



# The ObReco-360°: a new ecological tool to memory assessment using 360° immersive technology

Luca Pieri<sup>1,2</sup> · Silvia Serino<sup>3</sup> · Pietro Cipresso<sup>4</sup> · Valentina Mancuso<sup>5</sup> · Giuseppe Riva<sup>3,6</sup> · Elisa Pedrolì<sup>5,6</sup>

Received: 23 September 2020 / Accepted: 7 April 2021 / Published online: 27 April 2021  
© The Author(s) 2021

## Abstract

One important feature of a neuropsychological test is its ecological validity, which defines how much patients' test scores are linked to real-life functioning. However, many of the currently available neuropsychological tools show low to moderate levels of ecological validity. Virtual reality (VR) emerged as a possible solution that might enhance the ecological value of standard paper-and-pencil tests, thanks to the possibility of simulating realistic environments and situations where patients can behave as they do in real life. Moreover, a recent kind of virtual environments, the 360° spherical photos and videos, seems to guarantee high levels of graphical realism and lower technical complexity than standard VR, despite their limitations concerning interactive design. In this pilot study, we tested the possible application of 360° technology for the assessment of memory, developing an adaptation of a standardized test. We focused on Free Recall and Recognition accuracies as indexes of memory function, confronting and correlating the performances obtained by the participants in the standard and in the 360° test. The results, even if preliminary, support the use of 360° technology for enhancing the ecological value of standard memory assessment tests.

**Keywords** Memory · Assessment · 360° video · Object recognition · Virtual reality

## 1 Introduction

In recent years, many criticisms have been raised regarding the ecological validity of the tools used in the neuropsychological assessment of cognitive functions. The construct “ecological validity” refers to the functional and predictive relationship between patients' performance on neuropsychological tests and real-life cognitive functioning (Sbordone

et al. 1996). Neuropsychological assessment is usually provided through psychometric tests in a paper-and-pencil modality: these tests require patients to perform different behavioral and cognitive tasks in a controlled setting, determining the evaluation of abstracts concepts, without a direct link to the ecological behavior (Parsons 2015; Parsons et al. 2017; Serino and Repetto 2018). Hence, there is an ongoing debate about the effectiveness of psychometric tests in assessing life-like abilities in their natural environment (Chaytor and Schmitter-Edgecombe 2003). In fact, neuropsychological tests are not always sensitive, as patients could present scores in the normal range in the clinical setting but show difficulties in daily situations, or vice versa (Mondini et al. 2016). Parsons (2015) argued that this discrepancy can be traced back to the different methodologies used for developing the assessment procedures. Indeed, the most widely used neuropsychological tests were conceived following a construct-led approach, aiming to measure abstract cognitive constructs (e.g., working memory) without an explicit interest in predicting real-life functional abilities. On the contrary, function-led approaches provided an alternative framework focused on building neuropsychological models derived from the observation of everyday

✉ Luca Pieri  
l.pieri3@campus.unimib.it

<sup>1</sup> Department of Psychology, University of Milano-Bicocca, Milan, Italy

<sup>2</sup> Mind and Behavior Technological Center, University of Milano-Bicocca, Milan, Italy

<sup>3</sup> Department of Psychology, Catholic University of the Sacred Heart, Milan, Italy

<sup>4</sup> Department of Psychology, University of Turin, Turin, Italy

<sup>5</sup> Department of Psychology, eCampus University, Novedrate, Italy

<sup>6</sup> Applied Technology for Neuro-Psychology Lab, IRCCS Istituto Auxologico Italiano, Milan, Italy

behaviors, allowing the development of more ecological assessment procedures (e.g., the Rivermead Behavioral Memory Test by Wilson et al. 1989). Thus, from a construct-driven approach, research has recently been moving to a function-led approach, recognizing the importance of conducting the evaluation in an ecological way by observing the cognitive and behavioral functioning in a real-life context (Parsons et al. 2017; Serino and Repetto 2018).

To this regard, virtual reality (VR) technology was used to develop function-led tools for neuropsychological assessment, thanks to the capability of simulating realistic environments which can be used as complex stimuli for investigating individuals' cognitive abilities in a more ecological setting (Pedroli et al. 2015; Serino et al. 2015; Riva et al. 2019). The simulation is based on experiencing 3D computer-generated scenarios through a standard desktop interface (non-immersive VR) or in a first-perspective view thanks to a head-mounted display (HMD) (immersive VR). With a specific software (e.g., Unity 3D©), it is possible to develop virtual environments (VEs) which resemble daily life settings and activities (e.g., doing grocery) with a high level of plausibility (Parsons 2015; Rizzo and Koenig 2017). The twenty-first century has been characterized by the widespread use of this technology in both assessment and rehabilitation of neuropsychological deficits with promising results (Larson et al. 2014; Howard 2017, Pedroli et al. 2018; Moreno et al. 2019). In fact, VR allows researchers and clinicians to deliver highly ecological stimuli and collect measures which are very close to those observed in naturalistic settings, obtaining better prognostic indexes of real-life functioning in a safe and controlled situation. This is particularly true for immersive VR, which can dramatically increase the ecological value of the stimuli delivery thanks to the level of immersion provided and the elicited sense of presence. Specifically, immersivity relies on the use of specific technologies capable of providing multisensorial feedback, synchronizing virtual body actions to real movements and isolating users from the physical environment (Slater and Wilbur 1997). The sense of presence, instead, is the illusion of "being there" and is accompanied by the feeling of being involved, absorbed and captured by the virtual experience that could elicit strong emotional responses (Heeter et al. 1995; Chirico et al. 2017).

