through user feedback. Moreover, we argue that refining interfaces, interactions, and systems is crucial for the optimal integration of mechanisms into learning processes. Our call to research involves the development of interfaces that enable flexible mechanisms, and the mechanisms interactions can benefit most from, and the algorithmic incorporation of user input. This aims to advance the adaptability and responsiveness of robotic systems in human-centric environments.**

I. Introduction

This paper postulates the provision of multiple options for interaction mechanisms to make the tutoring experience more enjoyable and improve learning performance.

So far in robot learning in interaction with human tutors, the underlying algorithms of the learning system provide restricted mechanisms that directly determine the kind of interface used for input and the (mostly artificial) interaction through which the human user can provide input to the system, see Fig. 1. Human tutoring interactions, however, are much more flexible, and the tutor uses this flexibility to scaffold the learner.

Scaffolding, originally introduced in the context of developmental psychology [42, 34], refers to the supportive approach that enhances learning capabilities by modifying the learner's environment or providing guidance tailored to the learner's current capability level and adjusting the support a human teacher gives constantly concerning the learner's task progression [40, 31, 17, 33, 7]. It has been observed to enable "development of task competence by the learner at a pace that would far outstrip [their] unassisted efforts" [42]. While the benefits of scaffolding have been extensively studied in educational settings in human interaction [10, 12, 4], research on enabling dynamic scaffolding in human-robot interaction (HRI) remains limited. To address the scaffolding needs of users in human-robot interaction (HRI), we suggest providing more flexible input mechanisms. .

II. Related Work

Scaffolding has long been recognized to improve the efficiency and effectiveness of robot training but also is viewed to have the potential of enhancing the collaboration between humans and robots [1], leading to more sophisticated and adaptable robotic systems [7, 8]. Borrowing from the educational theory of scaffolding, in the context of Human-Robot Interaction (HRI) and Artificial Intelligence (AI), "scaffolding" refers to a supportive framework or set of interventions designed to facilitate and enhance the learning process by

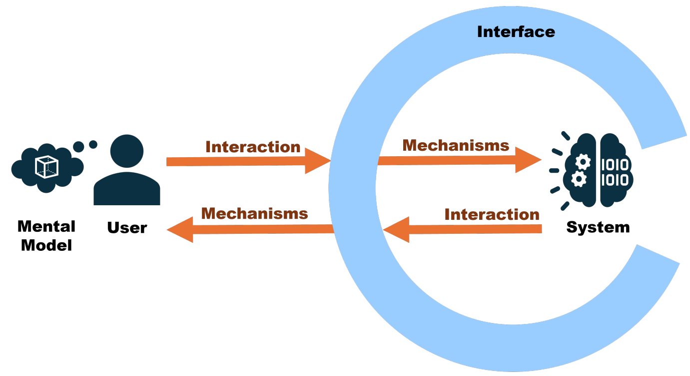

Fig. 1. The figure illustrates a typical one-to-one human-machine interaction, representing the involved components.

leveraging human feedback [25] and involves structuring the interaction between the user and the robot in a way that gradually builds up the robot's capabilities and decision-making skills [6].

However, this promising approach remains largely undervalued and underexplored in current research in HRI and AI, possibly because human scaffolding behavior cannot yet be dealt with, as it is highly variable and depends on the feedback of the learning system based on interfaces and the provided input mechanisms [38, 39].

Few works have adopted scaffolding to serve both as a strategy for robot interaction with humans and with other robots [19]. In intelligent tutoring systems, robots can act as tutors and adapt their interactions based on the learner's responses and progress [20, 37, 14, 32]. In robot learning, a human helps a robot to complete a task by giving instructions and feedback, guiding it through the task environment and breaking the task down into simpler and more manageable components [25, 20, 27] so that the robot's capabilities and decision-making skills increase over the course of the interaction time [6].

Despite its benefits, scaffolding behavior in human tutors is underestimated or often even misidentified as noise and not further processed [8, 21, 41]. This inefficient use of information could be avoided by providing users with flexible input mechanisms and corresponding interfaces, enabling them to effectively transfer their scaffolding. More flexible interactions with different interfaces and mechanisms would allow addressing user needs directly and intuitively.

