

Robot Assistants for Perimetry: A Study of Patient Experience and Performance

Allison M. McKendrick¹, Astrid Zeman^{1,2}, Ping Liu¹, Dilek Aktepe¹, Illham Aden¹, Daisy Bhagat¹, Kieren Do¹, Huy D. Nguyen¹, and Andrew Turpin³

¹ Department of Optometry and Vision Sciences, The University of Melbourne, Parkville, VIC, Australia

² Brain and Cognition Department, KU Leuven, Leuven, Belgium

³ School of Computing and Information Systems, The University of Melbourne, Parkville, VIC, Australia

Correspondence: Allison M. McKendrick, Department of Optometry & Vision Sciences, The University of Melbourne, Parkville 3010, VIC, Australia. e-mail: allisonm@unimelb.edu.au

Received: 17 September 2018

Accepted: 6 May 2019

Published: 28 June 2019

Keywords: perimetry; robot; feedback; social robot; healthcare robot

Citation: McKendrick AM, Zeman A, Liu P, Aktepe D, Aden I, Bhagat D, Do K, Nguyen HD, Turpin A. Robot assistants for perimetry: a study of patient experience and performance. *Trans Vis Sci Tech.* 2019;8(3):59, <https://doi.org/10.1167/tvst.8.3.59>

Copyright 2019 The Authors

Purpose: People enjoy supervision during visual field assessment, although resource demands often make this difficult. We evaluated outcomes and subjective experience of methods of receiving feedback during perimetry, with specific goals to compare a humanoid robot to a computerized voice in participants with minimal prior perimetric experience. Human feedback and no feedback also were compared.

Methods: Twenty-two younger (aged 21–31 years) and 18 older (aged 52–76 years) adults participated. Visual field tests were conducted using an Octopus 900, controlled with the Open Perimetry Interface. Participants underwent four tests with the following feedback conditions: (1) human, (2) humanoid robot, (3) computer speaker, and (4) no feedback, in random order. Feedback rules for the speaker and robot were identical, with the difference being a social interaction with the robot before the test. Quantitative perimetric performance compared mean sensitivity (dB), fixation losses, and false-positives. Subjective experience was collected via survey.

Results: There was no significant effect of feedback type on the quantitative measures. For younger adults, the human and robot were preferred to the computer speaker ($P < 0.01$). For older adults, the experience rating was similar for the speaker and robot. No feedback was the least preferred option of 77% younger and 50% older adults.

Conclusions: During perimetry, a social robot was preferred to a computer speaker providing the same feedback, despite the robot not being visible during the test. Making visual field testing more enjoyable for patients and operators may improve compliance and attitude to perimetry, leading to improved clinical outcomes.

Translational Relevance: Our data suggest that humanoid robots can replace some aspects of human interaction during perimetry and are preferable to receiving no human feedback.

Introduction

Visual field assessment is a longstanding and ongoing cornerstone of the diagnosis and management of glaucoma.^{1,2} Unfortunately, patient and operator experience of visual field testing can be negative. Threshold visual field testing typically requires approximately 5 to 10 minutes to perform per eye, is conducted in dimly lit conditions, and can be stressful for patients.³ Indeed, one study identified visual field testing as the least preferred test used in

glaucoma care from the patients' perspective.⁴ Focus group discussion of patient experience of visual field testing has revealed that people appreciate having a staff member present throughout the test.³ However, for the operator, sitting in a dark room invigilating visual field tests is repetitive and not particularly engaging. Despite patient preference for staff support during visual field assessment, in many clinical settings resource constraints do not permit this to occur.

Data are limited regarding whether supervision of

visual field testing by staff is strictly necessary. Previous research demonstrates that reliable test results often can be obtained in the absence of staff being present throughout the test.^{5,6} Considering quantitative results alone potentially neglects important aspects of patient experience that may affect their desire to attend for appointments, and their engagement with their clinical care. There is evidence that making clinical experiences more enjoyable increases patient well-being and engagement in their healthcare; for example, the well-studied hospital clown projects.⁷ However, increased human interaction is expensive to resource, and can have limitations in environments where language diversity creates additional communication barriers. In the context of perimetry, where testing is performed in a controlled and relatively structured environment, perhaps there is a role for technology to provide feedback during the testing.

We assessed participant acceptance and preference for two methods of providing such feedback. The first was via simple voice commands from an audio speaker (a disembodied voice), and the second via a physically present, humanoid social robot. Our social robot was a Nao robot (Softbank Robotics, Tokyo, Japan), which measures 58 cm in height, weighs 4.3 kg, and has 25 degrees of freedom of movement. Previous research on social robotics demonstrates that the physical presence of a humanoid robot results in greater persuasive ability and more positive responses from people when compared to video of the same robot or other virtual agents.⁸ Indeed, humanoid social robots can invoke social responses from people. For example, a recent study showed that people hesitate to turn off a humanoid robot that is pleading to be left turned on.⁹ We also included two additional conditions that are typical of clinical practice: (1) feedback during the test by a trained human operator and (2) no feedback once the test had commenced, including the operator leaving the room. We predicted that people would least prefer the condition where there was no feedback during the test; however, the main aim of the experiment was to determine whether feedback via a humanoid robot connoted advantages over disembodied, voice-alone feedback. While previous social robotic research suggests a difference in response to these two forms of nonhuman feedback, the fact that during a visual field test the participant can neither look directly at, nor engage directly with the robot might create a situation where the difference between these two forms of agent are minimized. We did not aim to determine the best type of feedback for perimetry, nor

did we aim to optimize the feedback. Rather, the experiment was principally designed to compare the social robot to the computer speaker.

