Music-based Assistive Feedback System for the Exploration of Virtual Environments in Individuals with Dementia

Luis Duarte Andrade Ferreira  
Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa  
Madeira Interactive Technologies Institute  
Funchal, Madeira  
luis.d.ferreira@m-iti.org

Mónica S. Cameirão  
Faculdade de Ciências Exatas e da Engenharia  
Madeira Interactive Technologies Institute  
Universidade da Madeira  
Funchal, Madeira  
monica.cameirao@m-iti.org

Sergi Bermúdez i Badia  
Faculdade de Ciências Exatas e da Engenharia  
Madeira Interactive Technologies Institute  
Universidade da Madeira  
Funchal, Madeira  
sergi.bermudez@m-iti.org

Abstract—Music is a powerful tool that shapes human experience since infancy with a known positive effect on health. As such, music has been used to aid in a variety of conditions among which Alzheimer's disease (AD). Although AD is known for disrupting competences related to activities of daily living, several learning paradigms have been applied to re-stimulate such competences. In this study, we explore the feasibility of a learning enhancement method that is based on musical distortions. We developed a quiz-type game where participants with dementia had to find the right answer. Two conditions were created, one condition exploited feedback based on musical distortion while the second condition was based on visual distortion. We applied a within-subjects study design, and measured the performance in terms of time and errors in 7 patients with dementia. The results support our hypothesis that music-based feedback can be an effective enhancer of task performance when compared to visual feedback in patients suffering dementia.

Keywords—Music, Music Distortion, Visual Distortion, Dementia, Alzheimer's Disease, Learning, Guided Feedback Systems

I. INTRODUCTION

Archaeological findings of musical instruments sustain the hypothesis that music has been part of the human experience since the Upper Palaeolithic Era [1], [2]. Since then, it seems that music is something that comes naturally to us. Children as early as two months react positively (or negatively) to such musical exposures depending on its pleasantness [3]. Certainly, music may not only lead to emotional responses [4], but may also induce physical (such as chills) [5] and chemical responses (such as Dopamine release) [6]. In addition, due to music's high processing demand, the brain uses different areas to process musical elements such as pitch, harmony and timbre [3]. Given that music appears to interact with many brain areas, it seems plausible that it can also be used to induce structural changes. For example, scientific literature suggests that toddlers who played music regularly have more developed brain areas - the corpus callosum - compared to non-musicians [6], [7].

In recent years, therapies based on music have been taken into consideration for treatment purposes [8]. Indeed, music-based therapies have been used on a variety of health conditions such are stroke [9], heart disease [10], Parkinson [11], among other conditions [12], [13], [14]. Music has been used as well in individuals with different subtypes of dementia, among which Alzheimer’s Disease (AD) (see [15] for review). AD is a subtype of dementia where individuals lose competences throughout the years [16]. AD is characterized by symptoms such as anxiety and depression [17], agitation [18], sleep deprivation [19], and memory deficits (explicit memory) [20], among others. Music therapy has been previously applied to aid individuals with AD by targeting some of its behavioural symptoms such as agitation [21]. For example, Ueda et al. shows that music is useful to control behavioural symptoms, and is recommended as complementary treatment to pharmaceutical approaches. Evidence shows that Alzheimer's patients became calmer after a session of music therapy as it stimulates a hormone called melatonin [22], [23]. Music can also be used to stimulate autobiographical memories [24], [25]. In fact, it is not rare that individuals with AD have a fairly well preserved musical memory [26]. It has been reported in a case study that a person suffering from AD when administrated the Distorted Tune Test was able to recognize not only music that was familiar to her, but also music distortions [27]. The patient had to identify melody distortions on a set of 26 well-known melodies. Surprisingly, the patient achieved an almost perfect score (25/26 correct).
To assist activities of daily living in individuals with dementia, other techniques such as Errorless learning, Trial-and-Error, Modelling Space Retrieval [28], Vanishing Cues [29], among others have been proposed (see [30] for review). However, the effectiveness of each learning method and the circumstances in which they should be applied is still debated by the scientific community, and further studies are needed [30]. Some of the learning methods have also been applied in a virtual reality serious game which aimed to stimulate activities of daily living [31]. This case study with one AD patient provided evidence that the participant was capable of learning activities related to cooking.

