



Efficacy of adaptive cognitive training through desktop virtual reality and paper-and-pencil in the treatment of mental and behavioral disorders

Joana Câmara^{1,2,3,5} · Rute Ferreira¹ · Liliana Teixeira¹ · Joana Nóbrega¹ · Carina Romeira¹ · Sergi Bermúdez i Badia^{3,4,5} · Ana Lúcia Faria^{3,5}

Received: 31 May 2020 / Accepted: 9 July 2021

© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

Cognitive deficits are a core feature of mental and behavioral disorders, leading to poor treatment adherence and functionality. Virtual reality (VR) methodologies are promising solutions for cognitive interventions in psychiatry once they provide greater ecological validity. This study assessed and compared two content-equivalent cognitive training (CT) interventions, delivered in desktop VR (Reh@City v2.0) and paper-and-pencil (Task Generator (TG)) formats, in patients with mental and behavioral disorders. 30 patients were randomly assigned to the Reh@City v2.0 group and the TG group. Both groups of patients underwent a time-matched 24-sessions intervention. Neuropsychological assessments were performed at baseline, post-intervention, and follow-up. A within-groups analysis revealed significant improvements in visual memory and depressive symptomatology after the Reh@City intervention. The TG group improved in processing speed, verbal memory, and quality of life (social relationships and environmental domains). Between groups, Reh@City led to a greater reduction in depressive symptomatology, whereas the TG group showed higher improvements in social relationships aspects of quality of life. At follow-up, previous gains were maintained and new improvements found in the Reh@City (global cognitive function, language, visuospatial and executive functions) and the TG groups (attention). The Reh@City significantly reduced depressive symptomatology, and the TG led to greater improvements in processing speed, abstraction, and social relationships domain of quality of life at follow-up. Both interventions were associated with important cognitive, emotional, and quality of life benefits, which were maintained after two months. Reh@City and TG should be considered as complementary CT methods for patients with mental and behavioral disorders.

Trial registration The trial is registered at ClinicalTrials.gov, number NCT04291586.

Keywords Cognitive training · Desktop virtual reality · Ecological validity · Psychiatric setting · Mental and behavioral disorders

✉ Joana Câmara
joana.fcâmara@gmail.com

¹ Casa de Saúde Câmara Pestana, Rua do Lazareto, 9060-378 Funchal, Madeira, Portugal

² Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra, Rua do Colégio Novo, 3000-115 Coimbra, Portugal

³ NOVA Laboratory for Computer Science and Informatics, Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal

⁴ Centro de Ciências Exatas e da Engenharia, Universidade da Madeira, Caminho da Penteada, 9020-105 Funchal, Portugal

⁵ Madeira Interactive Technologies Institute, Universidade da Madeira, Caminho da Penteada, 9020-105 Funchal, Madeira, Portugal

1 Introduction

Around 792 million people worldwide live with a mental disorder (Ritchie and Roser 2018). In Europe, it is estimated that each year 164.8 million people are affected by mental disorders, with anxiety and major depression being two of the most common conditions (Wittchen et al. 2011). In 2010 the American Psychiatric Association estimated that neuropsychiatric disorders, which encompass mental and behavioral disorders and neurological disorders, are the third leading cause of disability worldwide, corresponding to over 10% of the global burden (American Psychiatric Association 2010). In a large-scale community epidemiological survey ($n = 73,441$) conducted in 15 countries, most respondents attributed higher levels of disability to mental disorders than to physical disorders (Ormel et al. 2008).

It is well established that cognitive impairment is a core clinical feature of most psychiatric conditions (Rock et al. 2014; Robinson et al. 2006; McIntyre et al. 2013). Although specific symptoms of psychiatric disorders can be mitigated by current pharmacological treatments (e.g., depression, anxiety, delusions), cognitive deficits tend to be persistent, occurring in both acute and remitted stages of the disorder (Millan et al. 2012). Additionally, mood symptoms (e.g., depression and anxiety) have the potential to exacerbate cognitive deficits (Iosifescu 2012). According to Reichnberg et al. (2009), all patients from 4 different diagnostic groups, namely schizophrenia ($n = 94$), schizoaffective disorder ($n = 15$), bipolar disorder ($n = 78$) and major depression ($n = 48$), exhibited multi-domain cognitive impairment (e.g., processing speed, attention, memory and executive functions). A more in-depth analysis revealed that 67% and 53% of people with schizoaffective disorder and schizophrenia, respectively, were classified as neuropsychologically impaired, whereas the degree of impairment was consistently lower among people with major depressive disorder (35%) and bipolar disorder (23%). Previous studies indicate that cognitive deficits are a robust predictor of treatment response and functional ability (e.g., activities of daily living (ADLs)) in various psychiatric conditions, and thus an essential target for intervention (Groves et al. 2018; Harvey 2011; McCleery and Nuechterlein 2019; Green 2006).

Cognitive training (CT) interventions, including both computerized and paper-and-pencil programs, have proven to be an effective method to improve cognitive, emotional, and functional outcomes in patients with psychiatric conditions (Grynszpan et al. 2011; Kim et al. 2018). Considering particularly computerized CT, a meta-analysis revealed small to large effects for computerized CT on cognitive function (e.g., attention, working memory, and

global cognitive function), depressive symptomatology, and daily functioning in patients with major depression (Motter et al. 2016). In another meta-analysis conducted among patients with schizophrenia-spectrum disorders, results indicated small to moderate effects of computerized CT in processing speed, memory, verbal fluency, verbal and visual learning. However, almost no effect on social cognition and functional outcome measures was found (Prikken et al. 2019). Unfortunately, inconsistent results in the field of CT are frequent and typically due to methodological issues, such as small sample sizes and diversity of CT programs in terms of target domains, duration, frequency, structure, types of exercises, and outcome measures.

In the last few years, virtual reality (VR) has emerged as a valuable approach in the context of neurological and psychiatric disorders. In a recent literature overview about the application of VR in the treatment of psychiatric disorders, Park et al. (2019) conclude that VR can be used with different therapeutic objectives, such as exposure therapy (traumatic and anxiety-inducing situations), social skills training and medication managing skills, in a wide range of disorders, namely phobia, posttraumatic stress disorder (PTSD), and schizophrenia. Another systematic review conducted by Cieřlik et al. (2020) indicates that VR approaches are also useful to the assessment and treatment of bulimia and binge eating disorders, neurodevelopmental disorders, and neurocognitive disorders. Focusing specifically on the application of VR-based approaches for CT purposes, few studies assess the impact of these kinds of interventions in cognitive and non-cognitive domains in patients with mental and behavioral disorders. VR-based interventions have innumerable advantages over traditional paper-and-pencil methods, namely: the systematic and hierarchical presentation of stimuli and challenges; adaptation and personalization of training content to the patient's cognitive profile; immediate feedback; gaming elements to enhance motivation and engagement; increased ecological validity and, possibly, greater transfer of gains to everyday life (Gamito et al. 2015; Parsons 2016). Previous systematic reviews highlight VR's positive impact on cognitive and non-cognitive outcomes in patients with neurological disorders (Coyle et al. 2015; Maggio et al. 2019). In addition, prior studies emphasize the beneficial effect of combining VR-based CT with more conventional CT methods on cognitive function of stroke and brain tumor patients (Kim et al. 2011; Yang et al. 2014). There are many VR environments (e.g., cities, malls, kitchens, supermarkets) devised for CT, which comprise simulations of ADLs that allow the training of multiple cognitive domains simultaneously with the aim of promoting, not only cognitive function but also functional abilities (Faria et al. 2016; Oliveira et al. 2020; Rand et al. 2005). However, most of

these VR-based CT interventions are administered in the context of neurological conditions.

For instance, a study from Zajac-Lamparska et al. (2019) evaluated the effectiveness of the VR-based CT using the GRADYS game in healthy older adults and older adults with mild dementia. Findings indicated that the GRADYS game was more effective in healthy older adults, even though both groups showed progress in the course of training, with mild dementia patients exhibiting less progress. Maier et al. (2020) conducted a study with chronic community-dwelling stroke patients to assess a VR program that provided multi-dimensional CT. Results revealed positive influences of the VR program in attention, spatial awareness, and depressive symptomatology. Gamito et al. (2015) assessed the effectiveness of a VR-based CT application that encompassed several daily activities on stroke patients in comparison with a waiting list control group. Results showed significant improvements in attention and memory in the VR-based CT intervention. Another example of a virtual environment is the Reh@City, which consists of a virtual city with several ADLs simulations clinically validated for stroke CT (Faria et al. 2016). Some Reh@City studies had the following aims: (1) to evaluate the efficacy of the Reh@City v1.0 compared to conventional rehabilitation (CT exercises administered by an occupational therapist) in several outcome measures; (2) to compare patients' performance in the Reh@City and in its paper-and-pencil version (Task Generator (TG)); and (3) to evaluate the short-term and long-term effectiveness of the Reh@City v2.0 compared to the TG, in various primary and secondary domains. In the first study, Reh@City v1.0 proved to be more effective than conventional rehabilitation, in overall cognitive functioning, processing speed, attention, memory, visuospatial abilities, executive functions, and self-reported general health status (e.g., strength and mobility, memory, emotion, social participation). Regarding the second study, Faria et al. (2019) compared stroke patients' performance in the Reh@City and the Task Generator (TG)—two content-equivalent CT tools based on the same difficulty adaptation progression framework.¹ The authors concluded that there was no effect of training methodology in overall patients' performance, which means that both groups of patients performed at the same level irrespective of the CT method employed. Finally, in a more recent study, Faria et al. (2020) performed a one-month longitudinal randomized controlled trial (RCT) with 42 chronic stroke patients comparing the Reh@City v2.0 with the TG. The results demonstrated that Reh@City v2.0, an ecologically valid intervention, showed higher effectiveness, with