Because experience and acceptance of technology may differ with age, we tested two groups of participants: younger (aged <32 years) and older (aged >50 years) adults. People who regularly undergo perimetry as part of their clinical care may have preexisting positive or negative biases associated with their previous experience of perimetry. Consequently, we recruited participants with minimal perimetric experience to avoid such confounds.

Methods

Participants

Twenty-two young (median age, 22; range, 21–31 years) and 18 older (median age, 63; range, 52–76 years) adults participated in the study. This sample size was similar to that used in research in other domains than eye care to compare human responses to humanoid versus nonhumanoid robotics.⁸ More formally, for a paired *t*-test analysis between the speaker and the humanoid robot, assuming a difference in experience rating between the robot and speaker of 0.5 (Likert scale of 0–5, methods described in detail below) and a standard deviation of 0.75, a sample size of 20 was required to achieve a power of 80% with a 2-sided level of significance between groups.

To minimize bias arising from previous expectations and experience of visual field testing, participants were recruited with minimal prior visual field testing experience. All of the younger participants were naive to perimetry. Some older participants had performed perimetry previously as part of routine ophthalmic examinations, but none was regularly required to perform perimetry as part of their clinical care. After all procedures were explained, informed written consent was obtained from all participants. The study adhered to the tenets of the Declaration of Helsinki and was approved by The University of Melbourne Human Research Ethics Committee.

All participants had normal ocular health for age and had best corrected visual acuity of 6/7.5 or better in the tested eye.

Testing Conditions and Equipment

Each participant underwent perimetric testing under four conditions: (1) No feedback – no feedback was provided in any form during the testing, (2)

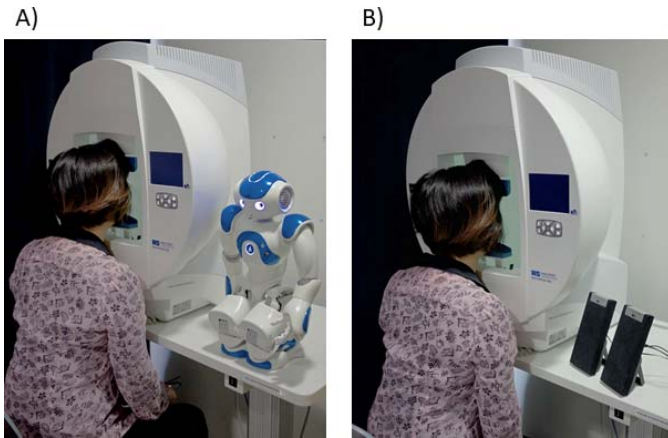


Figure 1. Illustration of the (a) Nao robot assistant and (b) computer speaker assistant used in the experiments. Note, during testing, neither could be seen by the participant, who was positioned on the head-chin rest of the perimeter.

Speaker – instructions and ongoing feedback were provided via a speaker without the presence of the human operator, (3) Robot – a humanoid robot was used to give instructions and ongoing feedback without the presence of the human operator during the test, and (4) Optometrist – an optometrist gave ongoing feedback to the participant during the testing.

An Octopus 900 perimeter (Haag-Streit Diagnostics, Koeniz, Switzerland) was used for the perimetric testing (Size III, white luminance increment targets, 200 ms) using the G-pattern (59 locations). The perimeter was controlled via an external computer using the Open Perimetry Interface.¹⁰ Thresholding was conducted using a ZEST procedure^{11,12} that terminated after four presentations. Terminating after a fixed number of trials enabled the test duration to be similar for all tests, and the number of fixation loss and false-positive checks to be consistent across conditions. Fixation losses were determined using the Heijl-Krakau blind spot method,¹³ and false-positive checks by interspersing trials with a stimulus presentation of 50 dB. Three quantitative performance measurements were collected for each test condition: (1) the measured sensitivity threshold in decibels (dB) for each tested location, (2) the false-positive rate (out of 23), and (3) fixation loss rate (out of 21 fixation checks).

