Check for updates

Virtual Reality for the Assessment of Everyday Cognitive Functions in Older Adults: An Evaluation of the Virtual Reality Action Test and Two Interaction Devices in a 91-Year-Old Woman

Andrea Chirico[1]*, Tania Giovannetti[2], Pietro Neroni[3,4], Stephanie Simone[2], Luigi Gallo[3], Federica Galli[1], Francesco Giancamilli[1], Marco Predazzi[5], Fabio Lucidi[1], Giuseppe De Pietro[3] and Antonio Giordano[6,7]

[1] Department of Psychology of Developmental and Socialization Processes, Sapienza University of Rome, Rome, Italy, [2] Psychology Department, Temple University, Philadelphia, PA, United States, [3] Institute for High Performance Computing and Networking, National Research Council, Naples, Italy, [4] Department of Engineering, Parthenope University of Naples, Naples, Italy, [5] Fondazione Il Melo Onlus, Gallarate, Italy, [6] Sbarro Institute for Cancer Research and Molecular Medicine, Center for Biotechnology, College of Science and Technology, Temple University, Philadelphia, PA, United States, [7] Department of Medical Biotechnologies, University of Siena, Siena, Italy

Performance-based functional tests for the evaluation of daily living activities demonstrate strong psychometric properties and solve many of the limitations associated with self- and informant-report questionnaires. Virtual reality (VR) technology, which has gained interest as an effective medium for administering interventions in the context of healthcare, has the potential to minimize the time-demands associated with the administration and scoring of performance-based assessments. To date, efforts to develop VR systems for assessment of everyday function in older adults generally have relied on non-immersive systems. The aim of the present study was to evaluate the feasibility of an immersive VR environment for the assessment of everyday function in older adults. We present a detailed case report of an elderly woman who performed an everyday activity in an immersive VR context (Virtual Reality Action Test) with two different types of interaction devices (controller vs. sensor). VR performance was compared to performance of the same task with real objects outside of the VR system (Real Action Test). Comparisons were made on several dimensions, including (1) quality of task performance (e.g., order of task steps, errors, use and speed of hand movements); (2) subjective impression (e.g., attitudes), and (3) physiological markers of stress. Subjective impressions of performance with the different controllers also were compared for presence, cybersickness, and usability. Results showed that the participant was capable of using controllers and sensors to manipulate objects in a purposeful and goal-directed manner in the immersive VR paradigm. She performed the everyday task similarly across all conditions. She reported no cybersickness and even indicated that interactions in the VR environment were pleasant and relaxing. Thus, immersive VR is a

feasible approach for function assessment even with older adults who might have very limited computer experience, no prior VR exposure, average educational experiences, and mild cognitive difficulties. Because of inherent limitations of single case reports (e.g., unknown generalizability, potential practice effects, etc.), group studies are needed to establish the full psychometric properties of the Virtual Reality Action Test.

INTRODUCTION

Performance-based tests, that evaluate the ability to perform everyday tasks in the laboratory/clinic, solve many of the limitations associated with the use of self- and informant-report questionnaires of everyday functioning in people with cognitive impairment (see Giovannetti et al., 2013 for a review). Performance-based, functional tests are objective, standardized, allow a systematic comparison between individuals and provide detailed information on behaviors during the natural performance of activities. The validity of performance-based measures is supported by studies showing expected differences between clinical groups and controls (Giovannetti et al., 2002, 2008a, 2018; Schwartz et al., 2002; Allain et al., 2014; Gold et al., 2015; Rycroft et al., 2018), significant (though modest) relations with cognitive tests (Giovannetti et al., 2002, 2008a, 2018; Schwartz et al., 2002; Kessler et al., 2007; Allain et al., 2014; Rycroft et al., 2018), and informant and clinician reports of functioning (Giovannetti et al., 2002, 2008b; Schwartz et al., 2002; Allain et al., 2014). Detailed analyses of errors and error-types afforded by performance-based tests of everyday function also have promoted theoretical frameworks to better characterize the breakdown of everyday function due to cognitive impairment (see Schwartz, 2006; Giovannetti et al., 2013; for a review). Despite their validity, objectivity and potential for characterization of functional difficulties, performance-based measures have not been widely adopted in clinics or research studies, because generally they require an extraordinary effort to administer and score, especially when used to assess minor difficulties.

Virtual reality (VR) technology has recently gained interest as an effective medium for administering different interventions in the context of healthcare (Cipresso and Serino, 2014; Chirico et al., 2016; Indovina et al., 2018). Several observational studies and a small number of controlled studies have found VR to be effective for a variety of health issues (Cipresso et al., 2016). VR also has been proposed to improve clinical assessments, as automated VR systems could dramatically reduce the time required for administration and scoring traditional performance-based assessments without sacrificing ecological validity. To date, efforts to develop VR systems for assessment of function in older adults have mostly relied on non-immersive systems (Cipresso et al., 2014). In 2014, Allain et al. (2014) reported results from the Virtual Kitchen (VK), a non-immersive activity that required participants to use a mouse to select and move target objects and avoid distractor objects on a computer screen to prepare a cup of coffee. In 2019 Giovannetti et al. (2019) reported preliminary data from a modified VK, called the Virtual

Kitchen Challenge (VKC), which included complex tasks to enable assessment of participants with mild cognitive difficulties and requires participants to use a touch screen interface instead of a mouse. Automated scores from the VKC were significantly associated with scores from the same tasks performed with real objects in a real kitchen.

Immersive VR systems also have been proposed to assess everyday function, as they have the advantage of creating a sense of realism or "presence" in the user. Presence is a multidimensional construct that describes the extent to which users believe and feel that they exist in the environment simulated by VR (e.g., kitchen; Diemer et al., 2015) rather than in their true physical location (e.g., clinic/lab; Witmer and Singer, 1998). Presence may be influenced by the quality of the visual scene, method of interaction/interface with the virtual environment, and other factors. Immersive VR assessments of everyday function that elicit a high degree of presence in the user might demonstrate greater ecological and predictive validity of everyday function than non-immersive tasks (Shahrbanian et al., 2012; Parsons, 2015). Although immersive systems afford greater "presence," they also introduce unique challenges. One challenge, which is particularly salient for older adults, is managing the interface between the user and the surrounding virtual environment, because the immersive context increases the complexity of the task. Using a head-mounted display (HMD), Nolin et al. (2013) and Banville et al. (2017, 2018) implemented an immersive VR task that required participants to use the computer keyboard and mouse to sort everyday objects – a task that would be quite easy for older adults in real-life. Results showed that that older participants took more time to navigate within the virtual environment and to complete the sorting task. Also, older participants were more variable in the time required to accomplish the sorting task as compared to younger participants. These findings underscore the importance of the comfort and ease of the interface, which should feel familiar to the user and optimize mobility. Many immersive VR hardware solutions have been introduced, such as data gloves or controllers, some with haptic feedback; however, they generally prove to be too expensive and require substantial set up time. New, low-cost and ready-to-use devices, such as advanced controllers, could keep costs and administration time low and promote presence in the user during the interaction (Caggianese et al., 2019).

