THE IMPACT OF A SOCIAL ROBOT'S VOICE ON CHILDREN'S LEARNING

Hsiu-Feng Wang and Shin-Lan Shiu
National Chiayi University, Taiwan
621, No.85 Wenlong, Mingsuin, Chiayi Hsien, Taiwan

ABSTRACT

This study employed a single factor experiment to investigate how different voices from a social robot affect the learning outcomes and motivation of 46 children. The learning motivation for the participants was measured using the ARCS questionnaire, which tested for four subcategories: attention, relevance, confidence, and satisfaction. The results showed that a social robot's voice significantly affected children's motivation to learn, with the children preferring a social robot with a child's voice. However, the voice of a social robot exhibited no significant impact on their learning outcomes. The findings of this experiment are of value to those involved in designing children's learning material that uses social robots and to those who work on interaction design with children.

KEYWORDS

Voice, Children, Social Robots, Learning Outcomes, Learning Motivation

1. INTRODUCTION

With the increase in Human-Robot Interaction (HRI) research, the use of robots has become more popular at home, in hospitals, and in schools. Examining what factors affect the way children engage with a robot has thus become an important issue. Research on the voice of a social robot is especially lacking in the current literature. As an interface, it receives far less attention than appearance and gesture (Tamagawa et al., 2011). The focus of this study is to investigate how a social robot's voice impacts children's learning outcomes and motivation. As one of the first attempts to address this relatively disregarded issue, we examined studies not only on social robots, but also pedagogical agents (PAs). The voice of a social robot and PAs are interfaces through which we "speak" to learners. Due to the similarity of these two research areas, we discussed these fields of research together in order to broaden the subject of research.

2. BACKGROUND OF THE STUDY

A number of theories have been developed to understand the interaction between learners and virtual humans (such as social robot or PAs). Social agency theory is one of the most prominent theories and was developed by Mayer, Sobko and Mautone (2003). The theory suggests pedagogical agents (PAs) that add social cues within multimedia presentations can lead to deeper cognitive processing during learning and result in a more meaningful learning outcome. Mayer (2014) suggests three of the most important social cues for PA design are conversational language, voice, and human-like gestures. The element of voice is one of the key features in social agency theory in its examination of cognitive load within multimedia presentations. Therefore, researchers have conducted a number of studies based on principles related to voice as they seek to understand its effect in more detail. Peng and Wang (2022) proposed a meta-analysis which reviewed papers from 2000 to 2019 on pedagogical agents and included a variety of subjects such as environment, learner, role, appearance, and social clues. The papers discovered that social cues, such as voice, had a verifiable impact on student learning.

Mayer and DaPra (2012) conducted three experiments to observe the ways in which a PA's social cues such as gesturing, facial expression, eye gaze, and human-like movement affected multimedia learning. Their results indicated learners performed better on a transfer test when a PA used social cues than when it did not. They also found that students performed better when PAs spoke using a human voice than when using a machine voice. It has now become a widely adopted concept that PAs speaking with a human voice provide learners with a more complete learning experience than PAs speaking with computer-generated voices (Atkinson et al., 2005; Craig & Schroeder, 2017). However, unlike computer-generated voices, human voices are intrinsically prosodic and convey lexical, semantic, and discourse information through sound (Akker & Cutler, 2003). Thus, the nature of a voice cannot be simply categorized as being only either human or machine. Further exploration into this area of research is thus required.

Davis et al. (2019) conducted an experiment to investigate the relationship between human voice conditions and learners. The researchers recruited 172 undergraduate students who were non-native speakers of English to evaluate how strong- and weak-prosodic human voices compared to a computer human voice in terms of the participants' perceived cognitive load, their evaluation of agent persona, and their information retention. With respect to the perceived cognitive load and their evaluation of the agent persona, the results of the study showed that the non-native speakers preferred a human voice with less prosodic elements when compared to a modern computer voice. A recent study by Wang and Li (2019) investigated how children are affected by voice and the PA's physical appearance. More specifically, the researchers examined how the PA's appearance affected the learning and motivation of 42 primary school children. The experiment employed three different appearances which included ones that looked like an adult tutor, a child companion, or an animal, as well as one that only displayed text. In the study, the children watched three different animated PAs presented and narrated by either an adult human's voice (the tutor) or a child's voice (the companion and animal). Their results indicated the children (11-12 year-olds) who watched and listened to the animated learning material presented by the child companion with the child's voice performed better in terms of learning outcome and intrinsic motivation than those who learned with the other two forms.

The above research demonstrates that a PA's (social robot's) use of a human voice and its prosody can influence the learner's perception of the material. However, it is uncertain if a voice from people of different ages—such as that of an adult or child—have any impact on the learner. Thus, more and deeper investigations need to be conducted in this area of study. For this study, we hypothesize that a social robot's voice has an effect on a child's learning outcomes and motivation.

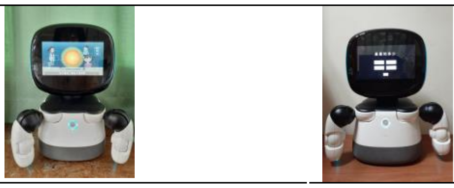
3. METHOD

3.1 Participants and Design

A total of 46 children from the same rural primary school in Chiayi, Taiwan participated in this study. Aged between 11 and 12 years old (M = 11.87, SD = .39), 22 (48.9%) of the children were male and 24 (51.1%) were female. All the children used computers regularly both at home and at school. They all had computer lessons at least twice per week, played computer games, and used the internet. All the participants had a similar level of understanding of the lesson topic (astronomy) to be given by the social robot. The experiment used a within-subject design to test the effects of a social robot's voice on the children's learning outcome and motivation. The experiment looked at both adult and children voices, with the two dependent variables being learning motivation and learning outcome.