A Nao robot (Softbank Robotics) was used to provide feedback in the robot condition (Fig. 1). The Nao has a range of sensors and actuators that enable a range of human type movements. The Nao has two loudspeakers, four microphones, two cameras, a

Table 1. Feedback Phrases Used in the Experiment

	Positive Phrases	Negative Phrases
1	Good work	Please pay attention
2	Excellent	Please focus
3	Well done	Keep looking at the center light
4	Good job	Remember to keep your eyes open
5	Great	Stay focused on the center light
6	Fantastic	Try not to fall asleep
7	Keep up the good work	Are you looking in the center?
8	Brilliant	
9	You're doing a great job	
10	Thanks for paying attention	

gyroscope, an accelerometer, and range sensors (two infrared [IR] and two sonar). The robot has an embedded computational core and connects externally via IEEE 802.11g WiFi or Ethernet. Choreographer software (Aldebaran Robotics, Paris, France) was used to program an introduction demonstration of the robot showcasing its functions, such as waving, providing a personalized greeting to the participants, and doing tai-chi. The voice for the speaker condition was created using Acapela Box (Acapela Group, Mons, Belgium), which is the same engine used to generate speech on the Nao robot. We selected the “Laura” voice among the English (United States) options, with a speech rate of +18 and voice shaping at 0. These settings provided a voice that was qualitatively similar to the Nao, and subjectively equally as pleasant to listen to, but that could be distinguished as a different voice.

During the robot and speaker conditions, verbal feedback responses were selected from a battery of possible responses. Custom software written in the R programming language¹⁴ was used to control feedback responses, selected according to progress during the perimetric test. Table 1 lists the possible feedback phrases that were used for the robot and speaker conditions in the same fashion. Triggers for verbal feedback included fixation loss and false-positive catch trials, in addition to progress throughout the test (a third, halfway, and two-thirds of the way). The same script was used as the basis for feedback from the optometrist; however, to better represent a typical

Table 2. Experience Questionnaire Conducted in Between Each Test Condition, Immediately After Testing

Statements for Experience Rating Immediately After Each Test Condition

1	I found the overall experience to be enjoyable.
2	I was distracted during the test.
3	I was able to maintain concentration during the test.
4	I felt sleepy or bored during the test.
5	The level of feedback provided was helpful.
6	I felt uneasy during the test.
7	I was clear what was expected of me during the test.

clinical environment, some variability in feedback was allowed. The only difference between the speaker and robot condition was a familiarization sequence that introduced the participants to the more humanoid functions of the robot before the perimetric testing. Throughout the test itself, neither the robot nor the speakers could be seen by the participant (see Fig. 1).

Participants were allocated to the four conditions using a randomized counterbalance to minimize the possible effects of learning and fatigue on the test performance and user experience. For all conditions, participants were provided with initial instructions from the optometrist and set up on the machine in the correct head and postural position. Because most participants were perimetrically naïve, a short introduction lasting approximately 1 to 2 minutes was provided before the first formal test condition, to help participants understand how to perform the test and to answer any initial questions before formal data collection.

There was a rest period of approximately 5 to 10 minutes between each visual field test while surveys were completed and the next condition was set up by the experimenter. Only one eye was tested. The total duration of actual visual field testing per participant was approximately 30 minutes (split into four tests with rest periods in between).

Questionnaires and Qualitative Response

After completing the perimetry test under each condition, participants were asked to rate their experience for a series of seven statements. Participants rated the extent to which they agreed or disagreed with each statement using a Likert scale. The ends of the scale for each item were labeled

“strongly disagree” and “strongly agree” and given the values of 1 and 5, respectively. Participants could mark anywhere along the scale. Reversed statements were included to minimize bias from the users’ answers. The seven statements included in the scale are provided in Table 2.

After completing all four perimetric tests, participants filled out a final user experience questionnaire in which they were asked to rank the test conditions according to their preference, describe what they liked and disliked about each condition, provide suggestions for improvements, and comment on any additional thoughts on the study.

Results

Quantitative Perimetric Outcomes

Figure 2 compares the quantitative data collected under each of the conditions for the younger (Fig. 2, left) and older (Fig. 2, right) groups. Sensitivity (Fig. 2, top) was calculated as an average across all 59 locations for each participant. All participants had intact, normal visual fields. A mixed, 2×4 , factorial analysis of variance (ANOVA) was used to test the effects of age group (older versus younger), test condition (no feedback, speaker, robot, and optometrist), and the interaction effects between the two factors on the average sensitivity. As expected, there was a significant main effect of age group, $F(1,152) = 22.77$, $P < 0.001$, partial $\eta^2 = 0.13$, with the older participants having lower sensitivity thresholds (comparison of data from Fig. 2, top right and top left). There was no significant main effect of test condition.

Figure 2 also shows the reliability indices returned for the two age groups, under the four test conditions. A mixed, 2×4 , robust ANOVA comparing the main effects of age group, test condition, and the interaction effects between the two revealed no main effect of age group, test condition, or interaction between the two for either false-positive responses or fixation loss.