Based on the above findings, we designed an experiment combining Trial-and-Error and Errorless Learning with music. We hypothesize that individuals with AD perform better when presented with music distortion based feedback than with a visual analogue condition. To test our hypothesis, we developed a quiz-type game where the participants had to select correct answers to questions, with the aid of the 2 feedback modalities. The goal is to guide AD participants to the correct answer, independently of their knowledge, through music or visual distortions. A similar approach was used in an upper limb rehabilitation study that combined music with error sonification, but in healthy participants [32]. With the proposed approach, our participants explore a virtual task while, at the same time, the exploration guides them in their decision-making.

II. METHODS

A. Participants

We recruited 8 participants (3 male and 5 female) at the Portuguese Alzheimer’s Association in Madeira. All participants were at an initial to moderate stage of dementia. Participants were included if they had (1) intact upper limb movements, (2) intact hearing, and (3) intact comprehension skills. Table I summarizes the characteristics of each participant. One participant (nr. 4) dropped out from the study due to a health-related problem not related to the study, and was not included in the analysis.

B. Experimental Design

We used a within-subjects experimental design. The experiment had two conditions: Music Distortion (MD) and Visual Distortion (VD). In the music distortion condition, a background music was manipulated with the addition of white noise, while in the visual distortion condition, visual elements of the game became distorted through visual motion. To avoid a learning effect, we developed multiple versions of the application with different questions. In total, 4 versions of the application were implemented (2 sets of questions x 2 types of feedback). Before the beginning of the study, all participants were randomly assigned to perform MD-1/VD-2, MD-2/VD-1, VD-1/MD-2 or VD-2/MD-1 while keeping conditions balanced. After finishing the experiment, participants were interviewed regarding their experience of the game.

C. Procedure

All participants (or legal guardian when applicable) signed an informed consent before commencing the experiment. After signing the consent, participants were seated in front of a table in a quiet room. The presence of health professionals was sometimes required to monitor participants in more advanced stages of dementia. In case of struggle, assistance was provided but without interfering with the purpose of the experiment; we would encourage them to look for alternative answers, in order to proceed to the next level. All participants went through three phases, (1) Tutorial phase, (2) Instruction phase and (3) Play phase, as described below.
D. Musique

The developed quiz-type game is called Musique and runs on a 9" Android Tablet (GALAXY, Samsung, South Korea). Musique was developed using the Unity 3D game engine (Unity Technologies, San Francisco, California). It consists of general knowledge questions that need to be answered correctly in order to progress to the next question. All answers exhibited in the game are represented as images (see Fig. 1). Both music and visual conditions have a well-known Madeiran song called "Bailinho da Madeira" playing in the background. As for the music condition, whenever participants tapped a wrong answer, the background sound was lowered and white noise reproduced. Regarding the visual distortions, we used the free plug-in iTween from the Unity asset store. The plug-in allowed us to manipulate image parameters such as position and speed (see Fig. 1). The speed parameter of this plugin was responsible for the intensity of the distortion. The prototype stores all responses and touch events in a .CSV (Comma Separated Values) file.

Musique was organized with the following phases:

- **Tutorial Phase**: The game has a tutorial so that participants can learn the game mechanics. The tutorial begins with verbal instructions from the game's system – “Press all images for 4 seconds”. Here, we taught participants that pressing an image for 4 seconds allows them to select an answer. The tutorial has 4 random images, and participants are required to tap all of them to continue to the next phase, the instruction phase. In addition, participants can repeat the tutorial by tapping a repeat button. No background music was being played during the tutorial and both visual and musical distortions were disabled during this phase. After selecting the four images, a button was displayed that allowed participants to continue to the next phase.

