People respond better to robots than computer tablets delivering healthcare instructions

Jordan A. Mann a, Bruce A. MacDonald b, I.-Han Kuo b, Xingyan Li b, Elizabeth Broadbent a,⇑

a Department of Psychological Medicine, The University of Auckland, New Zealand
b Department of Electrical and Computer Engineering, The University of Auckland, New Zealand

A R T I C L E   I N F O
Article history:
Available online 19 November 2014

Keywords:
Robot
Computer tablet
Healthcare behaviours
Relaxation
Attitude
Human–robot interaction

A B S T R A C T

The population of the world is ageing, particularly in developed countries. As the population’s age increases, the healthcare workforce is becoming progressively unable to meet the high healthcare demands of the elderly population. Increasingly, technology is being used to solve this dilemma. Using a sample from the general population (n = 65), this study examined how people interacted with either a robot or a tablet computer delivering healthcare instructions. During this interaction, the robot/tablet asked them several health-related questions, and to perform limited physical tests and a relaxation exercise. Results showed participants had more positive interactions with the robot compared to the computer tablet, including increased speech and positive emotion (smiling), and participation in the relaxation exercise. Further results showed the robot was rated higher on scales of trust, enjoyment, and desire for future interaction. This suggests that robots may offer benefits over and above computer tablets in delivering healthcare. These results further demonstrate that the physical nature of technology is important in determining responses to healthcare interactions.

1. Introduction

The double-edged sword of modern medicine is that as people are living longer, the emergence and prevalence of age-related diseases have increased (LaCroix, 2013). Thus, the use of health system resources has also increased. In the United States the cost of long-term care doubled from 1990 to 2001 (LaCroix, 2013). This has become increasingly important as the world’s population is aging rapidly (Lutz, Sanderson, & Scherbov, 2008). Nearly one in five people in the United States will be over the age of 65 by 2030, and the population over 85 will more than triple by 2050 (Vincent, Velkoff, & Census Bureau, 2010). European countries are faced with a similarly aging population and are concerned about not being able to meet the demand for healthcare services (Rechel et al., 2013). In the United States alone, it is expected there will be a shortage of 400,000 registered nurses by 2020 (Murray, 2002).

Technology is increasingly being used to help solve this dilemma. Assistive technologies such as robots have been developed for monitoring and providing physical assistance with activities of daily living (ADLs) (Dario, Guglielmelli, Laschi, & Teti, 1999; Noury, 2005). Robots have been shown to improve generic ADL’s in stroke rehabilitation, and dementia care (Mehrholz, Hadrich, Platz, Kugler, & Pohl, 2012). Other robots have been used for guiding people around assisted-living facilities and providing medication reminders to increase adherence (Stafford et al., 2010). Companion robots such as Paro (a baby seal) have been developed and initial evidence suggests they can improve patients’ moods (Wada, Shibata, Saito, & Tanie, 2004), lead to a higher quality of life (Shibata & Wada, 2011), and reduce loneliness (Robinson, MacDonald, Kerse, & Broadbent, 2013a) in elderly living in assisted-living care. A downside of robots is that they may be more expensive than other digital devices and thus the relative advantages of each need to be considered.

Many recent smartphone and tablet apps (applications) have been designed with the goal of improving health behaviours. In 2009, it was reported that over 300 million apps were downloaded and this number increased to five billion in 2010 (Boulos, Wheeler, Tavares, & Jones, 2011). As of 2010, there were over 7000 healthcare related apps available at the Apple App Store for patients (Kailas, Chong, & Watanabe, 2010). Healthcare apps have been shown to improve patients’ adherence to treatment, teach sexual education, and assist patients in the management of disabilities and chronic conditions such as diabetes (Arsand et al., 2012; Boulos et al., 2011; Lim, Hocking, Hellard & Aitken, 2008).
Smartphones and tablets are also used by clinicians and medical students, with a survey in the United Kingdom showing a high prevalence in these populations (Payne, Wharrad, & Watts, 2012).