Various mechanisms and interfaces can be used in human-robot interactions. Commonly used input mechanisms include demonstrations [30], preference-based inputs, and scalar inputs [18]. Interfaces such as graphical user interfaces (GUIs) [16] can facilitate these mechanisms. However, many current applications restrict users to specific interfaces and mechanisms, lacking the flexibility to respond to individual and changing user needs. There are only a few works about the use of multiple mechanisms and interfaces for the same scenario [28, 35]. Therefore, we are calling for more flexibility in mechanisms and interfaces for scaffolding in human-robot and human-AI interactions. To achieve this, we identify several potential options.

III. Components

Here, we discuss possible implementations and benefits of flexible input mechanisms.

a) Mental Model: First, we consider the mental model that the user has. Mental models are individual cognitive representations of the world and self based on experiences and interactions with the environment [11]. The mental model influences how users interact with the system and can lead to interaction problems when mismatched [22, 26]. It has been demonstrated that scaffolding depends on the tutor monitoring the learner [38]. Systems, thus, should show their learning state to the human user [38]. Introducing transparency for the system's operations, which calibrate the user's mental model, leads to an improvement in learning outcomes [15]. Through observation, users can adapt their scaffolding behavior, which affects the input mechanisms used, as found in initial studies [5].

b) User: We refer to the individual who operates a system as the user. Users can possess varying levels of prior knowledge regarding the system and the interface, ranging from experts to lay persons. Various forms of scaffolding that users may wish to apply during interactions with children and robots were already identified [29, 36], like the mediation of the relevance of the demonstration: users want to convey the most important aspects of a task, focusing on the actions and results that they consider to have a significant impact on the learner. Users also want to scaffold the learner's attention to those critical elements, they intend to scaffold the motivation of the robot and they want to provide timely and constructive feedback. It has been suggested that learning systems could actively elicit needed information from human tutors [39]. When users are given information about the robot's learning state, its current understanding and appropriate input mechanisms, they are better able to provide useful guidance. Flexible input mechanisms should therefore not only serve the robot, but also help users to structure their explanations to provide better guidance to the learner.

c) Interaction: In our paper, we define interaction as the action performed to input or output data through an interface. There are unique scaffolding opportunities and dependencies inherent in the design and use of interfaces and interactions. For example, different input mechanisms, such as demonstrations or textual feedback, which allow for different forms of scaffolding, work better with different interactions. Each method is more intuitive in its respective application area than the other, thereby using the best interaction enhances the scaffolding that users can provide to the system. Typing in VR can be cumbersome and imprecise, and therefore a regular keyboard might be the better option here [13]. However, VR allows for accurate tracking of demonstrations. In fact, there are many possible ways and combinations of scaffolding. Which interaction is best for which mechanism, interface, and scaffolding need is still an open question and its investigation promises drastic improvement in learning [5]. The interaction also influences the mechanisms used, as some interactions are more suitable for specific mechanisms, such as VR for demonstrations. Conversely, the choice of a mechanism can affect the type of interaction employed. Therefore, it remains an open question what needs to be optimized to achieve the best human-robot interaction: the mechanisms, the interactions, or a combination of both.

d) Interface: In our paper, we define an interface as the connection between the system and the user. Even when a system is capable of processing scaffolding, these capabilities must be effectively presented to the user for interaction. Many researched scaffolding methods lack conventional representation, and often several methods may be applicable, which can influence user behavior. For example, scalar evaluations of a robot's learning state could be represented with stars, hearts, or emoticons. It is crucial to investigate which representation best aligns with the users' scaffolding intentions and most effectively and intuitively conveys this information to the system. Since this transmission takes place via the input mechanism, the problem can be rewritten in such a way that we need flexible input mechanisms that are effectively presented to the user. However, the question remains of how the representation affects learning success and which interface works best for which input mechanism.

e) Mechanisms: In this paper, we refer to mechanisms as the potential processing methods for input, as well as their possible implementation. It has been shown that offering users direct input mechanisms within the interface to scaffold the system is advantageous [24, 23, 3, 9, 2]. We have also found that the combination of such scaffolding mechanisms with appropriate system integration offers significant value for learning [5]. However, it remains unclear which mechanisms are most suitable for scaffolding, e.g., positive mechanisms via guidance or negative mechanisms via corrections. How scaffolding feedback is integrated into the system needs to be carefully considered, ensuring that it aligns with the desired learning outcomes and interaction dynamics. The most effective approach in these scenarios is still largely unexplored. Initial studies have shown that combining input mechanisms and allowing users flexibility has a significant positive impact on learning [5]. Yet, optimal combinations, their limits, and variations between users remain open questions.

f) System: In this paper, we define the system as encompassing both the hardware and software with which users interact. As mentioned before, the input mechanisms for the users are crucial to support the robot's learning. However, the systems must also provide feedback to the users, acting as output mechanisms, so that the users can understand the robot's learning state and adjust their scaffolding accordingly. The output mechanisms should be designed to convey the robot's learning progress to the user. This feedback helps users to make their inputs more effective, improve the robot's learning process and support scaffolding. Most state-of-the-art systems use AI and machine learning to enable adaptive behavior and identify appropriate behavior patterns. By using these techniques, a learning system could draw more information from the user's feedback and non-verbal behavior, adapting to individual behavior patterns to make the interaction and concept learning more efficient [25, 27]. To fully realize the potential of in- and output mechanisms, it is essential to focus on increasing the flexibility of mechanisms within these systems. Mechanisms can only be implemented correctly if the system provides the necessary structure. Effective scaffolding depends on this fit, as mismatched systems may fail to analyze inputs correctly, undermining the mechanism's functionality. Since systems must fit specific mechanisms, it is crucial to investigate which systems are best suited for each mechanism to maximize scaffolding efficiency.

IV. Summary

In summary, we believe that flexible input mechanisms should be actively incorporated more extensively into human-machine interactions, especially in robotics. They make the interaction more efficient and natural. We assume that this is the case because users apply their scaffolding behavior to the provided input mechanisms. Thus, scaffolding provides a valuable tool in areas with limited feedback. There are still many potential opportunities for integrating scaffolding with flexible input mechanisms into these systems.

Acknowledgments

This work was funded by the German Research Foundation (DFG) under grant TRR 318/1 2021 - 438445824 and by SAIL, funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia under the grant no NW21-059A. The authors are with the Medical School Ostwestfalen-Lippe and the Center for Cognitive Interaction Technology, Bielefeld University, Bielefeld, Germany.

References

[1] Charoula Angeli and Nicos Valanides. Developing young children's computational thinking with educational robotics: An interaction effect between gender and scaffolding strategy. *Computers in human behavior*, 105: 105954, 2020.

[2] Christian Arzate Cruz and Takeo Igarashi. A survey on interactive reinforcement learning: Design principles and open challenges. In *Proceedings of the 2020 ACM Designing Interactive Syst. Conf.*, pages 1195–1209, 2020.

[3] Andrea Bajcsy, Dylan P Losey, Marcia K O'Malley, and Anca D Dragan. Learning from physical human corrections, one feature at a time. In *Proc. of 2018 ACM/IEEE Int. Conf. on Human-Robot Interaction*, pages 141–149, 2018.

[4] Arthur Bakker, Jantien Smit, and Rupert Wegerif. Scaffolding and dialogic teaching in mathematics education: Introduction and review. *Zdm*, 47:1047–1065, 2015.

[5] Helen Beierling and Anna-Lisa Vollmer. The power of combined modalities in interactive robot learning. *arXiv preprint arXiv:2405.07817*, 2024.

[6] Aude G Billard, Sylvain Calinon, and Rüdiger Dillmann. Learning from humans. *Springer handbook of robotics*, pages 1995–2014, 2016.

[7] Cynthia Breazeal. Learning by scaffolding. *Thesis Proposal*, pages 1–106, 1998.