- **Instruction Phase**: In the instruction phase, participants were briefed about the game's goal. The instructions were provided by the game. If participants had any doubts regarding the game's goal, the experimenter would repeat the instructions. Both game and experimenter advised participants to pay attention to musical/visual cues. However, they were never told what these cues were, nor how and when would present themselves throughout the experience.

- **Play Phase**: Both music and visual versions have 10 levels; each level has 1 question and 4 possible answers, in which only one answer is correct. Hence, the probability in selecting the wrong answer is 75%, whereas selecting the correct answer is only 25%. By making some of the questions very challenging we made sure that participants would have to rely on the music/visual distortions at some point of the experience.

At the beginning of the game, participants had to listen to a question that was introduced verbally and in written form by the game. If participants did not interact with the device for more than 15 seconds, the volume of the background music would lower, and the system would repeat the question again. Next, participants had to select with the finger one of the four possible answers. While the finger presses an answer, a four-second countdown - in form of a pie chart - is triggered and displayed above the answer. During these four seconds, if the answer is wrong, musical/visual distortions are activated, allowing participants to reconsider their decision. If the participant would remove the finger from the answer before the countdown reached zero, he/she would be allowed to choose an alternative answer. When the countdown reached 0, participants received additional feedback confirming whether the answer was right or wrong (see example in Fig. 1). This feedback was presented as a human voice saying "Very Good" if correct, or "Ohhhh, Try again" if wrong.

E. Statistical Analysis

We used the IBM SPSS statistics version 24 (IBM, New York, United States of America) for the statistical data analysis. Non-parametric methods were used as the size of our sample is small. To compare the measurements of the
dependent variables in the two conditions, we used the Wilcoxon signed rank test.

F. Interviews

After the experience, we made a short semi-structured interview about the game. The main purpose was to gather additional information regarding the overall experience of the experiment. Because we were advised by health professionals against using a questionnaire with scales with these patients, we developed questions based on Yes/No answers with follow up questions. We asked participants if they (1) were generally satisfied with the activity, (2) recognized the music playing in the background, (3) felt active during the experience, and (4) perceived the musical and visual distortion during the game. All responses were voice recorded and later transcribed.

<table>
<thead>
<tr>
<th>TABLE II. DEPENDENT VARIABLE MEASUREMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Total number of errors committed in both conditions by all the participants</td>
</tr>
<tr>
<td>Total number of partial answers committed in both condition by all the participants</td>
</tr>
<tr>
<td>Total time (seconds) needed to complete both conditions by all the participants</td>
</tr>
</tbody>
</table>

Fig. 2. Boxplots of (A) Total time spent in the music condition (MD) and visual condition (VD), and (B) Number of partial answers in the music condition (MD) and visual condition (VD). * indicates a statistical significant difference between conditions (p < 0.05).

III. RESULTS

All seven participants finished the experimental trial. We compared (1) the number of total errors, (2) the number of partial answers and (3) the overall time spent by the participants in each condition. Fig 2. illustrates the total time spent by each participant, and the partial answers made by the participants in both music (MD) and visual (VD) distortion conditions. Performance metrics in both conditions are presented in Table II. For the following analysis, we considered the total number of errors, partial answers and total time required to solve the task in each condition.

A. Performance metrics

We extracted 3 different metrics that are used as the dependent variables in our experiment:

- **Errors**: we measured the total errors per participant in both music and visual condition in order to evaluate the performance. We have an error if these two conditions are fulfilled: (1) the answer is incorrect and (2) the selection countdown timer reaches 0.

- **Partial Answers**: to evaluate which feedback system was more effective in preventing errors, we measured the count of partial answers per participant in both music and visual condition. We have a partial answer if the participant removes the finger from any answer before the countdown reaches 0.

- **Time**: we measured the total time (seconds) spent by each participant during both music and visual conditions. We accumulated the times for all levels of each condition,
starting at the beginning of each level and stopping when participants completed the selection of the right answer.