Previous research in this field has compared computers against humans (King, Friedman, Marcus, Castro, Napolitano, & Ahn, 2007; King et al., 2014). King et al. (2007) showed that an automated telephone-based computer system led to increases in physical activity levels similar to that achieved by humans. Follow up to this study showed that the effects in both interventions were still similar even at 6 months post-intervention (King et al., 2014). Other research in this field has shown that participants developed stronger alliances with a robot than a typical computer and used the robot for longer when recording dietary behaviours (Kidd & Breazeal, 2008). Further research has demonstrated the benefits of using a physically-embodied robot versus a software agent/avatar (Kiesler, Powers, Fussell, & Torrey, 2008). Software agents/avatars are enticing as they require less space to operate, and have lower costs of development and maintenance. However, people engaged with, and developed stronger affinities to the physically-embodied robot than the projected avatar. Additionally, participants interacted with robots in a similar style to the way they interacted with other people (Kiesler et al., 2008). To our knowledge, no studies to date have examined the differences in the therapeutic alliance, adherence to instructions, and device ratings between patients interacting with each device. This study hypothesised that participants allocated to the robot group would have a stronger alliance, higher adherence, and higher device ratings than participants in the tablet group.

2. Methods

The design was a randomised control trial (Robot vs. Tablet Computer). At the onset of the experiment, participant IDs were randomised using a random number generator into either a robot group (n = 34), or a tablet computer group (n = 31). On arrival participants entered a room with either a robot or computer tablet group, with both devices running identical software. A tablet was chosen over other forms of technology (e.g. a smartphone) to maintain the same screen size in both groups. Furthermore, a tablet allows a larger font to be used and has greater potential for future use in a geriatric population. The robot or computer tablet asked participants to perform some exercise and relaxation tasks. The aim was to compare the therapeutic alliance, adherence to instructions, and device ratings between patients interacting with each device. This study hypothesised that participants allocated to the robot group would have a stronger alliance, higher adherence, and higher device ratings than participants in the tablet group.

2.1. Participants

Volunteers 16 years of age and older were recruited from the Auckland region via departmental e-mails and advertisements posted at the University of Auckland. Subjects were entered into a draw to win a $200 gift voucher for participation. There were 65 subjects during the study, consisting of 43 females (66%) and 22 males (34%). Participants’ ages ranged from 19 to 62, with an average of 30 years (standard deviation (SD) 11.5 years). The majority of participants were New Zealand European (77%), followed by Indian (6%), Chinese (5%), Canadian European (3%), and other (9%).

Approval for the study was obtained from The University of Auckland Human Participants Ethics Committee (Reference Number 8740).

2.2. Materials

The robot condition used Yujin Robot’s iRobiQ robot. It stands 450 mm tall, is 320 mm wide and weighs 7 kg. It contains an Intel based internal computer to run its software, which was programmed by the Healthbots team at the University of Auckland Department of Electrical & Computer Engineering. Physically, the robot has a head with LED lights to make facial expressions, and two arms that are used for gestures. It also contains a 7-in. touchscreen LCD display in its body, which was used to receive input and display menus. It used a gender non-specific voice for communicating with participants (see Fig. 1).

The tablet computer condition used an ASUS Google Nexus 7 tablet (see Fig. 2). The tablet is 198.5 mm tall, and 120 mm wide. This tablet was chosen as it contains the same 7-in. sized touchscreen as the iRobiQ robot. The tablet ran Google’s Android Operating System version 4.1. In order to run the same software programme and use the identical voice, we used the remote desktop application Splashtop 2.

The exercise bike used was the Elite E10 Sport Exercycle. The study used the bike’s built-in program to increase magnetic resistance every minute for the first 10 min. The experimenter adjusted the bike seat for each participant.

Two booklets containing measures used in the study are detailed in Table 1. The blood pressure cuff used was Pulseeor’s CardioScope II. It used the standard adult sized upper arm cuff.