In this way, it is possible to build more complex tasks, evaluating both global functioning but also a specific function (Keefe et al. 2016). Several authors have exploited VR to develop technological tools to conduct a neuropsychological assessment (Parsons 2015; Negu et al. 2016; Rizzo and Koenig 2017). Matheis et al. (2007) pioneered the use of VR for memory evaluation developing a 3D virtual office in which patients had to learn and then recall some objects. Results showed that this protocol successfully discriminated between experimental and control groups,

also showing a positive correlation between scores in the virtual test and in the standard one. A more immersive tool for the neuropsychological evaluation of memory has been developed by Ouellet et al. (2018). In this test, subjects could navigate in a virtual supermarket in which they had to memorize the grocery list and look for that products. The authors then concluded that this task is a valid and flexible tool to measure memory functioning in an ecological life-like context.

One further possibility for neuropsychological assessment can be offered by a recently emerged kind of VEs: the 360° immersive photos and videos (i.e., 360°-VR). With 360° cameras is possible to record a circular fisheye view of the surroundings, which can be later experienced as an immersive VE using an HMD. This approach started to be implemented in neuropsychological assessment only in recent studies, showing promising results for what concerns the evaluation of executive functions (Serino et al. 2017; Realdon et al. 2019). When compared to VR tests including model-based VEs (i.e., computer-generated environments), the implementation of 360° immersive photos and videos does not require high-level technical skills to be mastered, and the equipment needed to record and then visualize 360° materials is also more affordable than standard VR set-up. Moreover, being captures of real-world scenarios, 360° environments also provide a higher visual realism, which can further increase the engagement of the participants. Notably, 360°-VR advantages are limited by the fact that currently this technology can be experienced only through a 3-degrees-of-freedom (3-DOF) exploration, which allows to track only the HMD orientation and rotation in the space. In other words, users can look at any portion of the VE but cannot move closer to the objects and interact explicitly with them (e.g., move them to another position). This limit in the interaction can reduce the overall ecological value provided by 360°-VR only to a visual fidelity level, and must be taken into account when developing such experiences. Thus, this kind of VEs appears more suitable for simulating all those situations which do not require a direct interaction with the scenario (e.g., a recognition task).

The simpler design which characterizes 360°-VR environments (i.e., limited interaction), hardware (e.g., 3-DOF HMDs) and development, make it suitable also for the neuropsychological assessment of patients with mild to severe impairments (Serino et al. 2017; Realdon et al. 2019), who can likely show difficulties when interacting with more sophisticated VR settings. Hence, 360°-VR can constitute a viable option to consider when thinking about VR-based test implementation, especially considering the cost–benefit value: For clinicians, this can mean having access to more sophisticated yet user-friendly and ecological assessment tools to support already existing techniques.

As regarding the memory assessment, with immersive 360°-VR scenarios people can perceive photorealistic environments in a first-person perspective view: This characteristic is very important for the assessment of the several aspects of memory and can improve the precision of the procedure (Matheis et al. 2007; Serino et al. 2017). Moreover, the photorealistic fashion which characterize the 360° VEs can further enhance their ecological value: In fact, Robertson et al. (2016) found that 360° videos resembling realistic natural environments provide a visual experience which is very close to the natural visual exploration of real-life scenarios. These results are also consistent with the literature finding that the degree of immersivity and realism affects memory coding processes (Sutcliffe et al. 2005; Slobounov et al. 2015; Makowski et al. 2017; Serino and Repetto 2018). In this regard, what makes 360°-VR a suitable tool to study memory functioning is (a) the possibility to adopt an immersive egocentric perspective when experiencing the 360° environments and (b) the high visual fidelity provided, linked to a better visual memory encoding (Robertson et al. 2016; Serino and Repetto 2018; Ventura et al. 2019). Then, given its capability to elicit visual exploration mechanisms similarly to those adopted in real environments, these new VEs appear to be promising for developing more ecological tools for the assessment of memory processes.

Thus, in the current pilot study, we wanted to evaluate the feasibility of a memory assessment protocol based on 360° VEs. During the development of our experimental protocol, we took inspiration from a specific task included in a traditional test, the Rivermead Behavioural Memory Test – III (RBMT-III) (Wilson et al. 2008). This test was built following a function-led approach, including several “real life” tasks (e.g., recalling details of a story or remembering names and faces) which provide a valid and ecological testing of different aspects of everyday memory (Smith et al. 2000; Efklides et al. 2002). Therefore, we considered the RBMT-III as an inspiration and a benchmark for the development of our 360° tool.

Secondly, we aimed to investigate the differences showed in the performances by the participants in our 360° tool and in the RBMT-III, calculating the correlations between the measures obtained with these two tests.

Finally, we performed an user experience (UX) evaluation to study the participants’ interaction with the used technology.

## 2 Materials and methods

In the current study, we explored the potential use of 360° video technology for memory assessment through a preliminary evaluation of a 360° adaptation of the Picture

Recognition sub-test included in the RBMT-III. We named this adaptation ObReco-360° (Object Recognition-360°).

### 2.1 Sample

The participants of the present study were enrolled among the outpatients coming from the Department of Medical Rehabilitation of Istituto Auxologico Italiano in Milan. The resulting sample of twenty-four people included nine females and fifteen males, with a mean age of 70.4 (SD = 8.5) and a mean of 9.7 (SD = 3.7) years of education. The exclusion criteria for the enrollment included the presence of severe internist, psychiatric and neurological impairments. Regarding the cognitive status, only the participants who obtained a corrected score above 18 points in the Mini Mental State Examination (MMSE) (Folstein et al. 1975) Italian Version (Measso et al. 1993) were considered for the recruitment.

The study was conducted in compliance with the Helsinki Declaration of 1975 (as revised in 2008) and received ethical approval by the Ethical Committee of the Istituto Auxologico Italiano. The demographic and neuropsychological data of the final sample are showed in Table 1.

### 2.2 Procedure

This pilot study involved randomized within-subject data collection. For this reason, participants were examined two times in a week to avoid learning effect and interferences between materials. Two different assessment protocols were administered: The standard one included only classic paper-and-pencil tests, while the 360° one also included the administration of the ObReco360° and two user-experience (UX) rating scales. During the 360° session all the participants were sitting on a turning chair, in order to freely explore the virtual environments using an Oculus Go© HMD.