Advanced controllers (hereafter controllers) include buttons and tactile surfaces that are manipulated by the participant. Controllers offer indirect tracking of the position and orientation of the participant's body. In contrast, egocentric sensors (hereafter sensors) are head-mounted small sensing devices used to detect and track the users' hands from images acquired

from the users' point of view, directly transforming hands and finger movements into interactions with virtual objects. Both controllers and sensors allow the user to see the movement of her/his hands while being immersed in a virtual environment. A recent study comparing the most frequently used controllers (HTC Vive Controllers) and sensors (Leap Motion) with three simple manipulation tasks (i.e., select, position and rotate virtual objects) in eight participants aged 30–40 years showed an advantage for Vive Controllers, which were more stable, accurate, and easier to learn than the Leap Motion sensor (Caggianese et al., 2019).

The aim of the present study was to evaluate the feasibility of a fully immersive VR environment for the assessment of everyday function in older adults. We present a detailed case report of an elderly woman (Tina) who was selected because she represents a typical older adult with no particular computer or technological expertise and an average level of education. Tina was observed while performing an everyday activity in an immersive VR context with two different types of interfaces (controller vs. sensor). VR performance was compared against performance of the same task with real objects outside of the VR system. Comparisons were made on several dimensions, including (1) quality of task performance (e.g., order of task steps, errors, use and speed of hand movements); (2) subjective impression (e.g., attitudes, presence, cybersickness, and usability), and (3) physiological markers of stress.

MATERIALS AND METHODS

The Participant

Tina is a 91-year-old, single women living independently in Northern Italy in a community-residence for older adults. Tina was born in Italy and is a native Italian speaker. At the time of the study she reported that she was functioning independently and had no current or past neurological or psychiatric disorders or other major medical illness (e.g., dementia, brain injury, schizophrenia, depression, etc.). She demonstrated no sensory or motor deficits that precluded interaction with a Head Mounted Display and controllers/sensors. Tina was recruited as a volunteer through an announcement made at her residence.

Procedure

The study was approved by the Ethical Committee in the Department of Psychology of Developmental and Socialization Processes at "Sapienza" University of Rome. All procedures were completed in a single 2- to 3-h session that included the following (in order of administration): (1) informed consent obtained by the participant, (2) screening interview, (3) training on the Virtual Reality Action Test (VRAT) with controllers, (4) testing on the VRAT with controllers followed by presence and attitudes questionnaires, (5) testing on the Real Action Test followed by presence and attitudes questionnaires (6) VRAT sensor training; (7) VRAT sensor testing followed by presence and attitudes questionnaire, and (8) cognitive tests and questionnaires of mood, anxiety and everyday function.

Performance-Based Functional Tests

The breakfast task was administered across all platforms: Real Action Test and Virtual Reality Action Test (with two different controllers). The breakfast task was selected because it has been widely studied as part of the Naturalistic Action Test (NAT), a performance-based test developed to evaluate the cognitive difficulties associated with the completion of everyday activities in people with neurologic impairment (Schwartz et al., 2002). The breakfast task requires participants to prepare a slice of toast with butter and jelly and a cup of coffee with milk and sugar while seated at a table containing a toaster, two knives, one spoon, butter in butter dish, sugar in a bowl, bottle of milk, mug filled with warm water, bread, instant coffee, jelly jar, and a napkin at the central workspace. The shape of the table and the spatial arrangement of objects was informed by procedures in the NAT manual[1].

The breakfast task was administered in real and two different VR conditions (described below). In each condition, Tina was instructed to complete the task in silence, as quickly as possible, and without making errors. She was asked to make her movements as clear as possible and to tell the examiner when she was finished. Performance was recorded for scoring. Physiological and kinematic data were obtained while the participant completed the breakfast task according to the procedures described below.

Real Action Test (RAT)

The RAT required the participant to complete the breakfast task without feedback using real objects. The participant performed the RAT while wearing a smart band and wireless controllers (described below) attached to her arms to acquire kinematic and physiological data (see **Figure 1**).

Virtual Reality Action Test (VRAT)

The VRAT is a VR version of the breakfast task designed to maximize ecological validity by simulating a real kitchen and household objects. In this respect, the VRAT environment is characterized by a high degree of realism, including accurate 3D
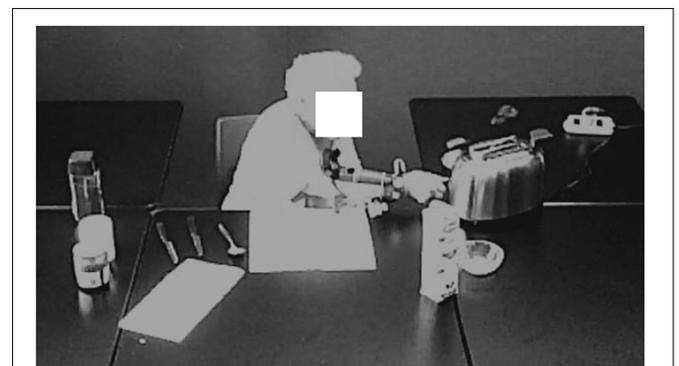
---

[1]https://mrri.org/wp-content/uploads/2016/01/NATManual.pdf

FIGURE 1 | The subject (Tina) performing the Real Action Test (RAT).

models and spatial audio. The VRAT includes automatic, real-time collection of movement data, as well as physiological and kinematic parameters (described below).

VRAT Apparatus and Controller Conditions

The VRAT system runs on a MSI Trident Gaming Desktop, with 8GB RAM and a GTX 1060 graphic card. The HTC Vive head mounted display[2] provides users with a fully immersive virtual environment. The HTC Vive visual system is based on two OLED displays for a total resolution of 2160 × 1200 pixels with a 110-degree FoV and a frequency of 90 Hz. The VR software was developed with Unity3D[3], a game development platform which provides native VR support.

Interaction in the VRAT was enabled through two different input devices: (1) *controllers* – the participant used HTC Vive controllers that provided tactile feedback through vibration; (2) *sensors* – a wearable egocentric sensor, the Leap Motion Controller[4], enabled interaction through movements of the participant's own hands. Performance with the two different devices were tested in different conditions.

Controllers: were worn during performance of the RAT and the VRAT-controller conditions. During the RAT, participants did not interact with the controller; it was used only to collet kinematic data. However, in the VRAT, the controller was used to interact with the VR environment while the participant was in a seated position using interaction metaphors similar to those used in real-life. To make the interaction as familiar and natural as possible, we implemented the Virtual Hand metaphor (Ruddle, 2005), in which the user's hand motions are directly mapped to the virtual hand movements. When the virtual hand reaches an object, the object is highlighted to inform the user through visual feedback that it is selected and interactable. To interact with a virtual object in the VRAT, the user is instructed to press the trigger button once the object is highlighted/selected. To end the interaction, the user is instructed to release the trigger. One advantage of the controller is that the participant is able to be tracked even when the user's hands are not visible within the user field of view, allowing a wider measurement area. Controllers also provide users with tactile feedback through vibrations of varying intensity. However, interactions with virtual objects occur through a tool that the user must always hold in the hands, even when they are not interacting with any object, reducing the naturalness of the interaction.

Virtual Reality Action Test sensors: were used during performance of the VRAT-sensor condition. In this condition, the participant interacted with virtual objects using Leap sensors by performing a pinch gesture (i.e., moving thumb and index fingers closer until they come into contact). To release the virtual object(s) the pinch gesture is relaxed. The Leap sensor allows the user to interact with virtual objects with their own hands, without having to wear gloves or hold controllers. Unlike the controllers, the sensor is able to track the main joints of the user's

hand and replicate them in the virtual environment, increasing the hand representation and the sense of presence. However, the interaction area is limited to the tracking area of the sensor and the user's field of view. The sensor is mounted in front of the headset; therefore, the user must keep their hands in their field of view to interact with virtual objects. Furthermore, tracking may fail if the hand is occluded by the user's other hand or an obstacle/object in the real world.