2.3. Procedure

Participants were invited to interact with healthcare technology for up to an hour at the medical school. They were told it would...
Measures in the experiment.

In the experiment, participants were advised by the robot or tablet to "take part in the exercises, including weighing themselves, balancing on one foot (each leg), and riding a stationary exercise bicycle. For all of the exercises, participants were advised by the robot or tablet to "take part in the exercises for as long as they were comfortable". While out of the room, the experimenter observed via live video/audio feed the number of times a participant spoke, smiled and touched the technology, and their participation in the exercises. At the end of the interaction, the experimenter re-entered the room and gave a final written questionnaire to assess the quality of their experience. All scales and measures used are shown in Table 1.

2.4. Data analysis

Data was entered into a Microsoft Excel Spreadsheet with an ordered participant ID as the unique identifier. An accuracy check of 13 subjects (20%) was then performed.

Data analysis was performed using SPSS Version 20.0 (SPSS Inc., Chicago, IL). Data were checked for normality using Kolmogorov-Smirnov tests and variables met the assumptions for parametric tests. Pearson’s chi-squared tests were used to test for differences between groups on categorical variables. Independent samples t-tests were performed to compare means between groups. A probability value of <0.05 was considered statistically significant.

3. Results

3.1. Baseline

In the robot group there were 34 participants, 8 male and 26 female. The tablet computer group consisted of 31 participants, 14 male and 17 female. There was no significant difference in gender between groups, $\chi^2(1, N = 65) = 3.39, p = 0.07$. Baseline scores for each group are displayed in Table 2 and indicate that there were no significant differences at baseline.

3.2. Observations

During the interactions, the robot group spoke to the device significantly more times than the tablet group. Similarly, the number of times the robot group smiled at the device was significantly greater than the tablet group.

The only significant differences between groups in observed adherence were for the relaxation exercise. All 34 people in the

Table 1

<table>
<thead>
<tr>
<th>Time points</th>
<th>Construct</th>
<th>Scale or observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Education level</td>
<td>At what level did you complete your highest education?</td>
</tr>
<tr>
<td></td>
<td>Computer knowledge</td>
<td>How good are you at using computers? (Novice (1) – Expert (8) scale)</td>
</tr>
<tr>
<td></td>
<td>Robot knowledge</td>
<td>How much do you know about robots? (Novice (1) – Expert (8) scale)</td>
</tr>
<tr>
<td></td>
<td>Computer programmer knowledge</td>
<td>How much do you know about programming computers? (Novice (1) – Expert (8) scale)</td>
</tr>
<tr>
<td>During interaction</td>
<td>Engagement</td>
<td>Observer rated: how many times did the participant speak to the technology?</td>
</tr>
<tr>
<td></td>
<td>Adherence</td>
<td>Observer rated: how many times did the participant smile at the technology?</td>
</tr>
<tr>
<td></td>
<td>Physiological response</td>
<td>Observer rated: did the participant participate in the breathing exercise? – uncross their aims, close their eyes, and breathe deeply</td>
</tr>
<tr>
<td></td>
<td>Attitudes towards the device</td>
<td>Recorded on device: how long did the participant balance on each leg?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recorded on device: how long did the participant ride the exercise bike?</td>
</tr>
<tr>
<td>After interaction</td>
<td>Personality of device</td>
<td>Blood pressure and heart rate taken by device</td>
</tr>
<tr>
<td></td>
<td>Trust in device</td>
<td>Robot Attitude Scale (Broadbent et al., 2009)</td>
</tr>
<tr>
<td></td>
<td>Enjoyment of device</td>
<td>(Better attitude (1) – worse attitude (8) scale, 11 items) Cronbach’s alpha ($\alpha$) = 0.92. How comfortable did you feel interacting with the robot/tablet? (not at all comfortable (0) – very comfortable (100) scale)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How well do you rate this robot/tablet? (very poor (0) – excellent (100) scale)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How much would you like to interact with this robot/tablet again? (Not at all (0) – very much (100) scale)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How accurate do you think the robot/tablet was in taking your blood pressure? (not at all (0) – very accurate (100) scale)</td>
</tr>
<tr>
<td></td>
<td>Ash’s personality checklist</td>
<td>(Asch, 1946) (20 items)</td>
</tr>
<tr>
<td></td>
<td>Trust in device</td>
<td>Trust in Medical Technology Scale – adapted from Trust in Physician Scale (Anderson &amp; Dedrick, 1990) (strongly agree – strongly disagree scale, 11 items) $\alpha = 0.76$.</td>
</tr>
<tr>
<td></td>
<td>Enjoyment of device</td>
<td>Social Interaction Scale (Berry &amp; Hansen, 1996) (not at all (1) – very much (8), 8 items) $\alpha = 0.75$. Quality of experience (Berry &amp; Hansen, 1996) (not at all (1) – very much (5), 8 items) $\alpha = 0.81$.</td>
</tr>
</tbody>
</table>
robot group participated in the relaxation, whereas only 26 of the 31 participants in the tablet group participated. Not only was the robot group more likely to participate in the relaxation exercise, but the group was also significantly more likely to follow directions given by the technology during the exercise. When asked by the device, subjects in the robot group uncrossed their arms more and closed their eyes more (see Table 3).

