THE LIMITATIONS AND CHALLENGES IN THE ASSESSMENT OF EXECUTIVE DYSFUNCTION ASSOCIATED WITH REAL-WORLD FUNCTIONING: THE OPPORTUNITY OF SERIOUS GAMES

D. Martínez-Pernía, L. Olavarría, V. Cabello, F. Henríquez, B, L. Alvarado, S. Barria, A. Antivilo, J. Velasquez, B. Fernández-Manjón, A. Ibáñez, M. Parra ; S. Gilbert; A. Slachevsky

ABSTRACT

Executive function -EF- is one of the human behavior's most complex cognitive domains. Its dysfunction is one main factor associated with an ineffective management in the activities of daily lives. Nowadays, there is a broad range of methods for detecting and evaluating executive dysfunction ranging from clinical interview to neuropsychological evaluation. Nevertheless, a critical issue of these assessments is the lack of correspondence of the neuropsychological test's results with real-world functioning. In this paper, we briefly discuss the contribution of current methods of evaluation of executive dysfunction (paper-and-pencil tests, naturalistic observation methods, and Information and Communications Technologies). We analyze what are the limitations of these methods to predict real-world performance: 1) A lack of appropriate instruments to investigate the complexity of real-world functioning (e.g., regulation of one's social behavior, decision-making involving emotional interpretation, multitasking in unstructured environments) 2) the vast majority of neuropsychological tests assess well-structured tasks, and 3) measurement of behaviors are based on simplistic data collection and statistical analysis. Finally, this paper proposes serious games as a new framework to improve the neuropsychological assessment of real-world functioning caused by executive dysfunction. We show how serious games offer an opportunity to develop more efficient tools to detect executive dysfunction in everyday life contexts. Serious games provide meaningful narrative stories and virtual or real environments that immerse the user in natural and social environments with social interactions. In those highly interactive game environments the player needs to adapt his/her behavioral performance to novel and ill-structured tasks (social, cognitive, and interpersonal) which are suited for collecting user interaction evidence. However, more research is still needed to implement serious games in everyday clinical practice

Keywords: Executive function, neuropsychological assessment, ecological validity, real-world functioning, Information and Communications Technologies, serious game.

1.- INTRODUCTION

Executive function (EF) is one of the most complex cognitive domains of human behavior, underlying goal-directed actions (Best & Miller, 2010; Burgess et al., 2006; Chan et al., 2008; Chaytor et al., 2006; Gioia & Isquith, 2004; Manchester et al., 2004). Muriel Lezak defined the EF as "*those mental capacities necessary for formulating goals, planning how to achieve them, and carrying out the plans effectively*" (Lezak, 1982). In pragmatic terms, EF acts when the person needs to take an active position in controlling their behavior, going from automatic processing to the formulation planning and execution of objectives (Stuss, 2009). EF comprises a wide range of cognitive subdomains and behavioral competencies, such as cognitive flexibility, the ability to move from one task to another ("shifting"), the inhibition of automatic responses not adequate to the context, the ability to keep in mind the information necessary to carry out the ongoing actions, the ability to modify behavior to adapt to the context, multitasking, problem-solving, resistance to interference, self-awareness, metacognition, and the capacity to address novelty (Burgess et al., 2000; Chan et al., 2008), among others.

The term EF is used in multiple ways in the literature. The confusion between "executive dysfunction", a psychological construct, and "frontal lobe dysfunction", an anatomical syndrome, represents the main difficulty in defining these syndromes. The two concepts have often been used interchangeably (Stuss, 2011). Nevertheless, it is now well established that EF depends multiple and distributed systems (Duncan et al., 2020)

Nowadays, there are different views of EF. For example, some authors differentiate EF to refer exclusively to the cognitive process, principally monitoring and task-setting, from self-regulation of behavior to the ability to monitor and modulate cognition, emotion, and behavior (Stuss, 2011, Berger et al., 2007). On the other hand, other authors have proposed to include in EF both cognitive processes (the cold component of EF) and self-regulation of behavior (the hot component of EF) (Chan, 2008). This manuscript will consider EF in this wide meaning (cognitive processes and self-regulation) because effective behavior management in daily life depends on integrating cognition and emotion.

In neuropsychiatric disorders, EF disorder, called "dysexecutive syndrome" or "executive dysfunction", has been described as main factor associated with functional impairment, which are manifested as problems with managing daily routine or emotional lability, experienced by persons with brain disorders independently of its etiology (Slachevsky *et al.*, 2009; Cieza, Anczewska, *et al.*, 2015; Cieza, Sabariego, *et al.*, 2015). Moreover, functional impairment due to executive dysfunction is observed in neurological disorders (Royall et al., 2002; Sabariego et al., 2015) such as brain tumor (Robinson et al., 2014), traumatic brain injury (Karr et al., 2014), cerebrovascular disease (Hua et al., 2014), Parkinson disease (Gasca-Salas et al., 2014), Multiple sclerosis (Phillips et al., 2014), Tourette syndrome (Eddy et al., 2012), attention-deficit/hyperactivity disorder (Sebastian et al., 2014), and dementia (Marshall et al., 2011; Moheb et al., 2017), but also in psychiatric disorders, for example schizophrenia (Xu et al., 2014), or obsessive-compulsive disorder (Nakao et al., 2014), among others. The transversality of the functional impact in such diverse neuropsychiatric conditions is described as "horizontal epidemiology" (Cieza et al., 2015)

Therefore, regardless of the pathology that produces it, executive dysfunction might cause people to be significantly affected in their ability to manage themselves in their daily context effectively. Thus, an adequate assessment of executive dysfunction might be a good predictor and indicator of the functional impairment (Adamit et al., 2015; Garrett et al., 2019; Romero-Ayuso et al., 2019), regardless of age (Chan et al., 2008; Hosenbocus & Chahal, 2012) or the severity of the underlying disease (McDonald et al., 2002). Even though there is a substantial number of EF assessment methods, there is consensus that there are numerous limitations and barriers to effectively predicting functional impact (Burgess et al., 1998, 2006).

This paper proposes serious games as a new framework to improve the detection of real-world functioning caused by executive dysfunction, providing meaningful narrative stories and virtual or real environments that immerse the user in natural and social environments with social interactions. We address three topics consecutively: First, we will briefly review the current methods of evaluation of executive dysfunction. Second, we will describe the main limitations of these methods to predict real-world performance. Finally, we will show how the use of Information and Communication Technologies (ICT), and more specifically, the use of serious games, may offer a novel opportunity to develop better tools to improve diagnosis of the executive dysfunction in everyday life contexts.

2.- CURRENT ASSESSMENT OF EXECUTIVE FUNCTION

Nowadays, there are different proposals for evaluating executive function associated with real-world functioning (Chan et al., 2008; Verdejo-García & Bechara, 2010), making it clear that there is no single instrument for this purpose. The assessment methods of executive dysfunction are the clinical interview and the traditional neuropsychological assessment. Currently, these two methods are the most informative to determine cognitive and behavioral sequelae of brain diseases (Bigler, 2016), given that brain imaging studies, by themselves, show a minimum value when predicting such deficits (Slachevsky, Reyes, Rojas, Silva, 2009). For this article, we are going to focus on neuropsychological assessment only (Table 1 summarizes some of the main evaluations of executive dysfunction).

Over time, the neuropsychological assessment of executive dysfunction has evolved into three different main methods. i) The first method, and probably the most used, is the paper-and-pencil tests, and there are numerous specially designed for the evaluation of executive function. While some of them are intended to measure a single subdomain of executive function, other ones seek to measure several aspects of this cognitive area (Burgess, 2003). Although clinicians widely use them, paper-and-pencil tests are not exempt from limitations. Firstly, this assessment requires the physical presence of an expert evaluator (Brearly et al., 2017; Valladares-Rodríguez et al., 2016), restricting access to neuropsychological services depending on the country's health system (Adjorlolo, 2015; Clement et al., 2001; Jacobsen et al., 2003). Another critical issue is the lack of correspondence of the test's results with everyday life, which is a condition every executive function test should meet to predict behavior (Burgess, 2003).

ii) Subsequently, instruments were developed based on naturalistic observation, which consists of an observational assessment while the patient performs specific daily life activities. Some of these are the Multiple Errands Test -MET- (Shallice and Burgess, 1991), the Executive Function Route-finding Task -EFRT- (Boyd and Sautter, 1993), the Executive Function Performance Test -EFPT- (Baum et al., 2008; Weiner et al., 2012), and the Functional Independence Measure -FIM- (Keith, 1987). This assessment method is generally developed in facilities specially designed to simulate real environments (Giovannetti et al., 2002) and focused on measuring the participant's activity and participation level rather than their impairment level (Poulin, Korner-Bitensky, & Dawson, 2013). These real-life assessments have shown to be more accurate than assessments developed in laboratory settings (Burgess et al., 2006; Debbie Rand et al., 2009). Nevertheless, they present some limitations, such as that they are time-consuming and it is challenging to obtain standardized measures among administration centers (Chevignard et al., 2000; Fortin et al., 2003).

