Investigating the Usability of a Socially Assistive Robotic Cognitive Training Task with Augmented Sensory Feedback Modalities for Older Adults

Nault, Emilyann ; Baillie, Lynne; Broz, Frank

DOI
10.1109/RO-MAN53752.2022.9900825

Publication date
2022

Document Version
Final published version

Published in
Proceedings of the 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)

Citation (APA)
Nault, E., Baillie, L., & Broz, F. (2022). Investigating the Usability of a Socially Assistive Robotic Cognitive Training Task with Augmented Sensory Feedback Modalities for Older Adults. In *Proceedings of the 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN): Social, Asocial, and Antisocial Robots* (pp. 735-742). (RO-MAN 2022 - 31st IEEE International Conference on Robot and Human Interactive Communication: Social, Asocial, and Antisocial Robots). IEEE. https://doi.org/10.1109/RO-MAN53752.2022.9900825

2022 31st IEEE International Conference on Robot and
Human Interactive Communication (RO-MAN)
August 29 - September 2, 2022. Naples, Italy.

2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) | 978-1-7281-8859-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/RO-MAN53752.2022.9900825

Investigating the Usability of a Socially Assistive Robotic Cognitive Training Task with Augmented Sensory Feedback Modalities for Older Adults*

Emilyann Nault[1] Lynne Baillie[1] and Frank Broz[2]

Abstract— **Cognitive training is effective at retaining cognitive function and delaying decline for typically ageing older adults, individuals with mild cognitive impairment, and persons with dementia. Technological resources can address limiting factors that inhibit engagement and access to this treatment. We investigated how a socially assistive robot-facilitated memory task with sensory feedback was received by older adults. The impact of unimodal and multimodal administration of auditory and haptic feedback using two robot embodiments (Pepper and Nao) was evaluated in terms of user performance, usability, and workload. In contrast to sensory feedback research, auditory feedback resulted in significantly higher task accuracy. This was, however, supported by previous work from neurological literature. Auditory feedback also received significantly higher usability, and this preference was validated by qualitative feedback from participants. Regardless of robotic embodiment, this study demonstrates an advantage for auditory feedback (over haptic and multimodal) in cognitive training activities for older adults.**

I. INTRODUCTION

In most parts of the world, the percentage of older adults, or individuals aged 65 and over, is projected to increase significantly by 2050, resulting in added pressure on the healthcare staff and system as a whole [1]. Engagement in cognitive training (CT) tasks can aid in retaining cognitive function and delaying cognitive decline along the ageing continuum: from typically ageing older adults [2] to those with mild cognitive impairment (MCI) and dementia [3], [4]. However, access to treatments such as CT remains a barrier, even for those with a formal diagnosis [5]. For those who can access treatment, engagement is often limited due to several factors [3], [5]. This emphasizes the need for other means of engaging older adults in cognitive activities.

In the age of the COVID-19 pandemic, the older adults not using technology due to preference or lack of access have become isolated further, particularly for those who are community dwelling or residing in long-term care facilities [6]. This increased isolation has resulted in a reported decline in function, cognition, and neuropsychiatric symptoms [7], [8], [9]. The pandemic has consequently resulted in a shift towards using technological interventions for a variety of healthcare purposes [10]. The use of technology to implement care provides additional safety benefits for both patients and staff through minimizing human-to-human contact.

Socially assistive robots (SARs) provide the added benefit over other technological interventions of being physically embodied, which assists in the social engagement of its users. This has been shown to have positive motivational effects with older adults [11], particularly in the rehabilitation space [12]. Additionally, while sensory impairment becomes more prevalent with age, it can be used to assist older adults with cognitive tasks [13], [14], [15]. However, previous work has mainly focused on sensory feedback outwith the context of cognitive decline in older adults (see Section II-C). These technologies can provide an engaging means for older adults to independently complete cognitive activities.

To investigate these areas of interest, a memory-based task was selected to exemplify the targeted approach taken by CT practices (Section II-A). It was evaluated with three combinations of sensory feedback and two SAR embodiments (Nao and Pepper). The work presented here is an expansion of a previous pilot study [16] through its inclusion of the target age group, introduction of a secondary SAR, and transition to a wearable haptic actuator that produces a stronger vibration. The following hypotheses were formulated to address the impact of sensory feedback administration and robot embodiment in regard to the memory task:

H1: Multimodal (simultaneous auditory and haptic) feedback will produce higher accuracy scores at the memory task compared to unimodal auditory and haptic feedback.

H2: Multimodal feedback will contribute to increased usability (evidenced by the System Usability Scale scores) and decreased workload (evidenced by the NASA TLX scores) compared to unimodal auditory and haptic feedback.

H3: Based on previous literature, the small tabletop robot embodiment (Nao) will be preferred over the larger freestanding robot embodiment (Pepper).

II. BACKGROUND

In this section, the need for alternate means of engaging older adults in cognitive tasks is discussed. Subsequently, the literature behind the technical components of the proposed solution is reviewed.

---

*This research is supported by the Engineering and Physical Sciences Research Council (EPSRC), Grant ID: EP/S023208/1

[1]Emilyann Nault and Lynne Baillie are with Mathematics and Computer Science, Heriot-Watt University, Edinburgh, United Kingdom `en27@hw.ac.uk, l.baillie@hw.ac.uk`

[2]Frank Broz with the Mathematics and Computer Science, Delft University of Technology, 2628 CD Delft, Netherlands `f.broz@tudelft.nl`

978-1-7281-8859-1/22/$31.00 ©2022 IEEE

A. Mild Cognitive Impairment & Dementia

While cognitive decline is a part of ageing, when it occurs at an increased rate from what is typical, this becomes classified as either MCI or dementia. MCI involves fewer areas of cognition and does not have as great an impact on daily life compared to dementia. CT is a nonpharmaceutical treatment which targets intact cognitive resources and supports areas of difficulty [3]. For instance, a pen and paper resource, Brainwave-R [17], is broken down into specific areas of neurological function: attention, visual processing, information processing, executive function, and memory.