Participants completed the RAT and both VRAT conditions while wearing a smart bracelet (Microsoft band 2) that was designed to obtain physiological measures of stress (described below).

Software Architecture

The system was designed as a multiplayer platform: one player is the participant, who performs the task within the virtual environment, and the other player is the examiner, who configures the test, and monitors, in real time, the scores and physiological parameters of the participant. The system includes a VR module that maps the data acquired by the HMD and input devices into the corresponding virtual actions within the virtual kitchen. The game logic of the breakfast task, including the physical features and behavior of each virtual element on the table, is coded in the VR module. An error checking module has been developed for automatically detecting an error by the participant. For each participant action during the task, the error checking module considers the virtual environment state, and through a specified set of rules, interprets the participant action as either an error or correct action. Each time the participant commits an error, it notifies the logger module. The logger module acquires data from various sources (error checking module, HMD, input devices) and synchronizes them under a single time value, making it possible to link all of the separate data streams (i.e., knowing the physiological state of the participant when she/he commits an error). All information is saved as. csv files at the end of the test. The examiner interface allows the examiner to manage the test from the control panel and view errors committed by the participant as well as physiological values in real time.

VR Training

Before each VR condition, the participant completed a brief training session with the system. Training included four mini-tasks that comprised elements of the breakfast task: (1) toast a slice of bread; (2) spread the jelly on toast; (3) add instant coffee to cup; (4) add milk to cup. The examiner controlled the presentation of each mini-task from a monitoring position.

Quality of Task Performance

Although the VRAT includes the error monitoring module, performance quality and accuracy on the RAT and two VRAT conditions were evaluated by trained coders who viewed recordings of the participant's performances. The following error scores were collected for each of the three conditions (RAT, VRAT-controller, VRAT-sensor):

Total overt errors – incorrect actions (commission), the failure to complete a step (omission), and off-task actions (additions)

---

were recorded and assigned a code according to the error taxonomy shown in **Table 1** (Schwartz et al., 2002).

Total micro-errors – subtle, inefficient but not overtly incorrect actions; this category of errors was added to the overt error taxonomy to improve detection of subtle, inefficient behaviors in healthy people and those with mild cognitive difficulties.

Clumsy-motor imprecision errors during the execution of an accurate task step.

Code sheets with an exhaustive list of overt/micro-errors were used to promote inter-rater reliability and are included in **Supplementary Material**.

In addition to errors, human coders evaluated video recordings for accomplishment, time to completion and the order of task steps as follows:

Accomplishment score – an accomplishment point was assigned for each task step of the breakfast task completed without error (range = 0–16).

Overall performance score – this score combines accomplishment score with the sum of a subset of key, overt errors (Schwartz et al., 2002).

Completion Time – was recorded in seconds; timing began when the first step was initiated and ended when the participant indicated that she was finished with the task.

Order of Task Steps – In addition to coding errors and completion time, the order in which the participant completed each task step was recorded to examine similarities/differences across the RAT and VRAT conditions.

Kinematic measures were obtained by the input devices used in the RAT and VRAT conditions. During the RAT and VRAT-controller conditions, the participant wore wireless controllers, and during the VRAT-sensor condition, the participants movements were recorded by Leap Motion. Kinematic data was obtained to measure the precise movements of both the right and left hands, with an accuracy in millimeters (100 Hz). Instantaneous velocity measures greater than three meters per second were excluded to avoid noisy data due to hand tracking problems in the VRAT-sensor condition. For each condition, the following kinematic measures were obtained:

- Total hand movement, in meters.
- Average speed of the hands, in meters per second, computed as total hand movement divided by completion time.

Subjective Impressions

Immediately following each condition (RAT, VRAT-controller, VRAT-sensor), the participant used a five-point scale to describe her reaction to the test condition on the following five items/dimensions: useless/useful, not pleasant/pleasant, boring/funny, tiring/resting, stressing/relaxing. Item scores were aggregated into a single score, ranging from 5 to 25, for which higher values indicated more positive attitudes about the test condition. This scale was created by the authors of the study according to procedures described by Ajzen (1991); see **Supplementary Material**.

Physiological Measures of Stress

To compare indicators of stress during each testing condition, physiological data were recorded via a smart bracelet (Microsoft band 2)[5] worn by the participant while completing the RAT and both VRAT conditions. Kubios software (Tarvainen et al., 2014) was used to obtain the following variables:

Heart rate (bpm, 1 Hz),
Galvanic Skin Response (kohms, 0,2/5 Hz),
R–R interval (i.e., time between heart beats; seconds, variable frequency),
skin temperature (degrees centigrade, 0,033 Hz).

To correct for artifacts, particularly in the measure of heart rate variability (RR interval), a threshold-based algorithm was applied that compares every RR interval value against a local average interval, obtained by median filtering the RR interval time series. RR interval values that differ from the local average of a specified threshold value (i.e., 0.45 s) are marked as artifact and replaced using cubic spline interpolation.

Physiologic variables (i.e., Heart rate, Galvanic Skin Response, Skin temperature) were used to calculate an index of cardiovascular system stress, called Baevsky's stress index (Baevsky and Berseneva, 2008). The Baevsky's stress index is strongly linked to sympathetic nervous activity and

---

TABLE 1 | Error Taxonomy Used Code Performance on the RAT and VRAT conditions.

Error type		Definition	Examples
Omission		Number of steps that are not performed	Does not add coffee grounds to coffee; does not add stamp to envelope
Commission	Substitution	Similar, alternate object is used in place of target object	Spreads butter on toast with spoon instead of knife
	Sequence	Anticipation of a step; steps or subtasks performed in reverse order	Butter on bread without toasting; applies jelly on bread then applies butter
	Perseveration	A step is performed more than once or for an excessive amount of time	Adds butter/jelly repeatedly to toast
Action-Additions		Performance of an action not readily interpreted as a task step	Puts toast in creamer
Micro-errors		Initiating and terminating an incorrect action before the error is completed by reaching for, touching or picking up an object	Reaches toward, touches or moves salt but never uses the salt in during the task
Clumsy		Correct step is performed but with difficulty due to motor imprecision	Coffee jar slips out of hand

---

[5]https://support.microsoft.com/it-it/help/4000323/band-hardware-sensors

increases during stressful situations. Physiologic data were stored on .csv files and although they may be combined with the test start time to synchronize physiological and kinematic information, for the current study, physiologic data were aggregated and averaged for each test condition to obtain an overall stress index per condition (RAT, VRAT-controller, VRAT-sensor).

VRAT Presence, Cybersickness, and Usability

The following questionnaires were administered immediately following performance on the VRAT-controllers and VRAT-sensors conditions.

Presence Questionnaire (PQ)

The Italian version of PQ was administered to the participant in this study (Scheuchenpflug et al., 2003). The PQ required the participant to use a seven-point scale to rate her experience with each condition on 28 items focused on the following features: Realism (7 items); Possibility to act (4 Items); Quality of interface (3 Items); Possibility to examine (3 items); Self-evaluation of performance (2 Items) (Witmer and Singer, 1998; Slater, 2002; Witmer et al., 2005). Strong internal reliability has been reported (0.88) for the total score.