**Table 1** The table shows the main demographic and neuropsychological data which characterize the sample of participants. The demographic indexes include age and years of education

Demographic and neuropsychological descriptives of the sample (N=24)					
	Min	Max	Mean	SD	Cut-Off
Age	58	87	70.3	8.5	//
Y.o.E	5	17	9.7	3.7	//
MMSE	20.7	30	26.1	2.3	18/30
FAB	8.7	18.4	14.2	2.7	13.5/18
BSRT	3.8	14.1	9.1	3	3.75/16

The neuropsychological data showed include the scores obtained in the Mini Mental State Examination, in the Frontal Assessment Battery and in the Babcock Story Recall Test

Y.O.E years of education, MMSE Mini Mental State Examination, FAB frontal assessment battery, BSRT Babcock story recall test

### 2.3 Neuropsychological assessment

The neuropsychological tests administered were the MMSE, the Frontal Assessment Battery (FAB) (Dubois et al. 2000) Italian Version (Appollonio et al. 2005), the Picture Recognition sub-test included in the RBMT-III Italian Version (Beschlin et al. 2013) and the Babcock Story Recall Test (BSRT) Italian Version (Spinnler and Tognoni 1987).

The MMSE is a brief screening test including thirty simple tasks (e.g., repeating and remembering words, copying a figure) oriented to a first-level assessment of different cognitive functions.

The FAB is a rapid test for the screening of executive functions which includes six tasks linked to frontal lobes activity, such as conceptualization (e.g., find common characteristics between objects) and inhibitory control (e.g., following rules given by the examiner).

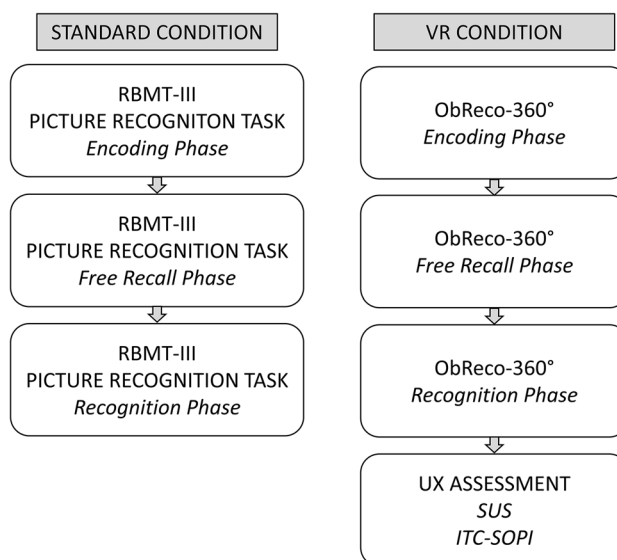
The Story Recall tests aim to assess short-term and long-term memory abilities, respectively, through an immediate or delayed recall of all the details contained in a brief tale.

### 2.4 RBMT-III picture recognition subtask

The Picture Recognition sub-test of the RBMT-III is divided in two phases. During the first phase (Encoding Phase), a set of 15 pictures representing common animate and inanimate objects (e.g., a clock, a chicken) are shown separately to the participants, who had to recognize and name each one of them. In the second phase (Recognition Phase), the participants must observe a total of 30 pictures including target items (i.e., the 15 pictures presented in the Encoding Phase) and distractors (i.e., 15 pictures non presented in the Encoding Phase): For each of these, they must answer yes if the picture was presented previously or no if it was not. The raw score obtained in the subtest is the number of pictures correctly recognized. In addition, before the Recognition Phase we included a Free Recall task, which required the participants to remember every object he/she could from those presented in the Encoding Phase. The raw score is defined by the number of object correctly reported. The flowchart of the procedure is presented in Fig. 1a.

### 2.5 UX measures

The UX assessment procedure included two questionnaires. The first instrument was the Independent Television Commission-Sense of Presence Inventory (ITC-SOPI) (Lessiter et al. 2001), a questionnaire including forty-four items which define a set of affirmations addressing the individual's feelings after the VR experience. Participants are asked to determine their degree of agreement with each of these affirmations using a five-points Likert scale ranging from "Strongly Agree" to "Strongly Disagree". The ITC-SOPI is



**Fig. 1** The figure shows the steps included in the two experimental conditions. The left diagram shows the steps included in the standard condition, while the left diagram reports the steps included in the VR condition. Abbreviations; RBMT-III, Rivermead Behavioral Memory Test-III; ObReco-360°, Object Recognition-360°; ITC-SOPI, Independent Test Commission-Sense of Presence Inventory; SUS, System Usability Scale

divided into 4 subscales: Sense of Physical Space (19 items), Engagement (13 items), Ecological Validity (5 items) and Negative Effects (6 items), each one linked a singular score.

The second instrument was the System Usability Scale (SUS) (Brooke et al. 1996), a questionnaire composed of ten sentences describing the user's feeling concerning the interaction with the product to evaluate. For each of these answer, the participant needs to define their degree of agreement using a five-points Likert scale ranging from "Strongly Agree" to "Strongly Disagree". The computed score ranges from 0 to 100.