Despite its effectiveness, treatments such as CT are not regularly offered. When it is provided, many factors can impede access and adherence to treatment including logistical considerations, denial, depression, difficulty motivating, and the cognitive impairment itself. In contrast, sustained and appropriate engagement is required to reach a therapeutic effect, therefore highlighting the need for tools to aid in providing rehabilitation to this population [3].

However, these techniques are not exclusive to those with MCI or dementia. For typically ageing older adults (individuals aged 65+ without an age-related cognitive impairment), CT has been shown to improve cognition [18], the benefits of which can persist over an extended period [19]. Smith et al. found improved cognition regarding memory and attention in typically ageing older adults with their adaptive computerized CT program [20]. Furthermore, engaging in leisure tasks that are cognitively stimulating (such as Sudoku) has been associated with higher levels of cognitive performance in typically ageing older adults [21]. In summary, previous literature indicates CT can aid in the prevention of cognitive decline for typically ageing older adults, thus highlighting the need for tools to foster engagement.

B. Socially Assistive Robots

SARs have the potential to improve the quality of life of older adults through social and cognitive engagement [22]. Ostrowski et al. investigated long-term social robot usage in the home. Over half of the older adults chose to keep the robot over 50 days, with 3 deciding to keep it over 2 years [23]. Kubota et al. interviewed clinicians and persons with MCI, and identified means of promoting engagement in rehabilitation to be transferred to a SAR, including taking breaks and using visual and auditory cues [24].

Compared to other means of CT, such as pen and paper [17] or screen-based tasks, physically embodied SARs allow for improved levels of engagement with older adults [25] over the short- and long-term [26]. When compared to a virtual agent for stroke rehabilitation, a SAR demonstrated increased engagement and performance in participants [27]. Improved performance in the form of sustained attention, increased accuracy, and decreased reaction time was also found with individuals with dementia when engaging with a SAR-facilitated cognitive music task for 6 months [28].

The current paper looks to investigate how Nao and Pepper, two humanoid robots with differing embodiments, are perceived in the context of a memory task. A study comparing the social influence between these robots found when each robot protested leaving a book in the room, Nao had higher compliance and was rated more positively [29].

C. Sensory Feedback

The classically defined five senses (vision, hearing, touch, taste, and smell), in addition to others such as proprioception, diminish with age [30]. Sensory feedback can be employed to convey information while reducing workload [13], [14]. Many assistive technologies already use varying forms of sensory feedback [31]. Furthermore, targeted treatments such as light or music therapy has had positive effects on cognition and related behaviors in those with Alzheimer's disease [32].

In the rehabilitation setting, sensory cues have been shown to improve performance and assist with processing of information. In post-stroke robotic rehabilitation, continuous auditory feedback was implemented resulting in improved engagement and performance [33]. Similar results were found for a robotic sleeve for upper limb stroke rehabilitation that incorporated auditory, haptic, and visual feedback to communicate different aspects of each exercise [34].

Some studies have suggested multimodal interactions can yield higher performance and preference over unimodal interactions. For instance, Qian et al. [35] found incorporating multimodal feedback (auditory and tactile, in this case) both improved walking behavior and was easier for older adults to process compared to unimodal feedback. However, this can be task dependent. One study with older adults assessed combinations of visual, auditory, and haptic feedback in the context of assisting with a computerized drag-and-drop task. The authors found unimodal visual and haptic feedback were not very useful [36]. This was attributed to the visual and motor nature of the task, thus suggesting adding stimuli that is the same modality as the task may mask the effects or increase the load of that sensory channel. Auditory feedback, alone, and in combination with other modalities, produced the most consistent improvement in performance across all participants. This aligns with Wickens and Liu's Multiple Resource Theory [13], developed in 1988, where the areas of the brain that process visual and spatial information (those required for the drag-and-drop task) use different resources than auditory processing. In summary, while multimodal applications have shown to be beneficial in relation to preference and reducing workload (e.g., [35]), this is highly dependent on the nature of the task.

The present study incorporates auditory and haptic feedback due to the previous literature that indicates both can improve performance at memory tasks. One study found providing auditory feedback in the form of musical notes while pianists learned a piece of music significantly improved their recall [37]. Another study found employing haptic feedback during the learning phase of visual or auditory stimuli, as well as during the recall phase, improved performance by 20% for participants who performed poorly at the task [38].

D. Summary of Previous Study

The experiment presented here expands upon a previous pilot study [16]. The study assessed the use of a memory task using the touch-based sensors on the SAR Pepper (Figure 1a) with auditory and haptic feedback. The system was evaluated with 9 young adults (M=25.5 years old), seven of which had previous experience with robots. Each participant, while seated, engaged in three rounds of the task, each with different feedback administration: auditory, haptic, and simultaneous auditory and haptic feedback (multimodal). The auditory feedback was a beep delivered through the robot's speakers. The wearable haptic actuator delivered a vibration to the user's wrist through a motor powered by an Arduino. The interaction was assessed in terms of task performance, usability, and workload using the measures presented in Section III-B.3. There were no significant differences in accuracy across conditions, and the overall workload was low for the system across feedback modalities. Auditory feedback was most preferred in terms of usability. This was attributed towards the limited output of the vibrotactile device, which was validated by participant feedback.

E. Literature Conclusion

The literature presented here supports the use of SARs to improve engagement and performance at cognitive activities for older adults. The addition of sensory feedback has the potential to improve performance and decrease workload by providing information, which can be particularly useful when an individual has one or more sensory deficits. To the authors knowledge, the use of a SAR-facilitated memory task with sensory feedback in the context of engaging older adults in cognitive tasks has not previously been investigated.

For this experiment, a memory task was chosen to represent the targeted CT strategy often utilized in practice. Whereas the embodiment of a SAR can impact engagement, two embodiments were evaluated to investigate whether one would be preferable towards improving engagement in the context of CT. Finally, the potential for sensory feedback to improve rehabilitation engagement and performance motivated us to examine this through the incorporation of auditory and haptic feedback.

III. METHODOLOGY

A. Participants

Participants consisted of independently living individuals aged 65 and over. They were recruited from community groups, professional connections, and independent-living housing associations. On average, the experiment lasted 1.5 hours. Each participant received a £10 voucher and were additionally entered into a prize draw for an additional £50 voucher. Twenty older adults participated, with one needing to be removed from the dataset due to technical difficulties, resulting in 19 participants total (M=72 years old, SD=5.7, 15 female, 4 male). Only one participant had previous experience with humanoid robots, and none of the participants had a tactile impairment. Five were diagnosed with an auditory impairment, 4 of whom stated they wore hearing aids. These participants were not excluded in order to retain a representative sample of typically ageing older adults. Additionally, the existence of an impairment or requiring a corrective device does not indicate that modality would no longer be useful for those individuals.

In the pilot study, because the young adults were seated, they had to reach to touch Pepper's head, and two even chose to bend and touch the feet using their hands. Therefore, the authors expected the older adults would find it easier to stand when interacting with the Pepper. The Timed Up and Go Test (TUG) was utilized to ensure none were at a risk of falling [39]. The TUG protocol is as follows: the individual is told to stand up from a chair with arm rests, walk 3 meters at their normal pace, turn around, walk back to the chair, and sit down. All participants met the CDC[1] requirement of completing this task in 12 seconds or less (M=8.5 seconds). Full ethical approval for this experiment was obtained by Heriot-Watt University.

B. Experimental Design

1) Memory Task: The memory task is the same as the previous pilot study [16]. To summarize, the SAR states a random sequence from the five potential body parts (Figure 1), followed by the word 'go'. Then the user must touch the body parts in the order listed. The number of items in each sequence is preset and increased over time. The activity continues until the user makes a mistake.

2) Socially Assistive Robots: The SARs Pepper[2] and Nao[3] were used in this study to determine how the robot's embodiment impacts the interaction. Pepper is 120cm tall, and Nao is 58cm when fully standing. These SARs have been studied previously with older adults and related ageing conditions [40], [41]. They were chosen due to their difference in embodiment (e.g., physical appearance, speech), as well as the similarities in the tactile sensors/bumpers (Figure 1) to have consistency in the task across the robots. Both SARs delivered identical dialogue and used the same animated speech feature provided in the robots' programing software, Choregraphe. The only exception to this was while engaging in task, during which time the robot was still.

3) Sensory Feedback: Feedback was provided when each body part was listed and as confirmation for each input. No control condition was provided because feedback was required to ensure the touch registered. The three sensory feedback conditions are as follows:

- **Auditory:** Beep delivered through the SARs speakers (800 Hz, 100ms).
- **Haptic:** Vibration delivered to the inside of the participants wrist (50% intensity, 100ms) (Figure 2).
- **Multimodal:** Both the beep and the vibration delivered simultaneously.

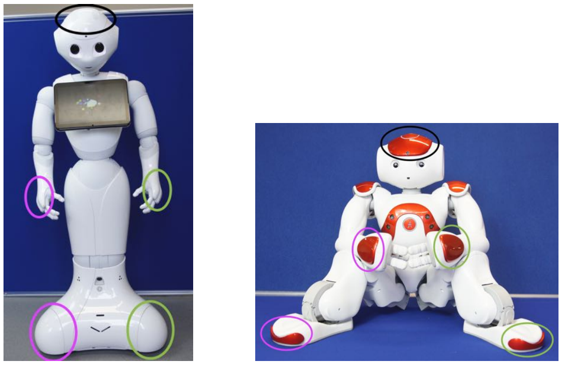
For this study we used Haptic Bracelets (Figure 2) [42], a wearable developed for gait rehabilitation [43], [44]. The

---

[1] TUG Protocol: https://www.cdc.gov/steadi/pdf/TUG_test-print.pdf

[2] Pepper: https://www.softbankrobotics.com/emea/en/pepper

[3] Nao: https://www.softbankrobotics.com/emea/en/nao



(a) Pepper*

(b) Nao*

Fig. 1: SARs with the body parts used in the activity highlighted. With both SARs, the head and hands have tactile sensors and the feet have bumpers. Black circle: head, pink circles: left hand and left foot, green circles: right hand and right foot.

* Image courtesy of the *Interactive and Trustworthy Technologies Group.*

device delivers haptic cues that are low in latency and high in strength (strength can be adjusted). It was designed to produce a 'sharp' cue, closer to a tap than a vibration.