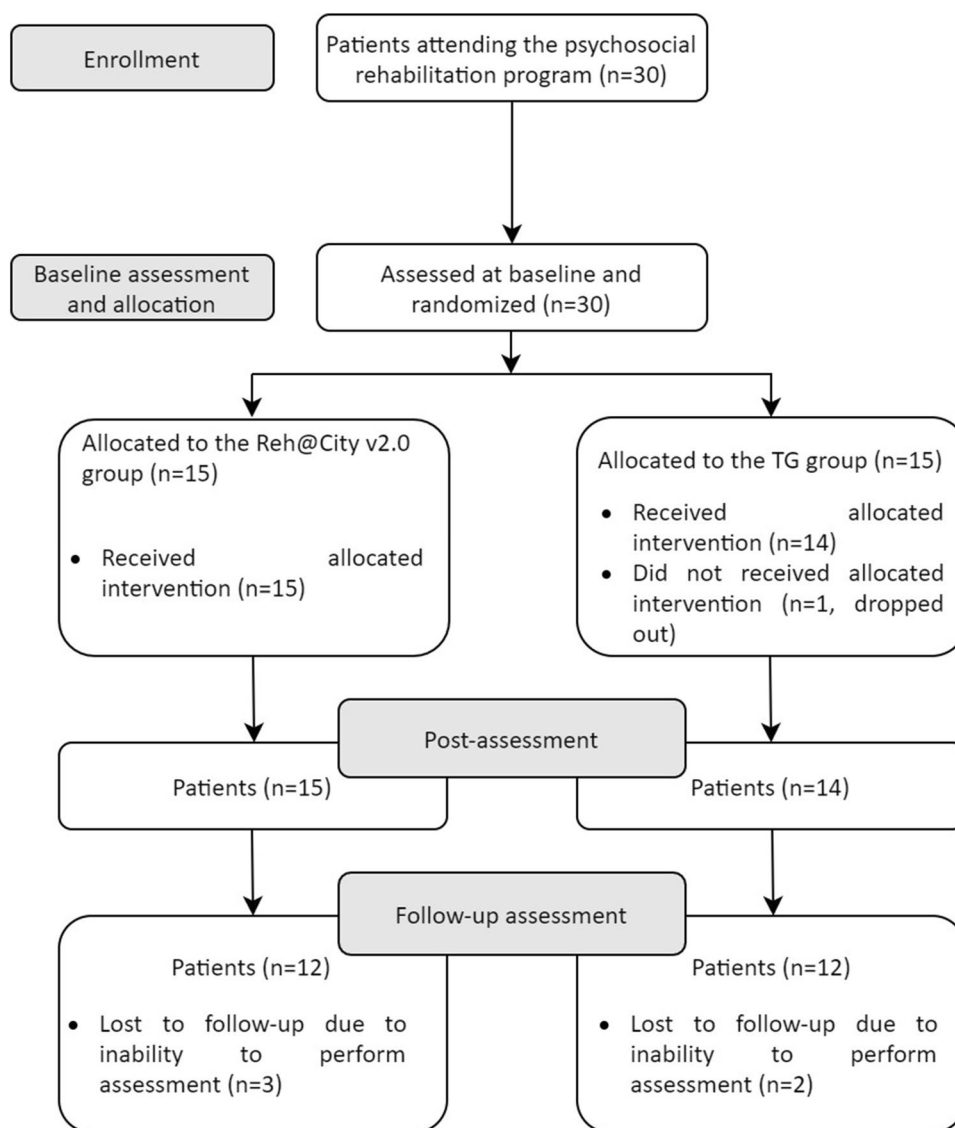
improvements in different cognitive domains (e.g., global cognitive functioning, attention, visuospatial abilities, executive functions, processing speed and verbal memory), and self-perceived cognitive deficits in everyday life, and the TG retained fewer cognitive gains (orientation, processing speed and verbal memory) that were maintained at follow-up.

As to the use of VR for CT purposes in psychiatric disorders, there are considerably fewer studies assessing its effectiveness, and the existing research is conducted mostly among schizophrenia patients. A pilot-study from Marques et al. (2008) evaluated the effectiveness of the Integrated Virtual Environment for Cognitive Rehabilitation (AVIRC) and an adaptation of the Virtual Environments for panic disorder (VEPD) in 14 patients with schizophrenia. The AVIRC consisted of a virtual city with streets, houses, shops, a church, and a supermarket, and the VEPD encompassed a coffee house, an urban context, and a supermarket. Their results suggested that the VR intervention led to improvements in several cognitive measures, such as processing speed, attention, perceptual organization, working memory, verbal comprehension, executive functions, and contributed to the reduction of relapse, re-hospitalization, and drop-out rates. Chan et al. (2010) examined the impact of a VR-based CT program adapted from the Interactive Rehabilitation Exercise System (IREX) in older adults with chronic schizophrenia. Results indicated that the VR group revealed significant improvements in global cognitive function (assessed with the Mini Mental State Examination (MMSE)). Similarly, Plagia et al. (2013) found that schizophrenia patients enrolled in a VR-based CT intervention showed an increase in global cognitive function, sustained attention, and executive function measures, as opposed to the control condition that only revealed gains in sustained attention.

Due to the scarcity of research in this field, we intend to contribute with evidence on the effectiveness of desktop VR-based CT compared to adaptive paper-and-pencil CT. Therefore, the main aims of the present study are to assess and compare the impact of two CT interventions, initially developed and clinically validated for stroke CT – the Reh@City v2.0 and the TG – in a heterogeneous sample of patients with mental and behavioral disorders from a long-term care psychiatric setting. To the best of our knowledge, this is the first two-month RCT investigating the impact of two content-equivalent CT tools, developed under the same adaptation and personalization framework, and delivered in different formats in this clinical population.

¹ The Reh@City and the TG, despite being content-equivalent methodologies, are delivered in different formats (desktop VR vs. paper-and-pencil) (Faria et al. 2019).

Fig.1 Flow diagram of the study intervention process



2 Methods

2.1 Participants and trial design

The recruitment of participants was carried out at Casa de Saúde Câmara Pestana (CSCP) (Funchal, Madeira), which is a female mental health institution that belongs to the Instituto das Irmãs Hospitaleiras do Sagrado Coração de Jesus. For this study, we defined the following inclusion criteria: (a) attending the psychosocial rehabilitation program; (b) having sufficiently preserved expressive and receptive language abilities in order to communicate difficulties when facing demanding tasks and following instructions; (c) maintaining visual and auditory acuity; and (d) being motivated to participate. Patients experiencing an acute psychiatric episode were excluded. All patients completed informed consent prior to participation.

The study was approved by the CSCP's Ethics Committee (reference number: 1/2020) and registered at ClinicalTrials.gov (number NCT04291586).

A total of 30 patients diagnosed with mental and behavioral disorders, according to the International Statistical Classification of Diseases and Related Health Problems 10th manual (ICD-10) (WHO 2010), met the inclusion criteria. They were further allocated to one of two interventions (Reh@City v2.0 or TG) by the psychologists involved in data collection (Fig. 1). The randomization process was conducted using a free web-based resource named Research Randomizer (<https://www.randomizer.org/>). Concerning patients' allocation to the CT interventions, 15 patients were randomly assigned to the Reh@City v2.0 group (3 were lost at follow-up) and 15 to the TG group (1 dropped-out after the baseline assessment and 2 were lost at follow-up).

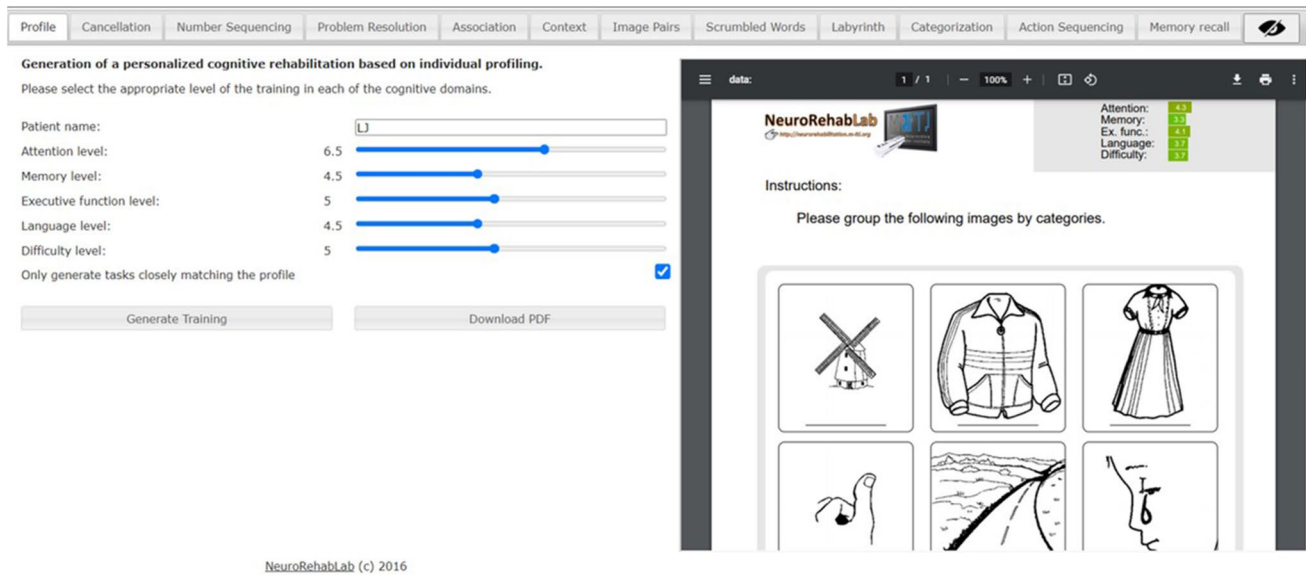


Fig. 2 TG parameters (left) and categorization CT task example (right)

