Children and Robots: A Preliminary Review of Methodological Approaches in Learning Settings

Vicky Charisi, Daniel Davison, Denis Reidsma and Vanessa Evers
Dept. of Electrical Engineering Mathematics and Computer Science
Human Media Interaction Group
University of Twente, The Netherlads
Email: {v.charisi, d.p.davison, d.reidsma, v.evers}@utwente.nl

***Abstract*—This review examines recent methodological approaches for the evaluation of child-robot interaction in learning settings. The main aim is to map existing work and to discuss potential future directions. We report representative studies, which have been thematically organized based on their research objectives. We then examine the evaluation methods that were used in these studies and we propose a conceptual framework based on the one hand on the themes that have emerged, namely the social interaction between the child and the robot, the social acceptance, possible emotional interaction, the learning process and the learning outcome and on the other hand on the corresponding measures. We use this framework to highlight current trends and needs for the field and to contextualize the methodological directions that we have adopted in a current project for child-robot interaction in learning settings. Finally, we discuss the challenges and future directions for evaluation methods in learning settings.**

I. Introduction

This review investigates the current approaches in evaluation methods for the growing field of child robot interaction (CRI) in learning settings. Children's interaction with technology has been extensively studied, and researchers in the field of child computer interaction have developed relevant methods of evaluation [21]. Although these methods can inform the methodological choices in research for CRI, the physical nature of robots and the social characteristics that are ascribed to them have indicated the need for additional methods for evaluation. This need has also emerged during our research studies in the context of a European Union funded project. During a series of studies we tried to evaluate the extent to which the EASEL robot is able to initiate, sustain and support a developmentally appropriate CRI for the purpose of teaching and learning, in which the robot has the role of learning companion. The aim of this review is to identify current trends in research studies for children and robots and to suggest a conceptual framework for the evaluation methods highlighting the current challenges and possible future directions.

II. Literature

In this section we summarize a series of representative studies for CRI and the corresponding methods. However, the degree of autonomy of the robot, the wildness of the interaction environment and the target group under investigation seem to have a catalytic role on the selection of the evaluation methods. In CRI, target groups may include typically developing children or children with special needs of a varied age range, who require developmentally appropriate methods. The robot's social characteristics and its role in the task may also affect the choice of the evaluation metrics. Previous work reports a variety of roles for robots interacting with children. Examples include a tutor, who gives hints or instructions which the child has to follow [30]; a peer or co-learner, which adopts empathetic behaviours towards the child [19]; a mediator, who facilitates the social interaction among the children [22]; or a tool for children to acquire specific learning goals such as programming skills [26]. In the following sections, we present an overview of some studies and we suggest a classification of their objectives.

A. A Preliminary Overview of Relevant Studies

This section aims to report a preliminary overview of relevant studies rather, than an exhaustive systematic review. As a starting point for the organization of this overview, we used the developmental paradigm of socio-cultural theory, according to which, children develop better when learning occurs in a social context. As a result, social interactions that take place during the learning process may influence the learning gains. In this case, children's learning companions include physical humanoid or non-humanoid robots.

Attributing socialness to inanimated objects is a well-investigated phenomenon that appears in young children. The evaluation of children's perspectives of robot's socialness has been mainly investigated with observational methods, semi-structured interviews and questionnaires [15]. Child's perceptions and interpretation of robots' social behaviour relates to the repertoire of its social acts. Multimodality creates the opportunities for richer and more effective interactions, however they are more complex to design and evaluate [3]. Heerink et al. [15] used observational methods to evaluate children's play with a robot and they developed a coding scheme of low-level micro-behaviours for systematic observations. However, children's social acceptance seems to be culture-specific. Intercultural studies have shown that children with different cultural backgrounds reacted differently towards robots [24]. For the evaluation of their acceptance by children, Shahid et al. combined self-report scores, perception tests and behavioural analyses and they found that children with different cultural backgrounds respond to the same social situations in a different way. Moreover, the social acceptance of a robot by a child has been investigated in long-term interactions. Children appear to be biased by the novelty effect when they interact with a robot for the first time. However, in long-term interactions, it has been found that social acceptance and effective interactions may be established, when the novelty effect diminishes [28]. Social acceptance in long-term interactions has been evaluated with the use of self-report measurement methods in combination with observational methods or objective methods of automated data.