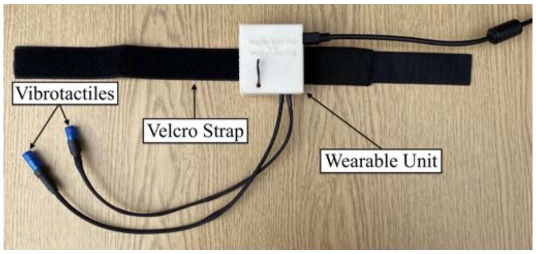


Fig. 2: Haptic Bracelets. Full technical specification can be found in [42], chapter 5.

C. Experimental Setup

Due to the COVID-19 pandemic, safety precautions were taken, including holding the experiment outdoors (Figure 3). This introduced external factors such as weather variation and people/vehicles passing by. To minimize distractors, a parking lot mostly enclosed by buildings was used and an awning was put up. The outdoor setting did not have an observable effect on the outcome.

D. Protocol

1) Memory Task Training: The experimenter demonstrated three preset rounds of the task using a script to ensure consistency. During the demonstration, it was emphasized that left and right refer to the participant's left and right, and not the robots (i.e., 'left foot' is the robot's right foot).

During this time, the level of volume was adjusted if needed. By default, the robots were set to maximum volume, although a few participants required it to be lowered to achieve better clarity. Subsequently, the SAR walked the participant through how to interact with its head, hands, and feet before engaging them in three rounds of the task.

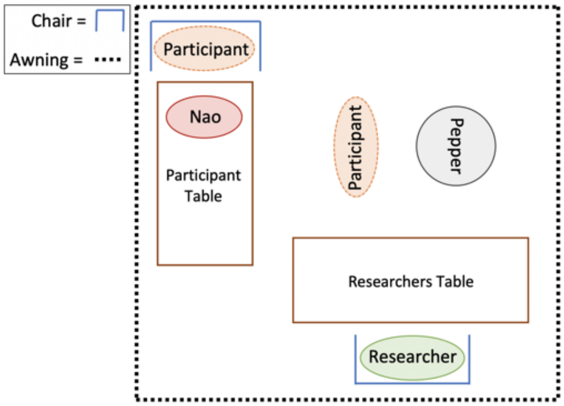


Fig. 3: Diagram of the outdoor setup. The two locations of each participant throughout the experiment are highlighted in orange: seated with Nao and standing in front of Pepper.

Fig. 4: A participant engaging in the memory task with Nao.

The robot repeated the word back to the user (e.g., 'left foot') so they could become familiar with the task without being pre-exposed to the experimental sensory feedback conditions. The participants were asked to repeat the task portion of the training module until they felt comfortable.

2) Experimental Conditions: This experiment used a within-subjects design, where each participant engaged with the task three times with each robot, once for each sensory feedback condition. A 2-minute break was provided in between each condition, along with a longer break of about 10 minutes between robots. The conditions were counterbalanced across participants by rotating which robot goes first and the 6 combinations of the feedback, resulting in a total of 12 permutations.

In some situations, participants were asked to repeat one of the experimental conditions to ensure they were properly assessed for their memory performance. This included if they made a mistake because a) they did not wait for the 'go', b) mixing up left and right for the robot's left and right, and c) unforeseen technical issues (e.g., software crashing). This occurred 28 times out of the 114 total rounds of the task, suggesting some participants may have benefited from further training before beginning the experimental conditions. These mistakes often happened early in the round, and whereas the conditions were counterbalanced, any learning effect would be spread across the experimental conditions.

3) Evaluation: After each iteration of the task, the following assessments were taken:

- **Accuracy Score:** The cumulative number of items the participant got correct before making a mistake.
- **System Usability Scale (SUS):** A standardized means of assessing the usability of a system. The phrase 'the/this system' was adjusted accordingly to account for the memory task and sensory feedback conditions [45].
- **NASA Task Load Index (TLX):** This standardized questionnaire assesses the perceived workload of the memory task [46].

IV. RESULTS

An a-priori power analysis showed 18 participants were required to achieve a power of 0.80, so the quantity of participants was sufficient to run the following statistical testing. The repeated measures ANOVA method was used to evaluate the impact of all levels of the two independent variables (SAR embodiment and sensory feedback administration) across all 3 areas of evaluation (accuracy, usability, and workload). The Bonferroni method of correcting pairwise comparisons was employed across all statistical tests unless otherwise stated.

A. H1 - Task Accuracy

In terms of task performance, there was a significant main effect of feedback on the accuracy score (F(2,36)=4.73, p=0.015, η_p^2=0.208) (Table I). The analysis showed the auditory condition had better performance than the multimodal (p=0.098) and haptic (p=0.063) conditions.

TABLE I: Summary of the significant main effects from this experiment with respect to the three evaluation measures. Asterisks represent significance with respect to the bolded value in the post hoc test. *$p \leq 0.05$, ***$p \leq 0.001$. (A = Auditory, H = Haptic, M = Multimodal.)

Accuracy Score	**A**	**H**	**M**
	15.4	12.1	13.1
SUS (out of 100)	**A**	**H**	**M**
	75.1	66.6***	67.4
NASA TLX (out of 100)	**Pepper**		**Nao**
	43.7		37.8*

There was also a significant interaction between the robots and feedback modality (F(2,36)=4.58, p=0.017, η_p^2=0.203) such that with Pepper, the multimodal condition had higher task accuracy than the haptic condition, and the opposite was true for the Nao; the multimodal condition had significantly lower task accuracy than the haptic condition.

B. H2 - SUS and NASA TLX

For the usability assessment through the SUS scores, the analysis demonstrated a significant main effect for feedback modality (F(1.44,25.93)= 4.49, p=0.031, η_p^2=0.199) (Table I). The haptic condition had significantly lower usability compared to the auditory condition (p=0.001).