2.2 Intervention protocol

The study was run between June 2019 and February 2020. We conducted a single-blind RCT comparing the impact of the two interventions – Reh@City v2.0 and TG. All patients involved in this study were also enrolled in the CSCP's psychosocial rehabilitation program. The ultimate goals of the CSCP's psychosocial rehabilitation program are to increase autonomy, by enhancing cognitive, emotional, social, and functional abilities, and to facilitate social reintegration. This program was devised for patients with clinically stable conditions, irrespectively of their mental and behavioral disorder, and focuses on helping them to learn or relearn essential skills for independent living (e.g., communication, social interaction, basic and instrumental activities of daily living (IADLs), and occupational attainment) so that they can reintegrate into the community. Besides, this program encompassed psychotherapy and monthly multidisciplinary meetings to assess patients' current situation and progress.

All patients went through a detailed neuropsychological assessment at baseline, post-intervention, and two-month follow-up. The baseline assessments were performed a week before the beginning of the intervention. Psychologists delivered two 30 min CT sessions per week for three months. After the 24th CT session, all participants were assessed within a week. To evaluate the long-term impact of both interventions, a follow-up neuropsychological assessment was performed two months after the end of both interventions. Each assessment session lasted approximately one hour and thirty minutes.

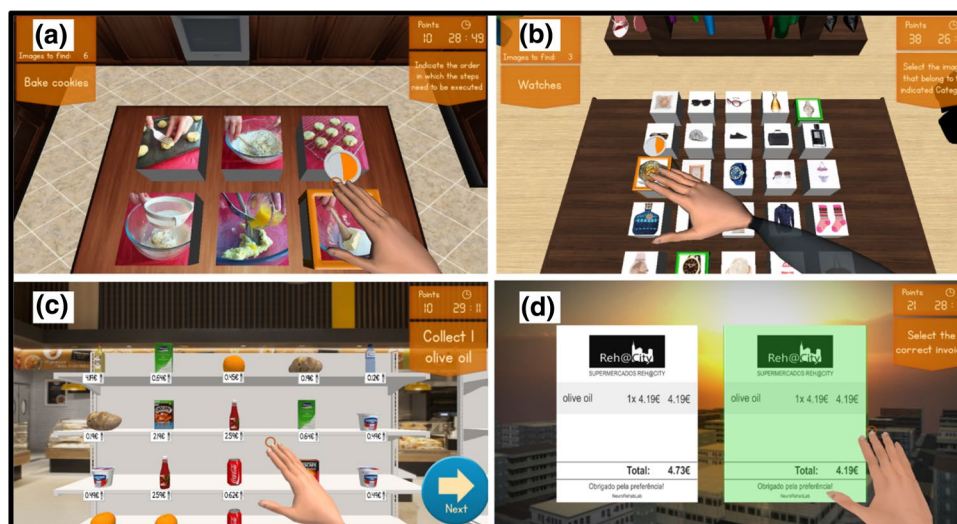
Before starting the intervention, all patients from both groups participated in a brief individual training session that lasted between 10 and 15 min, in order to get familiar with the CT tools. Concerning patients in the Reh@City v2.0 group, the initial session was devoted to a short training in the Reh@City v2.0 platform, where patients interacted with the software using the joystick. As to the patients in the TG group, the initial training session consisted of performing some of its paper-and-pencil tasks. After the initial training session, the two groups underwent 24 time-matched sessions of CT supervised by certified psychologists.

2.3 Intervention description

2.3.1 Task Generator: paper-and-pencil intervention

The TG is a free web-based tool that generates personalized paper-and-pencil CT tasks in a PDF format that are tailored to the user's cognitive profile. The TG comprises 11 different CT tasks, namely: cancellation, numeric sequences, problem solving, association, comprehension of contexts, image pairs, word search, mazes, categorization, action sequencing, and memory of stories and pictures. These tasks were selected, and models for their personalization were created through a participatory design process involving rehabilitation experts (Faria et al. 2018). In order to personalize the CT program through the TG, the psychologist only needs to access the TG website. In the website, it is necessary to complete the following steps to personalize the training content: (1) perform task parameterization (i.e., definition of the attention level, memory level, executive function level,

Fig. 3 Reh@City v2.0 types of tasks: **a** baking cookies at home; **b** collecting jewelry items in the store; **c** buying groceries in the supermarket; **d** paying the supermarket's bill



language level, and difficulty level); (2) click the generate training button; and (3) click the download PDF button. The generated PDF file contains all 11 personalized CT tasks. Psychologists only need to print the CT tasks and hand them over to patients (Fig. 2) (Faria and Bermúdez i Badia 2015).

2.3.2 Reh@City v2.0: desktop VR-based intervention

The Reh@City v2.0 consists of a virtual city with streets, sidewalks, and buildings, where participants are required to perform several CT tasks that resemble common ADLs, in eight different locations (e.g., pharmacy, supermarket, bank, home) (Paulino et al. 2019). In this virtual city, participants are presented with specific errands they need to run, such as: baking cookies at home (Fig. 3a); collecting jewelry items in the store (Fig. 3b); buying groceries in the supermarket (Fig. 3c); and paying the supermarket's bill (Fig. 3d).

During task performance, participants can access different information by pressing specific buttons on the keyboard, namely: task instructions (e.g., go to the post-office); a mini-map and/or a street arrow, illustrating the optimal navigation route to reach the intended location; time and the point counters, which are used as visual feedback elements related to the accomplishment of task objectives. Regarding the point system, participants accumulate points whenever they complete an intermediate task (+ 1) and the overall task (+ 20) and lose points (-1) when mistakes are committed, or a help button used. To enhance the relatedness of the VR tasks to the real world, the eight locations display billboards and products that are found in Portugal's retail stores. Reh@City v2.0 tasks are ecologically valid desktop VR versions of the same paper-and-pencil tasks that compose the TG (see Table 1). Despite the absence of a desktop VR version of the comprehension of contexts and the word search tasks, we did not remove these two paper-and-pencil tasks. Specific

computational models were developed after a participatory design study involving 20 rehabilitation experts to generate both interventions' content and to adjust difficulty parameterization (Faria et al. 2018). Removing CT content at this point could hamper the construct validity and methodological foundations of both tools. Faria et al. (2019) compared both interventions and found that patients had a similar performance irrespectively of the CT method employed. In former studies (Faria et al. 2016, 2020), all tasks were kept. Therefore, these two paper-and-pencil tasks were maintained to preserve the validity of the comparison.

2.4 Personalization procedure

Both TG and Reh@City v2.0 tools, besides sharing the same CT content, are personalized in the same manner, i.e., through the definition of five parameters, namely attention, memory, language, executive functions, and overall difficulty level. To personalize both interventions, we administered the Montreal Cognitive Assessment (MoCA) to all patients with the goal of determining their global cognitive profile. Then, different parameters were established as follows:

Attention parameter: MoCA's attention domain score (0–6 points);

Memory parameter: MoCA's memory domain score (i.e., delayed recall and spatial and temporal orientation) (0–11 points);

Executive functions parameter: MoCA's visuospatial, executive and abstraction domains scores (0–7 points);

Language parameter: MoCA's naming and language domains scores (0–6 points);

Difficulty parameter: MoCA's total score (0–30 points).

Table 1 Paper-and-pencil tasks (TG) correspondence to desktop VR tasks (Reh@City v2.0)

Tasks	TG	Reh@City v2.0
Cancellation	Identify a target stimulus (e.g., letter, number of symbol) among distractors	Identify a target item at the supermarket, pharmacy, and post-office, according to a shopping list
Numeric sequences	Calculate the missing numbers in a sequence considering a specific pattern (i.e., addition or subtraction)	Calculate the missing PIN numbers at the ATM
Problem solving	Solve calculations involving addition, subtraction, and multiplication, or solve arithmetic word problems presented	Select the correct invoice at the supermarket
Association	Pair associated images	Find matching cards in a memory game at the park
Comprehension of contexts	Analyze a target image and find the correct descriptions by answering true or false affirmations	Not applicable
Image pairs	Memorize a set of image pairs and then recall them after a 30 min interval	Find matching cards in a memory game at the park
Word search	Find target words that can be placed horizontally, vertically, or diagonally among random letters	Not applicable
Mazes	Find the route from the start to the exit of a labyrinth	Find the shortest navigation route to arrive to a given location in the virtual city
Categorization	Identify the category of different images	Select target items from a clothing store according to a given category (e.g., shoes, sunglasses, women's clothing)
Action sequencing	Organize a set of actions to perform a given activity	Select, in the proper order, the steps needed to accomplish an activity of daily living at home
Memory of stories or pictures	Memorize information about a story or picture and recall it after a few minutes by answering true or false questions	Memorize verbal or visual information from a newspaper or magazine at the kiosk and then answer true or false questions, regarding the previous information, when reaching the next location

All CT tasks have a different multi-domain cognitive profile, defined through the various parameters, namely the attention, memory, language, executive functions, and difficulty parameters. This means that the generated CT tasks do not train a specific cognitive domain but tackle all the above-mentioned domains to different extents, according to the established difficulty level.