User Experience Ratings Collected After Each Test

For the experience rating statements collected after each test (see Table 2 for the list of statements), there was no significant difference between groups, conditions, or significant interaction between group and condition, for items 1, 2, 3, and 6. For the item 4, “I felt sleepy or bored during the test,” younger adults

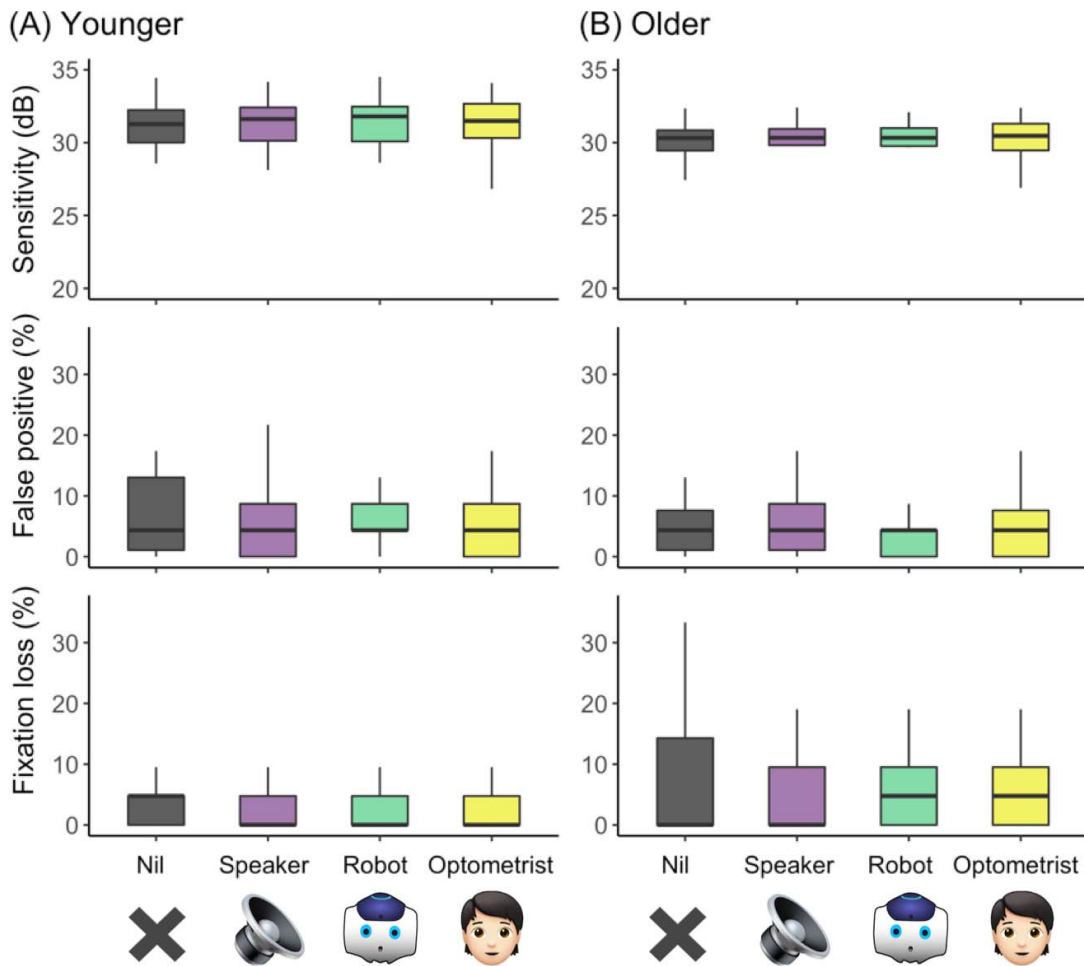


Figure 2. Quantitative output from the perimeter for (A) the younger and (B) the older groups. The boxes represent first (25th), second (median), and third (75th) quantiles. The upper and lower whisker extends from the upper and lower hinge respective to the largest value no further than $1.5 \times$ interquartile range from the hinge.

rated the question significantly higher than older adults (main effect of age group: $F(1,152) = 5.51$, $P = 0.02$, partial $\eta^2 = 0.03$). For item 5, “The level of feedback provided was helpful,” there was a significant interaction between the effects of age group and test condition ($F[3,152] = 3.72$, $P = 0.01$, partial $\eta^2 = 0.07$). The Tukey HSD post hoc tests revealed that the young adult participants felt the level of feedback provided in the speaker ($P = 0.04$), robot ($P = 0.008$), and optometrist ($P = 0.0003$) was more helpful than in the no feedback condition, whereas the older adults were ambivalent to the feedback method (including no feedback). For test item 7, “I was clear what was expected of me during the test,” there was also a main effect of age ($F[1,152] = 6.84$, $P = 0.009$, partial $\eta^2 = 0.04$) with the younger adults rating the question significantly higher (better clarity) than the older adults.

Exit User Experience Questionnaire

Figures 3 and 4 show the participant rating of test experience and test preference obtained from the final questionnaire conducted at the end of all tests, for the younger and older participants respectively. The left parts of Figures 3 and 4, show the distribution of responses to the question: “On a scale of 1 to 5 (1 being bad, 5 being good) how would you rate your experience with the speaker/robot/human assistant?” The right side shows percentage of participants who rated the various conditions first (Figs. 3B, 4B) or last (Figs. 3C, 4C) preference.