While sensory feedback modality did not significantly impact workload (p=0.706), the breakdown of the various categories from the NASA TLX can be seen in Figure 5.

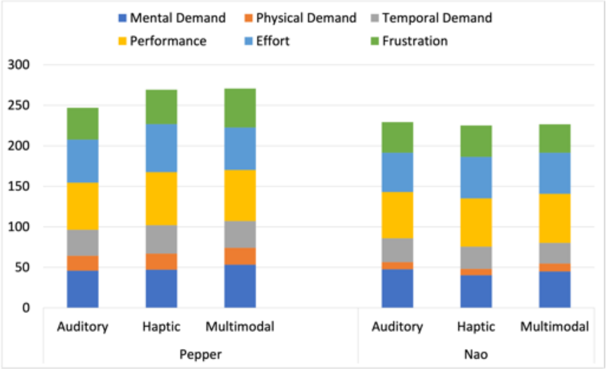


Fig. 5: RAW NASA TLX Scores.

C. H3 - SAR Embodiment

In regard to workload, there was a significant main effect for robot embodiment (F(1,18)=7.35, p=0.014, η_p^2=0.290), where the Nao had significantly lower workload compared to the Pepper (Table I). However, this significance did not carry over to task performance (p=0.644) or usability (p=0.570).

D. Supplementary Analysis: Impact of Age

The participants' age ranged from 65 to 82 years old, and it was observed by the researcher that the older participants struggled more with understanding and performing the task. As a result, an additional exploratory statistical test was run on the three evaluation measures to determine whether age was a contributing factor. Upon consultation with an academic expert in the field of ageing, it was decided to split the group down the median of 70 years old, mainly due to the small sample size. The older group had significantly lower task accuracy (M=11.6) compared to the younger group (M=15.3) (F(1,17)= 3.65, p=0.044). However, age did not have a significant impact on the usability scores (p=0.458) or workload (p=0.073).

E. Observations & Qualitative Feedback

The following section is based on observations and informal discussions with participants. The resulting notes were reviewed using the constant comparative method [47]. The themes that emerged through the analysis were comments and observations regarding the memory task, the feedback modalities, and the SARs.

Participants developed techniques to assist in remembering the sequence including hovering their hands over the body parts or using their eyes to track the sequence. Some participants struggled to wait for the robot to say 'go' before touching the sensors. This would often result in them making a mistake during the round. Others wanted the task to move faster so they could have an easier time remembering the sequence. There was a 750 ms pause between the list of words and they needed to wait to ensure they heard/felt the feedback before moving on to the next item. A few participants also found the task 'boring' and 'monotonous.'

The qualitative response regarding the feedback modalities was consistent with the results, where most of those who expressed a preference preferred the auditory feedback. One individual preferred the language-based feedback in the training session. One participant perceived the vibration was more indicative of an incorrect input. Some also highlighted the importance of having feedback options that can be adjusted depending on the users' needs and preferences.

Most participants who expressed a preference for robotic embodiment preferred Nao over Pepper. Some conveyed dissatisfaction with the elevated physicality required to engage in the task with Pepper (having to bend to touch the Pepper's hands and balance on one foot to press the feet). This is consistent with the above results in terms of workload. One participant even stated they just 'didn't like the look of Pepper.' One individual felt Nao's 'toy-like' nature would make them more willing to engage. However, a few were put off by the jerky motion of Nao returning to it's still position before starting the task (see Figure 4).

Eight of the participants commented they had trouble understanding the robots. While the volume was adjusted, some still struggled with the quality and clarity of the speech. For instance, difficulty differentiating between the words 'head' and 'hand' was a common occurrence. Two participants reported the Nao was easier to hear compared to the Pepper.

F. Comparison with the Pilot Study

Table II provides a direct comparison of the results between the pilot (summary in Section II-D) and usability study to be mainly referenced in the Discussion.

The young adults in the pilot study had higher average accuracy, rated higher usability, and reported lower workload compared to the older adults across all forms of sensory feedback (Table II). In both experiments, there was a favorability towards auditory feedback across the assessment measures (although for the young adults the difference between haptic and auditory workload scores was negligible (Table II)).

TABLE II: Comparative results between this study and the pilot study [16]. The areas shaded in gray indicate a significant main effect was found. Asterisks represent significance with respect to the bolded value. ***$p \leq 0.001$. (A = Auditory, H = Haptic, M = Multimodal.)

	Pilot Study			Current Usability Study		
Sensory Feedback	**A**	**H**	**M**	**A**	**H**	**M**
Accuracy Score	19.0	18.7	18.4	15.4	12.1	13.1
SUS (out of 100)	83.3	78.1	74.7	**75.1**	66.6***	67.4
NASA TLX (out of 100)	34.9	34.8	36.0	40.1	41.9	41.7

V. DISCUSSION

The first hypothesis (***H1***) predicted multimodal feedback would produce superior task performance over the unimodal conditions. In practice, the auditory condition resulted in the highest accuracy scores. Similar findings were discovered in both [36] and [38] regarding haptic feedback. This is also neurologically supported via Wickens and Liu's Multiple Resource Theory [13] discussed in Section II-C. It states visual and spatial information is processed separately from auditory processing. As discussed in Section IV-E, even though the sequence was delivered auditorily, some participants employed spatial and visual information to remember the sequence. Whereas the task already required touch-based interaction, it is possible the haptic feedback overwhelmed that sensory channel, rather than 'spreading' the information across separate sensory channels to decrease workload.