The Wilcoxon Signed ranks test indicated that the total time necessary to complete all 10 questions in the MD condition (Mdn = 231.21 sec.) was significantly lower compared to the VD condition (Mdn = 315.627 sec.), $Z = -2.197$, $p = 0.028$, $r = -0.6$. However, the difference in the total number of errors made in the MD condition (Mdn = 7.0) compared to the VD condition (Mdn = 10.0) was non-significant, $Z = -1.792$, $p = 0.073$, $r = -0.5$. Interestingly, the number of partial answers made in the MD condition (Mdn = 3.0) was significantly lower when compared to the VD condition (Mdn = 9.0), $Z = -2.207$, $p = 0.027$, $r = -0.6$.

**B. Decision making process**

In order to assess differences in decision making due to the provided feedback, we computed the transition probabilities from game state to game state for both feedback conditions. Fig. 3 presents all game states as a Markov chain, showing the probabilities of patients’ actions depending in which state they are. For instance, Fig 3A shows us that the probability of patients of choosing a correct answer as first choice in the music distortion is of 46%. Similarly, when patients have previously selected a partial answer, with a 53% of probability they will perform a partial answer in the next action. Hence, this analysis allows interpreting patients’ behaviour in terms of probability of their choices in a dynamical system.

Overall, we can see that the probability of answering correct right away varies between 30-46%, of making a mistake varies between 21-26% and of partially selecting an answer is significantly higher in the VD condition (Mdn = 4.0) with 49% than in the MD condition (Mdn = 2.0) with 28%, $Z = -2.271$, $p = 0.023$, $r = -0.5$. Interestingly, we can observe that when a partial selection is performed, the probability of performing another partial selection is of 53%, a behaviour that is consistent with a search strategy exploiting the presented feedback.

**C. Interviews**

When asked about their general satisfaction of the activity, all participants said that they liked the experience. When asked whether they were able to recognize the music that was playing in the background, participants number 1, 2, 5, 7 and 8 were able to remember the song’s name. In addition, we asked them how active they felt (if they thought the game was difficult, easy or confusing). In general, the participants reported that the overall experience was easy and not confusing. Yet, participants number 1 and 7 felt sometimes nervous when more difficult questions came up. Finally, we asked them if they noticed the musical and visual distortions. Interestingly, only participant number 3 said to have noticed the distortions and was able to explain the relationship between distortions and wrong answers. Participant 5 recognized the distortions in both conditions, but was not able to understand the relationship between the potential errors and the musical/visual distortions. In addition, participant 1 noticed and understood the meaning of the distortions in the
music condition but not in the visual condition, while participant 7 noticed the distortion of the images but not the distortion of the music. The rest of the participants (2, 6 and 8) did not notice any distortions in either condition.

IV. DISCUSSION

Data gathered in this study support our initial research hypothesis, as the general performance of the subjects in the music distortion condition was better than in the visual distortion condition. Despite the small sample size, we observed statistical differences in terms of time and partial answers, but not in terms of total errors committed. Moreover, the number of total errors in the visual distortion condition was also higher than in the music distortion condition, although this difference was not statistically significant. Thus, from this evidence we can conclude that, in this particular experiment, visual distortion was not as effective as a guidance system compared to the musical distortion.

Despite the heterogeneity of our sample in terms of types of dementia and cognitive stages, when we analyze individual cases we can observe that only participant 3 was faster in the visual distortion condition than in the music distortion condition. In terms of total errors accumulated, participant 3 made more errors in the musical distortion condition than in the visual condition. Participant 6 had the same performance in both conditions. Interestingly, regarding the number of partial answers, participant 6 made more partial answers in the visual distortion condition, while participant 2 tied. Hence, despite a clear statistical preference in favour of music, our data also suggest that this may be participant dependent. We also analysed the participant's general behaviour through a Markov diagram. Although there is only a single significant statistical difference in decision making process between conditions, participants in the MD condition had higher chance to go from a partial answer to a correct answer when compared to their performance in the VD condition. In some cases, participants were exploring all possible answers before tapping the correct answer, evidenced by a very high probability of consecutive partial answers. Thus, one possible interpretation of such behaviour is that participants were developing strategies in order to find the right answer. In addition, we also noticed that participants in the VD condition, went from partial answer to wrong answer more often than in the MD condition. A possible explanation for the lower effectiveness of the visual condition could be that the visual distortions were too intrusive, which could lead to confusion. Nevertheless, our observations indicate that sometimes participants also ignored the distortions in both conditions and selected wrong answers.