For the younger observers, there was a significant difference in the ranking of experience among the three assisted conditions (Kruskal-Wallis 1-way ANOVA on ranks: $H = 13.88$, $df = 2$, $P < 0.001$). Post hoc analysis revealed no significant difference in median experience between the robot and optometrist

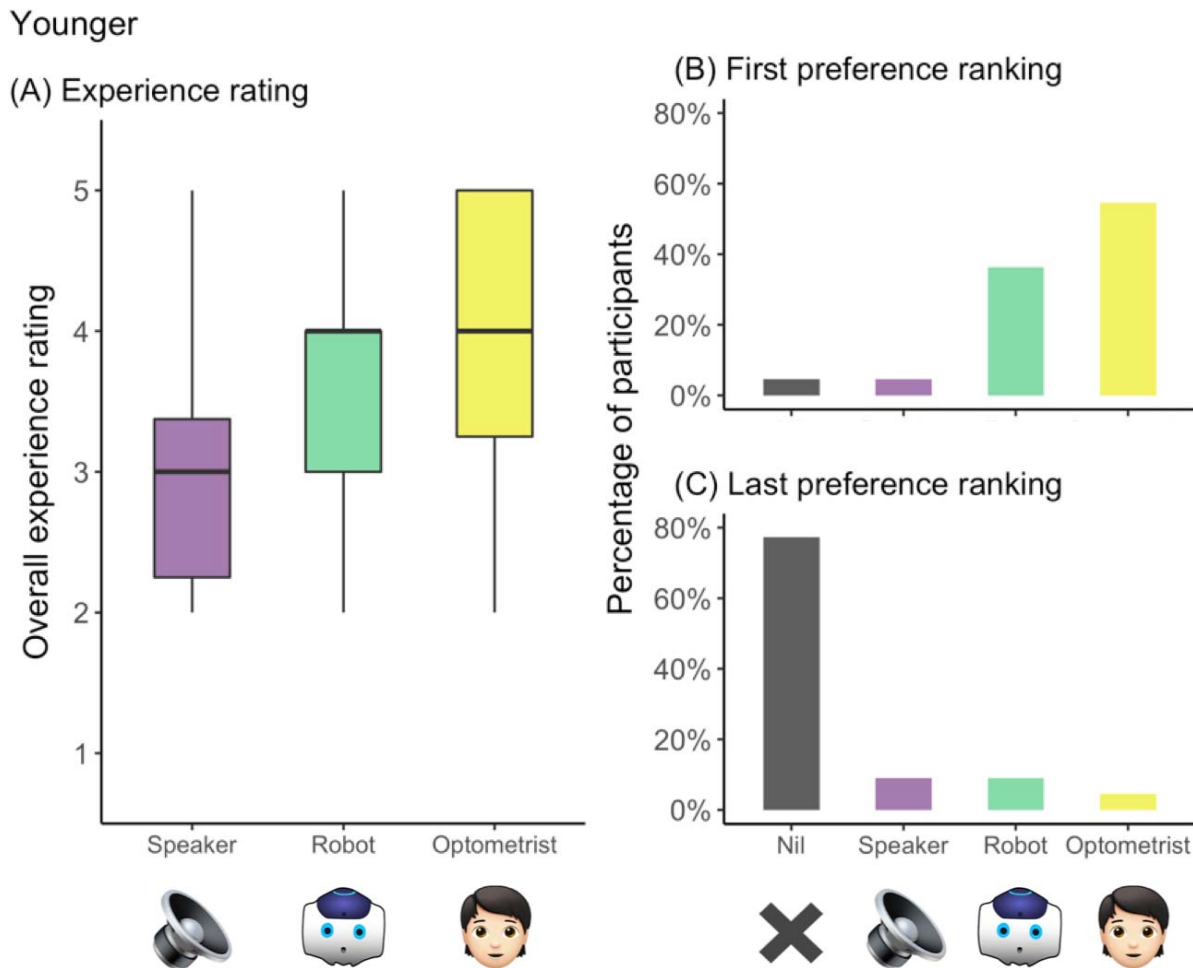


Figure 3. Younger adult exit questionnaire data. (A) distribution of overall experience ratings for each of the perimetry assistant conditions. Participants were asked to rate their experience on a scale of 1 (bad) to 5 (good). (B) the percentage of participants that chose each of the tests as their first preference; (C) the percentage of participants who chose each of the tests as their last preference.

($P = 0.23$), but a more positive experience for the optometrist relative to the speaker ($P = 0.002$), and the robot relative to the speaker ($P = 0.001$).

For the older participants, the data were sufficiently normally distributed to meet assumptions of a parametric repeated measures ANOVA, which revealed no significant difference in mean experience rating between the test conditions ($F[2,17] = 3.0$, $P = 0.06$).

Open-Ended Comments

To explore further the subjective response to the robot compared to the speaker, we compared qualitative open-ended responses to the question “How do the two electronic assistants compare?” Within the younger group, 14 statements were more positive towards the robot than the speaker, four were neutral, and three positive towards the speaker. For

the older adult group, consistent with the preference ratings shown in Figure 3, statements were more balanced with six neutral, eight positive towards the robot, and five favoring the speaker. Samples of statements are provided in Table 3 for younger (Y) and older (O) participants. All statements are provided in Appendix A, including responses to questions regarding the most and least preferred aspects of each of the supervisor types (human, robot, speaker). Many comments (either positive or negative) were related to the perceived quality of the information content and delivery provided by the robot/speaker. Note the experimental design was such that the feedback during the test was the same for the robot and speaker conditions, and the electronic voices highly similar, to ensure that the main difference was the humanoid features of the robot.