Consistent with previous work, after every training session, a mean performance score was calculated using a 0–100% scale, and the difficulty for the next set of tasks was established with the following reasoning: (a) the difficulty parameter was decreased by 0.5 if the mean performance was below 50%; (b) the difficulty parameter was increased by 0.5 if the mean performance was higher than 70%; and (c) the difficulty was maintained if the mean performance ranged from 50 to 70% (Faria and Bermúdez i Badia 2018). In the TG, the personalization and adaptation process were done manually. The Reh@City v2.0 presents additional features, when compared to the TG, namely the automatic normalization process of the different parameters and adjustment of the desktop VR CT tasks based on participant's performance, as well as the ability to save participant's performance from session to session (Faria et al. 2016).

2.5 Experimental setup

2.5.1 Task Generator (TG)

The TG consists of an online application that is freely available at <https://neurorehabilitation.m-iti.org/TaskGenerator/>. In this website, it is possible to personalize the CT tasks content through the definition of the five aforementioned parameters (obtained from MoCA). After this step, it is possible to generate a set of 11 personalized CT tasks, presented in a PDF file, that are to be downloaded and then printed. After printed, patients completed the tasks with the assistance of certified psychologists.

2.5.2 Reh@City v2.0

The Reh@City v2.0 was implemented using the Unity 3D game engine and installed on a PC. Patients worked on a tabletop, facing an LCD monitor, and interacted with the virtual environment through a joystick handle with two buttons, one for selection and another for help. This simplified user interface (PC and joystick) was considered more suitable for our sample for several reasons: (a) patients' clinical diagnosis [patients' with chronic mental and behavioral disorders are more prone to display a low

tolerance to external stimuli (e.g., HMD)]; (b) patients' chronic consumption of psychotropic medication and associated unwanted adverse effects (e.g., dizziness, nausea, and increased fall risk); and (c) patients' potential low digital literacy.

2.6 Outcome measures

2.6.1 Primary outcome measures: global cognitive functioning, processing speed, sustained and selective attention, verbal memory, visual memory, and executive functions

The Montreal Cognitive assessment was used to assess global cognitive functioning (Freitas et al. 2011). Moreover, we selected specific processing speed, attention, memory, and executive function instruments. To assess processing speed, we used the Digit Symbol and Symbol Search subtests of Wechsler Adult Intelligence Scale III (WAIS III) (Wechsler 1997a). The assessment of sustained and selective attention was performed with Toulouse-Piéron, which is a widely used cancellation ten-minute test (Toulouse and Piéron 1986). To assess verbal memory and visual memory, we used the Verbal Paired Associates I subtest of Wechsler Memory Scale-III (WMS-III) (Wechsler 1997b) and the Rey-Osterrieth Complex Figure Test (RCFT) (Rey 1998), respectively. Finally, executive functions were assessed with Semantic and Phonemic Verbal Fluency Tests (Cavaco et al. 2013).

2.6.2 Secondary outcome measures: depressive symptomatology and quality of life

We assessed the presence and severity of depressive symptomatology with the Beck Depression Inventory II (BDI-II), which is a 21-item self-report rating inventory (Beck et al. 1996). Quality of life was evaluated with the World Health Organization Quality of Life abbreviated (WHOQOL-BREF) self-report 26-item questionnaire that assesses four quality of life domains, namely: physical, psychological, social relationships and environmental (Canavarró et al. 2006).

2.7 Statistical analysis

The statistical analysis was performed using the *Statistical Package for the Social Sciences version 26* (SPSS Inc., Chicago IL, USA). Normality of the data was assessed with the Shapiro–Wilk test. Normally distributed continuous variables (age and years of schooling) were presented as mean and standard deviation, and categorical variables (diagnosis) as frequency and percentage. Differences between

both groups in terms of demographic and clinical data were assessed with the independent samples t-test and the Fisher's exact test, respectively. Considering that most of the neuropsychological assessment data were not normally distributed, non-parametric tests were run to evaluate within and between-groups differences in the different assessment moments. Thus, non-normally distributed continuous variables were presented as median and interquartile ranges (IQR). The Wilcoxon signed-rank test was used to analyze within-groups differences over time, whereas the two-tailed Mann–Whitney test was employed to assess between-groups differences from baseline to post-intervention and from baseline to follow-up. Effect sizes (r) estimates were calculated ($r = Z/\sqrt{N}$) and interpreted as follows: 0.2 = small, 0.5 = medium, and 0.8 = large (Cohen 1988). In all statistical analysis, p -values ≤ 0.05 are reported and considered statistically significant.

3 Results

3.1 Sample description

The demographic and clinical characteristics of the sample are reported in Table 2. All 29 patients were female, had an average age of 55.93 (SD = 11.57) years old, and an average of 5.55 (SD = 4.24) years of schooling. Most patients (51.72%) had a schizophrenia diagnosis, while the remaining presented distinct clinical conditions (e.g., mental retardation, recurrent depressive disorder, schizoaffective disorder). In addition, digital literacy was informally assessed by simply asking patients if they had previous computer experience and internet navigation skills. We found that the majority of our sample had low digital literacy, with only 4 (13.79%) out of 29 patients (2 from the Reh@City v2.0 group and 2 from the TG group) having used computers prior to this study and displaying basic Internet skills.

According to the Shapiro–Wilk test, data were normally distributed in both groups for age ($SW_{\text{Reh@City v2.0}} = 0.923$, $p = 0.214$; $SW_{\text{TG}} = 0.919$, $p = 0.212$) and years of schooling ($SW_{\text{Reh@City v2.0}} = 0.896$, $p = 0.082$; $SW_{\text{TG}} = 0.937$, $p = 0.387$). There were no significant differences between both groups in demographic (Age: $t(27) = 1.030$, $p = 0.312$; Schooling: $t(27) = -0.282$, $p = 0.780$) and clinical variables (Diagnosis: Fisher's exact test = 5.424, $p = 0.592$). Regarding the neuropsychological assessment baseline scores, the Mann–Whitney test indicated that there were no statistically significant differences between the Reh@City v2.0 and the TG groups in any of the primary and secondary outcome measures at baseline (see the supplementary material).

Table 2 Demographic and clinical characteristics of the sample

	Reh@City v2.0 (<i>n</i> = 15)	Task generator (<i>n</i> = 14)	Statistical test and <i>p</i> -value
Age	54.07, SD = 11.4	49.64, SD = 11.99	$t = 1.030, p = .312$
Schooling (years)	5.33, SD = 4.08	5.79, SD = 4.54	$t = -.282, p = .780$
Diagnosis			
Schizophrenia	9 (60%)	6 (42.9%)	FET = 5.424, $p = .592$
Schizoaffective disorder	1 (6.7%)	1 (7.1%)	
Recurrent depressive disorder	0	2 (14.3%)	
Mental retardation (mild)	3 (20%)	4 (28.6%)	
Obsessive compulsive disorder	1 (6.7%)	0	
Bipolar affective disorder	0	1 (7.1%)	
Mental and behavioral disorders due to the use of alcohol	1 (6.7%)	0	

Table 3 MoCA scores (presented as Medians and IQR) at baseline, post-intervention, and follow-up

	Reh@City v2.0			TG		
	Baseline (<i>n</i> = 15)	Post (<i>n</i> = 15)	FU (<i>n</i> = 12)	Baseline (<i>n</i> = 14)	Post (<i>n</i> = 14)	FU (<i>n</i> = 12)
Total	11 (12)	13 (9)	14.5 (11)	10 (15)	15.5 (13.7)	15 (15.5)
Visuospatial/executive	2 (3)	3 (3)	3 (2.5)	3 (3.3)	3.5 (4)	2.5 (3.7)
Naming	2 (1)	2 (1)	2.5 (2)	2 (1)	3 (1)	2.5 (2)
Attention	1 (2)	2 (2)	2.5 (3.8)	1 (5)	3 (1)	1 (5)
Language	0 (0)	0 (1)	1 (1.8)	0 (1.3)	1.5 (5)	1 (1.5)
Abstraction	0 (0)	0 (1)	0 (1)	.50 (1.3)	0 (-1)	.50 (2)*
Memory	0 (0)	0 (3)	0 (3)	0 (2)	1.5 (2)	1 (-3)
Orientation	5 (2)	6 (2)	6 (1.7)	(3)	5 (2.2)	5 (1.7)

Within-groups significant differences are represented in bold and between-groups differences indicated with *

3.2 Primary outcome measures

3.2.1 MoCA–global cognitive functioning

Table 3 describes the global cognitive functioning, as assessed by MoCA, of both groups in the three assessment moments. A within-groups analysis revealed that only the Reh@City v2.0 group showed statistically significant improvements in MoCA's total score [Baseline: Mdn = 11, IQR = 12; FU: Mdn = 14.5, IQR = 11 ($W_{(12)} = 58.000, Z = -2.239, p = 0.025, r = 0.65$)] at follow-up. In addition, the analysis of MoCA's domains revealed significantly higher scores in the language domain [Baseline: Mdn = 0, IQR = 1; FU: Mdn = 1, IQR = 1.8 ($W_{(12)} = 21.000, Z = -2.264, p = 0.024, r = 0.65$)] and in the visuospatial/executive domain [Pre: Mdn = 2, IQR = 3; FU: Mdn = 3, IQR = 2.5 ($W_{(12)} = 26.000, Z = -2.050, p = 0.040, r = 0.59$)] at follow-up for the Reh@City v2.0 group. A between-groups analysis indicated that the TG group improved significantly more than the Reh@City group v2.0 in the MoCA's abstraction domain [Baseline: Mdn = 0.50, IQR = 1.3; FU: Mdn = 0.50, IQR = 2 ($U = 75.50,$

$Z = -1.446, p = 0.038, r = 0.30$)] at follow-up (see the supplementary material).

3.2.2 Digit symbol and symbol search–processing speed

Table 4 describes the Digit Symbol and Symbol Search performance for both groups at baseline, post-intervention, and follow-up. Results showed that the TG group had significantly higher scores in the coding and incidental learning trials both at post-intervention [Coding: Baseline: Mdn = 20, IQR = 21; Post: Mdn = 30.5, IQR = 25.5 ($W_{(14)} = 94.500, Z = -2.641, p = 0.008, r = 0.70$)], [Incidental learning: Baseline: Mdn = 8, IQR = 11.2; Post: Mdn = 10.5, IQR = 11 ($W_{(14)} = 85.000, Z = 2.047, p = 0.041, r = 0.54$)] and at follow-up [Coding: Baseline: Mdn = 20, IQR = 21; FU: Mdn = 31.5, IQR = 24.2 ($W_{(12)} = 0.000, Z = -2.937, p = 0.003, r = 0.85$)], [Incidental learning: Baseline: Mdn = 8, IQR = 11.2; FU: Mdn = 15, IQR = 7.5 ($W_{(12)} = 73.000, Z = 2.680, p = 0.007, r = 0.77$)]. There was only a between-groups difference in the coding trial of the Digit Symbol at follow-up, with significantly higher scores in the TG group when compared to the Reh@City v2.0 group ($U = 74.50, Z = -1.997, p = 0.046,$

Table 4 Digit Symbol and Symbol Search scores (presented as Medians and IQR) at baseline, post-intervention, and follow-up

	Reh@City v2.0			TG		
	Baseline (<i>n</i> = 15)	Post (<i>n</i> = 15)	FU (<i>n</i> = 12)	Baseline (<i>n</i> = 14)	Post (<i>n</i> = 14)	FU (<i>n</i> = 12)
Coding	20 (21)	21 (20)	24.5 (19.5)	20 (24.8)	30.5 (25.5)	31.5 (24.2)*
Incidental learning	10 (13)	8 (13)	13 (13.7)	8 (11.2)	10.5 (11)	15 (7.5)
Copy	37 (38)	43 (22)	59.5 (37.7)	52 (47.8)	49.5 (45.5)	61.5 (58)
Symbol search	6 (8)	8 (10)	6 (5.8)	8 (11)	9.5 (7.5)	9.5 (12.8)

Within-groups significant differences are represented in bold and between-groups differences indicated with *

Table 5 Toulouse-Piéron scores (presented as Medians and IQR) at baseline, post-intervention, and follow-up

	Reh@City v2.0			TG		
	Baseline (<i>n</i> = 15)	Post (<i>n</i> = 15)	FU (<i>n</i> = 12)	Baseline (<i>n</i> = 14)	Post (<i>n</i> = 14)	FU (<i>n</i> = 12)
Work efficiency	0.13 (54)	23 (81)	21 (126.3)	27 (83.3)	46.5 (104.8)	71.5 (108.8)
Dispersion index	76.7 (110.13)	54.6 (112.3)	41.9 (93.9)	45.3 (120.7)	20.92 (96.5)	28.7 (80.1)

Within-groups significant differences are represented in bold

Table 6 Verbal paired associates I scores (presented as Medians and IQR) at baseline, post-intervention, and follow-up

	Reh@City v2.0			TG		
	Baseline (<i>n</i> = 15)	Post (<i>n</i> = 15)	FU (<i>n</i> = 12)	Baseline (<i>n</i> = 14)	Post (<i>n</i> = 14)	FU (<i>n</i> = 12)
Immediate recall	1 (1)	1 (1)	1.5 (2)	1 (2.3)	1.5 (2.3)	1.5 (2.8)
Retention	0 (75)	0 (100)	16.7 (87.5)	0 (90.6)	50 (73.3)	3.5 (58.3)
Recognition	22 (6)	23 (7)	24 (2.7)	21.5 (8.4)	24 (3.2)	24 (4.2)

Within-groups significant differences are represented in bold

$r = 0.41$). No within- or between-group differences were found in the Symbol Search subtest (see the supplementary material).

3.2.3 Toulouse-Piéron–sustained and selective attention

Table 5 presents the performance of the two groups in Toulouse-Piéron at baseline, post-intervention, and follow-up. Only the TG group demonstrated a statistically significant increase in work efficiency at follow-up [Baseline: Mdn = 27, IQR = 83.3; FU: Mdn = 71.5, IQR = 108.8 ($W_{(12)} = 74.000$, $Z = -2.752$, $p = 0.006$, $r = 0.79$)]. No significant between-groups differences were found in work efficiency and dispersion index in any assessment moment (see the supplementary material).

3.2.4 Verbal paired associates I–verbal memory

The Verbal Paired Associates I scores for both groups at baseline, post-intervention, and follow-up are illustrated

in Table 6. Only the TG group demonstrated a significant improvement in the recognition trial [Baseline: Mdn = 21.5, IQR = 8.4; Post: Mdn = 24, IQR = 3.2 ($W_{(14)} = 33.000$, $Z = -2.113$, $p = 0.035$, $r = 0.56$)] at post-intervention. No between-groups differences were identified in any of the three assessment moments (see the supplementary material).

3.2.5 Rey-Osterrieth complex figure test–visual memory

Table 7 illustrates the copy and immediate recall trials of the RCFT for both groups at baseline, post-intervention, and follow-up. The Reh@City v2.0 group presented significantly higher scores in the immediate recall trial in both post-intervention [Baseline: Mdn = 3.5, IQR = 9; Post: Mdn = 8.5, IQR = 11 ($W_{(15)} = 96.500$, $Z = -2.767$, $p = 0.006$, $r = 0.71$)] and follow-up [Baseline: Mdn = 3.5, IQR = 9; FU: Mdn = 6, IQR = 8.7 ($W_{(12)} = 71.000$, $Z = -2.512$, $p = 0.012$, $r = 0.73$)]. No between-groups differences were found in any of the three assessment moments (see the supplementary material).

Table 7 Rey-Osterrieth Complex Figure Test scores (presented as Medians and IQR) at baseline, post-intervention, and follow-up

	Reh@City v2.0			TG		
	Baseline (<i>n</i> = 15)	Post (<i>n</i> = 15)	FU (<i>n</i> = 12)	Baseline (<i>n</i> = 14)	Post (<i>n</i> = 14)	FU (<i>n</i> = 12)
Copy	21 (27)	21 (22.5)	18 (24.6)	17.3 (23.5)	22.8 (22.9)	20.8 (18.7)
Immediate recall	3.5 (9)	8.5 (11)	6 (8.7)	6.5 (6.6)	8.5 (10.5)	8.3 (15.1)

Within-groups significant differences are represented in bold

Table 8 Semantic and Phonemic Verbal Fluency scores (presented as Medians and IQR) at baseline, post-intervention, and follow-up

	Reh@City v2.0			TG		
	Baseline (<i>n</i> = 15)	Post (<i>n</i> = 15)	FU (<i>n</i> = 12)	Baseline (<i>n</i> = 14)	Post (<i>n</i> = 14)	FU (<i>n</i> = 12)
Semantic Verbal Fluency (animals)	8 (8)	9 (5)	9.5 (5.8)	10 (6.5)	8.5 (7.5)	10.5 (9.8)
Phonemic Fluency total	12 (19)	12 (12)	14 (7)	13 (26.8)	13.5 (5.3)	12.5 (27.5)
Letter M	4 (6)	4 (5)	4 (3.5)	2.5 (8)	3.5 (8)	3 (8)
Letter R	4 (6)	4 (3)	4 (4.5)	5 (9)	4 (8.8)	5 (8)
Letter P	4 (5)	5 (4)	5 (2.5)	6 (9.5)	4.5 (8.3)	4 (9.3)

Table 9 Beck Depression Inventory II scores (presented as Medians and IQR) at baseline, post-intervention, and follow-up

	Reh@City v2.0			TG		
	Baseline (<i>n</i> = 15)	Post (<i>n</i> = 15)	FU (<i>n</i> = 12)	Baseline (<i>n</i> = 14)	Post (<i>n</i> = 14)	FU (<i>n</i> = 12)
BDI-II	10 (9)	6 (9)*	3.5 (7.5)*	3 (12.8)	4 (6.8)	2 (5.7)

Within-groups significant differences are represented in bold and between-groups differences indicated with *

3.2.6 Semantic and phonemic verbal fluency–executive functions

The Semantic and Phonemic Verbal Fluency scores for both groups in the various assessment moments are listed in Table 8. No within- or between-group differences were identified at baseline, post-intervention, and follow-up.

3.3 Secondary outcome measures

3.3.1 Beck depression inventory II–depressive symptomatology

The BDI-II scores in both groups at baseline, post-intervention, and follow-up are presented in Table 9. Only the Reh@City v2.0 group showed a statistically significant reduction in the BDI-II scores both at post-intervention [Baseline: Mdn = 10, IQR = 9; Post: Mdn = 6, IQR = 9 ($W_{(15)} = 3.500$, $Z = -2.949$, $p = 0.003$, $r = 0.76$)] and follow-up [Baseline: Mdn = 10, IQR = 9; FU: Mdn = 3.5, IQR = 7.5 ($W_{(12)} = 0.000$, $Z = -2.812$, $p = 0.005$, $r = 0.81$)]. Between-groups

comparison showed consistently a statistically significant decrease in the BDI-II scores at both post-intervention ($U = 45.00$, $Z = -2.655$, $p = 0.007$, $r = 0.49$) and follow-up ($U = 17.50$, $Z = -3.171$, $p = 0.002$, $r = 0.65$) for the Reh@City v2.0 group (see the supplementary material).

3.3.2 World health organization quality of life abbreviated version–quality of life

Table 10 shows the WHOQOL-BREF scores for both groups in the three assessment moments. Only the TG group demonstrated significant improvements in the social relationships and in the environmental domains of the WHOQOL-BREF, at both post-intervention [Social relationships: Baseline: Mdn = 66.7, IQR = 18.8; Post: Mdn = 75, IQR = 8.3 ($W_{(14)} = 40.000$, $Z = -2.092$, $p = 0.036$, $r = 0.56$)], [Environmental domain: Baseline: Mdn = 67.2, IQR = 18; Post: Mdn = 75, IQR = 16.5 ($W_{(14)} = 80.000$, $Z = -2.418$, $p = 0.016$, $r = 0.65$)], and follow-up [Social relationships: Baseline: Mdn = 66.7, IQR = 18.8 FU: Mdn = 70.9, IQR = 8.3 ($W_{(12)} = 40.000$, $Z = -2.099$, $p = 0.036$, $r = 0.61$)]

Table 10 WHOQOL-BREF scores (presented as Medians (IQR)) at baseline, post-intervention, and follow-up

	Reh@City v2.0			TG		
	Baseline (<i>n</i> = 15)	Post (<i>n</i> = 15)	FU (<i>n</i> = 12)	Baseline (<i>n</i> = 14)	Post (<i>n</i> = 14)	FU (<i>n</i> = 12)
WHOQOL-BREF	72.6 (12.8)	72.3 (14.3)	69.5 (13.1)	73.14 (20.4)	74.7 (16.3)	78.3 (14.8)
Physical domain	67.9 (17.9)	75 (21.4)	71.4 (14.3)	82.1 (21.5)	80.4 (17)	82.1 (20.5)
Psychological domain	75 (17.2)	79.2 (23.2)	75 (20.7)	70.8 (28.1)	70.8 (16.6)	77.1 (21.6)
Social relationships domain	75 (19.7)	66.7 (16.7)	66.7 (25)	66.7 (18.8)	75 (8.3)*	70.9 (8.3)*
Environmental domain	71.9 (15.6)	71.9 (8.8)	71.9 (14.4)	67.2 (18)	75 (16.5)	78.2 (15.5)

Within-groups significant differences are represented in bold and between-groups differences indicated with *

[Environmental domain: Baseline: Mdn = 67.2, IQR = 18; FU: Mdn = 78.2, IQR = 15.5 ($W_{(12)} = 54.000$, $Z = -2.710$, $p = 0.007$, $r = 0.78$)]. We also found a significant between-groups difference in the social relationships domain at both post-intervention ($U = 55.00$, $Z = -2.217$, $p = 0.026$, $r = 0.41$) and follow-up ($U = 39.00$, $Z = -2.147$, $p = 0.032$, $r = 0.44$) with improvements in the TG group and decline in the Reh@City v2.0 group (see the supplementary material).

4 Discussion

VR applications to CT have been growing in the last few years, with encouraging results in neurological and psychiatric populations. VR methods provide a more ecological training experience through the performance of simulations of familiar ADLs. Also, VR simulations of ADLs are a rather global way of training cognitive functions, while traditional CT methods typically tackle domain-specific processes. In this two-month RCT, we aimed to evaluate and compare the impact of two content-equivalent CT approaches, delivered in a desktop VR format (Reh@City v2.0) and a paper-and-pencil format (TG), in a sample of institutionalized patients with chronic mental and behavioral disorders. We intended to contribute with additional research on the impact of two CT methodologies, grounded on the same adaptation and personalization framework, in an understudied and clinically diverse sample of patients from a long-term care psychiatric facility. This study highlighted both CT interventions' cognitive and non-cognitive benefits, suggesting that the combination of both interventions could lead to better outcomes. This is important since traditional and technology-based CT interventions should be viewed as complementary training approaches.

4.1 Primary outcome measures

The within-group analysis, from pre- to post-intervention, revealed statistically significant improvements in the Reh@City v2.0 group in visual memory (immediate recognition trial of RCFT). At two-month follow-up, the desktop VR

group maintained previous gains in visual memory, and new improvements were found in global cognitive function and other specific cognitive domains, namely language, visuospatial abilities, and executive functions, as assessed by MoCA. Concerning the TG group, there were improvements in processing speed (coding and incidental learning trials of Digit Symbol) and verbal memory (recognition trial of Verbal Paired Associates I). At follow-up, gains were maintained by the TG group, and new improvements identified in sustained and selective attention (work efficiency index of Toulouse-Piéron). The between-groups comparison showed greater gains in processing speed and abstraction (MoCA) in the TG group.

These findings are consistent with previous research. Visual memory improvements after VR-based CT have been reported in Zając-Lamparska et al. (2019), which assessed the effects of the GRADYS VR system in a sample of older adults living without and with mild dementia. In this study, older adults without dementia had more improvements in visual memory and visuospatial processing than older adults with mild dementia, despite the latter also showed positive changes in easier visual memory tasks (i.e., copy task of RCFT). Additionally, in Gamito et al. (2015), despite the absence of statistically significant changes, data of the VR group indicated an improvement in visual memory at follow-up. Also, the positive impact of VR application to CT in global cognitive function is well documented in prior studies with schizophrenia (Chan et al. 2010; Plagia et al. 2013) and stroke patients (Faria et al. 2016, 2020). In Faria et al. (2016) study, which compared the Reh@City v1.0 intervention with conventional CT, stroke patients in the desktop VR group exhibited significantly higher improvements, not only in global cognitive function, but also in more specific cognitive domains, such as attention, memory and visuospatial abilities, at post-intervention. Similar results were found in Faria et al. (2020), where stroke patients in the Reh@City v2.0 group showed significantly higher improvements in the MoCA global cognitive function and its attention, visuospatial, and executive functions domains at post-intervention. Our results corroborate Reh@City v2.0 positive impact on global cognitive function, visuospatial abilities,

and executive functions. Interestingly, these changes only appeared at follow-up, two months after the intervention. We also found an additional improvement in language that is consistent with Marques et al. (2008) study in schizophrenia patients.

TG's positive impact is also in line with previous findings in stroke patients (Faria et al. 2020), namely the processing speed (Symbol Search (WAIS III)), sustained attention (TMT A execution time), and verbal memory improvements (retention trial of Verbal Paired Associates I). Nonetheless, we found greater improvements in the TG group's performance in processing speed and abstraction (MoCA) at follow-up. TG's data must be considered with caution as results can be related to the similar nature of the TG's CT tasks and the cognitive assessment measures. Given that the TG's paper-and-pencil CT tasks tackle specific cognitive processes, which is also a common feature of cognitive instruments, it is reasonable to assume that there may be a more expressive translation of training gains to the assessment measures. However, studies show that these gains may not be necessarily representative of functional improvements in ADLs (Coyle et al. 2015; Prikken et al. 2019). On the other hand, the Reh@City v2.0 offers a multidimensional CT that may have a larger impact on broader domains (e.g., global cognitive function) and possibly a greater transfer of training gains to everyday life, which is the ultimate goal of a rehabilitation program. Most of our participants are inpatients and only basic ADLs are routinely assessed at the institution. As such, we could not assess the transfer of training gains to more advanced functional outcomes (i.e., instrumental activities of daily living).

It is also important to note that, since both interventions resulted in differential gains in cognitive function, it could be useful to combine desktop VR CT and adaptive paper-and-pencil CT in order to potentiate cognitive improvements. As reported by Kim et al. (2011) and Young et al. (2014), the combined use of VR and computerized conventional CT resulted in additional gains in cognition, both in stroke and brain tumor patients, in comparison with computerized CT alone.

4.2 Secondary outcome measures

Apart from cognitive functions, we evaluated the impact of both interventions in emotional and quality of life measures. The within- and between-groups analysis, from pre to post-intervention, indicated that the Reh@City v2.0 group showed a statistically significant reduction in depressive symptomatology. The TG group revealed improvements in social relationships and environmental domains of quality of life, being that only the social relationships domain improvement was significant in the between-groups analysis. At two-month follow-up, both groups maintained the

previous gains, with the Reh@City v2.0 group exhibiting a higher reduction in depressive symptomatology and the TG revealing greater improvements in the social relationships' domain of quality of life.

The improvement in depressive symptomatology after the desktop VR-based CT intervention is consistent with a study performed by Maier et al. (2020) that compared adaptive conjunctive cognitive training (ACCT) in VR with conventional at-home CT in chronic stroke patients. Results indicated that the ACCT VR intervention was associated with lower levels of depressive symptomatology and that improvements in particular cognitive domains, namely attention and memory, predicted lower levels of depression. Also, a systematic review from Coyle et al. (2015) exploring the effects of computerized CT and VR-based CT for individuals with an increased risk of cognitive decline (e.g., mild cognitive impairment and dementia) confirmed the beneficial impact of VR interventions in psychological outcomes, such as depressive symptomatology and anxiety. In the case of our study, we speculate that the depressive symptomatology reduction observed in the Reh@City v2.0 group could be due to improvements in cognitive outcome measures and, eventually, to the resemblance of the desktop VR-based CT tasks with everyday life. Regarding the first motif, evidence suggests that depression is correlated with cognitive function (Harvey 2011; Iosifescu 2012); in this sense, we can hypothesize that the global cognitive functioning improvements in the desktop VR group may explain depressive symptomatology amelioration, or, conversely, the desktop VR-based CT intervention itself could have impacted mood, which resulted in cognitive gains. Either way, this finding emphasizes the need of devising interventions that consider both cognitive and emotional aspects of human functioning, due to its indissociable nature. The second motif is associated with the naturalistic features of the desktop VR-based intervention, i.e., the existence of ADLs simulations using popular Portuguese brands, which increase the similarity of the training setting with daily life. This fact could have positively influenced motivation and engagement in the desktop VR group, and therefore reduced depressive symptomatology, given that these patients experienced a sense of relatedness with day-to-day functioning, which could have attributed a higher significance to this type of training. In addition, replicating ADLs in an institutionalized setting could lead to stronger emotional engagement due to the recreation of daily living outer environments, of which patients are normally deprived. Concerning the TG, positive effects in the social relationships domain of quality of life may be related to the fact that this intervention involved more interaction with the psychologist, which could, in fact, have strengthened the therapeutic relationship. The Reh@City v2.0 group required less help from

the psychologist during sessions, which resulted in greater independence and autonomy from the psychologist while executing the tasks. On the other hand, patients in the TG typically require more assistance from the psychologist while executing the task, since there is not a feedback system in the paper-and-pencil CT. This social interaction element, particularly present in the paper-and-pencil intervention, represents a crucial part of every therapeutic process and is known to contribute to patients' improvement.

4.3 Limitations

Our results must be interpreted with caution due to some limitations that should be acknowledged. Firstly, there was considerable heterogeneity in terms of clinical diagnosis in both groups. Since having a clinically stable condition, independently of the formal diagnosis, was one of the main criteria for admission on the CSCP's psychosocial rehabilitation program, the included patients had quite different clinical diagnosis. Consequently, it was not possible to create homogenous patient groups. Perhaps, patients diagnosed with a certain type of mental and behavioral disorder could have benefited more from the interventions; however, we cannot ascertain that. Besides, we did not control important clinical information, such as medication regimen, chronicity of illness or institutionalization, and relapse history. Concerning specifically the medication regimen, all patients were prescribed different medications according to their diagnosis and medical history. As such, we did not measure the influence of the pharmacological treatments in patients' performance and, therefore, cannot exclude the effects of medication, nor of the aforementioned clinical variables, in our patients' response to the cognitive interventions. It is also important to note that only participants were blinded to the interventions. Psychologists were aware of the participant's assignment to both treatment conditions, having then performed all the assessments and interventions. Being aware of this information may increase the likelihood of bias, since this knowledge might influence psychologists' behavior toward participants, even if unconsciously. Furthermore, the Reh@City v2.0 and the TG were originally designed for a different clinical condition—stroke—what poses challenges when transferring and assessing these interventions to other patient populations. In addition, despite consisting of two content-equivalent CT methods the Reh@City v2.0 and the TG imply different interaction patterns with the CT training method (i.e., desktop VR vs. paper-and-pencil), and, also, with the psychologist. The type of CT content (i.e., ADLs virtual simulations) displayed in the Reh@City v2.0 platform may be more engaging and motivating for patients living in a long-term psychiatric setting, deprived of performing these types of activities in the community setting. In the TG, the printed CT tasks do not consist of

everyday life tasks and often contain artificial stimuli (e.g., a cancellation task with squares, and circles). We assume that this type of content can be less motivating because it is harder for the patient to understand the significance of the training, i.e., how will a cancellation task with squares and circles be helpful for the improvement of attentional ability in their daily life? As to the psychologist and patient interaction, there is also a difference in both approaches. In the Reh@City v2.0 intervention, the patient is seated in front of a PC and the system provides the instructions and performance feedback, while in the TG it is the psychologist who gives both instructions and feedback, resulting in more interaction with the participants. For the reasons described above, we cannot exclude the influence of these two types of interaction (training method and psychologist-patient relationship) in our results. Regarding the neuropsychological assessment protocol, it would have been useful to include a measure of IADLs to assess the transfer of cognitive gains to more complex ADLs. Finally, there may have been learning effects given that most neuropsychological assessment measures lacked parallel versions for multiple assessments, except for MoCA. Nonetheless, the same neuropsychological assessment protocol was administered to both groups, so, irrespectively of the existence of learning effects, this comparison is still valid.

5 Conclusions

To the best of our knowledge, this is the first longitudinal RCT investigating the impact of two content-equivalent cognitive interventions, delivered in different formats—desktop VR (Reh@City v2.0) and paper-and-pencil (TG)—, in a sample of patients with mental and behavioral disorders from a long-term care psychiatric setting. Overall, patients from both groups revealed differential positive changes in primary and secondary outcome measures in the different assessment moments, which may suggest that the combined use of these two interventions could lead to higher short and long-term benefits in the assessed domains. In terms of cognitive gains, the Reh@City v2.0 group, at post-intervention, showed improvements in visual memory, and at follow-up, besides maintaining former gains, revealed additional improvements in global cognitive functioning, language, visuospatial abilities, and executive functions. The TG group exhibited processing speed and verbal memory improvements at post-intervention. At two-month follow-up, this group maintained processing speed gains and revealed new improvements in sustained and selective attention. Considering emotional and quality of life outcomes, Reh@City v2.0 contributed to lower levels of depressive symptomatology and the TG resulted in improvements in social relationships and environmental aspects of quality of life at post-intervention and

follow-up. When comparing both interventions, we have found that the desktop VR-based CT intervention consistently led to a higher reduction in depressive symptomatology, observed at both post-intervention and follow-up, while the adaptive paper-and-pencil CT intervention was superior in social relationships aspects of quality of life, processing speed and abstraction at follow-up. These results are quite important in the context of mental and behavioral disorders, once these conditions often result in cognitive and functional impairment that substantially affect mood and quality of life. Improving these two last outcomes through CT methods emphasizes the indissociable relationship between cognition and emotional domains, indicating that CT has a broader impact in different aspects of human functioning.

Future research in this field should take into account the aforementioned limitations and further explore the influence of desktop VR-based CT methods in mood symptoms (e.g., depression and anxiety), since these are a common clinical feature of several mental and behavioral disorders, known to aggravate cognitive deficits and, therefore, compromising quality of life and functional abilities. Also, given the differential benefits of both interventions, it is important to investigate the combined use of desktop VR and conventional CT methods in cognitive and non-cognitive outcomes, due to their potential complementary nature and potential cumulative benefits.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10055-021-00559-6>.

Acknowledgements This work was supported by the Fundação para a Ciência e Tecnologia (FCT) through the Belief Revision applied to Neurorehabilitation (BRaNT) project (PTDC/CCICOM/30990/2017), the Development and clinical validation of an adaptive cognitive training virtual tool for stroke patients project (SFRH/BD/145919/2019), which is a PhD Grant awarded to Joana Câmara, and the MACBIOIDI projects (MAC/1.1.b/098), and by the NOVA Laboratory of Computer Science and Informatics (UID/CEC/04516/2019).

Declarations

Conflict of interest The authors have no conflict of interest to disclose.

References

- American Psychological Association (2010) Global Mental Health. <https://www.psychiatry.org/psychiatrists/international/global-mental-health>. Accessed 24 April 2020
- Beck AT, Steer RA, Brown GK (1996) Manual for the beck depression inventory-II. Psychological Corporation, San Antonio
- Canavarro MC, Serra AV, Simões MS, Pereira M, Gameiro S, Quartilho MJ, Rijo D, Carona C, Paredes T (2006) Estudos psicométricos do instrumento de avaliação da qualidade de vida da Organização Mundial de Saúde (WHOQOL-Bref) para Português de Portugal. *Psiquiatr Clin* 27:41–49
- Cavaco S, Gonçalves A, Pinto C, Almeida E, Gomes F, Moreira I, Fernandes J, Teixeira-Pinto A (2013) Semantic fluency and phonemic fluency: regression-based norms for the portuguese population. *Arch Clin Neuropsychol* 28:262–271. <https://doi.org/10.1093/arclin/act001>
- Chan CLF, Ngai EKY, Leung PKH, Wong S (2010) Effect of the adapted virtual reality cognitive training program among Chinese older adults with chronic schizophrenia: a pilot study. *Int J Geriatr Psychiatry* 25:643–649. <https://doi.org/10.1002/gps.2403>
- Cieślak B, Mazurek J, Rutkowski S, Kiper P, Turolla A, Szczepańska-Gieracha, (2020) Virtual reality in psychiatric disorders: a systematic review of reviews. *Complement Ther Med* 52:1–15. <https://doi.org/10.1016/j.ctim.2020.102480>
- Cohen J (1988) Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates, New York
- Coyle H, Traynor V, Solowij N (2015) Computerized and virtual reality cognitive training for individuals at high risk of cognitive decline: systematic review of the literature. *Am J Geriatr Psychiatry* 23:335–359. <https://doi.org/10.1016/j.jagp.2014.04.009>
- Faria AL, Bermúdez i Badia S (2015) Development and evaluation of a web-based cognitive task generator for personalized cognitive training: a proof of concept study with stroke patients. In: Proceedings of the 3rd 2015 workshop on ICTs for improving patients rehabilitation research techniques - REHAB '15. Presented at the 3rd 2015 Workshop, ACM Press, Lisbon, Portugal, pp. 1–4
- Faria AL, Andrade A, Soares L, Bermúdez i Badia S, (2016) Benefits of virtual reality based cognitive rehabilitation through simulated activities of daily living: a randomized controlled trial with stroke patients. *J NeuroEngineering Rehabil* 13:96. <https://doi.org/10.1186/s12984-016-0204-z>
- Faria AL, Pinho MS, Bermúdez i Badia S, (2018) Capturing expert knowledge for the personalization of cognitive rehabilitation: study combining computational modeling and a participatory design strategy. *JMIR Rehabil Assist Technol* 5:e10714. <https://doi.org/10.2196/10714>
- Faria AL, Paulino T, Bermúdez i Badia S (2019) Comparing adaptive cognitive training in virtual reality and paper-pencil in a sample of stroke patients, in: 2019 international conference on virtual rehabilitation (ICVR). Presented at the 2019 international conference on virtual rehabilitation (ICVR), IEEE, Tel Aviv, Israel, pp. 1–7. doi: doi: <https://doi.org/10.1109/ICVR46560.2019.8994746>
- Faria AL, Pinho MS, Bermúdez i Badia S (2020). A comparison of two personalization and adaptive cognitive rehabilitation approaches: a randomized controlled trial with chronic stroke patients. Accepted for publication. *J Neuroeng Rehabil*<https://doi.org/10.1186/s12984-020-00691-5>
- Freitas S, Simões MR, Alves L, Santana I (2011) Montreal cognitive assessment (MoCA): normative study for the Portuguese population. *J Clin Exp Neuropsychol* 33:989–996. <https://doi.org/10.1080/13803395.2011.589374>
- Gamito P, Oliveira J, Coelho C, Morais D, Lopes P, Pacheco J, Brito R, Soares F, Santos N, Barata AF (2015) Cognitive training on stroke patients via virtual reality-based serious games. *Disabil Rehabil* 39:385–388. <https://doi.org/10.3109/09638288.2014.934925>
- Green MF (2006) Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. *J Clin Psychiatry* 67:36–42
- Groves SJ, Douglas KM, Porter RJ (2018) A Systematic review of cognitive predictors of treatment outcome in major depression. *Front Psychiatry* 9:382. <https://doi.org/10.3389/fpsy.2018.00382>
- Grynszpan O, Perbal S, Pelissolo A, Fossati P, Jouvent R, Dubal S, Perez-Diaz F (2011) Efficacy and specificity of computer-assisted cognitive remediation in schizophrenia: a meta-analytical study. *Psychol Med* 41:163–173. <https://doi.org/10.1017/S0033291710000607>

- Harvey PD (2011) Mood symptoms, cognition, and everyday functioning: in major depression, bipolar disorder, and schizophrenia. *Innov Clin Neurosci* 8:14–18
- Iosifescu DV (2012) The relation between mood, cognition and psychosocial functioning in psychiatric disorders. *Eur Neuropsychopharmacol* 22:S499–S504. <https://doi.org/10.1016/j.euroneuro.2012.08.002>
- Kim BR, Chun MH, Kim LS, Park JY (2011) Effect of virtual reality on cognition in stroke patients. *Ann Rehabil Med* 35:450. <https://doi.org/10.5535/arm.2011.35.4.450>
- Kim EJ, Bahk YC, Oh H, Lee WH, Lee JS, Choi KH (2018) Current status of cognitive remediation for psychiatric disorders: a review. *Front Psychiatry* 9:461. <https://doi.org/10.3389/fpsy.2018.00461>
- Maggio MG, Latella D, Maresca G, Sciarone F, Manuli A, Naro A, De Luca R, Calabrò RS (2019) Virtual reality and cognitive rehabilitation in people with stroke: an overview. *J Neurosci Nurs* 51:101–105. <https://doi.org/10.1097/JNN.0000000000000423>
- Maier M, Ballester BR, Leiva Bañuelos N, Duarte Oller E, Verschure PFMJ (2020) Adaptive conjunctive cognitive training (ACCT) in virtual reality for chronic stroke patients: a randomized controlled pilot trial. *J NeuroEngineering Rehabil* 17:42. <https://doi.org/10.1186/s12984-020-0652-3>
- Marques A, Queirós C, Rocha N (2008) Virtual reality and neuropsychology: a cognitive rehabilitation approach for people with psychiatric disabilities. In: Proceedings of the 7th ICDVRAT with ArtAhabilitation, Maia, Portugal, pp. 39–46.
- McCleery A, Nuechterlein K (2019) Cognitive impairment in psychotic illness: prevalence, profile of impairment, developmental course, and treatment considerations. *Dialogues Clin Neurosci* 21:239–248. <https://doi.org/10.31887/DCNS.2019.21.3/amccleery>
- McIntyre RS, Cha DS, Soczynska JK, Woldeyohannes HO, Gallagher LA, Kudlow P, Alsuwaidan M, Baskaran A (2013) Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions: Review: Cognitive Deficits and Functional Outcomes in MDD. *Depress Anxiety* 30:515–527. <https://doi.org/10.1002/da.22063>
- Millan MJ, Agid Y, Brüne M, Bullmore ET, Carter CS, Clayton NS, ..., Young LJ (2012) Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. *Nat Rev Drug Discovery* 11(2):141–168. <https://doi.org/10.1038/nrd3628>
- Motter JN, Pimontel MA, Rindskopf D, Devanand DP, Doraiswamy PM, Sneed JR (2016) Computerized cognitive training and functional recovery in major depressive disorder: A meta-analysis. *J Affect Disord* 189:184–191. <https://doi.org/10.1016/j.jad.2015.09.022>
- Oliveira J, Gamito P, Lopes B, Silva AR, Galhordas J, Pereira E, Ramos E, Silva AP, Jorge Á, Fantasia A (2020) Computerized cognitive training using virtual reality on everyday life activities for patients recovering from stroke. *Disabil Rehabil Assist Technol*. <https://doi.org/10.1080/17483107.2020.1749891>
- Ormel J, Petukhova M, Chatterji S, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Bromet EJ, Burger H, Demyttenaere K, de Girolamo G, Haro JM, Hwang I, Karam E, Kawakami N, Lépine JP, Medina-Mora ME, Posada-Villa J, Sampson N, Scott K, Üstün TB, Von Korff M, Williams DR, Zhang M, Kessler RC (2008) Disability and treatment of specific mental and physical disorders across the world. *Br J Psychiatry* 192:368–375. <https://doi.org/10.1192/bjp.bp.107.039107>
- Park MJ, Kim DJ, Lee U, Na EJ, Jeon HJ (2019) A literature overview of virtual reality (VR) in treatment of psychiatric disorders: recent advances and limitations. *Front Psychiatry* 10:1–9. <https://doi.org/10.3389/fpsy.2019.00505>
- Parsons TD (2016) *Clinical neuropsychology and technology: what's new and how can we use it*. Springer International Publishing, Switzerland
- Paulino T, Faria AL, Bermudez i Badia S (2019) Reh@City v2.0: a comprehensive virtual reality cognitive training system based on personalized and adaptive simulations of activities of daily living. In: 2019 5th experiment international conference (Exp.at'19). Presented at the 2019 5th experiment conference (exp.at'19), IEEE, Funchal (Madeira Island), Portugal, pp. 292–297. doi: <https://doi.org/10.1109/EXPAT.2019.8876539>
- Plagia F, Cascia C, Rizzo R, Sideli L, Barbera AFD (2013) Cognitive rehabilitation through Neurovr training. *Annu Rev Cyber Therapy Telemed*. <https://doi.org/10.3233/978-1-61499-282-0-110>
- Prikken M, Konings MJ, Lei WU, Begemann MJH, Sommer IEC (2019) The efficacy of computerized cognitive drill and practice training for patients with a schizophrenia-spectrum disorder: A meta-analysis. *Schizophr Res* 204:368–374. <https://doi.org/10.1016/j.schres.2018.07.034>
- Rand D, Katz N, Shahar M, Kizony R, Weiss PL (2005) The virtual mall: A functional virtual environment for stroke rehabilitation. *Annu Rev CyberTherapy Telemed* 3:193–198
- Reichenberg A, Harvey PD, Bowie CR, Mojtabai R, Rabinowitz J, Heaton RK, Bromet E (2009) Neuropsychological function and dysfunction in schizophrenia and psychotic affective disorders. *Schizophr Bull* 35:1022–1029. <https://doi.org/10.1093/schbul/sbn044>
- Rey A (1998) *Teste de cópia de figuras complexas*. Hogrefe Editora, Lisboa
- Ritchie H, Roser M (2018) *Mental Health*. <https://ourworldindata.org/mental-health>. Accessed 24 April 2020
- Robinson LJ, Thompson JM, Gallagher P, Goswami U, Young AH, Ferrier IN, Moore PB (2006) A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *J Affect Disord* 93(1–3):105–115. <https://doi.org/10.1016/j.jad.2006.02.016>
- Rock PL, