The benefits of emotional stimuli in a virtual reality cognitive and motor rehabilitation task
Assessing the impact of positive, negative and neutral stimuli with stroke patients

Ana Lúcia Faria
Madeira-ITI
Universidade de Coimbra
Coimbra, Portugal
ana.faria@m-iti.org

Mónica S. Cameirão
Madeira-ITI
Universidade da Madeira
Funchal, Portugal
monica.cameirao@m-iti.org

Teresa Paulino
Madeira-ITI
Funchal, Portugal
teresa.paulino@m-iti.org

Sergi Bermudez i Badia
Madeira-ITI
Universidade da Madeira
Funchal, Portugal
sergi.bermudez@m-iti.org

Abstract—VR-based methods for stroke rehabilitation have mainly focused on motor rehabilitation, but there is increasing interest towards the integration of cognitive training for providing more ecologically valid solutions. However, more studies are needed, especially in the definition of which type of content should be used in the design of these tools. One possibility is the use of emotional stimuli, which are known to enhance attentional processes. According to the Socio-emotional Selectivity Theory, as people age, this emotional salience arises for positive and neutral, but not for negative stimuli. Conversely, negative stimuli can be better remembered. In this study, we investigated the impact of using emotional stimuli with positive, negative and neutral valence in a VR cognitive and motor attention task. Ten stroke patients participated in a within-subjects experiment with four conditions based on the type of stimuli: abstract (control condition), positive, negative and neutral. The main task consisted of finding a target stimulus, shown for only two seconds, among fourteen neutral distractors. Eye movements were recorded with an eye-tracking system to investigate differences between conditions and in search patterns. Subsequently, a recall task took place and the patients had to identify all the target images among a valence-matched number of distractors. Our results corroborate the attention salience effect of positive and neutral stimuli in the VR task performance. Although we found no statistically significant differences between conditions in the recall task, there was a trend for recalling more negative images. This negative advantage comes at the expense of significantly more wrongly identified images/false memories for negative stimuli. Finally, we performed an analysis in which we relate performance scores with well-established cognitive assessment instruments, which supports the use of this approach both for assessment and rehabilitation purposes.

Keywords— emotional stimuli; stroke; virtual reality; cognitive and motor rehabilitation

I. INTRODUCTION

According to the World Health Organization, fifteen million people worldwide suffer a stroke each year, leaving 5 million survivors permanently disabled [1]. For those who survive, reducing the impact of post-stroke impairments is a major goal [2]. It has been estimated that more than 70% of individuals experience some degree of cognitive decline in the first few weeks following stroke and that more than one third remain cognitively impaired even 1 year post stroke [3]. These cognitive impairments have a direct influence on patients’ quality of life, being associated with greater rates of institutionalization [4], higher health-care costs [5] and reduction of the ability to understand task instructions or plan self-directed activities during the motor rehabilitation process. Among the most frequent sequelae, post-stroke patients commonly present decreased executive functioning, mental slowing, and impairment of goal formulation, initiation, planning, organizing, sequencing, executing, abstracting, and attention [2] being at risk of developing dementia [3]. Due to the role of stroke in the development of cognitive impairment, dependence in Activities of Daily Living (ADL) and dementia [6], early detection and intervention should be a target. Although several screening tools are available, specific deficits are only detectable with more complete neuropsychological assessments, which are rarely performed [7]. Consequently, there is a need to develop and employ more sensitive assessment tools combined with intensive rehabilitation programs based on customized treatments [8].

Virtual Reality (VR) as a rehabilitation tool

Traditional rehabilitation is labor and resource-intensive and at times can be demotivating due to its repetitiveness [9]. In contrast, enriched training tasks based in VR simulations can be used to provide meaningful and repetitive practice with augmented feedback [10]. Such enriched virtual environments have the potential to optimize rehabilitation by manipulating practice conditions that explicitly engage motivational, cognitive, motor control and sensory feedback-based learning mechanisms [11]. Moreover, training of cognitive and motor domains by means of VR approaches is gaining clinical acceptance in the therapy of patients post-stroke because of the motivational aspects as well as the capacity to promote sustained movement practice [12]. However, most VR rehabilitation approaches are generally dedicated either to motor or cognitive rehabilitation aspects. Nevertheless, given the dual motor and cognitive components of ADL, a combined motor and cognitive VR approach could provide a more ecologically valid training [13], [14].
The role of emotional stimuli in rehabilitation

Despite the wealth of evidence concerning the value of VR for rehabilitation of stroke patients [10], there is surprisingly little research about the type of content being used (neutral, abstract, meaningless, emotional, etc). From the literature we know that emotional stimuli are remembered better and more vividly than non-emotional stimuli [15]. This phenomenon, known as the emotional enhancement of memory, has been replicated across a range of paradigms and stimulus types. These emotional enhancements in memory are, at least partly, due to the increased attention directed toward emotional items at encoding [16]. However, not all aspects of memory are enhanced by emotionality patterns. Emotional items are often remembered at the expense of their contexts, this is, peripheral features of visual scenes are remembered less when an emotional item is present in the scene than when only non-emotional items are present [17]. The rationale behind this phenomenon is that emotional items with a high valence tend to capture attention and to get prioritized in the processing chain [18]. Since more attentional resources are directed towards the emotional components, people seem more likely to encode the emotional components of the scene and less likely to encode neutral contextual information. However, the processing of emotional stimuli seems to be also affected by age. In fact, the Socio-emotional Selectivity Theory [19] states that there is an age-associated motivational shift towards emotional goals. It states that when emotional material is attended to, is weighed more heavily, processed more deeply, and better remembered than non-emotional material. In addition, recent evidence shows that the recall of emotional information is disproportionately positive as people age. When recalling previously presented positive, negative, and neutral images, the recognition and recall of negative images declines linearly with age across younger, middle-aged, and older adults [20]. These results are consistent with eye-tracking research that showed that when a negative and a neutral picture are displayed together, both young and older adults initially glance at the negative picture but young adults looked for a longer time at the negative picture [21]. The same study found no age differences when the two pictures were positive and neutral. It has also been found that emotional stimuli can also induce more false memories than non-emotional stimuli [22]. This apparent paradox, that negative emotions can simultaneously improve and impair memory by its high valence, has been repeatedly found in memory recall experiments [23]. One hypothesis is that negative valence causes a narrowing of attention such that spatial and temporal information associated with the emotional item are better attended to and later remembered, while peripheral information is likely to be forgotten. An example is the weapon focus effect, where there is enhanced memory for a weapon in a scene but reduced memory for details of the background [25]. This focus may lead to selective memory for emotional components [24], although some researchers argue that selective memory of emotional content is not strongly related to attention at encoding [26].

The work presented here has multiple objectives: 1) addresses the need of having more specific and accessible assessment and rehabilitation tools that overcome some of the limitations of traditional methods; 2) proposes a VR rehabilitation task that integrates both cognitive and motor domains for increased ecological validity; and 3) investigates the impact of emotional stimuli in the assessment and performance of motor-cognitive rehabilitation training. That is, to determine how valence of stimuli affects task performance (attention) and recall (memory).

II. METHODS

A. Participants

Participants were recruited at the Nélio Mendonça and João Almada Hospitals (Madeira Health Service, Portugal), based on the following inclusion criteria: ischemic stroke; no vision problems; capacity to be seated; non-aphasic and with sufficient cognitive ability to understand the task instructions as assessed by the clinicians. The sample consisted of ten (7 female, 3 male) middle-aged (M=54.2 years old, SD=9.2) stroke patients (1 right hemisphere, 8 left hemisphere and 1 cerebellum), with a mean of 16.6 ± 19.5 months post-stroke, with a mean schooling of 8.1 ± 5.8 years. 60% had some computer literacy. The study was approved by the Madeira Health Service Ethical Committee and all the participants gave previous informed consent.