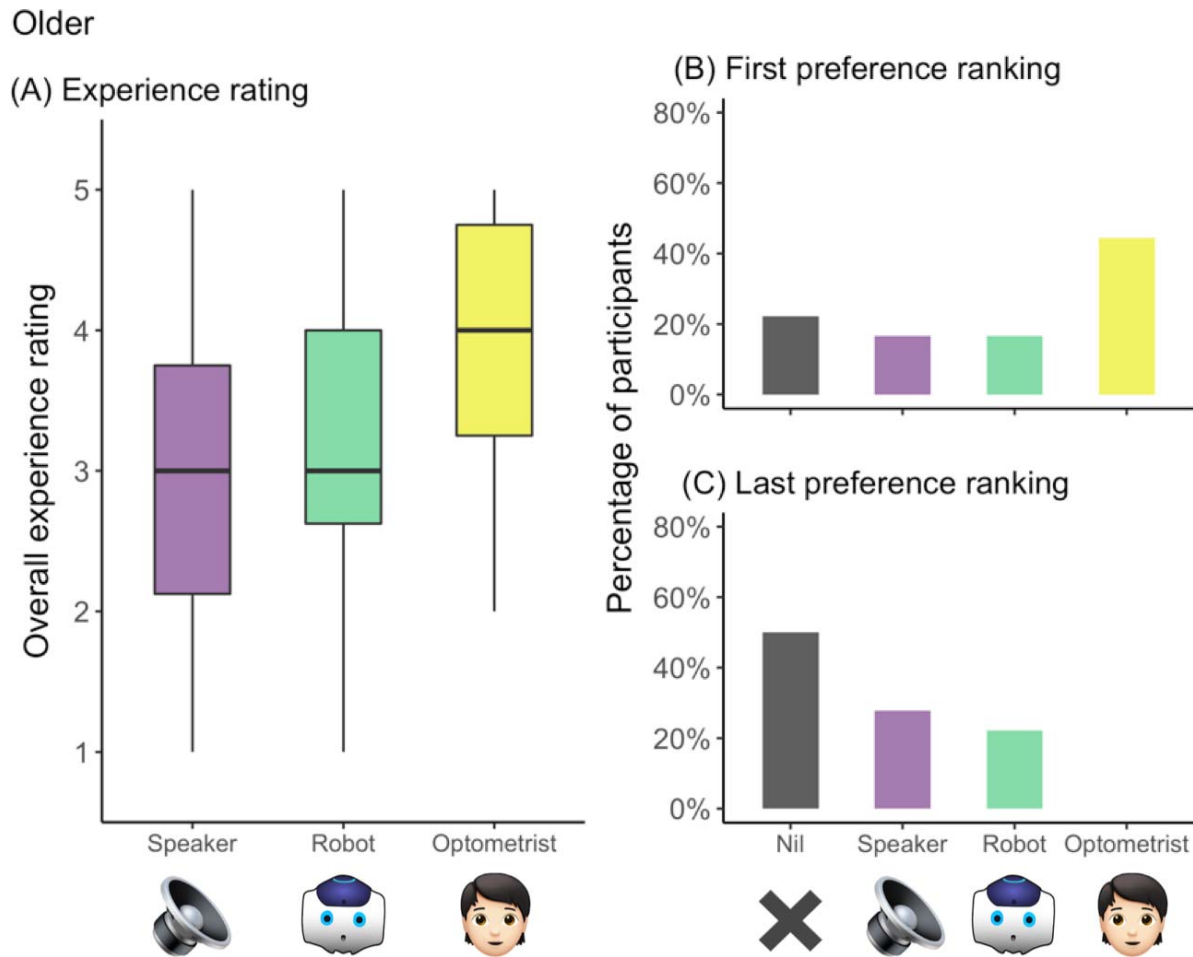


Figure 4. Older adult exit questionnaire data. (A) distribution of overall experience ratings for each of the perimetry assistant conditions. Participants were asked to rate their experience on a scale of 1 (bad) to 5 (good). (B) The percentage of participants who chose each of the tests as their first preference. (C) The percentage of participants who chose each of the tests as their last preference.

Discussion

Given the amount of time and resourcing devoted to visual field testing from healthcare provider and patient standpoints, it is important to create a testing environment that encourages the derivation of as useful information as possible from the testing. It is increasingly recognized that patient-centric engagement in healthcare improves adherence and positive outcomes. Consequently, processes that improve patient experience of visual field assessment also may derive overall benefits to their glaucoma and other ophthalmic management. The challenge is to derive such processes within the context of limited healthcare resources. Our data demonstrate that humanoid robots can provide feedback during visual field testing that is well accepted, and indeed enjoyed, by many participants. While feedback could be

provided by a computer voice alone, consistent with previous research on social robotics, our data show that many people responded more positively to a humanoid social robot, despite the fact that the perimeter obscures that robot completely during testing.