The interaction effect regarding task performance aligns with the differing levels of physicality required between the SARs. This result suggests having multimodal feedback with the Pepper, where the level of physical interaction and movement was far greater, was more helpful compared to the haptic feedback alone.

With respect to the usability and workload scores (***H2***), multimodal feedback was predicted to have the highest usability and lowest workload compared to the unimodal feedback conditions. Three of the administrations resulted in above average SUS scores (68+): the auditory condition in both robots and the haptic condition for the Nao robot [48]. That being said, similar to the task performance outcome, auditory feedback elicited higher usability scores compared to haptic and multimodal feedback. The strength of the significance between auditory and haptic SUS scores was surprising (Table I). One contributing factor could be that this experiment required the user to wear an external device to access the haptic feedback, therefore making it more cumbersome compared to the auditory feedback delivered directly from the robot. While no significant workload differences were found across the feedback modalities, the means suggest slightly lower workload for auditory feedback over the other conditions (see Table II).

Interestingly, ***H1*** and ***H2*** depict an advantage of auditory feedback despite 5 of the 19 participants having an auditory impairment. This suggests the existence of a sensory impairment does not necessarily signify it would not be useful, and in some cases it can be preferred. This may be an important design consideration going forward in this area of research.

The third supplementary hypothesis (***H3***) regarding the SAR embodiment was predicted correctly. There was a preference, both statistically and qualitatively, for the Nao robot. While there was no significant main effect for robot embodiment in terms of task performance and usability, the Nao did have significantly lower workload compared to the Pepper. While not significant, the breakdown of the NASA TLX subsections (Figure 5) clearly indicates Pepper had higher physical workload across all sensory feedback conditions, which is consistent with the Pepper requiring the user to stand, reach, and balance. This could also have been influenced by some participants having an easier time understanding the Nao's speech.

The comparison between the two age groups resulted in significantly decreased performance in those 71 and over. While this can be partially explained by the fact that cognitive task performance decreases with age [49], other factors such as familiarity and comfort with technology could have also influenced this result. Consequently, the tentative conclusion can be drawn that the relative age of older adults may not significantly impact the usability and workload of such a system.

A. Pilot and Usability Study - Comparative Analysis

Decreased accuracy with older adults compared to young adults would be expected considering age impacts cognitive task performance [49]. This finding, along with the decrease in reported usability and increase in workload for older adults, may have been impacted by previous experience. Jacko et al.'s study, discussed in Section II-C, found experience with the task medium had an impact on the reception of sensory feedback. Consequently, this outcome could be partially attributed to the difference in robot experience between the two groups (i.e., the young adults with robot experience may have had an advantage when engaging in the task, contributing to higher accuracy, higher usability, and lower workload scores).

The consistency in the results towards a preference for auditory feedback was surprising. The pilot study attributed this to the lack of vibrational strength (discovered through participant feedback). However, the current study did not receive any comments relating to the strength of the haptic feedback, and the outcome was consistent. This suggests this trend may be due to other factors such as how the feedback is administered and the cognitive resources required by this particular task. There is substantial support in the literature that young adults and older adults should be designed for separately [50], [30]. While the importance of user-centered design should not be overlooked, this tentative early finding suggests similarities across age may exist regarding the effectiveness of sensory feedback.

B. Future Work

Future work should assess whether these results generalize to other robot embodiments and CT activities. The memory task received feedback regarding its repetitiveness and monotony. In the future, the task will be expanded to incorporate other cognitive activities based on those used in CT. The next stage of this work is to hold a Participatory Design workshop with older adults and therapists to determine how standardized CT tools can be made more engaging through the addition of a SAR and sensory feedback.

Previous literature has shown differences in perception of haptic feedback between genders [51]. Unfortunately, the participant pool in this study did not have enough members of each gender to support statistical testing, but this is something to consider in subsequent work. Additionally, the wearable nature of the haptic actuator in this experiment may have influenced the resulting accuracy and usability scores. Therefore, future work could investigate the impact of providing sensory feedback through a device worn by the user compared to feedback which is delivered externally. Going forward, it would also be beneficial to assess a wider range of sensory feedback, (e.g., visual), including other categories within auditory and haptic feedback.

VI. CONCLUSIONS

This study investigated how a SAR-facilitated memory task with various sensory feedback modalities would be perceived by older adults with the aim of informing the design of a cognitive training system. The preference for auditory feedback (quantitatively and qualitatively) aligns with the previous pilot study with young adults [16], despite some participants having an auditory impairment. The assessment of two robot embodiments resulted in the smaller tabletop robot being preferred and having significantly lower workload, which indicates how the embodiment of a SAR can impact a CT-based interaction. This publication contributes to the growing body of work aiming to slow the rate of cognitive decline by providing other means for older adults to access and engage with cognitive training activities.

ACKNOWLEDGMENT

The authors would like to thank the Hanover Housing Association and Heriot-Watt University's Ageing Lab for their assistance in recruitment.

REFERENCES

[1] W. He, D. Goodkind, P. R. Kowal *et al.*, "An aging world: 2015," 2016.

[2] R. Naqvi, D. Liberman, J. Rosenberg, J. Alston, and S. Straus, "Preventing cognitive decline in healthy older adults," *CMAJ*, vol. 185, no. 10, pp. 881–885, 2013.

[3] J. Choi and E. W. Twamley, "Cognitive rehabilitation therapies for alzheimer's disease: a review of methods to improve treatment engagement and self-efficacy," *Neuropsychology review*, vol. 23, no. 1, pp. 48–62, 2013.

[4] S. Belleville, "Cognitive training for persons with mild cognitive impairment," *International Psychogeriatrics*, vol. 20, no. 1, pp. 57–66, 2008.

[5] M. Cations, K. E. Laver, M. Crotty, and I. D. Cameron, "Rehabilitation in dementia care," *Age and Ageing*, vol. 47, no. 2, pp. 171–174, 2018.

[6] A. Seifert, S. R. Cotten, and B. Xie, "A double burden of exclusion? digital and social exclusion of older adults in times of covid-19," *The Journals of Gerontology: Series B*, vol. 76, no. 3, pp. e99–e103, 2021.

[7] B. Lara, A. Carnes, F. Dakterzada, I. Benitez, and G. Piñol-Ripoll, "Neuropsychiatric symptoms and quality of life in spanish patients with alzheimer's disease during the covid-19 lockdown," *European journal of neurology*, vol. 27, no. 9, pp. 1744–1747, 2020.

[8] M. Canevelli, M. Valletta, M. T. Blasi, G. Remoli, G. Sarti, F. Nuti, F. Sciancalepore, E. Ruberti, M. Cesari, and G. Bruno, "Facing dementia during the covid-19 outbreak," *Journal of the American Geriatrics Society*, 2020.

[9] M. Suzuki, M. Hotta, A. Nagase, Y. Yamamoto, N. Hirakawa, Y. Satake, Y. Nagata, T. Suehiro, H. Kanemoto, K. Yoshiyama *et al.*, "The behavioral pattern of patients with frontotemporal dementia during the covid-19 pandemic," *International psychogeriatrics*, vol. 32, no. 10, pp. 1231–1234, 2020.

[10] E. Mantovani, C. Zucchella, S. Bottiroli, A. Federico, R. Giugno, G. Sandrini, C. Chiamulera, and S. Tamburin, "Telemedicine and virtual reality for cognitive rehabilitation: a roadmap for the covid-19 pandemic," *Frontiers in neurology*, vol. 11, p. 926, 2020.

[11] J. Fasola and M. J. Mataric, "Robot exercise instructor: A socially assistive robot system to monitor and encourage physical exercise for the elderly," in *19th International Symposium in Robot and Human Interactive Communication*. IEEE, 2010, pp. 416–421.

[12] A. Tapus and A.-M. Vieru, "Robot cognitive stimulation for the elderly," in *International Work-Conference on the Interplay Between Natural and Artificial Computation*. Springer, 2013, pp. 94–102.

[13] C. D. Wickens and Y. Liu, "Codes and modalities in multiple resources: A success and a qualification," *Human factors*, vol. 30, no. 5, pp. 599–616, 1988.

[14] C. D. Wickens, "Multiple resources and mental workload," *Human factors*, vol. 50, no. 3, pp. 449–455, 2008.

[15] D. Warnock, M. McGee-Lennon, and S. Brewster, "Multiple notification modalities and older users," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 1091–1094.

[16] E. Nault, L. Baillie, and F. Broz, "Auditory and haptic feedback in a socially assistive robot memory game," in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 369–371.

[17] K. B. Malia, *Brainwave-r: Cognitive strategies and techniques for brain injury rehabilitation*. PRO-ED, 2002.

[18] A. S. Berry, T. P. Zanto, W. C. Clapp, J. L. Hardy, P. B. Delahunt, H. W. Mahncke, and A. Gazzaley, "The influence of perceptual training on working memory in older adults," *PloS one*, vol. 5, no. 7, 2010.

[19] S. L. Willis, S. L. Tennstedt, M. Marsiske, K. Ball, J. Elias, K. M. Koepke, J. N. Morris, G. W. Rebok, F. W. Unverzagt, A. M. Stoddard *et al.*, "Long-term effects of cognitive training on everyday functional outcomes in older adults," *Jama*, vol. 296, no. 23, pp. 2805–2814, 2006.

[20] G. E. Smith, P. Housen, K. Yaffe, R. Ruff, R. F. Kennison, H. W. Mahncke, and E. M. Zelinski, "A cognitive training program based on principles of brain plasticity: results from the improvement in memory with plasticity-based adaptive cognitive training (impact) study," *Journal of the American Geriatrics Society*, vol. 57, no. 4, pp. 594–603, 2009.

[21] N. Ferreira, A. Owen, A. Mohan, A. Corbett, and C. Ballard, "Associations between cognitively stimulating leisure activities, cognitive function and age-related cognitive decline," *International journal of geriatric psychiatry*, vol. 30, no. 4, pp. 422–430, 2015.

[22] L. Pu, W. Moyle, C. Jones, and M. Todorovic, "The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies," *The Gerontologist*, vol. 59, no. 1, pp. e37–e51, 2019.

[23] A. K. Ostrowski, C. Breazeal, and H. W. Park, "Mixed-method long-term robot usage: Older adults' lived experience of social robots," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, 2022, pp. 33–42.

[24] A. Kubota, D. Cruz-Sandoval, S. Kim, E. W. Twamley, and L. D. Riek, "Cognitively assistive robots at home: Hri design patterns for translational science," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, 2022, pp. 53–62.

[25] M. Manca, F. Paternò, C. Santoro, E. Zedda, C. Braschi, R. Franco, and A. Sale, "The impact of serious games with humanoid robots on mild cognitive impairment older adults," *International Journal of Human-Computer Studies*, vol. 145, p. 102509, 2021.

[26] L. M. Beuscher, J. Fan, N. Sarkar, M. S. Dietrich, P. A. Newhouse, K. F. Miller, and L. C. Mion, "Socially assistive robots: measuring older adults' perceptions," *Journal of gerontological nursing*, vol. 43, no. 12, pp. 35–43, 2017.

[27] V. Vasco, C. Willemse, P. Chevalier, D. De Tommaso, V. Gower, F. Gramatica, V. Tikhanoff, U. Pattacini, G. Metta, and A. Wykowska, "Train with me: a study comparing a socially assistive robot and a virtual agent for a rehabilitation task," in *International Conference on Social Robotics*. Springer, 2019, pp. 453–463.

[28] A. Tapus, C. Tapus, and M. J. Mataric, "The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia," in *2009 IEEE international conference on rehabilitation robotics*. IEEE, 2009, pp. 924–929.

[29] S. Thunberg, S. Thellman, and T. Ziemke, "Don't judge a book by its cover: A study of the social acceptance of nao vs. pepper," in *Proceedings of the 5th International Conference on Human Agent Interaction*, 2017, pp. 443–446.

[30] R. Mazuch, "Sense-sensitive design for the ageing," *Architectural Design*, vol. 84, no. 2, pp. 108–111, 2014.

[31] D. Warnock, M. McGee-Lennon, and S. Brewster, "Older users, multimodal reminders and assisted living technology," *Health informatics journal*, vol. 18, no. 3, pp. 181–190, 2012.

[32] H. Yang, Y. Luo, Q. Hu, X. Tian, and H. Wen, "Benefits in alzheimer's disease of sensory and multisensory stimulation," *Journal of Alzheimer's Disease*, no. Preprint, pp. 1–22, 2021.

[33] G. Rosati, F. Oscari, D. J. Reinkensmeyer, R. Secoli, F. Avanzini, S. Spagnol, and S. Masiero, "Improving robotics for neurorehabilitation: enhancing engagement, performance, and learning with auditory feedback," in *2011 IEEE International Conference on Rehabilitation Robotics*. IEEE, 2011, pp. 1–6.

[34] M. Sousa, J. Vieira, D. Medeiros, A. Arsenio, and J. Jorge, "Sleevear: Augmented reality for rehabilitation using realtime feedback," in *Proceedings of the 21st international conference on intelligent user interfaces*, 2016, pp. 175–185.

[35] H. Qian, R. Kuber, A. Sears, and E. Murphy, "Maintaining and modifying pace through tactile and multimodal feedback," *Interacting with Computers*, vol. 23, no. 3, pp. 214–225, 2011.

[36] J. Jacko, V. K. Emery, P. J. Edwards, M. Ashok, L. Barnard, T. Kongnakorn, K. P. Moloney, and F. Sainfort, "The effects of multimodal feedback on older adults' task performance given varying levels of computer experience," *Behaviour & Information Technology*, vol. 23, no. 4, pp. 247–264, 2004.

[37] S. Finney and C. Palmer, "Auditory feedback and memory for music performance: Sound evidence for an encoding effect," *Memory & cognition*, vol. 31, no. 1, pp. 51–64, 2003.

[38] S. Kuznetsov, A. K. Dey, and S. E. Hudson, "The effectiveness of haptic cues as an assistive technology for human memory," in *International Conference on Pervasive Computing*. Springer, 2009, pp. 168–175.

[39] S. Mathias, U. Nayak, and B. Isaacs, "Balance in elderly patients: the" get-up and go" test." *Archives of physical medicine and rehabilitation*, vol. 67, no. 6, pp. 387–389, 1986.

[40] O. Pino, G. Palestra, R. Trevino, and B. De Carolis, "The humanoid robot nao as trainer in a memory program for elderly people with mild cognitive impairment," *International Journal of Social Robotics*, vol. 12, no. 1, pp. 21–33, 2020.

[41] L. Paletta, M. Fellner, S. Schüssler, J. Zuschnegg, J. Steiner, A. Lerch, L. Lammer, and D. Prodromou, "Amigo: Towards social robot based motivation for playful multimodal intervention in dementia," in *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*, 2018, pp. 421–427.

[42] T. Georgiou, *Rhythmic Haptic Cueing for Gait Rehabilitation of Hemiparetic Stroke and Brain Injury Survivors*. Open University (United Kingdom), 2018.

[43] T. Georgiou, S. Holland, and J. van der Linden, "Rhythmic haptic cueing for gait rehabilitation of people with hemiparesis: Quantitative gait study," *JMIR Biomedical Engineering*, vol. 5, no. 1, p. e18649, 2020.

[44] T. Georgiou, R. Islam, S. Holland, J. Van Der Linden, B. Price, P. Mulholland, and A. Perry, "Rhythmic haptic cueing using wearable devices as physiotherapy for huntington disease: Case study," *JMIR Rehabilitation and Assistive Technologies*, vol. 7, no. 2, p. e18589, 2020.

[45] J. Brooke *et al.*, "Sus-a quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

[46] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*. Elsevier, 1988, vol. 52, pp. 139–183.

[47] B. G. Glaser, A. L. Strauss, and E. Strutzel, "The discovery of grounded theory; strategies for qualitative research," *Nursing research*, vol. 17, no. 4, p. 364, 1968.

[48] J. Sauro, *A practical guide to the system usability scale: Background, benchmarks & best practices*. Measuring Usability LLC, 2011.

[49] C. T. Liverman, K. Yaffe, D. G. Blazer *et al.*, "Cognitive aging: Progress in understanding and opportunities for action," 2015.

[50] M. Zajicek and S. Brewster, "Design principles to support older adults," 2004.

[51] M. Akazue, M. Halvey, L. Baillie, and S. Brewster, "The effect of thermal stimuli on the emotional perception of images," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 4401–4412.