B. Materials

Emotional stimuli pictures were selected from the International Affective Picture System (IAPS) [27]. This widely used picture set consists of photographs of people, animals, objects, and scenes that have been originally rated for valence and arousal through the Manikin self-assessment, which depicts nine values along the dimensions of affective valence (ranging from unpleasant to pleasant) and arousal (ranging from calm to excited). 182 images were selected and allocated to the VR and recall tasks. The categorization of the images as positive, negative and neutral was based on the valence and arousal standard scores provided by the IAPS. Our final selection of neutral, negative and positive images had valence scores in the range of 4.5-5.5, 1.66-2.58 and 7.53-8.34, respectively. A Friedman test confirmed that valences were significantly different across conditions ($\chi^2(2)=28.0$, p<0.001). Arousal of the images was kept neutral with scores between 4.5 and 5.5.

The VR cognitive and motor task was designed as an adaptation in VR of the Toulouse Piéron (TP) task (TP-VR) [13]. This task was extended to incorporate also emotional stimuli as targets. In this VR task, a target stimulus is presented in the center of the screen for 2 seconds. Immediately after, in a 3D environment, fifteen cubes with images are displayed in a 3x5 grid structure on top of a table (Figure 1).
From these fifteen images, one is the target and fourteen are distractors randomly selected from a set of 92 neutral images (in case of emotional stimuli) or from the set of 8 TP symbols. Through a virtual representation of the arm, the patient has to select the target cube. When the hand is placed over a cube, highlighting it, a 6 seconds timer is activated; the selection is completed when these 6 seconds have elapsed. The interaction with the computer was made through 2D arm movements with a camera-based color tracking software (AnTS) [28]. The VR environment has a built-in calibration function that is able to compute the active range of motion of the patient, normalizing the motor effort required to the skill set of the patient [29].

The trial ends with a computerized version of the self-assessment Manikin, in which the patient indicates how he/she felt about the target image rating it from 1 to 9, where 1 is very sad and 9 is very happy. Eye movements were recorded using the Tobii® EyeX eye-tracking system.

C. Protocol

Subjects were seated at a 40 cm distance from a 22” high-resolution display monitor (1920x1068 pixels). Before starting the experiment, participants went through an average of five training trials only with abstract stimuli (TP). The training was intended to provide a clear understanding of the VR task and valence rating, as well as to get used to the natural user interface (AnTS) [30]. A within-subjects design was used with four experimental conditions corresponding to 4 types of stimulus (positive, negative, neutral, and abstract). The abstract and meaningless stimuli of the TP task were used as control condition, since this is a well-established attention assessment tool. The experiment entailed 56 trials with the 4 different conditions being presented in a random order. Performance, time to completion, ratings and eye-tracking data were saved in CSV files.

A recall test took place 30 minutes after the end of the VR task. It included 90 images from which 30 had a positive valence (14 targets + 16 distractors), 30 had a negative valence (14 targets + 16 distractors) and 30 had a neutral valence (14 targets + 16 distractors) (Figure 2).

Images were randomized and distributed within ten A3 pages (9 images per page). Participants indicated the ones they thought they had previously seen in the VR task (Figure 3).

D. Assessment

In order to have a comprehensive cognitive profile (Table 1), a general clinical assessment took place after the VR experiment. The cognitive screening was made through the Montreal Cognitive Assessment (MoCA) [31], which provides a total score and subscores for the following domains: Executive Functions, Naming, Attention, Language, Reasoning, Memory and Orientation. The Line Bisection (LB) test [32] was used to assess possible neglect. A short version of the Toulouse Piéron (TP) test [33] was also included to compare the performance on paper-and-pencil and the performance on the virtual environment. To assess stroke functional impairments we used the following subscales of the Stroke Impact Scale 3.0 (SIS v3.0) [34]: Arm, Hand, Cognitive, Communication and Recovery. Acknowledging the effects that different states of mood (e.g. depression) can have on cognition, we assessed depressive symptomology through the Geriatric Depression Scale – 30 (GDS-30) [35]. The tests were delivered in the following order: screening scale, LB, TP, SIS 3.0 and GDS-30.

**TABLE I. CLINICAL PROFILE OF THE PARTICIPANTS**

<table>
<thead>
<tr>
<th>Test</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Montreal Cognitive Assessment</td>
<td>17.40</td>
<td>6.38</td>
</tr>
<tr>
<td>Line Bisection</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Toulouse Piéron</td>
<td>58</td>
<td>33.93</td>
</tr>
<tr>
<td>Stroke Impact Scale - arm</td>
<td>53.60</td>
<td>24.17</td>
</tr>
<tr>
<td>Stroke Impact Scale - hand</td>
<td>41.40</td>
<td>34.88</td>
</tr>
<tr>
<td>Stroke Impact Scale - cognitive</td>
<td>78.50</td>
<td>13.08</td>
</tr>
<tr>
<td>Stroke Impact Scale - communication</td>
<td>78.40</td>
<td>25.36</td>
</tr>
<tr>
<td>Stroke Impact Scale - recovery</td>
<td>61.10</td>
<td>23.87</td>
</tr>
<tr>
<td>Geriatric Depression Scale - 30</td>
<td>8.80</td>
<td>6.37</td>
</tr>
</tbody>
</table>
E. Data analysis

From the raw data, we extracted for each condition the average task completion time, average length of the eye-gaze search (in pixels), % of correct target selection and the mean valence of the images as rated by the participants. For each variable, the normality of the distribution was assessed using the one sample Kolmogrov-Smirnov test. Because most distributions deviated from normality, non-parametric statistical tests were used for analysis. For assessing the overall difference between experimental conditions, a Friedman test was used. For pairwise comparisons, the Wilcoxon’s T matched pairs signed ranks test was used. The Spearman ρ was used to search for meaningful correlations between variables. Data were analyzed using Matlab and the Statistical Package for the Social Sciences 20 (SPSS.20).

III. RESULTS

A. Is performance on paper equivalent to VR?

In the original TP paper-and-pencil version, targets are always visible during performance. The paradigm was slightly simplified in the VR version to address the memory component and make it compatible with the presentation of emotional pictures (there is only one target symbol per trial, which is only visible for 2 seconds). Consequently, a comparison of the TP paper-and-pencil performance (M=58.00%, SD=33.93) with the TP-VR performance (M=80.72%, SD=18.76) revealed significantly better performance in VR when compared to traditional TP (W_{10}=40, Z=-2.073, p:<.05) (Figure 4).

B. Does emotional content affect task performance?

In order to assess the impact of using emotional content in the VR task, we compared the performance in the 4 conditions in terms of the mean percentage of correct responses and completion times. In terms of performance, a Friedman’s test revealed significant differences across the positive (M=87.51%, SD=12.05), negative (M=66.28%, SD=32.09) and neutral (M=88.96%, SD=8.59) valence stimuli (χ²(2) = 6.0, p =.050) (Figure 5). Specifically, pairwise comparisons showed a lower performance for negative stimuli as compared to positive (W_{10}=15.0, Z=-2.023, p<.05) and neutral (W_{10}=20.0, Z=-1.997, p<.05). A comparison of performance between neutral and abstract (TP) stimuli showed no significant differences (W_{10}=6.00, Z=-1.682, p=.93).

Concerning the time to completion, no significant differences were found across positive (M=15.45 sec, SD=5.17), negative (M=16.11 sec, SD=6.94) and neutral (M=16.91 sec, SD=6.50) stimuli conditions. Consistent with these data, no differences were found in the eye pattern search length and dispersion across the 4 conditions.

C. Does emotional content impact memory recall?

Concerning the Recall Task, we did not find significant differences between the overall recall performance for each emotional stimuli condition (χ²(2) =.000, p =.1000). However, there was a trend for recalling better negative stimuli (M=70.87%, SD=25.07) than positive (M=67.62%, SD=24.99) or neutral (M=67.29%, SD=34.03). Interestingly, when considering the wrongly identified images (false memories) we found significant differences (χ²(2)= 7.583, p <.05), being the number of false memories for negative stimuli (M=10.645, SD=14.46) significantly higher than for neutral (M=2.510, SD=6.06; W_{10}=.000, Z=-2.207, p<.05) but not for positive (M=7.51, SD=8.24) (Figure 6).
This means that cognitive state (MoCA) and emotional content (positive, negative and neutral) by our patients than those originally provided by IAPS.