Finally, with the arrival of the computer, information and communication technologies (ICT) were developed, highlighting three primary assessment methods: Computer-based neuropsychological assessment; Virtual reality; and Serious Games.

1) <u>Computer-based neuropsychological assessment:</u> This is an assessment procedure similar to the paper-and-pencil tests but administered by computational

support. It is focused on identifying cognitive impairment (Edwards et al., 2014). Some examples of computerized versions applied to assess EF are the Tower of Hanoi (Mataix-Cols & Bartrés-Faz, 2002), the Category Test (Choca & Morris, 1992), the Raven Progressive Matrices Test (Williams & McCord, 2006), and the CANTAB (Sahakian & Owen, 1992), which have shown a high correlation with the paper-and-pencil tests (Choca & Morris, 1992; Hoskins et al., 2010; Mataix-Cols & Bartrés-Faz, 2002; Sahakian & Owen, 1992; Schatz & Browndyke, 2002; Williams & McCord, 2006).

2) Virtual Reality (VR): This is a technological assessment that simulates real-life situations with the potential to have higher ecological validity (Parsey & Schmitter-Edgecombe, 2013; Parsons, 2011). This evaluation allows the user to immerse themselves in interactive three-dimensional environments that reproduce natural environments and situations that directly measure the functional limitations caused by neuropsychological impairments. For example, some of the tasks that have been used to detect executive dysfunction associated with a functional impairment through VR are: to perform an errand task in a three-floor building (McGeorge et al., 2001), to buy fruit in a store(V-Store) (Lo Priore et al., 2003), to make a purchase in a supermarket (VAP-S, VMall, VMET) (Klinger et al., 2004; D. Rand et al., 2005; Debbie Rand et al., 2009), to prepare soup in a kitchen setting (Zhang et al., 2001), and the development of different tasks both in the house and the city (Kourtesis et al., 2020), among others.

3) Serious games (SG): A third alternative for assessing executive dysfunction through the ICTs is the use of serious games. SG are defined as specialized digital games whose main purpose is not entertainment (Michael & Chen, 2005; Robert et al., 2014). At present, the use of video games for purposes other than mere enjoyment has been well received in the field of rehabilitation (Legouverneur et al., 2011; Manera et al., 2015; Martínez-Pernía et al., 2017) and as a virtual training simulator to improve real-life skills (Vallejo et al., 2017). Besides, there have been several SG developments to assess executive dysfunction. For instance, Levy and colleagues (Levy et al., 2019) created a virtual environment with a seven-aisle grocery story (V-mart) where participants are required to perform

different tasks (time management, budgeting, planning) of increasing complexity. Another example is the multitasking in the city test -MCT- (Jovanovski et al., 2012a, 2012b) where participants have to perform specific errands while moving through the city composed of different services, shops, and the participant's home. Finally, another example is the Non-immersive Virtual Coffee task (NI-VCT), where participants have to make a cup of coffee with a coffee machine by selecting from a virtual workbench the items necessary to achieve the task (Besnard et al., 2016). Furthermore, different studies developed with SG have shown both good psychometric properties and ecological validity, measured with naturalistic event-based tasks and in the real-word, in the assessment of the executive dysfunction, such as the Video Assessment of Prospective Memory -VAPM- (Clune-Ryberg et al., 2011), the Virtual Interactive Shopper -VIS (Hadad et al., 2012), the Virtual Library Task -VLT- (Renison et al., 2012), the Virtual Reality automated Teller Machine -VR-ATM- (Fong, et al., 2010), and the Virtual Reality-Based Prospective Memory Training Programme -VRPM- Assessment Scenario (Yip & Man, 2013).

Despite the advances during last years, neuropsychological assessments (traditional and technological) still have substantial limitations, which are reflected in a relatively low capacity to predict the impact of executive dysfunction in daily life activities (Chaytor et al., 2006; Chaytor & Schmitter-Edgecombe, 2003; McGeorge et al., 2001; Sbordone, 1996; Spooner & Pachana, 2006; Tupper & Cicerone, 1990; Zhang et al., 2001).

Table 1

Paper-and-pencil tests	Naturalistic observation	ICT

Test based on specific cognitive domains: Rey Complex Figure Test (Osterrieth, 1944) Rivermead Behavioral Memory Test (Wilson et al., 1985) Test of Everyday Attention (Robertson et al., 2001) Wisconsin Card Sorting Test (WCST) (1996) Trail Making Test (TMT) (1958) Neuropsychological batteries: Wechsler Adult Intelligence Scale (WAIS) (1939) Luria-Nebraska Battery (Golden et al., 1985) Questionnaires: Behavioral Assessment of the Dysexecutive Syndrome (Wilson et al., 1996). Behavioral Dysexecutive Syndrome Inventory (BDSI) (2010) The Neuropsychiatric Inventory (NPI) (1994)	Multiple Errands Test -MET- (Shallice and Burgess, 1991) Executive Function Route-finding Task -EFRT- (Boyd and Sautter, 1993) Executive Function Performance Test -EFPT- (Baum et al., 2008; Weiner et al., 2012) Functional Independence Measure -FIM- (Keith, 1987).	Computer-based neuropsychological assessment: Tower of Hanoi (Mataix-Cols & Bartrés-Faz, 2002) the Category Test (Choca & Morris, 1992) Raven Progressive Matrices Test (Williams & McCord, 2006) CANTAB (Sahakian & Owen, 1992), Virtual Reality V-Store (Lo Priore et al., 2003) Virtual Action Planning-Supermarket -VAP-S- (Klinger et al., 2004) Virtual Mall (Rand et al., 2005) Virtual MET (Rand et al., 2009) Serious Games Virtual Reality Grocery (V-Store) (Levy et al., 2019) The multitasking in the city test -MCT- (Jovanovski et al., 2012a, 2012b) Non-immersive Virtual Coffee task (NI-VCT) Besnard et al., 2016).

Table 1: Tests to detect executive dysfunction

3.- LIMITATIONS OF CURRENT NEUROPSYCHOLOGICAL INSTRUMENTS FOR THE EVALUATION OF EXECUTIVE DYSFUNCTION IN THE DETECTION AND PREDICTION OF FUNCTIONAL IMPAIRMENT

Despite recent progress in the development and refinement of neuropsychological instruments to evaluate executive dysfunction, considerable challenges remain. We summarize three key issues below.

i) Content of evaluation

While there has been a long history of assessments for investigating specific aspects of executive functioning, such as cognitive flexibility or set-shifting ability (e.g., Grant & Berg, 1948; Reitan, 1958), other aspects of executive functioning have traditionally received far less attention. Examples of these relatively-neglected areas include contextual control (Ibañez & Manes, 2012), prospective memory (Fish et al., 2010), and multitasking (Burgess et al., 2000). Each of these could be associated with functional impairment associated with executive dysfunction; however the relative paucity of appropriate measures for assessment makes this hard to evaluate. It should also be noted that executive dysfunction is also likely to play a role in domains that are not traditionally considered to fall within the definition of executive function, for example, aspects of social cognition (Adolphs, 1999; Amodio & Frith, 2006), emotion regulation (Goldin et al., 2008), and creativity (Gonen-Yaacovi et al., 2013). A lack of appropriate instruments to investigate these aspects of cognition can also lead to a failure to evaluate functional impairment appropriately. Furthermore, even when these domains have been evaluated, they are typically considered in isolation. This contrasts with real-world situations in which cognitive demands across multiple domains must be integrated (e.g., tasks that involve both social and executive demands). As a result, the standard tools for evaluating executive dysfunction could mask difficulties that only become apparent when multiple demands, across distinct cognitive domains, need to be integrated (Mesulam et al., 2014).