### 2.6 ObReco-360°

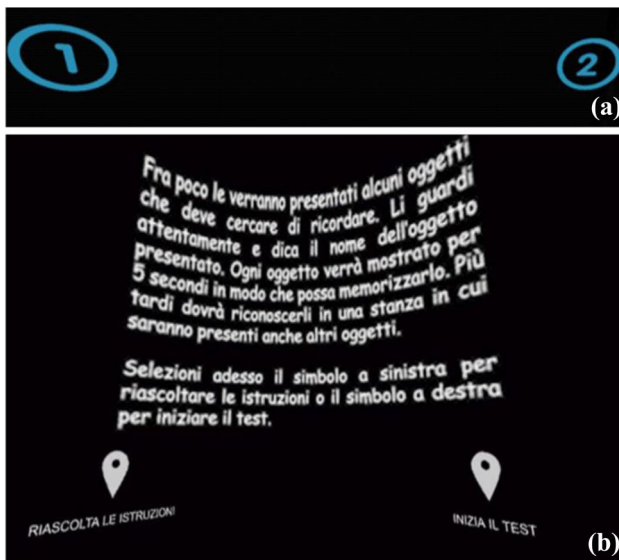
The ObReco-360° is a novel neuropsychological assessment tool, developed using 360° immersive photo and video as VEs and derived from the Picture Recognition sub-test of the RBMT-III. The included VEs were recorded using an omnidirectional video camera, the Ricoh Theta S(c), which can record spherical photos with a resolution of 5376 × 2688 pixels and spherical videos with a resolution on 1920 × 1080 pixels. The final version of the ObReco-360° consists in a custom Android application which can be sideloaded on an Oculus Go© headset. The application was developed using the InstaVR© software, which allowed to organize the virtual environments in a single experience. The ObReco-360° test includes four different phases: the Familiarization Phase,



the Encoding Phase, the Free Recall Phase and the Recognition Phase.

The Familiarization Phase is aimed to make the participants comfortable with the experience and to detect possible side-effects linked to VR exposure (e.g., dizziness, nausea). Here, the participants find themselves in a black room with a floating icon showing the number one on the center. Then, they are asked to point and select the icon with the Oculus Go© controller, in order to show a text message, “search for the number 2”, which shows the instruction for the task. The procedure is the same for numbers from 2 to 4, which are positioned in the four cardinal points around the participants: when they finally find and select the number 4 (Fig. 2a), the second virtual environment is loaded.

The Encoding Phase represents the starting point of the proper test: The first scenario includes a 3D wall showing the instructions for the task (Fig. 2b), which are also presented in auditory modality. Then, the participants can choose whether to playback the instructions or go to the test phase. In the test phase, participants must pay attention to different objects presented by a virtual clinician (Fig. 3a). The objects are randomly placed in an office room; the target objects are 10 mixed with other 17 non-target ones. During the video, the clinician moves around in the room and presents the target objects closely to the camera for 5 s; in the meanwhile, the participants must name the object showed. At the end of this task, the participants are invited to take off



**Fig. 2** **a** The figure shows a screenshot of the target numbers in the familiarization phase. Participants need to point and select the circles showing numbers to expand a text label indicating to search the subsequent number **b** Screenshot of the instructions as they are presented at the beginning of the task. The two options on the low sides allowed the participants to choose between instructions playback or test start



**Fig. 3** **a** The figure shows a set of screenshots of the virtual clinician showing the 10 objects to recognize from the participants’ point of view. From the upper left to the bottom right there are a pan, an apple, a cup, a doll, a plant, a telephone, a stapler, a key, an umbrella and a torch. **b** A panoramic view of the virtual room with all the 10 target and 17 non-target objects

the headset and join the “real” clinician for a 10-min long session of non-interferent tests.

Then, the next step is represented by the Free Recall phase. The tasks simply require the participant to remember the 10 objects presented 10 min earlier in the Encoding Phase. The raw score is the number of objects correctly reported.

After wearing back the headset, the Recognition Phase begins. Again, the scenario includes a visual and auditory presentation of the instructions for the task, which asks the participants to search an immersive 360° photo of the same room included in the first task to find and nominate all the ten objects previously showed, located among other 17 non-target objects (Fig. 3b).

## 2.7 Data analysis

We organized all the data collected in a Windows Excel sheet and computed different indexes, both for the standard and VR assessment protocol. For the Free Recall tasks, we computed the accuracy percentages observed in the performances. For the Recognition tasks, we computed three

different scores: the Hit Rate (HR, the proportion of yes responses to old items), the False Alarm Rate (the proportion of yes responses to new items) and the discrimination score PR (i.e., hits -false alarms; Snodgrass and Corwin 1988). All these scores are reported in percentages.

Then, we performed four Wilcoxon signed-rank tests to compare the Free Recall and Recognition scores in both classic and 360° mode, investigating the statistical significance of the detected differences in performances. Finally, we performed a second statistical analysis, aimed to explore the presence of significant correlations between

the computed scores in the two conditions. All the analyses were performed using JASP (Version 0.14.1.0).

### 3 Results

#### 3.1 Task performance

The descriptives of the accuracy performances showed by the participants on Free Recall and Recognition tasks in the two modalities are shown in Table 2. The results indicate that for the Free Recall tasks, participants performed better after the 360° presentation than after the standard one in terms of accuracy percentages (Fig. 4a), and the difference is statistically significant ( $W = 14.5, p < 0.001$ ). For what concerns the Recognition indexes, the participants performed better in recognizing the objects after the standard presentation than after the 360° one (Fig. 4b), and the observed difference is statistically significant both for the HR ( $W = 131, p < 0.05$ ) and PR ( $W = 186, p < 0.05$ ) scores between the two conditions. No significant differences were detected in confronting the percentages of FARs. The Wilcoxon Signed-Rank Tests statistics are shown in Table 3.