A further potential advantage of automated assistants in healthcare is the ability to provide consistent instructions, potentially in multiple languages. Previous studies have shown that measured visual field sensitivity can vary markedly depending on the instructions provided.¹⁵ The provision of consistent instructions is more difficult in busy clinical environments with time pressures and multiple operators creating possible points of failure. An alternate approach is to provide patient training videos, which have been demonstrated to improve visual field reliability when taking the test for the first time.¹⁶ Recorded explanations and expectations of

Table 3. Sample Positive and Negative Open-Ended Responses to the Question: “How Do the Two Electronic Assistants Compare?”**Robot preferred**

- “Very similar but the Robot was better as I felt I was actually being supervised (Y)”
- “Robot was more cheerful than the computer (Y)”
- “Robot assistant was amusing and made the experience more enjoyable (Y)”
- “The Robot is more engaging, makes me more inclined to focus on the test, gives better feedback (Y)”
- “The Robot was a lot more encouraging than the computer (Y)”
- “I think the Robot put me in better mood, hence, the whole experience is more pleasant (Y)”
- “The robot felt better because it made me smile, made a practical test seem more enjoyable (O)”
- “Robot seems to have more personal touch than computer (O)”
- “Robot different. Speaker overbearing. (O)”

Speaker preferred

- “I have similar opinions for both of them, except that the Robot feels strange - having a small object looking over you (Y)”
- “Electronic device was better, less distracting (Y)”
- “Felt that computer was more reliable and on task; robot was fun (O)”
- “I am used to hearing electronic device words and can more easily ignore them if I want than for the robot that I am less familiar with (O)”
- “Sam was less effective for me (O)”^a

^a The robot was introduced as “Sam.”
Y, younger patients; O, older patients.

perimetry provided in the patients’ native languages also are beneficial.¹⁷

In this study, we did not use the robot (or computer speaker) to initially explain the test because our experimental design involving four tests in a row (with questionnaires interspersed) would have created significant repetition. A more realistic future clinical scenario might involve initial instruction and preparation for the test in communication with the robot, perhaps alignment on the machine by a human operator, and then invigilation and feedback during the test by the robot. Some changes to current perimeters would be required to fully enable robotic set-up on the machines, and this may not be desirable

because the human interaction when introducing the test enables a human touch-point for patients that is likely to be of value. The robot assistant could be used to ensure full instructions and feedback are given to all patients during the test, freeing up time for the human operators to engage in more personalized discussion and care before and after the assessment.

Our younger adult group was more positively disposed to the robot than the older adults. Indeed, 36% of younger adults actually rated the robot as their first preference, ahead of a human operator. We did not collect further demographic data on our participants, such as the use and acceptance of technology in their daily lives, educational status, and cultural background, all of which may affect their acceptance of robot assistants.^{18,19} A better understanding of the psychology and influencing factors on robot acceptance could be applied on an individual basis to preferentially allocate human versus robot assistance when both are available.

Many of our older adults had undergone perimetry at some point previously as part of their routine ophthalmic care. In contrast, none of the younger adults had previously undergone a visual field test, so had no prior expectations. Within the older group, there was no obvious effect of past perimetric experience influencing responses. Eight older participants had no prior perimetric experience and these were fairly balanced with their choice of test first preference (no feedback = two, speaker = two, robot = one, optometrist = three). Two older adults had substantial prior experience of visual field assessment (self-reported as approximately 10 previous visual field tests). One of these preferred receiving no feedback at all during the test, and the other preferred the presence of an optometrist throughout.

There are a number of limitations of our study, and additional research that would be required before implementation in clinical practice. The primary question addressed here was whether people engage more with a robot than identical voice-only feedback, even when they cannot see the robot after a brief introduction. Our results indicated this to be the case for many people. This is a human-computer interaction phenomenon that may inform future design of perimetric instruments and has broader significance beyond the ophthalmic perimetric situation. We made no attempt to design the optimal human, robot, or speaker feedback for perimetry, but this is an interesting area for further study, and requires different experiments from those within this study.

We did not test people with significant perimetric

experience, nor any glaucoma patients. It is not clear that there is a hypothesis-driven reason to suggest that patients with glaucoma should have a different response to the robot/speaker comparison than their age-matched peers due to glaucoma per se. They may have different biases in their responses to humans or to no feedback due to prior interactions in clinics, and different levels of either anxiety or complacency with visual field assessment, but this is a separate issue that requires an alternate study design to explore. Many current visual field testing algorithms in clinical practice result in longer test times in those with reduced sensitivity. In our experiments, we wished to avoid any differential test duration lengths between the older and younger participants, so fixed the number of trials to be identical for all people in the experiment. Consequently, it is unclear whether longer test durations may further influence the preference of observers for specific perimetric assistants. Although we did not test people with manifest visual field loss, it is worth noting that arguably the majority of patients undergoing visual field assessment across the eye care sector also do not have visual field loss or only early damage, and possibly have minimal perimetric experience (i.e., the people that we included in this study). For example, in the Australian context where data are readily available, 60% of visual field Medicare items are claimed through optometry and 40% from ophthalmology (available in the public domain at: http://medicarestatistics.humanservices.gov.au/statistics/mbs_item.jsp). While data on the number of distinct individuals tested are not revealed in these government statistics, restrictions on the number of times that these items can be billed annually suggests that the division in terms of individuals tested would be even further skewed towards optometry.

Except for four older participants, our participants displayed a strong preference to receive feedback during visual field testing rather than no feedback. Quantitative perimetric outcomes were not altered with feedback in our participants who all had healthy normal vision for age; however, their attitude and enjoyment of the test was improved. The real impact of patient attitude to perimetry requires further study; however, it is expected that positive clinical experiences will positively affect their adherence and attitude to their clinical care. Technologic advances in social robotics create new opportunities within healthcare settings. It is plausible that some of the responses, positive and negative, to our humanoid robot may be influenced by novelty, and perhaps may

shift over time as people become more accustomed to social robotics in day-to-day scenarios. Nevertheless, our research demonstrated that social robots can be well accepted to assist with visual field testing. Instruction and invigilation during perimetry is important to patients and should not be overlooked in the various research endeavors directed to improved measurement of visual performance.

Acknowledgments

Supported by the Australian Research Council Linkage Project: LP150100815 (to authors AMM and AT).

Disclosure: **A.M. McKendrick**, Haag-Streit AG (R); **A. Zeman**, None; **P. Liu**, None; **D. Aktepe**, None; **I. Aden**, None; **D. Bhagat**, None; **K. Do**, None; **H.D. Nguyen**, None; **A. Turpin**, Haag-Streit AG (R)

References

1. Wu Z, Medeiros FA. Recent developments in visual field testing for glaucoma. *Curr Opin Ophthalmol*. 2018;29:141–146.
2. Camp AS, Weinreb RN. Will perimetry be performed to monitor glaucoma in 2025? *Ophthalmology*. 2017;124:S71–S75.
3. Glen FC, Baker H, Crabb DP. A qualitative investigation into patients' views on visual field testing for glaucoma monitoring. *BMJ Open*. 2014;2014:e003996.
4. Gardiner SK, Demirel S. Assessment of patient opinions of different clinical tests used in the management of glaucoma. *Ophthalmology*. 2008; 2008:2127–2131.
5. Van Coevorden RE, Mills RP, Chen YY, Barnebey HS. Continuous visual field test supervision may not always be necessary. *Ophthalmology*. 1999;106:178–181.
6. Johnson LN, Aminlari A, Sassani JW. Effect of intermittent versus continuous patient monitoring on reliability indices during automated perimetry. *Ophthalmology*. 1993;100:76–84.
7. Dionigi A, Canestrari C. Clowning in health care settings: the point of view of adults. *Eur J Psychol*. 2016;12:473–488.
8. Li J. The benefit of being physically present: a survey of experimental works comparing copre-

- sent robots, telepresent robots, and virtual agents. *Int J Human-Computer Studies*. 2015;77:23–37.
9. Horstmann AC, Bock N, Linhuber E, Szczuka JM, Straßmann C, Krämer NC. Do a robot's social skills and its objection discourage interactants from switching the robot off? *PLoS One*. 2018;13:e0201581.
 10. Turpin A, Artes PH, McKendrick AM. The Open Perimetry Interface: an enabling tool for clinical visual psychophysics. *J Vis*. 2012;12:11:1–5.
 11. King-Smith P, Grigsby S, Vingrys A, Benes S, Supowit A. Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation, and practical implementation. *Vision Res*. 1994;34:885–912.
 12. Turpin A, McKendrick AM, Johnson CA, Vingrys AJ. Development of efficient threshold strategies for frequency doubling technology perimetry using computer simulation. *Invest Ophthalmol Vis Sci*. 2002;43:322–331.
 13. Anderson DR, Patella VM. *Automated Static Perimetry*, 2nd ed. St. Louis, Missouri: Mosby; 1999.
 14. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria; 2013. Available at: URL <http://www.R-project.org/>.
 15. Kutzko KE, Brito CF, Wall M. Effect of instructions on conventional automated perimetry. *Invest Ophthalmol Vis Sci*. 2000;41:2006–2013.
 16. Sherafat H, Spry PG, Waldock A, Sparrow JM, Diamond JP. Effect of a patient training video on visual field test reliability. *Br J Ophthalmol*. 2003; 87:153–156.
 17. Neshet R, Ever-Hadani P, Epstein E, Stern Y, Assia E. Overcoming the language barrier in visual field testing. *J Glaucoma*. 2001;10:203–205.
 18. Broadbent E, Stafford R, MacDonald B. Acceptance of healthcare robots for the older population: review and future directions. *Int J Soc Robot*. 2009;1:319–330.
 19. Broadbent E. Interactions with robots: the truths was reveal about ourselves. *Annu Rev Psychol*. 2017;68:627–652.