Ms. Robot Will Be Teaching You: Robot Lecturers in Four Modes of Automated Remote Instruction

Jamy Li, Wendy Ju
Center for Design Research, Stanford University, Stanford, CA, USA
{jamy, wendyju} @stanford.edu

Abstract

Methods and materials are described for employing a human-shaped robot as a lecturer in automated remote instruction. Video segments from the stimuli of a 2 (participant substrate: VR or non-VR) x 2 (robot embodiment: copresent or screen) balanced between-participants experiment are provided. In each condition, a robot delivers the content for a lecture on the nutrition of carbohydrates. The robot uses identical speech and body movement while the same set of slides plays on an adjacent computer, thereby controlling for such factors as educational content, robot appearance and robot size. The experiment employs Aldebaran Robotics' 25-degrees-of-freedom Nao as the robot and the Oculus Rift as the immersive VR system. The lecture speech and slides were obtained with permission from a Mandarin Chinese-language online course and translated into English. The set-up for different delivery modes for automated remote instruction are illustrated using a robot delivering foreign language online content. These methods support the design and evaluation of robots that perform the role of lecturer.

Introduction

Experimental stimuli were prepared to assess human-robot interaction in which a learner views a lecture delivered by a human-shaped robot. Stimuli addressed four modes of delivery: immersive virtual reality with a 3D model of a robot (VR), immersive virtual reality with a video of a robot (VR*), a collocated robot without virtual reality (CR) and a video of a 3D model of a robot (CR*). Robots can be effective educational aids (Howley, Kanda, Hayashi, and Rosé 2014; Shirouzu and Miyake 2013), particularly as a means of remotely delivering foreign language content that has been translated. The substrate (i.e., immersive virtual reality or unmediated reality) and embodiment (virtual model or video) of these robots can be varied (Li 2015; Milgram, Takemura, Utsumi, and Kishino 1995). Methods and materials are described.

Robot Voice

Mandarin-language lecture slides and videos were obtained for a single 17-minute lecture from the edX course entitled, "PekingX: 18000123x Nutrition Around You," with permission from the instructor. Speech was transcribed and translated by a research assistant fluent in both English and Mandarin Chinese. Audio files were created by inputting transcribed text into http://www.fromtexttospeech.com/. We selected the "Heather" voice and slow speed to match the human lecturer's gender and speech tempo. During quality assessment of audio files, translated text was modified to serve as better input to the TTS engine (e.g., "let's" to "letz") and newlines were added to approximate the pausing in the original lecture. Timing and speed of the text-to-speech output were then edited in Audacity™ to better match the human lecturer's prosody.

Robot Movement

Gestures were created using the playback mode on Aldebaran Robotics' Nao. A member of the research team separately modeled movements for the head, right arm and left arm of the robot based on visual inspection of the human lecturer's movement, then combined these into a single animation sequence.

Creation of Stimuli in Four Modes

Stimuli for four modes of automated remote instruction were created: immersive virtual-reality 3D model of a robot (VR), immersive virtual-reality video of a robot (VR*), collocated robot not delivered using virtual reality (CR) and animated robot not delivered using virtual reality (CR*). Experimental stimuli varied for each of the four conditions but were developed to control for the appearance of the environment from the user's perspective. For the VR condition, an independent developer constructed a virtual version of the experiment room and wrote an interpreter to translate movements from the robot's behavior file into a virtual-reality 3D model of the robot; these were made into a scene displayed on the Oculus Rift™ virtual reality headset. For VR*, the same scene was used as in VR, except the robot was filmed and the resulting video displayed on a computer screen in the VR environment. For CR, the robot's behavior was executed on the robot itself while in front of the user (no virtual-mediation was used). For CR*, the robot was replaced by a computer screen displaying a 3D model of the robot (taken from the VR condition). All conditions featured a 27-inch iMac computer displaying slides with a keyboard and mouse (logos were removed in VR conditions).

Future Work

We are currently running this experiment with 40 participants from the student population at Stanford University. We hope this work illustrates how robots can be used in automated remote instruction for foreign language content. We also hope our results will demonstrate how exploratory assessment can be used to evaluate user perception of design variations in a robot that mediates face-to-face communication in virtual and real environments.

Acknowledgment

We thank Xu Yajun, Pamela Hinds, Nik Martelaro, Paulo Muggler Moreira and Kellie Sun for assistance with experimental materials and setup. This research was funded by the Institute for Research in the Social Sciences (IRiSS).

References

Howley, I.; Kanda, T.; Hayashi, K.; and Rosé, C. 2014. Effects of Social Presence and Social Role on Help-Seeking and Learning. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*, 415–422. New York, New York: ACM Press.

Li, J. 2015. The Benefit of Being Physically Present: A Survey of Experimental Works Comparing Copresent Robots, Telepresent Robots and Virtual Agents. *International Journal of Human-Computer Studies* 77: 23–37.

Milgram, P.; Takemura, H.; Utsumi, A.; and Kishino, F. 1995. Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum. In Proceedings SPIE Conference on Telemanipulator and Telepresence Technologies, 2351(34), 282–292. Boston, MA: International Society for Optics and Photonics.

Shirouzu, H.; and Miyake, N. 2013. Effects of Robots' Revoicing on Preparation for Future Learning. In Proceedings of Computer-Supported Collaborative Learning. Gothenburg, Sweden: University of Gothenburg.