D. Do the patients’ ratings of the images differ from those of the original IAPS ratings?

Overall, there was no significant difference between the original IAPS ratings ($M=5.024$, $SD=2.37$) and the participants reported ratings ($M=5.012$, $SD=1.78$). Nevertheless, when we analyzed separately the 3 emotional stimuli, we found significant differences between the ratings for the positive (IAPS=7.85, Patients=6.68; $W_{14}=1.000$, $Z=-3.233$, $p=.001$) and negative (IAPS=2.15, Patients=3.47; $W_{14}=99.000$, $Z=2.919$, $p<.01$), but not for neutral stimuli. This indicates a more moderate rating of the emotional content in the images by our patients than those originally provided by IAPS.

E. How does the patient profile affect their performance in the cognitive-motor task with emotional content?

A regression model was computed for task performance ($F(2,7)=10.721$, $p=.007$). According to the model, performance can be predicted on the basis of the overall cognitive state (MoCA) and depressive symptomology (GDS-30) with an $R^2$ of .754:

$$\text{Performance} = 69.519 + 1.689^*(\text{MoCA}) - 11.999^*(\text{GDS-30})$$

This means that in average, participants’ performance increased 1.683 for each MoCA point and decreased approximately 12 for each GDS-30 point. In summary, a less depressive symptomology and a better cognitive status lead to higher overall performance results in the VR task (all conditions).

When we perform a correlation analysis, we can observe a general positive effect of MoCA and its subdomains in various metrics of the VR task performance (Table II).

The overall performance with emotional stimulus (excluding TP) correlates with the MoCA total score ($r=.756$, $p=.011$) and all its subdomains with the exception of Language and Reasoning, what makes task performance a very representative evaluation of the overall cognitive functioning. Moreover, this analysis also reveals that task performance with positive stimulus does exclusively correlate with the Attention ($r=.821$, $p=.50$) and Memory ($r=.820$, $p=.004$) subdomains. Finally, we found that Memory ability does not seem to affect performance when using negative stimuli. Regarding the recall test, we have found that only the recall of neutral stimulus seems to be a good predictor of the MoCA total score ($r=.634$, $p=.049$). Also, higher Memory scores correlate with improved neutral ($r=.822$, $p=.004$) and negative ($r=.681$, $p=.030$) recall performance with less overall errors ($r=.699$, $p=.024$), and specifically for negative stimuli ($r=.639$, $p=.047$). Overall recall errors also correlate negatively with the Executive Functions score ($r=-.685$, $p=.029$). Finally, depression, as assessed by GDS-30, correlates exclusively with the amount of errors obtained for neutral stimuli ($r=-.667$, $p=.035$).

IV. DISCUSSION AND CONCLUSION

Here we presented a novel VR rehabilitation task that integrates both cognitive and motor domains, designed to investigate the impact of emotional stimuli in the assessment and performance of motor-cognitive rehabilitation training. In this study, we assessed the impact of emotional stimuli in a group of ten stroke patients. We used 4 types of stimuli: abstract (TPVR), positive, negative and neutral. The abstract condition performance, which entailed the TP symbols, was compared with its paper-and-pencil counterpart, revealing a