3.3. Technology ratings

Participants in the robot group rated the technology significantly better than the tablet group. Compared with the tablet group, they also responded significantly more positively when asked “How much would you like to interact with the technology again?”. Furthermore, participants thought the robot was more accurate when measuring their blood pressure. When the groups rated their views on each technology using the Robot Attitude Scale, the robot group perceived the technology more positively than did the tablet group.

3.3.1. Enjoyment

As part of the Social Interaction Scale participants were asked to rate their enjoyment interacting with each technology. Participants rated the robot significantly more enjoyable than the tablet. Furthermore, they thought the interaction with the robot was more smooth, natural, and relaxed, than the tablet. Similarly, participants also viewed the robot interaction as less forced and awkward.

3.3.2. Trust

As part of the Trust in Medical Technology Scale, participants in the robot condition responded they were significantly more likely to trust the technology’s judgement than in the tablet condition. This is also shown by participants in the robot group viewing the technology more likely to keep their private information confidential. However, there was no significant difference in any of the scale’s other nine items.

3.3.3. Personality

As shown in Table 4, pairwise comparisons indicated there were significant differences in the ratings of each technology’s personality in eight of the 20 characteristics: generous–ungenerous, sociable–unsociable, popular–unpopular, unreliable–reliable, important–insignificant, ruthless–humane, imaginative–hard headed, and strong–weak. Compared to the computer tablet, the robot was seen as more generous, sociable, popular, reliable, important, humane, imaginative, and strong.

4. Discussion

This is the first study to show that using a robot to promote healthcare behaviours can have advantages over a tablet computer. Participants found the robot was more enjoyable to interact with than the computer tablet, and reported higher desires to interact with the robot again in the future. Furthermore, people were more likely to trust a robot’s advice and viewed it as more accurate when taking blood pressure. This study also showed that people viewed robots as less likely to breach confidentiality. These findings suggest that participants formed stronger relationships with the robot, than others did with the tablet computer. It also appears that this relationship was formed more readily with the robot as participants made these evaluations after only one interaction.