Child-robot social interaction often leads to children's emotional response and eventually to a possible emotional bonding with the robot. In such a case, emotion recognition is associated with the child's ability to distinguish the various emotional expressions in facial, gestural, vocal and verbal displays in robots and to understand their social meaning [12]. Beck et al.[2] measured children's perceptions of robot's emotional states by systematically altering its head positions of 6 emotional key poses. They used questionnaires to ask children to ascribe the robot's posture into a certain emotional state. For emotional long-term interaction, models of adaptive emotion expression have been developed [29] with the use of video analysis and questionnaires to find that children react more expressively to a robot, which exhibits emotional expressions in an adaptive way.

Long-term repeated studies have used standardized validated existing tools to evaluate the robot's effects. Kanda et al. [17], for example, used continuous automated methods over time to investigate the process of children's adaptation to interactions with robots. Long-term studies have investigated children's learning processes when a robot is used as a tool for children to learn specific skills by using qualitative approaches such as design-based approaches [26]. Other methods for the evaluation include observational methods of children's task-related activity over time [17] or of the collaboration among children when a robot has the role of the facilitator [11], automatically generated data, measurement of the length of utterances and qualitative categorization of the content of children's utterances in order to investigate the gradual change of children's behaviour [25].

Summative evaluation methods have been widely used to assess children's learning gains. Tests before and after the intervention as well as task performance during the intervention reveal the effectiveness of the interactions in the learning task [20]. Holmes [16] used interactive agents to support learning by explaining and he used a pretests and posttests set up, which consisted of both multiple choice/short answer type questions. While summative evaluation methods are objective methods, which allow for generalizations, there are limitations in investigating clear relations and causalities when the intervention happens in a real school setting, where multiple factors may affect the result of the evaluation.

B. A Classification of the Objectives in Learning Settings

The examination of the research goals of the above-mentioned studies led us to their classification into five interconnected thematic constructs that include social, emotional and cognitive features:

i. Social interaction, which describes the short term and long term interaction between the child and the robot, such as verbal and non-verbal communication.

ii. Social acceptance, which includes children's attitudes and perspectives about the socialness of the robot and indicate their willingness to accept it as a social agent.

iii. Emotional interaction, which refers to long-term interaction that may trigger children's emotional engagement and eventual bonding with the robot.

iv. Learning process, which the child follows, and is affected by her interaction with the robot in the context of a specific learning task.

v. Learning outcome, which is defined as the fulfillment of a specific learning goal, in the form of a cognitive achievement or as a desired skill or competence.

The evaluation methods that are used in each of these categories have been organized in the following conceptual framework.

III. The Conceptual Framework

In this section, we take the above-mentioned studies as a starting point to expand the discussion about the evaluation methods that are commonly used in this field. When it is necessary, we provide additional input from the field of child-child interaction. A large body of quantitative methods seek to investigate correlations or causations in child-robot interaction, while qualitative ethnographic-inspired methods seek for a deep understanding of phenomena under investigation in naturalistic settings. Four main categories of methods that are commonly used for the evaluation of CRI are as follows:

1) Self-assessment Methods: Self-assessment methods have been used to investigate participants' perceptions of the phenomenon under investigation. However, children's abilities for reflecting on themselves and their capabilities of accurate verbal expression of their thoughts are under development, which raises several challenges. Attempts to overcome these difficulties include indirect approaches, such as the use of pictorial tasks. Harter and Pike [14] built the pictorial scale of perceived competence and social acceptance for young children. In the field of child-robot interaction, pictorial tasks have been used as a starting point for semi-structured interviews, using stimuli and settings that are meaningful for children [4]. Measures of conversational engagement, the frequency of the use of specific words as well as content analysis of their responses are the main methods for analysis of semi-structured interviews. Additionally, questionnaires that are already validated from human-human interaction have been slightly modified in order to be appropriate for children. Alves-Olivera et al [1] for example, were inspired by the scale of Bhattacherjee and Premkumar to build a questionnaire to measure children's expectations and satisfaction for their interaction with robots. Finally, while the use of large-scale surveys seems to inform research in human-robot interaction with adults, it does not seem to be common among the methods of self-assessment for children. Borgers et al. [7] discuss the influence of cognitive development on response quality, which requires appropriate questions for children. Okita and Schwartz [23] addressed this challenge by asking the children only simple questions. However, oversimplification of the questions limits the scope and the goals of the research. Despite the challenges in self-assessment methods with children and the ways that they evaluate their interaction with a robot, it has been shown that children can and do provide reliable responses if questioned in a manner that they can understand and about events that are meaningful to them [18].

2) Behavioural Observations: Behavioural observations have been used in order to gain understandings from the child's behavioural cues during her interaction with the robot. Ethnographic approaches, for example, use behavioural observations to discover high-level patterns and regularities in real settings [13]. Descriptive and exploratory methods, such as design-based research, have been used for the identification of newly researched fields [26]. Most of these studies use behavioural observations to assess CRI in real-life settings with the analysis of high-level behaviours; however they have often been critisized by scholars that follow different epistemological paradigms. Behavioural observations have also been used in microgenetic and experimental studies, such as in [27]. In these cases, the researchers analyse low-level behaviours, such as touch, speech and gesture, in order to achieve more objective metrics.

3) Psycho-physiological Metrics: Psycho-physiological metrics for the evaluation of CRI include automated measures that seek to detect or identify a specific human psychological statement or signals. The evaluation of participant's responses to a robot occurs through these signals, which are used for the development, implementation of real-time control and modification of robot behaviours, such as social signal processing (for a review [5]). Brain-based methods, sociometrics and psychometrics have been extensively used as objective tools to measure psychological statements. Examples include the Emotional Understanding Score (EUS), which measures participants' ability to correctly recognize and label characters emotional states [19].

4) Task performance measurements: Lastly, task performance is used to evaluate the effectiveness of interventions with robots. More specifically, pre tests and post tests as well as standardized tests or measureent of the time spent on a specific activity are examples of the objective measurements of participants' task performance [13].

IV. Examples from EASEL Project

In the previous chapters, we reported representative studies from CRI and we discussed their methodological approaches and some of their limitations. In this context, we refer to a series of studies for the evaluation of CRI within the EASEL project. Here, we report the methods from 6 of our studies [10] [9] [8] [6] [31] [30], in which we used 21 measures, clustered to 9 different methods covering all the 4 methodological types. In some cases, one measure is able to relate with multiple aims, which verifies our description of the objectives of the evaluation, which appear interconnected.

As far as the selection of the methodologies is concerned, while one of the most common approaches is the use of questionnaires, in the case of young children we faced several challenges. In order to address these challenges, we tried a variety of approaches inspired by the relevant literature, such as pictorial tasks. One of the initial indications of the effectiveness of this approach is that when we combined the pictorial task with semi-structured interviews, we observed that children used the visual stimulus for their better understanding of the questions and their elaborated reasoning. In addition, while children tend to polarize their responses trying to give socially desirable answers, by giving a second chance to the child to make a selection, they avoid polarization. While we had some initial indications of the effectiveness of this approach, they are not yet validated. More research is needed for the validation of those approaches. In the case of behavioural observations, we chose to annotate low-level behaviours, which did not need any interpretation from the annotator. In a later stage, low-level behaviours are grouped into a higher level in order to inform our research questions. Psycho-physiological metrics were used in four out of five EASEL goals, namely the child-robot social interaction, the social acceptance, the emotional interaction and aspects of the learning process. Now we have the positive indications of the effectiveness of this methodological approach, we can use it for the future work. Lastly, task performance metrics include metrics of pre test and post test. However, given our initial definition of learning, we will include these tests in longitudinal studies with cross-sectional design that may allow for such testing.

V. Conclusion and Future Directions

This preliminary review allowed the development of a conceptual framework for the investigation of current trends regarding the objectives and the corresponding methods in studies for CRI in learning settings. More specifically, while the objectives across studies seem homogeneous, there is a substantial lack of studies regarding the long term CRI such as the examination of the learning process with a robotic learning companion. This gap in the literature could possibly relate (i) to the practical challenges in researching in the classroom for long term, (i) to the technological readiness of the robotic systems and (iii) to the development of novel methodological approaches that allow the mapping of the process in such an environment. This requires learning tasks that are clear and compact enough for using them in short episodes of learning, but at the same time adaptable enough to allow repeated sessions over a longer period able to keep the interest of the children. Additionally, while there are successful attempts with novel ways of evaluation, validation of these methods is needed. A considerable number of studies try new methodological approaches, which have a lot of potential, such as the pictorial tasks and the human-likeness thermometer. However, their validation and their use across cases, it would be catalytic for the advance of the field of CRI, since cross-studies examinations would be possible. Furthermore, this review reveals an imbalance between the use of experimental well-controlled studies and naturalistic ones. Despite the increasing number of studies in naturalistic settings, often the setting tends to be artificial. This requires modification of the tasks in a way that the setup can function more independently as a natural element. To sum up, existing studies already have yielded insights into the interplay between robots, children, and learning; however there is still much room for more advanced techniques and methods for the evaluation of the complex phenomenon of child robot interaction that are developmentally appropriate for children in different age ranges, in order for researchers to obtain a better view of the longer term impact of social robots on children's development.

ACKNOWLEDGMENT

This project has received funding from the European Union Seventh Framework Programme (FP7-ICT-2013-10) as part of EASEL under grant agreement n 611971.

REFERENCES

[1] ALVES-OLIVERA, P., RIBEIRO, T., PETISCA, S., DI TULLIO, E., MELO, F. S., AND PAIVA, A. An empathic robotic tutor for school classrooms: cosidering expection and satisfaction of children as end-users, 2015.

[2] BECK, A., CANAMERO, L., DAMIANO, G., SOMMAVILLA, G., TESSER, F., AND COSI, P. Children interpretation of emotional body language displayed by a robot, 2011.

[3] BELPAEME, T., BAXTER, P., READ, R., WOOD, R., CUAYA HUITL, H., KIEFER, B., RACIOPPA, S., KRUIJFF-KORBAYOVA, I., ATHANASOPOULOS, G., ENESCU, V., LOOIJE, R., NEERINEX, M., DEMIRIS, Y., ROS-ESPINOZA, R., BECK, A., CAN AMERO, L., HIOLLE, A., LEWIS, M., BARONI, I., NALIN, M., COSI, P., PACI, G., TESSER, F., SOMMAVILLA, G., AND HUMBERT, R. Multimodal Child-Robot Interaction: Buiding Social Bonds. *Journal of Human-Robot Interaction 1*, 2 (2012), 33–53.

[4] BERAN, T. N., RAMIREZ-SERRANO, A., KUZYK, R., FIOR, M., AND NUGENT, S. Understanding how children understand robots: Perceived animism in child–robot interaction. *International Journal of Human-Computer Studies 69*, 7–8 (2011), 539–550.

[5] BETHEL, C. L., SALOMON, K., MURPHY, R. R., AND BURKE, J. L. Survey of psyhophysiology measurements applied to human- robot interaction, 2007.

[6] BLANCAS, M., VOULOUTSI, V., GRECHUTA, K., AND VERSCHURE, P. Effects of the robot's role on human-robot interaction in an educational scenario, 2015.

[7] BORGERS, N., DE LEEUW, E., AND HOX, J. Children as respondents in survey research: cognitive development and response quality. *Bulletin de Methodologie Sociologique 66* (2000), 60–75.

[8] CAMERON, D., COLLINS, E., CHUA, A., FERNANDO, S., MCAREE, O., MARTINEZ-HERNANDEZ, U., AITKEN, J. M., BOORMAN, L., AND LAW, J. Help! I can't reach the buttons: facilitating helping behaviours towards robots, 2015.

[9] CAMERON, D., FERNANDO, S., COLLINS, E., MILLINGS, A., MOORE, B. R., SHARKEY, A., EVERS, V., AND PRESCOTT, T. Presence of life-like robot expressions influences children's enjoyment of human-robot interactions in the field, 2015.

[10] CAMERON, D., FERNANDO, S., MILLINGS, A., MOORE, B. R., SHARKEY, A., AND PRESCOTT, T. Children's age influences their perceptions of a humanoid robot as being like a person or machine, 2015.

[11] CHANDRA, S., ALVES-OLIVEIRA, P., LEMAIGNAN, S., SEQUEIRA, P., PAIVA, A., AND DILLENBOURG, P. Can a child feel responsible for another in the presence of a robot in a collaborative learning activity? *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on* (2015), 167–172.

[12] COHEN, I., LOOIJE, R., AND NEERINCX, M. Child's recognition of emotions in robot's face and body, 2011.

[13] CONINX, A., BAXTER, P., OLEARI, E., BELLINI, S., BIERMAN, B., HENKEMANS, B., CAN AMERO, L., COSI, P., ENESCU, V., ESPINOZA, R. R., HIOLLE, A., HUMBERT, R., KIEFER, B., KRUIJFF-KORBAYOVA, I., LOOIJE, R., MOSCONI, M., NEERINCX, M., PACI, G., PATSIS, G., POZZI, C., SCCHITELLI, F., SAHLI, H., SANNA, A., SOMMAVILLA, G., TESSER, F., DEMIRIS, Y., AND BELPAEME, T. Towards long-term social child-robot interaction: using multi-activity switching to engage young users. *Journal of Human Robot Interaction in press* (2015).

[14] HARTER, S., AND PIKE, R. The pictorial scale of perceived competence and social acceptance for young children. *Child Development 55*, 6 (1984), 1969–1982.

[15] HEERINK, M., DIAZ, M., ALBO-CANALS, J., ANGULO, C., BARCO, A., CASACUBERTA, J., AND GARRIGA, C. A field study on perception of social presence and interactive behavior with a pet-robot with primary school children, 2012.

[16] HOLMES, J. Designing agents to support learning by explaining. *Computers and Education 48* (2007), 523–547.

[17] KANDA, T., NISHIO, S., ISHIGURO, H., AND HAGITA, N. Interactive humanoid robots and androids in children's lives. *Children, Youth and Environments 19*, 1 (2009), 12–33.

[18] KELLET, M., AND DING, S. Middle childhood. In *Doing research with children and young people*, S. Fraser, V. Lewis, S. Ding, M. Kellett, and C. Robinson, Eds. Sage, London, 2004, pp. 161–174.

[19] LEITE, I., CASTELLANO, G., PEREIRA, A., MARTINHO, C., AND PAIVA, A. Empathic robots for long-term interaction: evaluating social presence, engagement and perceived support in children. *International Journal of Social Robotics 6* (2014), 329–341.

[20] LEYZBERG, D., SPAULDING, S., AND SCASSELLATI, B. Personalizing robot tutors to individuals' learning differences , 2014.

[21] MARKOPOULOS, P., READ, J., HOŸSNIEMI, J., AND MACFARLANE, S. Child computer interaction: advances in methodological research. *Cognition, Technology & Work 10*, 2 (2007), 79–81.

[22] MOCHIZUKI, T., MITATE, Y., TATENO, Y., WAKIMOTO, T., MIYATA, Y., NAKAHARA, J., AND MIYAKE, N. Robot as a Learning Partner for Promoting Proactive Discussion in peer Groups: a Case Study for Career Development, 2013.

[23] OKITA, S. Y., AND SCHWARTZ, D. L. Young children's understanding of animacy and entertainment robots. *International Journal of Humanoid Robotics 3*, 3 (2006), 393–412.

[24] SHAHID, S., KRAHMER, E., AND SWERTS, M. Child–robot interaction across cultures: How does playing a game with a social robot compare to playing a game alone or with a friend? *Computers in Human Behavior 40* (2014), 86–100.

[25] SHORT, E., SWIFT-SPONG, K., GRECZEK, J., RAMACHANDRAN, A., LITOIU, A., GRIGORE, E. C., FEIL-SEIFER, D., SHUSTER, S., LEE, J. J., HUANG, S., LVONISOVA, S., LITZ, S., LI, J., RAGUSA, G., SPRUIJT-METZ, D., MATARIC, M. J., AND SCASSELLATI, B. How to train your dragonbot: Socially assistive robots for teaching children about nutrition through play, 2014.

[26] SULLIVAN, A., AND UMASCHI BERS, M. U. Robotics in the early childhood classroom: learning outcomes from an 8-week robotics curriculum in pre-kindergarten through second grade. *International Journal of Technology and Design Education* (2015).

[27] TAKAYAMA, L., DOOLEY, D., AND JU, W. Expressing thought: improving robot readability with animation principles, 2011.

[28] TANAKA, F., MOVELLAN, J. R., FORTENBERRY, B., AND AISAKA, K. Daily HRI evaluation at a classroom environment: reports from dance interaction experiments, 2006.

[29] TIELMAN, M., NEERINCX, M., MEYER, J.-J. C., AND LOOIJE, R. Adaptive emotional expression in robot-child interaction, 2014.

[30] VOULOUTSI, V., MUNOZ, M. B., GRECHUTA, K., LALLEE, S., DUFF, A., YSARD LLOBET PUIGBO, J., AND VERSCHURE, P. F. M. J. A new biomimetic approach towards educational robotics: the Distributed Adaptive Control of a Synthetic Tutor Assistant, 2015.

[31] WIJNEN, F., CHARISI, V., DAVISON, D., VAN DER MEIJ, J., REIDSMA, D., AND EVERS, V. Inquiry learning with a social robot: can you explain that to me?, 2015.