#### 3.2 Correlation Analysis

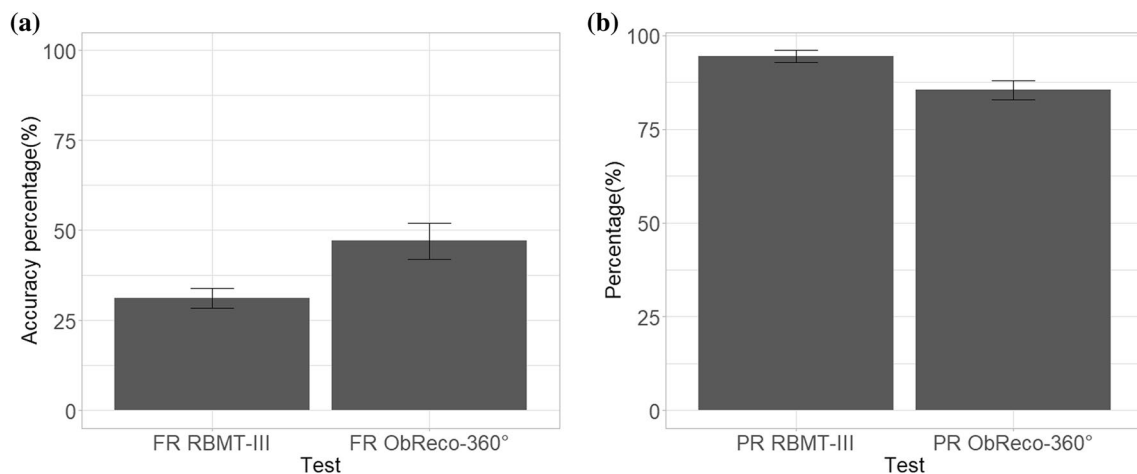
We also performed a statistical analysis to search for possible significant correlations between the accuracy scores obtained by the participants on the free recall and recognition indexes in both modalities. The results show a statistically significant correlation ( $r = 0.64, p < 0.010$ ) between the Free Recall task in standard and 360° modality (Table 4).

**Table 2** The upper part of the table shows the descriptives of the accuracy obtained by the participants in the free recall tasks (FR) in the standard (RBMT-III) and VR (ObReco-360°) condition

	Mean	SD	SE
Free recall tasks performance descriptives ( $n = 24$ )			
FR RBMT-III	31	13.8	2.8
FR ObReco-360°	47	24.4	4.9
Recognition tasks performance descriptives ( $n = 24$ )			
HR RBMT-III	96.4	7.6	1.6
HR ObReco-360°	89.6	10.4	2.1
FAR RBMT-III	2	3.1	0.7
FAR ObReco-360°	4.2	5.6	1.1
PR RBMT-III	94.4	7.7	1.6
PR ObReco-360°	85.4	12.5	2.6

The lower part of the table shows the descriptives concerning the three indexes (HR, FAR and PR) regarding the recognition performance in the two conditions (RBMT-III vs ObReco-360°). All the descriptives are reported in percentages

*FR* free recall, *HR* hit rate (HR, the proportion of yes responses to old items), *FAR* false alarms rate (the proportion of yes responses to new items), *PR* discrimination score (hits -false alarms)



**Fig. 4** The bar plots show the performance differences in the free recall (FR) and the discrimination tasks (PR) for each modality. **a** Accuracy difference in the free recall (FR) task between the standard

(RBMT-III) and VR (ObReco-360°) conditions. **b** Percentage differences in the discrimination scores (PRs) between standard (RBMT-III) and VR (ObReco-360°) conditions

**Table 3** The table reports the statistics of the Wilcoxon Signed-Ranked Tests performed to analyze the significance of the differences observed between the standard (RBMT-III) and VR (ObReco-360°) conditions in the free recall scores (FR) and in the recognition indexes scores (HR, FAR and PR)

Wilcoxon signed-rank test statistics			
		Statistic	<i>p</i>
FR RBMT-III	FR ObReco-360°	14.5	<0.001***
HR RBMT-III	HR ObReco-360°	131	0.010**
FAR RBMT-III	FAR ObReco-360°	33	0.231
PR RBMT-III	PR ObReco-360°	186	0.003**

FR free recall, HR hit rate (HR, the proportion of yes responses to old items), FAR false alarms rate (the proportion of yes responses to new items), PR discrimination score (hits -false alarms)

\*Statistically significant at  $p < 0.05$ ; \*\*Statistically significant at  $p < 0.010$ ; \*\*\*Statistically significant at  $p < 0.001$

### 3.3 UX analysis

The descriptives of the scores given by the participants in the 4 scales of the ITC-SOPI and in the SUS are showed in Table 5.

## 4 Discussion

With this explorative study, we wanted to test the feasibility of using 360° technology in the neuropsychological assessment of memory. The rationale of the study was based on the ongoing scientific debate concerning the ecological validity of standard paper-and-pencil tests (Chaytor and Schmitter-Edgecombe 2003; Parsons 2015) and on the evidence suggesting the implementation of VR in the neuropsychological assessment (Rizzo and Koenig 2017). Moreover, we wanted

**Table 5** The table shows the descriptives of the scores given by the participants in each of the 4 scales of the ITC-SOPI and in the SUS

UX descriptives				
	Min	Max	Mean	SD
SPS(ITC-SOPI)	1.8	4.3	3.1	0.7
E (ITC-SOPI)	2.5	4.3	3.4	0.6
EV (ITC-SOPI)	2.6	5	4	0.6
NE (ITC-SOPI)	1	3	1.4	0.5
SUS	57.5	97.5	73.5	11

SOP sense of physical space, E engagement, EV ecological validity, NE negative effects

to make a further step developing a pilot tool using 360° naturalistic VEs, given their photorealistic features allowing a more ecological memory encoding (Robertson et al. 2016; Serino and Repetto 2018), especially when compared to computer-generated VEs. Considering previous results from literature (Serino et al. 2017; Realdon et al. 2019) we expected to find some correlation confronting the memory performances showed by the participants in the standard subtest of the RBMT-III and in the ObReco-360°: in particular, as previously suggested (Serino and Repetto, 2018) we focused on two memory function indexes, namely the free recall and recognition accuracy. Additionally, we studied the UX ratings given by our participants to the ObReco-360° in order to find possible difficulties related to technological asset used.

The results showed that participants obtained low scores on the Free Recall tasks included in the two conditions, showing a better performance after the ObReco-360° Encoding Phase (FR RBMT-III = 31%, FR ObReco-360° = 47%). For what instead concerns the Recognition performance, the pattern of results is

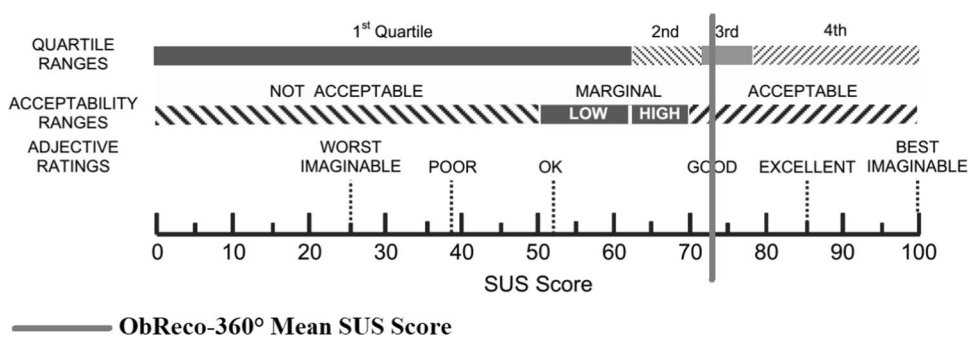
**Table 4** The table shows the correlation coefficients computed between the scores obtained on the free recall (FR) and the recognition (R) tasks in the standard (RBMT-III) and VR (ObReco-360°) conditions

Pearson’s correlations					
Variable		FR RBMT-III	FR ObReco 360°	PR RBMT-III	PR ObReco-360°
FR RBMT-III	Pearson’s <i>r</i>	–			
	<i>p</i> -value	–			
FR ObReco-360°	Pearson’s <i>r</i>	0.640**	–		
	<i>p</i> -value	0.001	–		
PR RBMT-III	Pearson’s <i>r</i>	0.320	0.373	–	
	<i>p</i> -value	0.146	0.073	–	
PR ObReco-360°	Pearson’s <i>r</i>	0.067	0.098	0.082	–
	<i>p</i> -value	0.766	0.649	0.705	–

FR free recall, HR hit rate (HR, the proportion of yes responses to old items), FAR false alarms rate (the proportion of yes responses to new items), PR discrimination score (hits -false alarms)

\*Statistically significant at  $p < 0.05$ ; \*\*Statistically significant at  $p < 0.010$ ; \*\*\*Statistically significant at  $p < 0.001$

**Fig. 5** The vertical line in the figure shows the position of the mean score of 73.5 obtained by the ObReco-360° according to the SUS rating comparison scale provided by Bangor et al. (2008). The SUS score marks the ObReco-360° as an “Acceptable” and “Good” tool



inverted: indeed, the participants showed high levels of accuracy in both conditions, performing slightly better in the RBMT-III condition (PR RBMT-III = 94.4%, PR ObReco-360° = 85.4%).

These results could be explained by the fact that in the ObReco-360 condition the participants needed to recall 10 objects, against the 15 of the RBMT-III subtask. We introduced this mismatch in the target items between the two conditions according to previous findings from the literature suggesting that immersive VR tasks are more cognitive demanding (Frederiksen et al. 2020; Harris et al. 2019), in order to avoid a possible floor-effect in the ObReco-360° Free Recall performance. However, we can also hypothesize that the photorealistic fashion of 360° VEs could have led and to a better visual encoding of the stimuli and to a lower cognitive load facilitating the subsequent recall of the objects, as also suggested by previous evidence linking immersivity and realism levels to memory encoding (Robertson et al. 2016; Makowski et al. 2017). Indeed, this result is also consistent with evidence coming from the work of Ventura et al. (2019), who showed that immersive 360°-VR scenario could elicit a better visual encoding and subsequently a better recall of the encoded items when compared to the same task performed in a non-immersive scenario. Then, 360°-VR scenarios might elicit a visual memory encoding which is very close to the one performed in everyday life, thus improving the ecological validity of the assessment procedure.

Regarding the Recognition tasks, the lower scores obtained by the participants in the ObReco-360° condition could be explained by the higher complexity which characterizes this task: in fact, the 360° scenario required the users to actively explore the environment in order to discriminate the target items from the distractors, as a results of the ecological fashion which characterized the task’s design. This may have allowed a slightly more sensitive and ecological assessment of Recognition memory when compared to the RBMT-III condition, where participants’ performance showed a more prominent ceiling-effect.

Moreover, the presence of a significant correlation ( $r=0.64$ ,  $p<0.10$ ) between the Free Recall tasks of in the

two conditions support the use of ObReco-360° as a measure of memory functioning.

On the UX assessment side, the scores obtained by the ObReco-360° in the 4 dimensions of the ITC-SOPI are consistent with the ones showed by Yildirim et al. (2019), which provided the ITC-SOPI benchmark scores for the 360° videos category and certificated the high ecological value provided by this technology. Additionally, the low mean score obtained by the ObReco-360° in the Negative Effects scale can be the result of the adoption of a fixed position in the VE and of the limited duration of the VR exposure (about 10 min). Moreover, considering the rating comparison scale proposed by Bangor et al. (2008), the SUS mean score of 73.5/100 ranks the usability of the ObReco-360° slightly above the third quartile, defining the tool as “Acceptable” and with a “Good” level of usability (Fig. 5).

## 5 Limitations and conclusion

The present work has several limitations. First, the sample is limited in its number and in its representativity: Considering the preliminary nature of the study, we primarily focused on the features of the technology, but further studies must include a larger sample size with different demographic characteristics. Then, the technological equipment used was entry-level: Currently, the 360° devices market offers much higher-quality omnidirectional cameras (e.g., Insta360 Pro 2(c)) and better all-in-one headsets (e.g., Oculus Quest 2(c)), which together can provide a more realistic experience and thus give a higher ecological value to the obtained measures. Future works are needed in order to clarify what advantages/disadvantages characterize 360°-VR when compared to model-based VR. For example, further information could be obtained confronting the same task (e.g., memory encoding) in three different modalities: paper-and-pencil, 360°-VR and model-based VR.

Summarized, even if limited by the explorative nature of the study, these preliminary results encourage the implementation of 360° technology in the development of ecological tests for memory assessment. Further work is needed to



improve the design of 360° experiences searching for tasks that resembles the challenges of daily life, working around the interactivity limitations that characterize this technology.

**Funding** Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement.

**Code availability** The materials used for building the application will be made available upon request.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethics approval** The study was approved by the Ethical Committee of Istituto Auxologico Italiano.

**Consent to participate** All participants signed a written informed consent before joining the study.

**Availability of data** Data will be made available upon request.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Appollonio I, Leone M, Isella V, Piamarta F, Consoli T, Villa ML, Forapani E, Russo A, Nichelli P (2005) The frontal assessment battery (FAB): normative values in an Italian population sample. *Neurol Sci* 26:108–116. <https://doi.org/10.1007/s10072-005-0443-4>
- Bangor A, Philip TK, James TM (2008) An empirical evaluation of the system usability scale. *Int J Hum Comput Int* 24:574–594. <https://doi.org/10.1080/10447310802205776>
- Beschin N, Urbano T, Treccani B (2013) Rivermead behavioural memory test, 3rd edn. Adattamento Italiano, Giunti Psychometrics, Firenze
- Brooke J (1996) SUS: a “quick and dirty” usability scale. In: Jordan P, Thomas W, Weerdmeester A, McClelland I (eds) Usability evaluation in industry. Taylor & Francis, London, pp 189–194
- Chaytor N, Schmitter-Edgecombe M (2003) The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. *Neuropsychol Rev* 13:181–197. <https://doi.org/10.1023/B:NERV.0000009483.91468.fb>
- Chirico A, Cipresso P, Yaden DB, Biassoni F, Riva G, Gaggioli A (2017) Effectiveness of immersive videos in inducing awe: an experimental study. *Sci Rep* 7:1–11. <https://doi.org/10.1038/s41598-017-01242-0>
- Dubois B, Slachevsky A, Litvan I, Pillon B (2000) The FAB: a frontal assessment battery at bedside. *Neurology* 55:1621–1626. <https://doi.org/10.1212/wnl.55.11.1621>
- Efklides A, Yiultsi E, Kangelidou T, Kounti F, Dina F, Tsolaki M (2002) Wechsler memory scale, Rivermead behavioral memory test, and everyday memory questionnaire in healthy adults and Alzheimer patients. *Eur J Psychol Assess*. <https://doi.org/10.1027//1015-5759.18.1.63>
- Folstein MF, Folstein SE, McHugh PR (1975) “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12:189–198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
- Frederiksen JG, Sørensen SMD, Konge L et al (2020) Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: a randomized trial. *Surg Endosc* 34:1244–1252. <https://doi.org/10.1007/s00464-019-06887-8>
- Harris D, Wilson M, Vine S (2019) Development and validation of a simulation workload measure: the simulation task load index (SIM-TLX). *Virtual Real* 24:557–566. <https://doi.org/10.1007/s10055-019-00422-9>
- Heeter C (1995) Communication research on consumer VR. In: Biocca F, Levy MR (eds) LEA's communication series. Communication in the age of virtual reality. Lawrence Erlbaum Associates Inc, New Jersey, pp 191–218
- Howard MC (2017) A meta-analysis and systematic literature review of virtual reality rehabilitation programs. *Comput Hum Behav* 70:317–327. <https://doi.org/10.1016/j.chb.2017.01.013>
- Keefe RSE, Davis VG, Atkins AS, Vaughan A, Patterson T, Narasimhan M, Harvey PD (2016) Validation of a computerized test of functional capacity. *Schizophr Res* 175:90–96
- Larson EB, Feigon M, Gagliardo P, Dvorkin AY (2014) Virtual reality and cognitive rehabilitation: a review of current outcome research. *NeuroRehabilitation* 34:759–772. <https://doi.org/10.3233/NRE-141078>
- Lessiter J, Freeman J, Keogh E, Davidoff J (2001) A cross-media presence questionnaire: the ITC-sense of presence inventory. *Presence (Camb)* 10:282–297. <https://doi.org/10.1162/105474601300343612>
- Makowski D, Sperduti M, Nicolas S, Piolino P (2017) “Being there” and remembering it: presence improves memory encoding. *Conscious Cogn* 53:194–202. <https://doi.org/10.1016/j.concog.2017.06.015>
- Matheis RJ, Schultheis MT, Tiersky LA, DeLuca J, Millis SR, Rizzo A (2007) Is learning and memory different in a virtual environment? *Clin Neuropsychol* 21:146–161. <https://doi.org/10.1080/13854040601100668>
- Measso G, Cavarzeran F, Zappalà GD, Lebowitz BH, Crook TJ, Pirozzolo F, Amaducci AL, Massari D, Grigoletto F (1993) The mini-mental state examination: normative study of an Italian random sample. *Dev Neuropsychol* 9:77–85. <https://doi.org/10.1080/87565649109540545>
- Mondini S, Mapelli D, Arcara G (2016) Semeiotica e diagnosi neuropsicologica: metodologia per la valutazione. Carocci, Roma
- Moreno A, Wall KJ, Thangavelu K, Craven L, Ward E, Dissanayaka NN (2019) A systematic review of the use of virtual reality and its effects on cognition in individuals with neurocognitive disorders. *Alzheimer's Dement Transl Res Clin Interv* 5:834–850. <https://doi.org/10.1016/j.trci.2019.09.016>
- Negu A, Matu SA, Sava FA, David D (2016) Virtual reality measures in neuropsychological assessment: a meta-analytic review. *Clin Neuropsychol* 30:165–184. <https://doi.org/10.1080/13854046.2016.1144793>