ii) Task structure

The vast majority of tests used for neuropsychological assessment are "well-structured" (for discussion of this concept, see Gilbert et al., 2010; Goel & Grafman, 2000). That is, the goal of the task, the set of appropriate responses, and the criteria to evaluate whether or not the goal has been achieved are clearly specified. This contrasts with "ill-structured" tasks, which typically lack well-defined criteria for evaluating whether the goal has been met. Ill-structured tasks may also be open-ended in the sense that it is not obvious at what point the task has been completed. It is up to the individual to set their own criteria for whether or not they have been successful, and when the task is complete. Although ill-structured tasks are rare in neuropsychological assessment, this type of task may be much more characteristic of everyday life than the standard assessment procedures used by neuropsychologists.

iii) Measurement of performance

The arguments above highlight the utility of developing new assessment methods that can provide rich multidimensional datasets in the context of ill-structured tasks. Such methods will also require new analysis techniques. This is for two key reasons. First, whereas traditional tasks such as the Trail Making Test yield only a small number of measures (e.g. completion time) which can be analyzed separately with a univariate approach, more complex datasets with a large number of measures will require multivariate methods. A second issue when it comes to ill-structured tasks is that, by definition, such tasks lack clear criteria according to which an individual's performance can be scored as "correct" or "incorrect". Instead, it is up to the individual to set their own criteria. In order to link functional outcomes or diagnostic categories with multidimensional datasets from ill-structured tasks, novel statistical approaches will be required. One promising approach is to use machine learning methods to reduce the dimensionality of complex datasets. For an example of such an approach applied to the Trail Making Test, see Dahmen et al. (2017).

Therefore, there are still critical deficiencies in detecting and predicting functional impairment associated with executive dysfunction. An analysis resulting from the insights in the previous paragraphs is related to the fact that these assessment methods consider the EF in a narrow sense. This means, on the one hand, that these methods focus on assessing some specific subdomains of the EF, while other subdomains are rarely considered part of real-world functioning (e.g., contextual control, decision-making intersubjective contexts, multitasking). On the other hand, and more critically, these evaluation methods prioritize the understanding of EF as cognitive processes depending on mechanistic and logical tasks, but they omit those cognitive processes influenced by social and interactional context. Nowadays, the neuropsychological evaluation assesses the performance of the person embedded in the environment (e.g., buying food in the supermarket, cooking in a kitchen, self-caring in the house). However, it does not assess how the social-interaction influences these processes (e.g., regulation of one's social behavior, decision-making involving emotional interpretation, multitasking in unstructured environments, resistance to social-context-interference, understanding explicit and implicit contextual keys, among others). Predicting real-world behavior based only on non-social tasks is a key deficiency that needs to be addressed in neuropsychological assessment; because the ability to interpret and coordinate behaviors in social situations is one of the main functions of prefrontal cortex and executive function (Hari & Kujala, 2009), and this ability enables appropriate behavior in daily life (Ibañez & Manes, 2012). Mesulam (1986) argues that conventional neuropsychological tests are insensitive to executive dysfunction because their diagnostic capacity places cognition as a phenomenon outside the person's daily activity and of social context. Therefore, traditional tests are limited for the evaluation of real-world dysexecutive function because their cognitive demands do not integrate fundamental cognitive and emotional processes into the carrying out actions in everyday life (Burgess et al., 2009), limiting the ability to infer the meaning of the performance of these tests to the personal and social conditions of the evaluated individual. In addition, Damasio (1994) points out that these evaluation methods lack ecological utility as "instruments of rationality", which means that the mind is understood in logical and abstract terms (words, numbers, objects, space, temporality). Some authors consider that the neuropsychological tests do not evaluate cognition according to the facilities or difficulties that a person presents in their daily lives but only as a disembodied cognition, separated from the cultural and interpersonal context (Martínez-Pernía, 2020; Damasio, 1994).

Therefore, and to improve the detection and prediction of functional impairment, we believe that a neuropsychological assessment that considers EF in a broad sense is necessary. That is an assessment that, in addition to assessing simultaneously the different components of EF (Chan et al., 2008; Schwartz et al., 1991; Shallice & Burgess, 1991; B. A. Wilson et al., 1996), assesses how these processes are influenced by the cultural context and social interaction. In the following section, we will show how information and communication technology (ICT), and more specifically, SG, offers an opportunity to develop tools to more efficiently diagnose executive dysfunction in everyday life contexts.

4.- FUTURE PERSPECTIVES: THE OPPORTUNITY OF SERIOUS GAMES

In the scientific community, there is a consensus to develop tools for the more efficient evaluation of executive dysfunction in real-life contexts and with a greater capacity to predict real-world functioning (Burgess et al., 1998, 2006). Given the recent analysis presented in the previous section, we consider that the two most promising lines of research related to ICTs are: on the one hand, immersive technologies (VR and augmented reality) and on the other hand, SG. However, the frontier between both technologies is not always straightforward (e.g., Corti et al., 2021), and besides, there are SG that use immersive technologies (e.g., Sánchez-Herrera-Baeza et al., 2020; Vogiatzaki, E., & Krukowski, 2014).

The main characteristic of immersive technologies is that they immerse the user in a 3D virtual world, either entirely in the VR or combined with the real world in augmented reality (Cipresso 2018). Immersive technologies, and especially VR, are being used in the detection of executive dysfunction, but above all, they are being applied extensively in the creation of new treatments and interventions in neurological disorders (Ghai et al., 2020; Saredakis et al., 2020; Stasolla et al., 2020). For instance, VR memory training improved cognitive abilities in participants navigating along with virtual environments with a head-mounted system (Optale et al., 2010). However, despite immersive technologies' evolution, they are still not completely mature and have not yet been widely adopted in society, making their application more costly and difficult in neuropsychological assessment (Rizzo, 2017). Moreover, VR still has critical issues related to body-machine interactions, social interactions, and technological developments that hinder its simple implementation in detecting real-world behavior. Below, we explain these three limitations.

1) Concerning body-machine interactions with VR, it has been shown that the performance of a test in a semi-closed environment (VR glasses) could produce anxiety, preventing participants (cognitively impaired and healthy) from completing the assessment (M. Rizzo et al., 2003; Stanney et al., 2002, Saredakis et al., 2019). Besides, the use of head-mounted display systems causes motion sickness (dizziness, nausea, headache, instability, fatigue), which endangers the health and safety of the participants (Parsons et al., 2018), affects behavioral and cognitive performances (Nalivaiko et al., 2015), and decreases the reliability of data (Kourtesis et al., 2019). Additionally, most VR produces a strong sense of stimulation and immersion to generate a vivid virtual environment, which may be unsuitable in some populations (e.g., people with heart diseases) or, being cautious in its use (e.g., elderly people) (Ning et al., 2020).

2) Concerning social interactions, realistic social interactions are not yet possible to implement in VR (Pan & Hamilton, 2018). This is caused because i) the avatars' appearance and behavior are not "truly life-like", and ii) although game users can interpret avatars' thoughts and emotions, participants never interact with them (Hermans et al., 2019).

3) Concerning the technological limitations of immersive systems, different factors complicate its clinical implementation: i) VR is a high-cost technology that still has high technical maintenance requirements (Parsey & Schmitter-Edgecombe, 2013), both in the devices needed for their deployment and in the development of the contents. In addition, the lack of widely accepted industry standards means that developments and devices have a shorter life cycle, making VR even more expensive. ii) VR development is still a fragile and complex process that requires high technical expertise, hampering the incorporation of clinical professionals (neurologists and neuropsychologists) (Martínez-Pernía et al., 2017). iii) Immersive systems have specific requirements that limit their application in the clinical field (i.e., lighting and large-scale game devices) (Rizzo et al., 2004; Werner et al., 2009).