3.2 Materials

The experiment used Kebbi social robots manufactured by NUWA Robotics specifically designed for children. The Kebbi social robots were used in the experiment to show the children videos of animated learning material and to conduct the interactive test. The narrations were done by a professional female child voice actor, as well as a professional female adult voice actor who was hired through a voice acting company. Three specialists designed the learning material and test questions for the experiment. Two of the specialists were school teachers, each with over 5 years of teaching experience, while the other specialist was an interaction designer with over 10 years of work experience. The specialists created two sets of learning material, both of which were animated videos suitable for children between the ages of 11 and 12. The specialists also checked and verified that the content and questions in the two sets were at a similar level of difficulty. The three specialists eventually settled on the theme “Understanding the Stars: An Intro to Astronomy” for the media content. The video was divided into three sections. Following each section, the participants needed to touch the social robot monitor to answer five questions. The test was formatted into multiple choice questions with 4 possible answers, and totaled 15 questions over the entire test.



(a) The learning material on the social robot (b) A question with four possible answers

Figure 1. Social robot used in the experiment

3.3 Procedure

On arrival at the lab, an experimenter collected participants’ consent forms signed in advance by their parents or guardians. Each participant was given their own social robot and worked alone. During the experiment, one participant was in the classroom at a time with one experimenter. The experimenter gave oral instructions and demonstrated how to operate the social robot. Participants were told they could interact with the social robot freely for three minutes to familiarize themselves with the social robots. Then, each participant started with a practice round. The practice round involved the children viewing the learning material on the social robot (Figure 1a) and answering two multiple-choice questions about each one.

After the practice round, participants could press “Start” to begin their experiment. Each child viewed the learning material on the social robot, and the material divided into three parts of approximately equal length. When finished viewing each part of video, participants answered multiple-choice questions (Figure 1b) that were displayed randomly. Participants needed to touch the screen on the robot to answer the questions. The participants viewed the material spoken in both an adult’s voice and a child’s voice. The order to which they viewed the two types of material was random. The children were told they could take as long as they wanted to answer the questions and should click the 'Done' button on the social robot when they finished each video. Upon finishing, they had to complete a motivation questionnaire. No child showed signs of fatigue during the experiment. Before leaving the room, each child filled out a demographic form and received a gift of stationary.

4. RESULTS

A paired sample t-test was carried out to evaluate the social robot’s two different voices with respect to the mean scores. The mean and standard deviation of the scores obtained in the tests are displayed in Table 1. It was shown a social robot’s voice did not have a significant effect on the learning outcome, t (45) = 1.230, p = .225. However, the results revealed the two different voices from the social robot did result in significantly different scores for the children’s learning motivation, t (45) = 2.746, p = .009. With regards to how the children rating their own motivation, they preferred the social robot with a child’s voice over the adult’s voice.

Table 1. The means and standard deviations for the learning outcome and motivation scores

	M	*SD*
Learning outcome		
children’ voice	54.57	8.03
adults’ voice	52.66	7.00
Learning motivation		
children’ voice	3.42	.624
adults’ voice	3.13	.511

5. CONCLUSION

The experiment examined how a social robot's voice affects children's learning outcome and motivation. We tested social agency theory which claims that social cues within multimedia presentations can lead to deeper cognitive processing during learning. We used human voice on social robot as social cues, divided it into an adult or child as a subcategory.

The results showed the social robot's voice had no significant impact on the children's learning outcomes, yet did affect the children's learning motivation. The children preferred the social robot to have the voice of a child over that of an adult as shown by their higher learning motivation with the social robot using the child's voice.

This is in line with social agency theory which suggests social cues can lead to a better learning experience. These results are also in agreement with those obtained by Wang and Li (2019), where they discovered 11- to 12 year-old children preferred PAs that looked and sounded like a child over something like an adult tutor and or an animal.

This experiment clearly shows that a social robot's voice significantly impacts children's learning motivation, and that 11- to 12-year-old children in particular prefer a child's voice over an adult. This research should be useful to educators, instructional designers, and social robot developers who create learning material for children. However, this experiment only focused on the voice of social robot. This is the limitation of this research. Further research is needed to determine how tone of voice (such as accent, gender, etc.) affects a child's learning when interacting with a social robot.

ACKNOWLEDGEMENT

This experiment was kindly supported by a grant from the National Science Council of Taiwan. (Contract number: NSC110-2410-H-415-022).

REFERENCES

Atkinson, R.K., Mayer, R.E. and Merrill, M.M. (2005). "Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology, 30*(1), 117–139.

Craig, S.D. and Schroeder, N.L. (2017). Reconsidering the voice effect when learning from a virtual human", *Computers & Education*, *114*, 193–205.

Davis, R.O., Vincent, J. and Park, T.J. (2019). Reconsidering the Voice Principle with Non-native Language Speakers", *Computers & Education*, Vol. 103605.

Mayer, R. E. (2014). *14 principles based on social cues in multimedia learning: Personalization voice image and embodiment principles.* The Cambridge handbook of multimedia learning.

Mayer, R. E., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology. Applied, 18*(3), 239–252. http://dx.doi.org/10.1037/a0028616

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95*(2), 419.

Tamagawa, R., Watson, C. I., Kuo, I. H., MacDonald, B. A., & Broadbent, E. (2011). The effects of synthesized voice accents on user perceptions of robots. *International Journal of Social Robotics, 3*, 253–262. https://doi.org/10.1007/s12369-011-0100-4.

Wang, H. F. and Li, Y. F. (2019). The impact of pedagogical agent on children's learning performance and intrinsic motivation. *Journal of Liberal Arts and Social Sciences, 15*(3), 257-281. (in Chinese)