- Ouellet É, Boller B, Corriveau-Lecavalier N, Cloutier S, Belleville S (2018) The virtual shop: a new immersive virtual reality environment and scenario for the assessment of everyday memory. *J Neurosci Methods* 303:126–135. <https://doi.org/10.1016/j.jneumeth.2018.03.010>
- Parsons TD (2015) Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Front Hum Neurosci* 9:1–19. <https://doi.org/10.3389/fnhum.2015.00660>
- Parsons TD, Carlew AR, Magtoto J, Stonecipher K (2017) The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychol Rehabil* 27:777–807. <https://doi.org/10.1080/09602011.2015.1109524>
- Pedroli E, Serino S, Cipresso P, Pallavicini F, Riva G (2015) Assessment and rehabilitation of neglect using virtual reality: a systematic review. *Front Behav Neurosci* 9:226. <https://doi.org/10.3389/fnbeh.2015.00226>
- Pedroli E, Serino S, Pallavicini F, Cipresso P, Riva G (2018) Exploring virtual reality for the assessment and rehabilitation of executive functions. *Int J Virtual Augment Real* 2:32–47. <https://doi.org/10.4018/IJVAR.2018010103>
- Realdon O, Serino S, Savazzi F, Rossetto F, Cipresso P, Parsons TD, Cappellini G, Mantovani F, Mendozzi L, Nemni R, Riva G, Baglio F (2019) An ecological measure to screen executive functioning in MS: the picture interpretation test (PIT) 360°. *Sci Rep* 9:5690. <https://doi.org/10.1038/s41598-019-42201>
- Riva G, Wiederhold BK, Mantovani F (2019) Neuroscience of virtual reality: from virtual exposure to embodied medicine. *Cyberpsychol Behav Soc Netw* 22:82–96. <https://doi.org/10.1089/cyber.2017.29099.gri>
- Rizzo AS, Koenig ST (2017) Is clinical virtual reality ready for prime-time? *Neuropsychology* 31:877–899. <https://doi.org/10.1037/neu000040>
- Robertson CE, Hermann KL, Mynick A, Kravitz DJ, Kanwisher N (2016) Neural representations integrate the current field of view with the remembered 360° panorama in scene-selective cortex. *Curr Biol* 26:2463–2468. <https://doi.org/10.1016/j.CUB.2016.07.002>
- Sbordone RJ (1996) Ecological validity: some critical issues for the neuropsychologist. In: Sbordone RJ, Long CJ (eds) *Ecological validity of neuropsychological testing*. Gr Press/St Lucie Press Inc, Delray Beach, pp 15–41
- Serino S, Repetto C (2018) New trends in episodic memory assessment: immersive 360° ecological videos. *Front Psychol* 9:1878. <https://doi.org/10.3389/fpsyg.2018.01878>
- Serino S, Morganti F, Di Stefano F, Riva G (2015) Detecting early egocentric and allocentric impairments deficits in Alzheimer's disease: an experimental study with virtual reality. *Front Aging Neurosci* 7:88. <https://doi.org/10.3389/fnagi.2015.00088>
- Serino S, Baglio F, Rossetto F, Realdon O, Cipresso P, Parsons TD, Cappellini G, Mantovani F, De Leo G, Nemni R, Riva G (2017) Picture interpretation test (PIT) 360°: an innovative measure of executive functions. *Sci Rep* 7:16000. <https://doi.org/10.1038/s41598-017-16121-x>
- Slater M, Wilbur S (1997) A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence (Camb)* 6:603–616. <https://doi.org/10.1162/pres.1997.6.6.603>
- Slobounov SM, Ray W, Johnson B, Slobounov E, Newell KM (2015) Modulation of cortical activity in 2D versus 3D virtual reality environments: an EEG study. *Int J Psychophysiol* 95:254–260. <https://doi.org/10.1016/j.ijpsycho.2014.11.003>
- Smith GV, Della Sala S, Logie RH, Maylor EAM (2000) Prospective and retrospective memory in normal ageing and dementia: a questionnaire study. *Memory* 8:311–321. <https://doi.org/10.1080/09658210050117735>
- Snodgrass JG, Corwin J (1988) Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J Exp Psychol* 117:34–50. <https://doi.org/10.1037//0096-3445.117.1.34>
- Spinnler H, Tognoni G (1987) Standardizzazione e taratura italiana di test neuropsicologici: Gruppo italiano per lo studio neuropsicologico dell'invecchiamento. *Masson Italia Periodici*, Milano
- Sutcliffe A, Gault B, Shin JE (2005) Presence, memory and interaction in virtual environments. *Int J Hum Comput Stud* 62:307–327. <https://doi.org/10.1016/j.IJHCS.2004.11.010>
- Ventura S, Brivio E, Riva G, Baños R (2019) Immersive versus non-immersive experience: exploring the feasibility of memory assessment through 360° technology. *Front Psychol*. <https://doi.org/10.3389/fpsyg.2019.02509>
- Wilson B, Cockburn J, Baddeley A, Hiorns R (1989) The development and validation of a test battery for detecting and monitoring everyday memory problems. *J Clin Exp Neuropsychol* 11(6):855–870. <https://doi.org/10.1080/01688638908400940>
- Wilson BA, Greenfield E, Clare L, Baddeley A, Cockburn J, Watson P, Tate R, Sopena S, Nannery R, Crawford JR (2008) *The Rivermead behavioural memory test*, 3rd edn. Pearson Education, London
- Yildirim Ç, Bostan B, İlker Berkman M (2019) Impact of different immersive techniques on the perceived sense of presence measured via subjective scales. *Entertain Comput*. <https://doi.org/10.1016/j.entcom.2019.100308>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.