Cybersickness Symptoms

The Virtual Reality Symptom Questionnaire (VRSQ), developed by Ames in 2005 (Ames et al., 2005), was administered immediately after the VRAT-controllers condition and the VRAT-sensors condition to evaluate symptoms of cybersickness, a type of motion sickness caused by exposure to VR. The questionnaire assesses eight general physical side effects (general discomfort, fatigue, boredom, drowsiness, headache, dizziness, concentration difficulties, and nausea) and five visual effects (tired eyes, aching eyes, eyestrain, blurred vision, and difficulties focusing) on a seven-point scale (0–6), with 0-scores indicating no symptoms and higher scores indicating more severe symptoms. In the validation study, only symptoms that met a minimum correlation coefficient value of 0.2 with the total score were included in the final measure. The Italian version of the VRSQ (Solimini et al., 2011) was used with the participant in this study.

System Usability Scale (SUS)

The SUS is a 10-item measure that required the participant to use a five-point scale ranging from strongly disagree to strongly agree to indicate the extent to which they agree/disagree with positive and negative statements about the VRAT-controller and VRAT-sensor conditions (Brooke, 1996). SUS responses were transformed to a single score ranging from 0 to 100, with higher scores reflecting more favorable usability. The SUS is considered a robust measure of system usability (Bangor et al., 2008), even with a small sample size (Tullis and Stetson, 2004). The Italian version of the SUS was used in this study (Borsci et al., 2009).

Mood, Anxiety, and Cognition

Questionnaires of mood and anxiety symptoms, disposition toward immersive tendencies, and cognitive and functional abilities as well as neuropsychological tests of global and specific cognitive abilities were administered by a trained psychologist (AC). When available, Italian validated versions of questionnaires/tests were used; other measures were translated using a back-translation procedure (see **Table 2**).

Analysis Plan

Descriptive analyses of questionnaires and cognitive tests were performed to characterize the participant. Cognitive test scores also were evaluated by calculating the standardized (Z) score for the participant relative to normative data, using samples that were comparable to the age and education level of the participant. The following formula was used to calculate the Z-score (participant's raw test score – mean of the normative sample/E.S. of the normative sample).

Descriptive data from the RAT, VRAT-controllers, and VRAT-sensors were obtained to compare performance across the testing conditions on measures of (1) the quality of task performance (e.g., errors, accomplishment, time to completion, order of task steps, errors, use and speed of hand movements, etc.); (2) subjective impressions (e.g., attitudes, presence, cybersickness, and usability), and (3) physiological markers of stress.

RESULTS

Characterization of the Participant
Mood Status

Tina's report of depression (Geriatric Depression Scale = 4) and anxiety (Geriatric Anxiety Scale = 12) symptoms was well within the non-clinical range (cut-off scores: GDI > 11; GAI > 17) (Yesavage et al., 1982; Segal et al., 2010; Gould et al., 2014; Galeoto et al., 2018; Gatti et al., 2018).

Cognitive Testing

Raw cognitive test scores along with age- and education-adjusted normative-based Z-scores are reported in **Table 3**. Tina's overall cognitive status, as measured by the MMSE was well within the range of healthy, non-demented people. Scores on most tests of specific abilities fell within the average range, including tests of verbal episodic memory, processing speed, executive functions, and verbal fluency. She performed in the high average range on a verbal test of executive function and in the low average range on a test of visual episodic memory (immediate and delayed free recall).

On questionnaires of cognitive and functional abilities, Tina reported no significant change in her cognitive abilities as compared to 10 years ago [The ECOG SF12 = 1.75, cut-off score = 2.30 (Farias et al., 2008)] and minimal functional difficulties within the normal range [FAQ (score = 6) and the ADL-PI (score = 22)].

On a questionnaire pertaining to one's personal disposition toward immersion (ITQ), Tina reported an average level of

TABLE 2 | Mood and neuropsychological tests performed to characterize the participant.

Variable	Test	Original scale citation	Italian scale used for the study	Validity/Reliability of the instrument
Depression	Geriatric Depression Scale (GDS)	Yesavage et al., 1982	Galeoto et al., 2018	Cronbach's Alpha scored 0.84 in the Italian validated study (Galeoto et al., 2018)
Anxiety	Geriatric Anxiety Scale	Segal et al., 2010	Gatti et al., 2018	Cronbach's Alpha of the Italian scale was = 0.88 (Gatti et al., 2018)
Cognitive abilities	The Everyday Cognition scale short form 12 (ECOG SF12)	Farias et al., 2011	Back Translation procedure has been made for the study purposes	E-Cog has been reported to have high internal consistency ($\alpha = 0.96$). Additionally, the scale demonstrates good test–retest reliability ($r = 0.82$) (Farias et al., 2011)
Functional activities	Functional Activity Questionnaire (FAQ)	Pfeffer et al., 1982	Stancati and Salussi, 2001	The scale has been reported to have high internal consistency ($\alpha > 0.90$) (Pfeffer et al., 1982)
Daily living activities	The Activities of Daily Living-Prevention Instrument (ADL-PI)	Galasko et al., 2006	Back Translation procedure has been made for the study purposes	Test–Retest reliability: was $r = 0.74$ (Galasko et al., 2006)
Education	Brief Intelligence Test (Test di Intelligenza Breve; TIB)	Colombo et al., 2002	Original scale is in Italian	Cronbach's Alpha scored 0.91 (Colombo et al., 2002)
Visual memory	Brief Visual Memory Test Revised (BVMT – R)	Benedict et al., 1996	Argento et al., 2016	Test–retest reliability coefficients ranged from 0.60 for Trial 1 to 0.84 for Trial 3 (Argento et al., 2016)
Verbal fluency	Category Fluency	Sivan and Benton, 1984	Novelli et al., 1986	Test–retest reliability coefficients for the scale was > 0.75 (Kingery et al., 2011)
Processing speed	Trail Making Test-Part B	Armitage, 1946	Gaudino et al., 1995	Validity of the test has been extensively discussed and confirmed (for an extensive review see Sánchez-Cubillo et al., 2009)
Working memory	Digit Span backward	Wechsler et al., 2008	Monaco et al., 2013	The test reliability scored 0.89 (Orsini and Pezzuti, 2015)
Processing speed and visual perception	Symbol search	Wechsler et al., 2008	Orsini and Pezzuti, 2015	The test reliability scored 0.88 (Orsini and Pezzuti, 2015)
Personal disposition toward immersion	Immersive Tendencies Questionnaire	Witmer and Singer, 1998	Scheuchenpflug et al., 2003	The scale reliability scored 0.81 (Witmer and Singer, 1998)

immersion in terms of ability to focus and to become deeply involved in activities (Witmer and Singer, 1998).

Comparisons Across the RAT, VRAT-Controllers, and VRAT-Sensors
Performance Quality

As shown in **Table 4**, Tina made few errors on the breakfast task across all conditions, with most errors on the VRAT-controllers condition. She made no clumsy errors on the RAT, but an equal number of clumsy errors on both VRAT conditions. The Overall Performance Score, which considers accomplishment and the performance of key overt errors was identical across the conditions. Time to completion, also shown in **Table 4**, revealed a longer completion time for the VRAT – controllers than the other two conditions.