[8] Cynthia Breazeal, Matt Berlin, Andrew Brooks, Jesse Gray, and Andrea L Thomaz. Using perspective taking to learn from ambiguous demonstrations. *Robotics and autonomous systems*, 54(5):385–393, 2006.

[9] Carlos Celemin and Jens Kober. Knowledge-and ambiguity-aware robot learning from corrective and evaluative feedback. *Neural Comput. Appl.*, pages 1–19, 2023.

[10] Morteza Dianatfar, Jyrki Latokartano, and Minna Lanz. Review on existing vr/ar solutions in human–robot collaboration. *Procedia CIRP*, 97:407–411, 2021.

[11] Dedre Gentner and Albert L Stevens. *Mental models*. Psychology Press, 2014.

[12] Talip Gonulal and Shawn Loewen. Scaffolding technique. *The TESOL encyclopedia of English language teaching*, pages 1–5, 2018.

[13] Jens Grubert, Lukas Witzani, Eyal Ofek, Michel Pahud, Matthias Kranz, and Per Ola Kristensson. Effects of hand representations for typing in virtual reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018, Tuebingen/Reutlingen, Germany, 18-22 March 2018*, pages 151–158. IEEE Computer Society, 2018. doi: 10.1109/VR.2018.8446250.

[14] Omer Gvirsman and Goren Gordon. Effect of social robot's role and behavior on parent-toddler interaction. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '24, page 222–230. Association for Computing Machinery, 2024. ISBN 9798400703225.

[15] Lukas Hindemith, Jan Philip Göpfert, Christiane B Wiebel-Herboth, Britta Wrede, and Anna-Lisa Vollmer. Why robots should be technical: Correcting mental models through technical architecture concepts. *Interaction Studies*, 22(2):244–279, 2021.

[16] Lukas Hindemith, Oleksandra Bruns, Arthur Maximilian Noller, Nikolas Hemion, Sebastian Schneider, and Anna-Lisa Vollmer. Interactive robot task learning: Human teaching proficiency with different feedback approaches. *IEEE Transactions on Cognitive and Developmental Systems*, 2022.

[17] Nurul Farhana Jumaat and Zaidatun Tasir. Instructional scaffolding in online learning environment: A meta-analysis. In *2014 International Conference on Teaching*