These issues are possible reasons why VR is still limited to an evaluation method in the experimental phase (Camara Lopez et al., 2016), where only a handful of investigations have established virtual environments' ecological and constructive validity (Besnard et al., 2016). Some authors have stated that the VR ecological validity could not be better than traditional evaluation methods (Chan et al., 2008). As Rizzo and Koening point out for VR technologies, "the majority of conducted studies are pilot trials without sufficient power or the study design needed to draw decisive conclusions about efficacy, transfer of gained skills to the daily life of clients, long-term outcomes, and cost-effectiveness." (Rizzo & Koening 2017, pp. 888). The literature confirms these aspects because most applications of immersive technologies are mainly exploratory experiments to verify the feasibility, with a limited number of subjects and conducted in controlled environments where there is extensive technical support (Kim, 2019; Park, 2019).

As an alternative to the use of immersive technologies to detect real-world functioning, we support the idea that the SG, and more specifically, the screen-based simulations, are potentially the most promising technology for neuropsychological assessment. SG are rapidly becoming an essential tool in the medical domain used with different goals such as improving health behaviors, training, and research (Wattanasoontorn et al., 2013). These real simulations are developed through a multimodal interaction (de Freitas & Liarokapis, 2011) using animations, texts, language, graphics, haptics, audios, and so on. (Arnab et al., 2011; Chicchi et al., 2018; Laamarti et al., 2014; Orozco et al., 2012). Within this realistic simulation, the player has to explore, make decisions, solve tasks, decide among alternative strategies, plan, sequence, interact socially, interpret the behaviors of others, and so on, instead of just physical challenges (Adams & Rollings, 2010) (Figure 1 shows an example of a serious game for the evaluation of executive dysfunction). SG are currently employed to support and improve the assessment of different functional and cognitive abilities and provide alternative solutions for patients' treatment, stimulation, and rehabilitation (Manera, 2016). SG have already demonstrated that they can overcome some of the limitations identified with immersive systems and non-technological VR and traditional pen-and-paper-based neuropsychological tests. Below, we will show some advantages of the SG for the detection of executive dysfunction with higher ecological validity. We will show how the use of SG allows overcoming some of the challenges of current neuropsychological assessment and consider the assessment of EF in a broad sense (Figure 2).

Figure 1

Figure 1: *In this serious game, the user has the objective of buying foods in the supermarket to cook a recipe. To do the task successfully, the user will need to select products to cook with, interact with vendors to buy the right product, tailor the purchase to available products, interact with other shoppers in uncomfortable or friendly social situations, adjust behavior based on social circumstances and social-emotional interpretation.*

Games usually provide highly interactive environments that keep the user engaged and motivated. Usually, games provide a narrative story and scenario that immerses the user in an environment where the player needs to attain a goal, respecting the game rules. All these characteristics can be used to create evaluation situations in social contexts that are more naturalistic and close to the daily living situations as they can also include social interaction with one or many game characters simultaneously. As the game has complete control, it is easier to include different stimuli related to social cognition, empathy, emotional regulation, morality, and emotional recognition in a natural way and incorporate meaningful tasks for the user. Therefore, SG allows for the inclusion of dimensions, all of which directly affect the modulation and processing of the EF. The screen-based simulations allow for accurate reproduction of tasks and procedures in real environments (Martínez-Pernía et al., 2017), generating an atmosphere as if the person is in the intersubjective social world (Lamberts et al., 2010; Poulin et al., 2013), and not just embedded in an environment (in a kitchen, in a city).

SG allow the generation of evaluation that include different stimuli in ill-structured tasks that are not always presented in the same form or sequence but that are presented in a way that is still meaningful for the user and overcome limitations of the neuropsychological assessment methodologies based on "well-structured" tasks. Exploration is a natural behavior in games. SG give the opportunity to generate different game dynamics in which the user requires continuous adaptation and innovation in the way he/she responds to the stimuli. This element is fundamental to adequately evaluate the executive function as a process that is evaluated in novel situations that the user has not faced before (Chang et al., 2008). SG are capable of generating evaluation processes where simultaneously a narrative of social interaction is generated and where the subject must act through this social influence and his executive function skills.

In SG, it is common to track user interaction in detail, obtaining evidence of the actual user behavior, and even adapt the game's behavior to those interactions. These circumstances make games particularly suitable for neuropsychological assessment. SG can be easily administered and can feature-rich interactive environments to evaluate complex neuropsychological constructs that are difficult to evaluate through traditional tests (Valladares-Rodriguez 2016). It is possible to use game analytics techniques, all the rich in-game user interaction data to estimate the parameters of the underlying cognitive processes, and use the parameters' values to estimate the user performance (Hagler 2014). SG can include advanced game analytics systems that allow the discovery of complex multidimensional patterns. Game analytics will allow a richer analysis of the data (including machine learning) than some ICT evaluation systems that focus on simpler data such as speed of response, the accuracy of response, or the number of errors (Alonso-Fernández et al., 2019; Valladares-Rodriguez et al, 2018).

SG is a robust and proven technology since the video game industry is the most important in the entertainment field and is present in most homes in the world. There are many game development platforms ranging from professional environments (e.g., Unity3D, Unreal) to authoring environments that simplify the creation of SG (e.g., eAdventure). SG environments such as eAdventure simplify the creation of the games by clinicians with a minimal background in computer science (Martinez-Pernía et al., 2017; Pérez-Colado et al., 2017). Also, those environments allow for the deployment of SG on different devices without requiring any additional development (or with minimal ones). For instance, the game can be used on a computer or a mobile device (e.g., tablet or mobile), simplifying the administration at the point of care or in other environments (Valladares-Rodriguez, 2019).

Figure 2

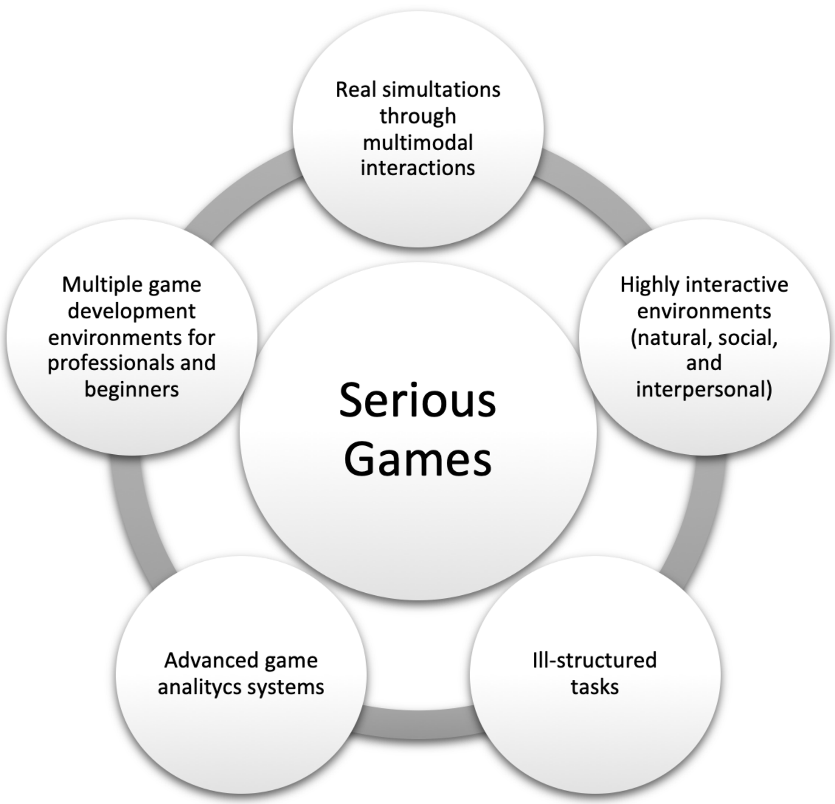

Figure 2: *SG offers the opportunity to improve the detection of executive dysfunction real-world contexts through real simulations based on multimodal interactions, highly*

interactive environments (natural, social, and interpersonal), ill-structured tasks, advanced game analytics systems, and multiple game development environments for professionals and beginners

Despite the advantages recently mentioned, SG do have their drawbacks. For instance, the participant's interaction with the non-immersive virtual environment does not directly correspond to how their sensorimotor system works in real life, requiring the participant to undergo a learning process to integrate their feeling of 'being there' (Slater, 2009; Corti et al., 2021). Another example is that SG are played on electronic devices in front of screens, potentially damaging their sight. Furthermore, it implies that some populations with poor eyesight could find it hard to interact on the screen (elderly people) (Ning et al., 2020).

However, by the arguments previously indicated, we consider that SG are a mature technology, with potentially high power as a predicting tool of executive dysfunction associated with real-world functioning. However, more research is still needed to implement SG that are reliable, validated in different environments, real-world simulations, and interactional social contexts, and thus ready to be used in everyday clinical practice.

CONCLUSION

Executive function is one of the most complex cognitive domains of human behavior, and its dysfunction has been described as one of the main factors associated with functional impairment. Nowadays, there is a broad range of methods for detecting and evaluating executive dysfunction ranging from clinical interview to neuropsychological evaluation. This diversity of methods suggests that current approaches do not account for EF's complexity. On many occasions, a dissociation is observed regarding the scores obtained in the tests -many times in normal ranges- and the severe difficulties in daily life. For instance, clinical observation of patients with good performance in traditional neuropsychological tests, but with different degrees of executive dysfunction in real-world functioning (Mesulam, 1986).

An innovative strategy is the use of information technology and communication (ICT). The development of technological methods increases ecological validity while keeping satisfactory control of experimental variables (Fan et al., 2021). In the present work, we propose that SG offer an opportunity to develop more efficient tools to detect executive dysfunction in everyday life contexts. They provide meaningful narrative stories and virtual or real environments that immerse the user in natural and social environments with social interactions, while the player needs to adapt his/her behavioral performance to the challenges of ill-structured tasks (social, cognitive, and interpersonal). Moreover, the user has to explore, make decisions, solve tasks, interact socially, interpret the behaviors of others, all of which are interspersed with each other and presented in different forms and sequences. We believe that SG are a well-balanced non-intrusive method for obtaining users' behavioral data in more naturalistic settings that can be used by clinical personnel as evidence to detect executive dysfunction in real-world performance.

Funding

DM-P was supported by ANID/Fondecyt/11190507. DM-P and AS were supported by grants from ANID/FONDAP/15150012. DM-P, LO, and AS were supported by ANID/FONDEF/ID 18I10113. AS were supported by ANID/Fondecyt/1191726, 1210176, and 1210195 and MULTI-PARTNER CONSORTIUM TO EXPAND DEMENTIA RESEARCH IN LATIN AMERICA [ReDLat, supported by National Institutes of Health, National Institutes of Aging (R01 AG057234), Alzheimer's Association (SG-20-725707), Tau Consortium, and Global Brain Health Institute], and Alzheimer's Association GBHI ALZ UK-20- 639295. AI is partially supported by grants from CONICET; ANID/FONDECYT Regular (1210195 and 1210176); FONCYT-PICT 2017-1820; ANID/FONDAP/15150012; Takeda CW2680521; Sistema General de Regalías (BPIN2018000100059), Universidad del Valle (CI 5316); Alzheimer's Association GBHI ALZ UK-20-639295; and the MULTI-PARTNER CONSORTIUM TO EXPAND DEMENTIA RESEARCH IN LATIN AMERICA [ReDLat, supported by National Institutes of Health, National Institutes of Aging (R01 AG057234), Alzheimer's Association (SG-20-725707), Rainwater Charitable foundation - Tau Consortium, and Global Brain Health Institute)]. The contents of this publication are solely the responsibility of the authors and do not represent the official views of these Institutions.

ZOTERO BIBLIOGRAPHY

Adamit, T., Maeir, A., Assayag, E. B., Bornstein, N. M., Korczyn, A. D., & Katz, N. (2015). Impact of first-ever mild stroke on participation at 3 and 6 month post-event: The TABASCO study. *Disability and Rehabilitation*, *37*(8), 667-673. https://doi.org/10.3109/09638288.2014.923523

Adjorlolo, S. (2015). Can Teleneuropsychology Help Meet the Neuropsychological Needs of Western Africans? The Case of Ghana. *Applied Neuropsychology: Adult*, *22*(5), 388-398. https://doi.org/10.1080/23279095.2014.949718

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, *3*(12), 469-479. https://doi.org/10.1016/S1364-6613(99)01399-6

Alderman, N., Burgess, P. W., Knight, C., & Henman, C. (2003). Ecological validity of a simplified version of the multiple errands shopping test. *Journal of the International Neuropsychological Society*, *9*(1), 31-44. https://doi.org/10.1017/S1355617703910046

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268-277. https://doi.org/10.1038/nrn1884

Arciniegas, D. B. (2013). Medical evaluation. *Management of Adults with Traumatic Brain Injury*, 25–73.

Assuras, S., & Levin, B. (2013). Special Considerations for the Neuropsychological Interview with Older Adults. En L. D. Ravdin & H. L. Katzen (Eds.), *Handbook on the Neuropsychology of Aging and Dementia* (pp. 3-9). Springer New York.

https://doi.org/10.1007/978-1-4614-3106-0_1

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different Contributions of the Human Amygdala and Ventromedial Prefrontal Cortex to Decision-Making. *Journal of Neuroscience*, *19*(13), 5473-5481. https://doi.org/10.1523/JNEUROSCI.19-13-05473.1999

Best, J. R., & Miller, P. H. (2010). A Developmental Perspective on Executive Function. *Child Development*, *81*(6), 1641-1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., & Rowland, J. A. (2017). Neuropsychological Test Administration by Videoconference: A Systematic Review and Meta-Analysis. *Neuropsychology Review*, *27*(2), 174-186. https://doi.org/10.1007/s11065-017-9349-1

Brown, L., Fish, J., Mograbi, D. C., Bellesi, G., Ashkan, K., & Morris, R. (2019). Awareness of deficit following traumatic brain injury: A systematic review of current methods of assessment. *Neuropsychological rehabilitation*, 1–35.

Burgess, P. W. (2003). Assessment of executive function. *Handbook of clinical neuropsychology*, 302–321.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M., Dawson, D. R., Anderson, N. D., Gilbert, S., Dumontheil, I., & Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, *12*, 194-209.

Burgess, P. W., Alderman, N., Volle, E., Benoit, R. G., & Gilbert, S. J. (2009). Mesulam's frontal lobe mystery re-examined. *Restorative Neurology and Neuroscience*, *27*(5), 493-506. https://doi.org/10.3233/RNN-2009-0511

Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, *38*(6), 848-863. https://doi.org/10.1016/S0028-3932(99)00134-7

Chan, R. C. K., Shum, D., Toulopoulou, T., & Chen, E. Y. H. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*(2), 201-216. https://doi.org/10.1016/j.acn.2007.08.010

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The Ecological Validity of Neuropsychological Tests: A Review of the Literature on Everyday Cognitive Skills. *Neuropsychology Review*, *13*(4), 181-197. https://doi.org/10.1023/B:NERV.0000009483.91468.fb

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology*, *21*(3), 217-227. https://doi.org/10.1016/j.acn.2005.12.002

Chevignard, M., Pillon, B., Pradat-Diehl, P., Taillefer, C., Rousseau, S., Le Bras, C., & Dubois, B. (2000). An Ecological Approach to Planning Dysfunction: Script Execution. *Cortex*, *36*(5), 649-669. https://doi.org/10.1016/S0010-9452(08)70543-4

Choca, J., & Morris, J. (1992). Administering the category test by computer: Equivalence of results. *Clinical Neuropsychologist*, *6*(1), 9-15. https://doi.org/10.1080/13854049208404112

Cieza, A., Anczewska, M., Ayuso-Mateos, J. L., Baker, M., Bickenbach, J., Chatterji, S., Hartley, S., Leonardi, M., Pitkänen, T., & PARADISE Consortium. (2015). Understanding the Impact of Brain Disorders: Towards a «Horizontal

Epidemiology» of Psychosocial Difficulties and Their Determinants. *PloS One*, *10*(9), e0136271. https://doi.org/10.1371/journal.pone.0136271

Clement, P. F., Brooks, F. R., Dean, B., & Galaz, A. (2001). A neuropsychology telemedicine clinic. *Military medicine*, *166*(5), 382–384.

Coenen, M., Cabello, M., Umlauf, S., Ayuso-Mateos, J. L., Anczewska, M., Tourunen, J., Leonardi, M., Cieza, A., & Consortium, P. (2016). Psychosocial difficulties from the perspective of persons with neuropsychiatric disorders. *Disability and Rehabilitation*, *38*(12), 1134-1145. https://doi.org/10.3109/09638288.2015.1074729

Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., McKeith, I., Saxby, B. K., & Kirkwood, T. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ pilot study. *Journal of the American Geriatrics Society*, *55*(10), 1630-1635. https://doi.org/10.1111/j.1532-5415.2007.01379.x

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain: Amazon.es: Antonio R. Damasio: Libros en idiomas extranjeros*. GP Putnam.

Damasio, A. R. (1995). Toward a Neurobiology of Emotion and Feeling: Operational Concepts and Hypotheses. *The Neuroscientist*, *1*(1), 19-25. https://doi.org/10.1177/107385849500100104

Edwards, J., Vess, J., Reger, G., & Cernich, A. (2014). The Use of Virtual Reality in the Military's Assessment of Service Members With Traumatic Brain Injury: Recent Developments and Emerging Opportunities. *Applied Neuropsychology: Adult*, *21*(3), 220-230. https://doi.org/10.1080/09084282.2013.796554

Engel, E., Peipert, A., & Ehrlich-Jones, L. (2018). Measurement Characteristics and Clinical Utility of the Awareness Questionnaire in Individuals With Traumatic

Brain Injury. *Archives of Physical Medicine and Rehabilitation*, *99*(4), 813–814.

Fish, J., Wilson, B. a, & Manly, T. (2010). The assessment and rehabilitation of prospective memory problems in people with neurological disorders: A review. *Neuropsychological rehabilitation*, *20*(2), 161-179. https://doi.org/10.1080/09602010903126029

FitzGerald, M. C. C., O'Keeffe, F., Carton, S., Coen, R. F., Kelly, S., & Dockree, P. (2019). Rehabilitation of emergent awareness of errors post traumatic brain injury: A pilot intervention. *Neuropsychological Rehabilitation*, *29*(6), 821-843. https://doi.org/10.1080/09602011.2017.1336102

Fortin, S., Godbout, L., & Braun, C. M. J. (2003). Cognitive Structure of Executive Deficits in Frontally Lesioned Head Trauma Patients Performing Activities of Daily Living. *Cortex*, *39*(2), 273-291. https://doi.org/10.1016/S0010-9452(08)70109-6

Franzen, M. D. (2000). *Reliability and Validity in Neuropsychological Assessment* (2.ª ed.). Springer US. https://doi.org/10.1007/978-1-4757-3224-5

Franzen, M. D., & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neuropsychological assessment. En *Ecological validity of neuropsychological testing* (pp. 91-112). Gr Press/St Lucie Press, Inc.

Garrett, S. L., Kennedy, R. E., Sawyer, P., Williams, C. P., Brown, C. J., & Allman, R. M. (2019). Association Between Executive Dysfunction and Instrumental Activities of Daily Living: Racial and Ethnic Differences Among Community-Dwelling Older Adults in the Southeastern US. *Journal of the National Medical Association*, *111*(3), 320-327. https://doi.org/10.1016/j.jnma.2018.11.002

Geytenbeek, M., Fleming, J., Doig, E., & Ownsworth, T. (2017). The occurrence of early impaired self-awareness after traumatic brain injury and its relationship with emotional distress and psychosocial functioning. *Brain Injury*, *31*(13-14), 1791-1798. https://doi.org/10.1080/02699052.2017.1346297

Gilbert, S. J., Zamenopoulos, T., Alexiou, K., & Johnson, J. H. (2010). Involvement of right dorsolateral prefrontal cortex in ill-structured design cognition: An fMRI study. *Brain research*, *1312*, 79-88. https://doi.org/10.1016/j.brainres.2009.11.045

Gioia, G. A., & Isquith, P. K. (2004). Ecological Assessment of Executive Function in Traumatic Brain Injury. *Developmental Neuropsychology*, *25*(1-2), 135-158. https://doi.org/10.1080/87565641.2004.9651925

Giovannetti, T., Libon, D. J., Buxbaum, L. J., & Schwartz, M. F. (2002). Naturalistic action impairments in dementia. *Neuropsychologia*, *40*(8), 1220–1232.

Goel, V., Grafman, J., Tajik, J., Gana, S., & Danto, D. (1997). A study of the performance of patients with frontal lobe lesions in a financial planning task. *Brain*, *120*(10), 1805-1822. https://doi.org/10.1093/brain/120.10.1805

Goel, Vinod, & Grafman, J. (2000). ROLE OF THE RIGHT PREFRONTAL CORTEX IN ILL-STRUCTURED PLANNING. *Cognitive Neuropsychology*, *17*(5), 415-436. https://doi.org/10.1080/026432900410775

Golden, C., Purisch, A. D., & Hammeke, T. A. (1985). *Luria-Nebraska Neuropsychological Battery: Forms I & II*.

Goldin, P. R., McRae, K., Ramel, W., & Gross, J. J. (2008). The Neural Bases of Emotion Regulation: Reappraisal and Suppression of Negative Emotion. *Biological Psychiatry*, *63*(6), 577-586. https://doi.org/10.1016/j.biopsych.2007.05.031

Gonen-Yaacovi, G., de Souza, L. C., Levy, R., Urbanski, M., Josse, G., & Volle, E. (2013). Rostral and caudal prefrontal contribution to creativity: A meta-analysis of functional imaging data. *Frontiers in Human Neuroscience*, *7*.

https://doi.org/10.3389/fnhum.2013.00465

Grafman, J., Schwab, K., Warden, D., Pridgen, A., Brown, H. R., & Salazar, A. M. (1996). Frontal lobe injuries, violence, and aggression: A report of the Vietnam Head Injury Study. *Neurology*, *46*, 1231-1238.

Grafman, Jordan, & Litvan, I. (1999). Importance of deficits in executive functions. *The Lancet*, *354*(9194), 1921-1923. https://doi.org/10.1016/S0140-6736(99)90438-5

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404-411. https://doi.org/10.1037/h0059831

Green, M. F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? *The American Journal of Psychiatry*, *153*(3), 321-330. https://doi.org/10.1176/ajp.153.3.321

Green, M. F., Kern, R. S., Braff, D. L., & Mintz, J. (2000). Neurocognitive Deficits and Functional Outcome in Schizophrenia: Are We Measuring the “Right Stuff”? *Schizophrenia Bulletin*, *26*(1), 119-136. https://doi.org/10.1093/oxfordjournals.schbul.a033430

Gulde, P., Schmidle, S., Aumüller, A., & Hermsdörfer, J. (2019). The effects of speed of execution on upper-limb kinematics in activities of daily living with respect to age. *Experimental Brain Research*, *237*(6), 1383-1395. https://doi.org/10.1007/s00221-019-05507-0

Hari, R., & Kujala, M. V. (2009). Brain Basis of Human Social Interaction: From Concepts to Brain Imaging. *Physiological Reviews*, *89*(2), 453-479. https://doi.org/10.1152/physrev.00041.2007

Henríquez, F., Martínez-Pernía, D., Forno, G., Nuñez-Huasaf, J., & Slachevsky, A. (2019).

Anatomía funcional del córtex prefrontal y modelos de función ejecutiva. En *Tratado de neuropsicología clínica: Bases conceptuales y técnicas de evaluación* (Labos, E., Slachevsky, A. Torralva, T, Fuentes, P, Manes, Vol. 1). Akadia.

Hosenbocus, S., & Chahal, R. (2012). A Review of Executive Function Deficits and Pharmacological Management in Children and Adolescents. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *21*(3), 223-229.

Hoskins, L. L., Binder, L. M., Chaytor, N. S., Williamson, D. J., & Drane, D. L. (2010). Comparison of Oral and Computerized Versions of the Word Memory Test. *Archives of Clinical Neuropsychology*, *25*(7), 591-600. https://doi.org/10.1093/arclin/acq060

Ibañez, A., & Manes, F. (2012). Contextual social cognition and the behavioral variant of frontotemporal dementia. *Neurology*, *78*(17), 1354-1362. https://doi.org/10.1212/WNL.0b013e3182518375

Jacobsen, S., Sprenger, T., Andersson, S., & Krogstad, J.-M. (2003). Neuropsychological assessment and telemedicine: A preliminary study examining the reliability of neuropsychology services performed via telecommunication. *Journal of the International Neuropsychological Society*, *9*(3), 472–478.

Kessels, R. P. C. (2019). Improving precision in neuropsychological assessment: Bridging the gap between classic paper-and-pencil tests and paradigms from cognitive neuroscience. *The Clinical Neuropsychologist*, *33*(2), 357-368. https://doi.org/10.1080/13854046.2018.1518489

Klinger, E., Chemin, I., Lebreton, S., & Marié, R. M. (2004). A virtual supermarket to assess cognitive planning,. *CyberPsychology & Behavior*, *7*(3), 292-293.

Kourtesis, P., Korre, D., Collina, S., Doumas, L. A. A., & MacPherson, S. E. (2020).

Guidelines for the Development of Immersive Virtual Reality Software for Cognitive Neuroscience and Neuropsychology: The Development of Virtual Reality Everyday Assessment Lab (VR-EAL), a Neuropsychological Test Battery in Immersive Virtual Reality. *Frontiers in Computer Science*, *1*. https://doi.org/10.3389/fcomp.2019.00012

Legouverneur, G., Pino, M., Boulay, M., & Rigaud, A.-S. (2011). Wii sports, a usability study with MCI and Alzheimer's patients. *Alzheimer's & Dementia*, *7*(4S_Part_14), S500-S501. https://doi.org/10.1016/j.jalz.2011.05.2398

Lezak, M. D. (1982). The Problem of Assessing Executive Functions. *International Journal of Psychology*, *17*(1-4), 281-297. https://doi.org/10.1080/00207598208247445

Lo Priore, C., Castelnuovo, G., Liccione, D., & Liccione, D. (2003). Experience with V-STORE: Considerations on Presence in Virtual Environments for Effective Neuropsychological Rehabilitation of Executive Functions. *CyberPsychology & Behavior*, *6*(3), 281-287. https://doi.org/10.1089/109493103322011579

Luria, A. R., Pribram, K. H., & Homskaya, E. D. (1964). An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelioma (meningioma). *Neuropsychologia*, *2*(4), 257-280. https://doi.org/10.1016/0028-3932(64)90034-X

Mahoney, D., Gutman, S. A., & Gillen, G. (2019). A Scoping Review of Self-Awareness Instruments for Acquired Brain Injury. *The Open Journal of Occupational Therapy*, *7*(2), 3.

Manchester, D., Priestley, N., & Jackson, H. (2004). The assessment of executive functions: Coming out of the office. *Brain Injury*, *18*(11).

https://www.tandfonline.com/doi/abs/10.1080/02699050410001672387

Manera, V., Petit, P.-D., Derreumaux, A., Orvieto, I., Romagnoli, M., Lyttle, G., David, R., & Robert, P. H. (2015). 'Kitchen and cooking,' a serious game for mild cognitive impairment and Alzheimer's disease: A pilot study. *Frontiers in Aging Neuroscience*, *7*. https://doi.org/10.3389/fnagi.2015.00024

Marcotte, T. D., & Grant, I. (2009). *Neuropsychology of Everyday Functioning*. Guilford Press.

Martínez-Pernía, D., González-Castán, Ó., & Huepe, D. (2017). From ancient Greece to the cognitive revolution: A comprehensive view of physical rehabilitation sciences. *Physiotherapy Theory and Practice*, *33*(2), 89-102. https://doi.org/10.1080/09593985.2016.1266720

Martínez-Pernía, D., Núñez-Huasaf, J., del Blanco, Á., Ruiz-Tagle, A., Velásquez, J., Gomez, M., Robert Blesius, C., Ibañez, A., Fernández-Manjón, B., & Slachevsky, A. (2017). Using game authoring platforms to develop screen-based simulated functional assessments in persons with executive dysfunction following traumatic brain injury. *Journal of Biomedical Informatics*, *74*, 71-84. https://doi.org/10.1016/j.jbi.2017.08.012

Mataix-Cols, D., & Bartrés-Faz, D. (2002). Is the Use of the Wooden and Computerized Versions of the Tower of Hanoi Puzzle Equivalent? *Applied Neuropsychology*, *9*(2), 117-120. https://doi.org/10.1207/S15324826AN0902_8

McDonald, B. C., Flashman, L. A., & Saykin, A. J. (2002). Executive dysfunction following traumatic brain injury: Neural substrates and treatment strategies. *NeuroRehabilitation*, *17*(4), 333-344. https://doi.org/10.3233/NRE-2002-17407

McGeorge, P., Phillips, L. H., Crawford, J. R., Garden, S. E., Sala, S. D., Milne, A. B., Hamilton, S. W., & Callender, J. S. (2001). Using Virtual Environments in the Assessment of Executive Dysfunction. *Presence: Teleoperators and Virtual Environments*, *10*(4), 375-383. https://doi.org/10.1162/1054746011470235

Mesulam, M. (1986). Frontal cortex and behavior—Mesulam. *Annals of neurology*, *19*(4), 320-325.

Michael, D. R., & Chen, S. L. (2005). *Serious Games: Games That Educate, Train, and Inform*. Muska & Lipman/Premier-Trade.

Osterrieth, P. A. (1944). Le test de copie d'une figure complexe; contribution a l'etude de la perception et de la memoire. *Archives de psychologie*.

Ownsworth, T., Fleming, J., Doig, E., Shum, D. H., & Swan, S. (2019). Concordance between the awareness questionnaire and self-awareness of deficits interview for identifying impaired self-awareness in individuals with traumatic brain injury in the community. *Journal of rehabilitation medicine*, *51*(5), 58–61.

Padala, K. P., Padala, P. R., Malloy, T. R., Geske, J. A., Dubbert, P. M., Dennis, R. A., Garner, K. K., Bopp, M. M., Burke, W. J., & Sullivan, D. H. (2012). Wii-Fit for Improving Gait and Balance in an Assisted Living Facility: A Pilot Study. *Journal of Aging Research*. https://doi.org/10.1155/2012/597573

Parsey, C. M., & Schmitter-Edgecombe, M. (2013). Applications of Technology in Neuropsychological Assessment. *The Clinical Neuropsychologist*, *27*(8), 1328-1361. https://doi.org/10.1080/13854046.2013.834971

Parsons, T. D. (2011). Neuropsychological Assessment Using Virtual Environments: Enhanced Assessment Technology for Improved Ecological Validity. En S. Brahnam & L. C. Jain (Eds.), *Advanced Computational Intelligence Paradigms in Healthcare 6. Virtual Reality in Psychotherapy, Rehabilitation, and Assessment* (pp.

271-289). Springer. https://doi.org/10.1007/978-3-642-17824-5_13

Parsons, T. D. (2014). Virtual Teacher and Classroom for Assessment of Neurodevelopmental Disorders. En A. L. Brooks, S. Brahnam, & L. C. Jain (Eds.), *Technologies of Inclusive Well-Being: Serious Games, Alternative Realities, and Play Therapy* (pp. 121-137). Springer. https://doi.org/10.1007/978-3-642-45432-5_7

Parsons, T. D. (2016). Neuropsychological Assessment 3.0. En T. D. Parsons (Ed.), *Clinical Neuropsychology and Technology: What's New and How We Can Use It* (pp. 65-96). Springer International Publishing. https://doi.org/10.1007/978-3-319-31075-6_5

Phay, A., Gainer, C., & Goldstein, G. (1986). Clinical Interviewing of the Patient and History in Neuropsychological Assessment. En T. Incagnoli, G. Goldstein, & C. J. Golden (Eds.), *Clinical Application of Neuropsychological Test Batteries* (pp. 45-73). Springer US. https://doi.org/10.1007/978-1-4684-4991-4_2

Prigatano, G. P., & Schacter, D. L. (1991). *Awareness of deficit after brain injury: Clinical and theoretical issues*. Oxford University Press.

Rand, D., Katz, N., Shahar, M., Kizoni, R., & Weiss, P. L. (2005). The virtual mall: A functional virtual environment for stroke rehabilitation. *Annual Review of CyberTherapy and Telemedicine*, *3*, 193-198.

Rand, Debbie, Rukan, S. B.-A., Weiss, P. L. (Tamar), & Katz, N. (2009). Validation of the Virtual MET as an assessment tool for executive functions. *Neuropsychological Rehabilitation*, *19*(4), 583-602. https://doi.org/10.1080/09602010802469074

Reitan, R. M. (1958). Validity of the Trail Making Test as an Indicator of Organic Brain Damage. *Perceptual and Motor Skills*, *8*(3), 271-276.

https://doi.org/10.2466/pms.1958.8.3.271

Riva, G. (2005). Virtual Reality in Psychotherapy: Review. *CyberPsychology & Behavior*, *8*(3), 220-230. https://doi.org/10.1089/cpb.2005.8.220

Rizzo, A. A., Schultheis, M., Kerns, K. A., & Mateer, C. (2004). Analysis of assets for virtual reality applications in neuropsychology. *Neuropsychological Rehabilitation*, *14*(1-2), 207-239. https://doi.org/10.1080/09602010343000183

Rizzo, A. “Skip”, & Kim, G. J. (2005). A SWOT Analysis of the Field of Virtual Reality Rehabilitation and Therapy. *Presence: Teleoperators and Virtual Environments*, *14*(2), 119-146. https://doi.org/10.1162/1054746053967094

Robert, P., König, A., Amieva, H., Andrieu, S., Bremond, F., Bullock, R., Ceccaldi, M., Dubois, B., Gauthier, S., Kenigsberg, P.-A., Nave, S., Orgogozo, J. M., Piano, J., Benoit, M., Touchon, J., Vellas, B., Yesavage, J., & Manera, V. (2014). Recommendations for the use of Serious Games in people with Alzheimer’s Disease, related disorders and frailty. *Frontiers in Aging Neuroscience*, *6*. https://doi.org/10.3389/fnagi.2014.00054

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (2001). *The test of everyday attention (TEA)*. *4*(1), 6.

Romero-Ayuso, D., Castillero-Perea, Á., González, P., Navarro, E., Molina-Massó, J. P., Funes, M. J., Ariza-Vega, P., Toledano-González, A., & Triviño-Juárez, J. M. (2019). Assessment of cognitive instrumental activities of daily living: A systematic review. *Disability and Rehabilitation*, *0*(0), 1-17. https://doi.org/10.1080/09638288.2019.1665720

Royall, D. R., Mahurin, R. K., & Gray, K. F. (1992). Bedside Assessment of Executive Cognitive Impairment: The Executive Interview. *Journal of the American*

Geriatrics Society, *40*(12), 1221-1226. https://doi.org/10.1111/j.1532-5415.1992.tb03646.x

Sabariego, C., Coenen, M., Ballert, C., Cabello, M., Leonardi, M., Anczewska, M., Pitkänen, T., Raggi, A., Mellor, B., Covelli, V., Świtaj, P., Levola, J., Schiavolin, S., Chrostek, A., Bickenbach, J., Chatterji, S., & Cieza, A. (2015). Determinants of Psychosocial Difficulties Experienced by Persons with Brain Disorders: Towards a 'Horizontal Epidemiology' Approach. *PLOS ONE*, *10*(12), e0141322. https://doi.org/10.1371/journal.pone.0141322

Sahakian, B. J., & Owen, A. M. (1992). Computerized assessment in neuropsychiatry using CANTAB: Discussion paper. *Journal of the Royal Society of Medicine*, *85*(7), 399-402.

Sbordone, R. J. (1996). Ecological validity: Some critical issues for the neuropsychologist. En *Ecological validity of neuropsychological testing* (pp. 15-41). Gr Press/St Lucie Press, Inc.

Sbordone, R. J. (2010). Neuropsychological Tests are Poor at Assessing the Frontal Lobes, Executive Functions, and Neurobehavioral Symptoms of Traumatically Brain-Injured Patients. *Psychological Injury and Law*, *3*(1), 25-35. https://doi.org/10.1007/s12207-010-9068-x

Sbordone, R. J. (2014). The Hazards of Strict Reliance on Neuropsychological Tests: Applied Neuropsychology: Adult: Vol 21, No 2. *Applied Neuropsychology Adult*, *21*(2), 98-107.

Sbordone, R. J., Seyranian, G., & Ruff, Ronald. M. (1998). Are the subjective complaints of traumatically brain injured patients reliable? *Brain Injury*, *12*(6), 505-515. https://doi.org/10.1080/026990598122467

Schatz, P., & Browndyke, J. (2002). Applications of Computer-based Neuropsychological Assessment. *The Journal of Head Trauma Rehabilitation*, *17*(5), 395–410.

Schmitter-Edgecombe, M., McAlister, C., & Weakley, A. (2012). Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: The day-out task. *Neuropsychology*, *26*(5), 631.

Schwartz, M. F., Reed, E. S., Montgomery, M., Plamer, C., & Mayer, N. H. (1991). The Quantitative Description of Action Disorganisation after Brain Damage: A Case Study. *Cognitive Neuropsychology*, *8*(5), 381-414. https://doi.org/10.1080/02643299108253379

Shallice, T. I. M., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*(2), 727–741.

Sharma, B., Bradbury, C., Mikulis, D., & Green, R. (2014). Missed Diagnosis of Traumatic Brain Injury in Patients with Traumatic Spinal Cord Injury. *Journal Rehabilitation Medicine*, *46*(4), 370-373. https://doi.org/info:doi/10.2340/16501977-1261

Sherer, M., Bergloff, P., Levin, E., High Jr, W. M., Oden, K. E., & Nick, T. G. (1998). Impaired awareness and employment outcome after traumatic brain injury. *The Journal of head trauma rehabilitation*.

Spooner, D. M., & Pachana, N. A. (2006). Ecological validity in neuropsychological assessment: A case for greater consideration in research with neurologically intact populations. *Archives of Clinical Neuropsychology*, *21*(4), 327-337. https://doi.org/10.1016/j.acn.2006.04.004

Stavros, Z., Fotini, K., & Magda, T. (2010). Computer based cognitive training for patients with mild cognitive impairment (MCI). *Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments*, 1–3.

https://doi.org/10.1145/1839294.1839319

Toglia, J., Askin, G., Gerber, L. M., Taub, M. C., Mastrogiovanni, A. R., & O'Dell, M. W. (2017). Association Between 2 Measures of Cognitive Instrumental Activities of Daily Living and Their Relation to the Montreal Cognitive Assessment in Persons With Stroke. *Archives of Physical Medicine and Rehabilitation*, *98*(11), 2280-2287. https://doi.org/10.1016/j.apmr.2017.04.007

Toglia, J., & Kirk, U. (2000). Understanding awareness deficits following brain injury. *NeuroRehabilitation*, *15*(1), 57–70.

Tupper, D. E., & Cicerone, K. D. (Eds.). (1990). *The Neuropsychology of Everyday Life: Assessment and Basic Competencies*. Springer US. https://doi.org/10.1007/978-1-4613-1503-2

Valladares-Rodríguez, S., Pérez-Rodríguez, R., Anido-Rifón, L., & Fernández-Iglesias, M. (2016). Trends on the application of serious games to neuropsychological evaluation: A scoping review. *Journal of biomedical informatics*, *64*, 296–319.

Vallejo, V., Wyss, P., Chesham, A., Mitache, A. V., Müri, R. M., Mosimann, U. P., & Nef, T. (2017). Evaluation of a new serious game based multitasking assessment tool for cognition and activities of daily living: Comparison with a real cooking task. *Computers in Human Behavior*, *70*, 500-506. https://doi.org/10.1016/j.chb.2017.01.021

Weybright, E. H., Dattilo, J., & Rusch, F. R. (2010). Effects of an Interactive Video Game (Nintendo Wii^{TM}) on Older Women with Mild Cognitive Impairment. *Therapeutic Recreational Journal*, *XLIV*(4), 271-287.

Williams, J. E., & McCord, D. M. (2006). Equivalence of standard and computerized versions of the Raven Progressive Matrices Test. *Computers in Human Behavior*, *22*(5), 791-800. https://doi.org/10.1016/j.chb.2004.03.005

Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. (1996). *Behavioural assessment of the dysexecutive syndrome.* Bury St Edmunds.

Wilson, B., Cockburn, J., & Baddeley, A. (1985). *The Rivermead behavioural memory test (RMBT)*. Thames Valley Test Company.

Zhang, L., Abreu, B. C., Masel, B., Scheibel, R. S., Christiansen, C. H., Huddleston, N., & Ottenbacher, K. J. (2001). Virtual Reality in the Assessment of Selected Cognitive Function After Brain Injury. *American Journal of Physical Medicine & Rehabilitation*, *80*, 597-604. https://doi.org/10.1097/00002060-200108000-00010