A qualitative analysis of the order in which steps were performed showed remarkable consistency. Task steps were performed in the following order across all three conditions: take bread, place bread in toaster, turn on toaster, wait for bread to

TABLE 3 | Characterization of the participant adjusted for age and education.

Test	Subtest	Raw score	Standard score (z-score)	Qualitative descriptor
MMSE		29/30		Within normal limits
BVMT[1]				
	Trial 1	3	−4.66	Impaired
	Trial 2	4	−3.79	Impaired
	Trial 3	8	−1.21	Low Average
	Learning trial	5	−0.18	Average
	Delayed recall trial	5	−2.95	Impaired
RAVLT[2]				
	Total score	45	1.52	Average
	Delayed recall	8	0.20	Average
	Recognition hits	15	2.25	Average
Symbol search[3]		19	1.67	Average
TMT-B[4]		298.21	0.48	Average
DIGIT SPAN backward[5]		5	3.37	High average
Fluency global[6] (Categories: car brand, animal, fruit)		52	2.52	Average
Questionnaires pertaining to cognition, everyday function, and immersive tendencies FAQ		6		
ECOG-Short Form		1.75		
ADL PI		22		
ITQ				
	Focus subscale	31		
	Involvement subscale	21		

BVMT, Brief Visual Memory Test Revised; RAVLT, Rey Auditory Learning Test; TMT-B, Trail Making Test-Part B; FAQ, Functional Activity Questionnaire; ECOG. SF-12, The Everyday Cognition scale short form 12; ADL-PI, The Activities of Daily Living-Prevention Instrument; ITQ, Immersive Tendencies Questionnaire. [1]Normative data for the Italian population (Argento et al., 2016); [2]Normative data (Carlesimo et al., 2002); [3]Normative data for the Italian population (Orsini and Pezzuti, 2015); [4]Normative data for the Italian population (Giovagnoli et al., 1996); [5]Normative data for the Italian population (Monaco et al., 2013); [6]Normative data for the Italian population (Novelli et al., 1986).

toast, remove bread from toaster, add butter to toast, add jelly to toast, add coffee to mug, add milk to mug, add sugar to mug. The final step of stirring the coffee was completed only in the RAT. Tina did not stir the virtual coffee mug in either the VRAT-controller or VRAT-sensor condition; this was coded as an overt (omission) error in both of the VRAT conditions.

Kinematic Results

Hand movements and average hand speed are also shown in **Table 4**. The same pattern of hand movement distance and speed was observed across all conditions – the right hand made more and faster movements than the left hand. There were few differences across conditions, except for a greater reliance on the right hand in the VRAT-controller condition.

A heatmap showing the paths of the right and left hand during each condition is shown in **Figure 2**. Note that the heatmap for the RAT was superimposed on a virtual display for presentation purposes only. The participant actually completed the RAT using real objects as shown in **Figure 1**. The heat maps illustrate subtle differences across conditions. In the RAT, the participant used both hands to perform the steps (i.e., using her left hand to grab the milk bottle, the butter dish and sugar bowl), with each hand performing tasks in the corresponding hemispace. In the VRAT conditions, particularly in the VRAT-controller condition, the participant used the dominant, right hand more frequently, even when completing subtasks in the opposite (left) hemispace.

Physiological Markers

As expected, the lowest stress index was obtained during the RAT (stress index = 4.1); followed by the VRAT-controller (stress index = 4.9) and VRAT-sensor (stress index = 6.2). This result suggests that the participant felt more comfortable with controllers rather than in the sensor condition without the controllers (**Table 4**).

Subjective Impressions

As shown in **Table 4**, Tina reported the most positive attitude toward the VRAT-controllers (Total = 25/25) and the RAT (Total = 24/25). She indicated the lowest score for the VRAT-sensor condition (16/25), as she reported that the VRAT-sensor condition was less "pleasant," "funny," "resting," and "relaxing" (each scored 3 out of 5).

Measures of presence, cybersickness, and usability were obtained after each of the VRAT conditions. As shown in **Table 4**, Tina reported a stronger feeling of presence in the VRAT-controllers than in the VRAT-sensors condition (PQ). Scores for each of the PQ subscales, except the "quality of interface" scale were all higher in the VRAT-controller condition (see **Table 4**). Tina reported no symptoms of cybersickness on VRSQ for either condition (Ames et al., 2005). Finally, Tina reported higher usability ratings for the VRAT-controllers than the VRAT-sensors condition.

DISCUSSION

This paper reports the detailed analysis of a 91-year old woman's (Tina) performance of a real (RAT) and immersive VR breakfast task (VRAT) to evaluate the feasibility of immersive VR for the assessment of everyday function in older adults. Two different VR interfaces were examined: controllers and sensors. Results showed similarities in performance quality, stress, and subjective reports between the RAT and both VRAT conditions, as well as positive ratings and no cybersickness for either VR condition. Taken together the results demonstrate the feasibility of immersive VR for function assessment in older adults and suggest the potential of the validity of this method.

Our results clearly demonstrate the feasibility of immersive VR for function assessment, even in an older adult with very limited computer experience, no prior VR exposure, average educational experiences, and mild cognitive difficulties.

TABLE 4 | Quality task performance, kinematic, physiological data and system usability.

	RAT	VRAT-controllers	VRAT-sensors
Performance analysis			
Accomplishment (%)	100	100	100
Total overt errors	0	1	1
Total micro-errors	3	6	0
Total clumsy errors	0	7	7
Overall Performance (Max = 6)	6	6	6
Completion Time	108.72	203.74	165.81
Kinematic analysis			
Total hand movement (m)			
Right hand	22.91	43.97	28.77
Left hand	11.3	6.71	11.6
Total hand speed (m/s)			
Right hand	0.21	0.21	0.17
Left hand	0.1	0.03	0.07
Physiological data – Mean (SD)			
Baevsky's stress index	4.1	4.9	6.2
Heart rate (bpm)	72.34 (1.54)	68.79 (3.63)	78.15 (2.29)
Galvanic skin response (kohms)	2467 (200)	4714 (820)	893 (44)
Skin temperature (celsius degree)	35.16 (0.07)	34.99 (0.13)	35.22 (0.02)
System usability			
System usability (SUS)		62.5/100	50/100
Sense of presence			
Sense of presence global score (PQ)		113/126	100/126
Realism – subscale		39/49	37/49
Possibility to act – PQ subscale		27/28	18/28
Quality of the interface – PQ subscale		16/21	18/21
Possibility to examine – PQ subscale		19/21	17/21
Self-Evaluation of the performance – PQ subscale		12/14	10/14

SUS, System Usability Scale; PQ, Presence Questionnaire.

The participant was capable of using controllers and sensors to manipulate objects in a purposeful and goal-directed manner in the VR paradigm. She reported no cybersickness and even indicated that interactions in the VR environment were pleasant and relaxing.