The findings of this study are likely to be due to the physical nature of the robot, which contained humanlike attributes including

<table>
<thead>
<tr>
<th>Measure</th>
<th>Robot group mean (SD)</th>
<th>Tablet group mean (SD)</th>
<th>Test statistic</th>
<th>Sig (P-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoke to technology</td>
<td>2.09 ± 3.8</td>
<td>0.4 ± 0.6</td>
<td>t(63) = 2.5</td>
<td>0.016</td>
</tr>
<tr>
<td>Smiled at technology</td>
<td>2.9 ± 0.61</td>
<td>0.61 ± 0.76</td>
<td>t(63) = 4.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Relaxation exercise participant (Yes:No)</td>
<td>34.0</td>
<td>26.5</td>
<td>χ²(1, N = 65) = 5.9</td>
<td>0.015</td>
</tr>
<tr>
<td>Uncrossed arms during relaxation exercise (Yes:No)</td>
<td>33.1</td>
<td>25.6</td>
<td>χ²(1, N = 65) = 4.5</td>
<td>0.033</td>
</tr>
<tr>
<td>Closed eyes during relaxation exercise (Yes:No)</td>
<td>20:11</td>
<td>20:11</td>
<td>χ²(1, N = 65) = 5.1</td>
<td>0.023</td>
</tr>
<tr>
<td>Time balanced on right leg (s)</td>
<td>190.6 ± 170.2</td>
<td>179.1 ± 153.0</td>
<td>t(63) = 0.28</td>
<td>0.77</td>
</tr>
<tr>
<td>Time balanced on left leg (s)</td>
<td>168.9 ± 112.1</td>
<td>181.5 ± 127.4</td>
<td>t(63) = -0.42</td>
<td>0.67</td>
</tr>
<tr>
<td>Time on exercise bike (s)</td>
<td>420.4 ± 304.9</td>
<td>376.3 ± 187.1</td>
<td>t(63) = -0.70</td>
<td>0.49</td>
</tr>
<tr>
<td>Robot Attitude Scalea</td>
<td>28.1 ± 12.8</td>
<td>35.7 ± 12.0</td>
<td>t(63) = -2.5</td>
<td>0.016</td>
</tr>
<tr>
<td>Technology rating</td>
<td>81.4 ± 17.5</td>
<td>57.4 ± 21.1</td>
<td>t(63) = 5.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Interact again</td>
<td>76.1 ± 22.5</td>
<td>49.4 ± 21.1</td>
<td>t(63) = 4.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Accuracy of technology</td>
<td>85.4 ± 16.5</td>
<td>69.5 ± 28.2</td>
<td>t(63) = 2.8</td>
<td>0.007</td>
</tr>
<tr>
<td>Was the interaction smooth, natural, and relaxed</td>
<td>6.1 ± 1.5</td>
<td>4.8 ± 1.7</td>
<td>t(63) = 3.2</td>
<td>0.002</td>
</tr>
<tr>
<td>How much would you like to interact again</td>
<td>5.9 ± 1.7</td>
<td>4.7 ± 1.8</td>
<td>t(63) = 2.9</td>
<td>0.005</td>
</tr>
<tr>
<td>Was the interaction forced, strained, and awkwarda</td>
<td>3.0 ± 1.6</td>
<td>3.9 ± 1.7</td>
<td>t(63) = -2.1</td>
<td>0.034</td>
</tr>
<tr>
<td>Trust technologies judgement</td>
<td>3.1 ± 1.0</td>
<td>2.5 ± 0.9</td>
<td>t(63) = 2.3</td>
<td>0.026</td>
</tr>
<tr>
<td>Will not keep information privatea</td>
<td>2.2 ± 1.1</td>
<td>3.1 ± 1.2</td>
<td>t(63) = -2.2</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Bold indicates that the effects observed are statistically significant at p < .05.

Lower number indicates more favourable rating.
facial expressions and body language. Participants may therefore have seen the robot as having other human-like attributes. This is supported by the personality ratings, which showed the robot was seen as more humane, imaginative and sociable than the tablet. These results are in agreement with initial studies examining robots versus simple computers and projected robots (Fasola & Mataric, 2013; Kidd & Breazeal, 2008; Kiesler et al., 2008).

With the exception of relaxation, the results did not show that the robot had a positive influence on exercise behaviours. These findings are unlike studies which compare robots to traditional computers (Kidd & Breazeal, 2008), and this may be because the advertisement stated that the participant had to perform some physical exercise. Therefore only people who wanted to exercise may have responded to the advertisement and would have turned up expecting to exercise. This is supported by the observation that most participants arrived wearing exercise clothing.

As evidenced by Andrews (1999), there is a difference between the efficacy, effectiveness, and efficiency of treatment options. This study demonstrated the efficacy of using a robot versus a computer tablet in promoting healthcare behaviours. The sample size of the study was comparable to similar studies that determine efficacy (Broadbent et al., 2010; Broadbent et al., 2011; Fasola & Mataric, 2013; Kidd & Breazeal, 2008). The effect sizes found for the majority of measures were large and therefore the power of the study was sufficient to detect these effects. For example, the size of the effect between the robot and the tablet for the question “How much would you like to interact with the technology again?” was 1.22 (Cohen’s d), and the sample size achieved power of 0.99 with alpha of 0.05.

To evaluate effectiveness, future studies should be performed in a clinical setting with patients. Evaluating the efficiency of using robots in healthcare is also needed as one of the limiting factors of using a robot is cost. The price of simple robots is declining, and may reach a price point that is cost-effective. At what price point this would occur is beyond the scope of this paper and a future cost-benefit analysis is recommended.

Increasingly robots are being used in healthcare throughout the entire lifespan. Whether this is by helping to reduce a child’s pain during a flu vaccine (Beran, Ramirez-Serrano, Vanderkooi, & Kuhn, 2013), or as an aide to an elderly patient with dementia (Robinson, MacDonald, Kerse, & Broadbent, 2013b), the findings of this study may be helpful as part of the solution to the ageing population dilemma. In a study by Kuo et al. (2009), attitudes and reactions to a healthcare robot were investigated to explore differences between middle-aged and older people. The study found that compared to middle-aged individuals, elderly people (65 years and older) had less experience with computers but similar attitudes and ratings towards the robot.

One area of particular importance in older people’s health is cognitive impairment. As the population ages, so will age-related illnesses such as dementia. Therefore healthcare robots may be used in this domain as well. The study by Robinson et al. (2013b) demonstrated that there is potential for health robots in patients with dementia, but further research is required to determine the best forms for this population.

4.1. Limitations

This study is limited by the generalisability of the results. The majority of participants were university students and the research was conducted in the laboratory. Therefore the results may not translate to an older population in the community where these health robots may be needed most in the future. Also, participants’ responses were evaluated in just one session and therefore we do not know if the differences between the two groups would be maintained in multiple sessions over time.

Another limitation was that we did not show both devices to the participants to see which one they would chose to use. However this was not the aim of the study. The aim of the study was to compare behaviours in response to the instructions provided by each device, and the results would have been biased had the participants been given a choice of device. The question asking how much people wanted to use the device again is often used as a measure of technology acceptance (Mathieson, 1991). The finding that participants in the robot group wanted to interact with the technology again more than the tablet group did, suggests that the robot was better accepted.

4.2. Future studies

Future studies could be conducted with an older sample and/or in the community. Previous research has shown that older age is still compatible with robot use, contrary to popular opinion (Broadbent et al., 2010). Future research could compare how people interact with each device in other healthcare domains, e.g. increasing adherence to medication, or stroke rehabilitation. Other research could look to compare avatar/video-formats on both technology types to determine if this would provide similar results. Likewise, comparing human-like avatars against robot avatars may be an area for future research.

5. Conclusions

Overall, this study found that using robots to promote healthcare behaviours appears to have some advantages over tablet computers in encouraging relaxation behaviours and in engaging users. The results of the study suggested that participants formed relationships with the robot more readily, and that these relationships were stronger than with the computer tablet. The findings can be attributed to the physical appearance of the robot, which appears to have an integral role in determining a user’s perceptions and adherence.

Acknowledgements

This research was funded as part of the University of Auckland’s summer studentship research programme in the Faculty of Medical and Health Sciences. Funders had no role in study design, collection,
We would like to thank our colleagues at the University of Auckland Healthbots research team for their support.

References