and Learning in Computing and Engineering, pages 74–77, 2014.
[18] Timo Kaufmann, Paul Weng, Viktor Bengs, and Eyke Hüllermeier. A survey of reinforcement learning from human feedback. *arXiv preprint arXiv:2312.14925*, 2023.
[19] Murat Kirtay, Erhan Oztop, Minoru Asada, and Verena V. Hafner. Trust me! i am a robot: an affective computational account of scaffolding in robot-robot interaction. In *2021 30th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 189–196, 2021.
[20] Murat Kirtay, Verena V. Hafner, Minoru Asada, and Erhan Oztop. Trust in robot–robot scaffolding. *IEEE Transactions on Cognitive and Developmental Systems*, 15(4):1841–1852, 2023.
[21] W Bradley Knox and Peter Stone. Reinforcement learning from human reward: Discounting in episodic tasks. In *2012 IEEE RO-MAN: The 21st IEEE international symposium on robot and human interactive communication*, pages 878–885. IEEE, 2012.
[22] Sau-lai Lee, Ivy Yee-man Lau, Sara Kiesler, and Chi-Yue Chiu. Human mental models of humanoid robots. In *Proceedings of the 2005 IEEE international conference on robotics and automation*, pages 2767–2772. IEEE, 2005.
[23] Mengxi Li, Alper Canberk, Dylan P Losey, and Dorsa Sadigh. Learning human objectives from sequences of physical corrections. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2877–2883. IEEE, 2021.
[24] Huihan Liu, Alice Chen, Yuke Zhu, Adith Swaminathan, Andrey Kolobov, and Ching-An Cheng. Interactive robot learning from verbal correction. *arXiv preprint arXiv:2310.17555*, 2023.
[25] Hermann Mayer, Darius Burschka, Alois Knoll, Eva U Braun, Rüdiger Lange, and Robert Bauernschmitt. Human-machine skill transfer extended by a scaffolding framework. In *2008 IEEE International Conference on Robotics and Automation*, pages 2866–2871. IEEE, 2008.
[26] Jonathan Mumm and Bilge Mutlu. Human-robot proxemics: physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 331–338, 2011.
[27] Nuno Otero, Joe Saunders, Kerstin Dautenhahn, and Chrystopher Nehaniv. Teaching robot companions: The role of scaffolding and event structuring. *Connect. Sci.*, 20:111–134, 09 2008.
[28] Dennis Perzanowski, Alan C Schultz, William Adams, Elaine Marsh, and Magda Bugajska. Building a multi-modal human-robot interface. *IEEE intelligent systems*, 16(1):16–21, 2001.
[29] Karola Pitsch, Anna-Lisa Vollmer, Katharina J Rohlfing, Jannik Fritsch, and Britta Wrede. Tutoring in adult-child interaction: On the loop of the tutor's action modification and the recipient's gaze. *Interaction Studies*, 15(1):55–98, 2014.
[30] Harish Ravichandar, Athanasios S Polydoros, Sonia Chernova, and Aude Billard. Recent advances in robot learning from demonstration. *Annu. Rev. Control Rob. Auton. Syst.*, 3:297–330, 2020.
[31] Joe Saunders, Chrystopher L. Nehaniv, and Kerstin Dautenhahn. Teaching robots by moulding behavior and scaffolding the environment. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, HRI '06, page 118–125. Association for Computing Machinery, 2006.
[32] Thorsten Schodde, Laura Hoffmann, Sonja Stange, and Stefan Kopp. Adapt, explain, engage—a study on how social robots can scaffold second-language learning of children. *J. Hum.-Robot Interact.*, 9(1), dec 2019.
[33] C. Addison Stone. Should we salvage the scaffolding metaphor? *Journal of Learning Disabilities*, 31(4):409–413, 1998.
[34] C Addison Stone. The metaphor of scaffolding: Its utility for the field of learning disabilities. *Journal of learning disabilities*, 31(4):344–364, 1998.
[35] Andrea Lockerd Thomaz, Cynthia Breazeal, et al. Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance. In *Aaai*, volume 6, pages 1000–1005. Boston, MA, 2006.
[36] Janneke Van de Pol, Monique Volman, and Jos Beishuizen. Scaffolding in teacher–student interaction: A decade of research. *Educational psychology review*, 22:271–296, 2010.
[37] Paul Vogt, Mirjam de Haas, Chiara de Jong, Peta Baxter, and Emiel Krahmer. Child-robot interactions for second language tutoring to preschool children. *Frontiers in Human Neuroscience*, 11, 2017.
[38] Anna-Lisa Vollmer, Katrin Solveig Lohan, Kerstin Fischer, Yukie Nagai, Karola Pitsch, Jannik Fritsch, Katharina J Rohlfing, and Britta Wredek. People modify their tutoring behavior in robot-directed interaction for action learning. In *2009 IEEE 8th international conference on development and learning*, pages 1–6. IEEE, 2009.
[39] Anna-Lisa Vollmer, Manuel Mühlig, Jochen J Steil, Karola Pitsch, Jannik Fritsch, Katharina J Rohlfing, and Britta Wrede. Robots show us how to teach them: Feedback from robots shapes tutoring behavior during action learning. *PloS one*, 9(3):e91349, 2014.
[40] Anna-Lisa Vollmer, Britta Wrede, Katharina J. Rohlfing, and Pierre-Yves Oudeyer. Pragmatic frames for teaching and learning in human–robot interaction: Review and challenges. *Frontiers in Neurorobotics*, 10, 2016.
[41] Paul Weng, Robert Busa-Fekete, and Eyke Hüllermeier. Interactive q-learning with ordinal rewards and unreliable tutor. 2013.
[42] David Wood, Jerome S Bruner, and Gail Ross. The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, 17(2):89–100, 1976.