Our results also suggest the potential validity of the VR paradigm, as overall performance and accomplishment scores were similar, and task steps were completed in exactly the
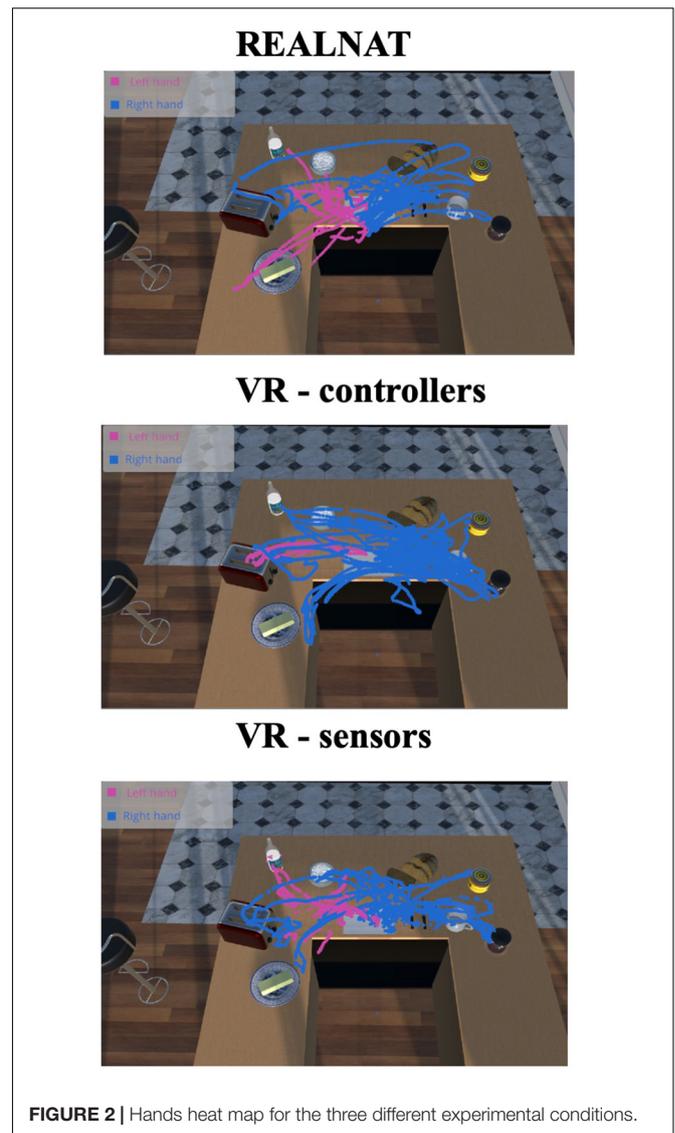
FIGURE 2 | Hands heat map for the three different experimental conditions.

same order across conditions, even though there were numerous opportunities for variation in the order of steps (e.g., coffee could have been made before toast and the order of cream and sugar and butter and jelly was not fixed). Kinematic analyses also were generally comparable between the real (RAT) and the VRAT-sensor condition, and the participant reported positive attitudes toward real (RAT) and both VRAT tasks. These similarities are striking considering that immersive VR was completely unfamiliar to the participant.

Some important differences between the real and VR paradigms were observed and should inform future research. For example, the participant required less time and demonstrated a lower stress index while completing the real breakfast task (RAT). She also demonstrated fewer clumsy errors in the real task as compared to both VRAT conditions. These differences suggest that the real condition was considerably easier for the participant. Training with the VR controllers and sensors was minimal in the present study, and the participant had no prior

experience with VR. Future studies that use VR with older adults should consider including more training to determine whether increased familiarity with the VR environment and practice with VR controllers/sensors may further reduce differences between real and virtual everyday task performance.

In contrast to past research with healthy participants showing advantages with controllers (Caggianese et al., 2019), our results do not clearly indicate which VR interface is best for the assessment of function in older adults, as each interface showed different strengths and weaknesses. When using the controllers, the participant made more micro-errors, and kinematic analyses showed a pattern of hand use that was dissimilar from performance on the real task, such that she appeared to favor her dominant (right) hand for completing the tasks in the VRAT-controller condition. However, she subjectively reported that she preferred the controllers, with higher ratings for usability and positive attitude toward the VRAT-controllers condition. Physiological indicators also reflected lower stress when she used the controllers (VRAT-controllers) than when she used the sensors (VRAT-sensors). By contrast, with the sensors, the participant showed a more natural pattern of use of the right and left hands (see kinematic data). Taken together, the results suggest that if problems in precisely controlling movements in the sensor interface could be addressed in future research, the sensor interface has potential to offer more accurate and naturalistic assessments of everyday function for older adults than controllers.

There are several limitations to acknowledge. First, the extent to which the results are influenced by order effects cannot be determined from this single case report. Future studies should control for and examine task order and practice effects on virtual and real everyday tasks. Future studies with more participants are needed to determine whether our results are generalizable and to establish the full psychometric properties of the VRAT.

In conclusion, our results support the feasibility of immersive VR as a tool to evaluate everyday function in older adults considering also the evaluated safety of the technology as suggested by a recent meta-analysis (Kourtesis et al., 2019). The results also provide guidance on considerations for VR interfaces (sensors vs. controllers). Because of its strong potential to offer objective, sensitive and standardized assessment of everyday function in older adults and a wide range of clinical populations future research on VR assessments is needed to identify optimal interfaces and procedures, compare the utility against non-immersive VR methods (Allain et al., 2014; Giovannetti et al., 2018), and ultimately establish the psychometric properties of immersive VR measures of everyday function. Moreover, the potential for immersive VR systems to offer interventions that might improve everyday functioning and promote independence should be explored (Banville et al., 2018; Foloppe et al., 2018).

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* 50, 179–211. doi: 10.1016/0749-5978(91)90020-T

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee – Department of Psychology of Developmental and Socialization Processes, Sapienza University of Rome, Rome, Italy. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in the manuscript.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00123/full#supplementary-material

Allain, P., Foloppe, D. A., Besnard, J., Yamaguchi, T., Etcharry-Bouyx, F., Le Gall, D., et al. (2014). Detecting everyday action deficits in Alzheimer's disease using a nonimmersive virtual reality kitchen. *J. Int. Neuropsychol. Soc.* 20, 468–477. doi: 10.1017/S1355617714000344

Ames, S. L., Wolffsohn, J. S., and Mcbrien, N. A. (2005). The development of a symptom questionnaire for assessing virtual reality viewing using a head-mounted display. *Optom. Vis. Sci.* 82, 168–176. doi: 10.1097/01.OPX. 0000156307.95086.6

Argento, O., Smerbeck, A., Pisani, V., Magistrale, G., Incerti, C. C., Caltagirone, C., et al. (2016). Regression-based norms for the brief visuospatial memory test-revised in Italian population and application in MS patients. *Clin. Neuropsychol.* 30, 1469–1478. doi: 10.1080/13854046.2016.1183713

Armitage, S. G. (1946). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychol. Monogr.* 60, 41–48. doi: 10.1037/h0093567

Baevsky, R. Ì, and Berseneva, A. P. (2008). Methodical Recommendations use kardivar System for Determination of the Stress Level and Estimation of the Body Standards of Measurements and Physiological Interpretation. Available at: https://pdfs.semanticscholar.org/74a2/ 92bfafca4fdf1149d557348800fcc1b0f33b.pdf (accessed August 16, 2019).

Bangor, A., Kortum, P. T., and Miller, J. T. (2008). An empirical evaluation of the system usability scale. *Int. J. Hum. Comput. Interact* 24, 574–594. doi: 10.1080/10447310802205776

Banville, F., Couture, J. F., Verhulst, E., Besnard, J., Richard, P., and Allain, P. (2017). "Using virtual reality to assess the elderly: The impact of human-computer interfaces on cognition," in *Lecture Notes in Computer Science*, (Berlin).

Banville, F., Lussier, C., Massicotte, E., Verhulst, E., Couture, J. F., Allain, P., et al. (2018). Lecture Notes in Computer Science, ln. *Validation of a Sorting Task Implemented in the Virtual Multitasking Task-2 and Effect of Aging.* (Berlin).

Benedict, R. H. B., Groninger, L., Schretlen, D., Dobraski, M., and Shpritz, B. (1996). Revision of the brief visuospatial memory test: studies of normal performance, reliability, and, validity. *Psychol. Assess* 8, 145–153. doi: 10.1037/ 1040-3590.8.2.145

Borsci, S., Federici, S., and Lauriola, M. (2009). On the dimensionality of the System usability scale: a test of alternative measurement models. *Cogn. Process* 10, 193–197. doi: 10.1007/s10339-009-0268-9

Brooke, J. (1996). "SUS - A quick and dirty usability scale Industrial usability evaluation," in *Usability Evaluation In Industry*, eds P. W. Jordan, B. T. I. L. McClelland, and B. Weerdmeester, (Boca Raton, LA: CRC Press).

Caggianese, G., Gallo, L., and Neroni, P. (2019). "The vive controllers vs. leap motion for interactions in virtual environments: a comparative evaluation," in *Smart Innovation, Systems and Technologies*, eds H. Robert, and J. C. Lakhmi, (Berlin: Springer).

Carlesimo, G. A., Buccione, I., Fadda, L., Graceffa, A., Mauri, M., Lorusso, S., et al. (2002). Normative data of two memory tasks: short-Story recall and Rey's Figure. *Nuova Riv. di Neurol.* 2, 1–13.

Chirico, A., Lucidi, F., De Laurentiis, M., Milanese, C., Napoli, A., and Giordano, A. (2016). Virtual reality in health system: beyond entertainment. a mini-review on the efficacy of VR during cancer treatment. *J. Cell. Physiol.* 231, 275–287. doi: 10.1002/jcp.25117

Cipresso, P., Albani, G., Serino, S., Pedroli, E., Pallavicini, F., Mauro, A., et al. (2014). Virtual multiple errands test (VMET): a virtual reality-based tool to detect early executive functions deficit in parkinson's disease. *Front. Behav. Neurosci.* 8:405. doi: 10.3389/fnbeh.2014.00405

Cipresso, P., and Serino, S. (2014). *Virtual Reality: Technologies, Medical Applications and Challenges.* Hauppauge, NY: Nova Science Publishers, Incorporated.

Cipresso, P., Serino, S., and Riva, G. (2016). Psychometric assessment and behavioral experiments using a free virtual reality platform and computational science. *BMC Med. Inform. Decis. Mak.* 16:37. doi: 10.1186/s12911-016-0276-5

Colombo, L., Sartori, G., and Brivio, C. (2002). Stima del quoziente intellettivo tramite l'applicazione del TIB (test breve di intelligenza). *G. Ital. Psicol.* 3, 613–637.

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y., and Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Front. Psychol.* 6:26. doi: 10.3389/fpsyg.2015.00026

Farias, S. T., Mungas, D., Harvey, D. J., Simmons, A., Reed, B. R., and Decarli, C. (2011). The measurement of everyday cognition: development and validation of a short form of the Everyday Cognition scales. *Alzheimer's Dement.* 7, 593–601. doi: 10.1016/j.jalz.2011.02.007

Farias, S. T., Mungas, D., Reed, B. R., Cahn-Weiner, D., Jagust, W., Baynes, K., et al. (2008). The measurement of everyday cognition (ECog): scale development

and psychometric properties. *Neuropsychology* 22, 531–544. doi: 10.1037/0894-4105.22.4.531

Foloppe, D. A., Richard, P., Yamaguchi, T., Etcharry-Bouyx, F., and Allain, P. (2018). The potential of virtual reality-based training to enhance the functional autonomy of Alzheimer's disease patients in cooking activities: A single case study. *Neuropsychol. Rehabil.* 28, 709–733. doi: 10.1080/09602011.2015. 1094394

Galasko, D., Bennett, D. A., Sano, M., Marson, D., Kaye, J., and Edland, S. D. (2006). ADCS Prevention Instrument Project: Assessment of instrumental activities of daily living for community-dwelling elderly individuals in dementia prevention clinical trials. *Alzheimer Dis. Assoc. Disord.* 20(4 Suppl. 3), S152–S169. doi: 10.1097/01.wad.0000213873.25053.2b

Galeoto, G., Sansoni, J., Scuccimarri, M., Bruni, V., De Santis, R., Colucci, M., et al. (2018). A psychometric properties evaluation of the Italian version of the geriatric depression scale. *Depress. Res. Treat.* 2018:1797536. doi: 10.1155/2018/ 1797536

Gatti, A., Gottschling, J., Brugnera, A., Adorni, R., Zarbo, C., Compare, A., et al. (2018). An investigation of the psychometric properties of the Geriatric Anxiety Scale (GAS) in an Italian sample of community-dwelling older adults. *Aging Ment. Heal.* 22, 1170–1178. doi: 10.1080/13607863.2017.1347141

Gaudino, E. A., Geisler, M. W., and Squires, N. K. (1995). Construct validity in the trail making test: what makes Part B harder? *J. Clin. Exp. Neuropsychol.* 17, 529–535. doi: 10.1080/01688639508405143

Giovagnoli, A. R., Del Pesce, M., Mascheroni, S., Simoncelli, M., Laiacona, M., and Capitani, E. (1996). Trail making test: normative values from 287 normal adult controls. *Ital. J. Neurol. Sci.* 17, 305–309. doi: 10.1007/BF01997792

Giovannetti, T., Bettcher, B. M., Brennan, L., Libon, D. J., Burke, M., Duey, K., et al. (2008a). Characterization of everyday functioning in mild cognitive impairment: a direct assessment approach. *Dement. Geriatr. Cogn. Disord.* 18, 787–798. doi: 10.1159/000121005

Giovannetti, T., Bettcher, B. M., Brennan, L., Libon, D. J., Kessler, R. K., and Duey, K. (2008b). Coffee with jelly or unbuttered toast: commissions and omissions are dissociable aspects of everyday action impairment in Alzheimer's Disease. *Neuropsychology* 22, 235–245. doi: 10.1037/0894-4105.22.2.235

Giovannetti, T., Libon, D. J., Buxbaum, L. J., and Schwartz, M. F. (2002). Naturalistic action impairments in dementia. *Neuropsychologia.* 40, 1220–1232. doi: 10.1016/S0028-3932(01)002299

Giovannetti, T., Richmond, L. L., Seligman, S. C., Seidel, G. A., Iampietro, M., Bettcher, B. M., et al. (2013). "A process approach to everyday action assessment," in *The Boston Process Approach to Neuropsychological Assessment: A Practitioner's Guide*, eds L. Ashendorf, R. Swenson, and D. Libon (Oxford University Press), 355–379.