<table>
<thead>
<tr>
<th>$r$</th>
<th>VR Task Performance</th>
<th>Recall Task Performance</th>
<th>Recall Task Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Perf TP</td>
<td>Perf total</td>
<td>Perf neg</td>
</tr>
<tr>
<td>MoCA total</td>
<td>.667*</td>
<td>.756*</td>
<td>.704*</td>
</tr>
<tr>
<td>MoCA: EF</td>
<td></td>
<td>.644*</td>
<td>.718*</td>
</tr>
<tr>
<td>MoCA: Nam</td>
<td>.652*</td>
<td>.670*</td>
<td>.639*</td>
</tr>
<tr>
<td>MoCA: Attent</td>
<td>.633*</td>
<td>.821**</td>
<td>.771*</td>
</tr>
<tr>
<td>MoCA: Mem</td>
<td>.787**</td>
<td>.698*</td>
<td>.820**</td>
</tr>
<tr>
<td>MoCA: Orien</td>
<td></td>
<td>.721*</td>
<td>.662*</td>
</tr>
<tr>
<td>GDS</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
higher performance in the TPVR version. This difference can be explained by a simpler task (with only one target) and the fact that the used natural user interface facilitates the interaction between the patient and VR environment, eliminating some of the constraints of a paper-and-pencil task.

The data from the emotional stimuli conditions were analyzed separately for the performance in the VR task, and for a later recall test. Performance data is consistent with the Socioemotional Selectivity Theory, showing that attention processes are less targeted to negative information in older adults. Indeed, the performance for neutral and positive stimuli was significantly better than for negative. This finding, leads us to consider that positive and neutral content might be better for attention rehabilitation in this population. The time for completion and eye search pattern data were not affected by the emotional stimulus conditions. The recall test results corroborate the study by Steimetz and Kensinger [26] which shows that selective memory for emotional information is not strongly related to attention at encoding. In fact, although we have observed a positive effect in the task performance itself with neutral and positive images, this finding was not replicated in the recall, where we did not find significant differences between positive, negative and neutral stimuli. Although it would be expectable that performing better for positive and neutral stimuli would also provide a better encoding, there was a trend for better recall in the negative images (the ones that provided worse performance results).

From our data we may conclude that negative stimuli had a salient effect but only at recall. Nevertheless, when we analyzed the wrongly identified images, we found that the negative distractors lead to significantly more false memories. This finding is also consistent with the literature, with negative valence content causing a narrowing of attention, with individuals disregarding details. This could also be observed in our study in some cases, when having an image of a starving person as a target and then having two different images of a starving person in the recall task, the participant tended to select both, even if completely different. Overall, our findings lead us to conclude that using both positive and neutral stimuli will provide better results in the training of attention. Yet, if we want increase task difficulty or train memory, negative stimuli should also be considered.

The overall VR task performance was found to be a good metric for cognitive status, being it proportional to the MoCA assessment and to the absence of depressive symptomology, as assessed by GDS-30. This finding takes us to consider that, besides a rehabilitation purpose, this task may also be valuable in providing information about the patient’s cognitive status. Finally, the relationship found between performance and the MoCA assessment gives us indications on how this results can be extrapolated to cognitive outcomes. More specifically, performance with positive stimuli could be used to specifically assess attention and memory domains as well as overall recall errors can be used to assess Executive Functions. The recall of neutral stimulus correlated significantly with MoCA total score and more strongly, with the Memory domain. This also means that recalling positive and negative valence content is an easier task, regardless of the MoCA score. Finally, higher recall errors for neutral stimuli could be related to a mood disorder. To conclude, in this study we have contributed towards improving assessment and rehabilitation through the use of emotional content on a VR training task. We showed how emotional content modulates performance in an attentional task, and provided guidelines on how emotional stimulus can also be used in the context of a VR task to assess different domains of cognitive performance.

ACKNOWLEDGMENTS

This work is supported by the European Commission through the RehabNet project - Neuroscience Based Interactive Systems for Motor Rehabilitation - EC (303891 RehabNet FP7-PEOPLE-2011-CIG), and by the Fundação para a Ciência e Tecnologia (Portuguese Foundation for Science and Technology) through SFRH/BD/97117/2013 Projeto Estratégico - LA 9 - 2015-2016 and Agência Regional para o Desenvolvimento da Investigação, Tecnologia e Inovação (ARDITI), (http://neurorehabilitation.m-iti.org).

REFERENCES