During the interview, participants reported that they did enjoy the experience and that they were, in general, comfortable in playing the game. Although two participants reported feeling anxious when more challenging questions appeared, they were still capable of continuing the game. The majority was able to remember the music correctly. In fact, the majority sang the lyrics of the song during the experiment. Nevertheless, when asked about the distortions, we were surprised by the fact that only participant number 3 was able to recognize and understand the meaning of the musical and visual condition. This participant was one of the two that were not able to recognize the name of the song, but still paid attention to its distortions. It should be noted that this patient, with frontotemporal dementia, had language impairments. Sometimes this required the assistance of the health professionals to confirm his responses. Some participants partially perceived the distortions. While participant 1 did perceive and understand the meaning of the distortions in the music condition, participant 7 only perceived distortions in the visual condition. The latter, when asked about her interpretation of such distortions, replied "It was shaking (distorted) because it was alive".

It is also possible that explicit memory impairments could have interfered in the recall during the interview. Explicit memory is "the ability to consciously and directly recall or organize recently processed information" [20]. However, it is also possible that implicit memory - which "reflects the unconscious effects of previous experiences on subsequent task performance, without conscious recollection" [20] - was used in our game. Thus, one possible explanation could be that participants exploited the information gathered by the distortions, but in an unconscious manner. This would be supported by some of our observations. There were moments in which participants were clearly exploring alternative answers before tapping the correct one without committing any errors, just partial selections. One particular case, participant 7, sometimes clearly reacted to either of the distortions presented in the musical and visual distortion conditions. Patient 7 reported "I don’t understand why this image is shaking". However, the opposite was also true; there were moments in the game where participants completely ignored the distortions and committed errors.

Despite its findings, this study has some limitations. Firstly, the sample size is limited to 7 participants. Secondly, we tested our hypothesis on a heterogeneous group with different types and stages of dementia. Consequently, performance among participants was also heterogeneous. Thirdly, we noticed that some participants were sometimes confused with the audio instructions of the prototype. Whenever a participant tapped a wrong answer, the audio feedback said "Ohhhhhhh, try again". However, some participants thought they had to re-try again to select the same answer. Although these were excluded from the analysis, we cannot be certain that the above stated limitations did not bias our results.

V. CONCLUSION AND FUTURE WORK

As far as we know, this is the first time that the advantages of music distortion feedback is studied in individuals suffering from dementia. Despite the heterogeneity of our sample, the results of the experiment are encouraging. Overall, participants were more efficient in the music distortion condition than in the visual condition in all measured parameters. However, due to some limitations of this study, further research is necessary to better understand the feasibility of the proposed approach. To this end, different strategies should be taken into consideration to adapt the feedback to participants. For example, instead of adding white noise to a current sound, we plan to manipulate musical distortions by changing musical parameters such as speed, harmony or pitch. In addition, we
need to improve audio feedback to make sure it cannot be misinterpreted.

ACKNOWLEDGMENT

This work is supported by the Portuguese Foundation for Science and Technology through the Augmented Human Assistance project (CMUP-ERJ/HCI/0046/2013), Projecto Estratégico - LA 9 - 2015-2016 and the ARDITI Agência Regional para o Desenvolvimento da Investigação, Tecnologia e Inovação) institution. We also want to thank the Portuguese Alzheimer Association, health professionals and participants for their availability and motivation during our visit. In addition, we would like to thank Ana Lucia Faria for contributions to the methodology and for the help provided in the field study in the Portuguese Alzheimer Association.

REFERENCES