Giovannetti, T., Yamaguchi, T., Roll, E., Harada, T., Rycroft, S. S., Divers, R., et al. (2018). The Virtual Kitchen Challenge: preliminary data from a novel virtual reality test of mild difficulties in everyday functioning. *Aging, Neuropsychol. Cogn.* 26, 823–841. doi: 10.1080/13825585.2018.1536774

Giovannetti, T., Yamaguchi, T., Roll, E., Harada, T., Rycroft, S. S., Divers, R., et al. (2019). The virtual kitchen challenge: preliminary data from a novel virtual reality test of mild difficulties in everyday functioning. *Aging, Neuropsychol. Cogn.* 26, 823–841. doi: 10.1080/13825585.2018.1536774

Gold, D. A., Park, N. W., Murphy, K. J., and Troyer, A. K. (2015). Naturalistic action performance distinguishes amnestic mild cognitive impairment from healthy aging. *J. Int. Neuropsychol. Soc.* 21, 419–428. doi: 10.1017/S135561771500048X

Gould, C. E., Segal, D. L., Yochim, B. P., Pachana, N. A., Byrne, G. J., and Beaudreau, S. A. (2014). Measuring anxiety in late life: a psychometric examination of the geriatric anxiety inventory and geriatric anxiety scale. *J. Anxiety Disord.* 28, 804–811. doi: 10.1016/j.janxdis.2014.08.001

Indovina, P., Barone, D., Gallo, L., Chirico, A., De Pietro, G., and Giordano, A. (2018). Virtual reality as a distraction intervention to relieve pain and distress during medical procedures. *Clin. J. Pain* 34, 858–877. doi: 10.1097/AJP. 0000000000000599

Kessler, R. K., Giovannetti, T., and MacMullen, L. R. (2007). Everyday action in schizophrenia: performance patterns and underlying cognitive mechanisms. *Neuropsychology.* 21, 439–447. doi: 10.1037/0894-4105.21.4.439

Kingery, L., Bartolic, E., Meyer, S., Perez, M., Bertzos, K., Feldman, H., et al. (2011). Test-Retest Reliability of the MMSE, ADAS-Cog, and executive function tests in a phase II mild to moderate Alzheimer's disease study. *Alzheimer's Dement* 7, S249. doi: 10.1016/j.jalz.2011.05.711

Kourtesis, P., Collina, S., Doumas, L. A. A., and MacPherson, S. E. (2019). Technological competence is a pre-condition for effective implementation of virtual reality head mounted displays in human neuroscience: a technological review and meta-analysis. *Front. Hum. Neurosci* 13:342. doi: 10.3389/fnhum. 2019.00342

Monaco, M., Costa, A., Caltagirone, C., and Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. *Neurol. Sci.* 13:342. doi: 10.1007/s10072-012-1130-x

Nolin, P., Banville, F., Cloutier, J., and Allain, P. (2013). Virtual reality as a new approach to assess cognitive decline in the elderly. *Acad. J. Interdiscip. Stud* 2, 612–616.

Novelli, G., Papagno, C., Capitani, E., Laiacona, M., Vallar, G., and Cappa, S. F. (1986). Tre test clinici di ricerca e produzione lessicale. Taratura su soggetti normali. . *Arch. Psicol. Neurol. Psichiatr* 47, 477–506.

Orsini, A., and Pezzuti, L. (2015). *WAIS IV Contributo Alla Taratura Italiana (70 - 90 anni)*, ed. O. S. Giunti, (Italy:OrganizzazioniSecialiFirenze).

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Front. Hum. Neurosci.* 9:660. doi: 10.3389/fnhum.2015.00660

Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Chance, J. M., and Filos, S. (1982).). Measurement of functional activities in older adults in the community. *J. Gerontol.* 37, 323–329. doi: 10.1093/geronj/37.3.323

Ruddle, R. (2005). 3D User interfaces: theory and practice - REVIEW. *Teleoperators Virtual Environ.* 14, 117–118.

Rycroft, S. S., Giovannetti, T., Divers, R., and Hulswit, J. (2018). Sensitive performance-based assessment of everyday action in older and younger adults. *Aging, Neuropsychol. Cogn.* 25, 259–276. doi: 10.1080/13825585.2017.1287855

Sánchez-Cubillo, I., Periáñez, J. A., Adrover-Roig, D., Rodríguez-Sánchez, J. M., Ríos-Lago, M., Tirapu, J., et al. (2009). Construct validity of the Trail making test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *J. Int. Neuropsychol. Soc.* 15, 438–450. doi: 10.1017/S1355617709090626

Scheuchenpflug, R., Ruspa, C., and Quattrocolo, S. (2003). "Presence in virtual driving simulators," in *Human Factors in the Age of Virtual Reality* , eds D. de Waard, K. A. Brookshuis, S. M. Breker, and W. B. Verwey, (Düren: Shaker), 143–148.

Schwartz, M. F. (2006). The cognitive neuropsychology of everyday action and planning. *Cogn. Neuropsychol.* 23, 202–221. doi: 10.1080/02643290500202623

Schwartz, M. F., Segal, M., Veramonti, T., Ferraro, M., and Buxbaum, L. J. (2002). The Naturalistic action test: a standardised assessment for everyday action impairment. *Neuropsychol. Rehabil.* 12, 311–339. doi: 10.1080/09602010244000084

Segal, D. L., June, A., Payne, M., Coolidge, F. L., and Yochim, B. (2010). Development and initial validation of a self-report assessment tool for anxiety among older adults: the Geriatric Anxiety Scale. *J. Anxiety Disord.* 24, 709–714. doi: 10.1016/j.janxdis.2010.05.002

Shahrbanian, S., Ma, X., Aghaei, N., Korner-bitensky, N., and Simmonds, M. J. (2012). Use of virtual reality (immersive vs. non immersive) for pain management in children and adults: a systematic review of evidence from randomized controlled trials. *Eur. J. Exp. Biol.* 2, 1408–1422.

Sivan, A. B., and Benton, A. L. (1984). Problems and conceptual issues in neuropsychological research in aging and dementia. *J. Clin. Neuropsychol.* 6, 57–63. doi: 10.1080/01688638408401196

Slater, M. (2002). Measuring presence: a response to the witmer and singer presence questionnaire. *Presence Teleoperators Virtual Environ.* 8, 560–565. doi: 10.1162/105474699566477

Solimini, A. G., Mannocci, A., and di Thiene, D. (2011). A pilot application of a questionnaire to evaluate visually induced motion sickness in spectators of tri-dimensional (3D) movies. *Ital. J. Public Health.* 12:779.

Stancati, A., and Salussi, F. (2001). *Scale di Valutazione e Malattie Reumatiche.* IBS.it: Assago, 1885.

Tarvainen, M. P., Niskanen, J. P., Lipponen, J. A., Ranta-aho, P. O., and Karjalainen, P. A. (2014). Kubios HRV - Heart rate variability analysis software. *Comput. Methods Programs Biomed.* 113, 210–220. doi: 10.1016/j.cmpb.2013.07.024

Tullis, T. S., and Stetson, J. N. (2004). "A comparison of questionnaires for assessing website usability," in *UPA 2004 Presentation*, (Boston, MA).

Wechsler, D., Coalson, D., and Raiford, S. (2008). *Wechsler Adult Intelligence Testin Fourth Edition Technical and Interpretive Manual.* San Antonio, TX: San AntonioPearson.

Witmer, B. G., Jerome, C. J., and Singer, M. J. (2005). The factor structure of the Presence Questionnaire. *Presence Teleoperators Virtual Environ.* 14, 298–312. doi: 10.1162/105474605323384654

Witmer, B. G., and Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence Teleoperators Virtual Environ.* 7, 225–240. doi: 10.1162/105474698565686

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1982). Development and validation of a geriatric depression screening scale: a preliminary report. *J. Psychiatr. Res.* 17, 37–49. doi: 10.1016/0022-3956(82)90033-4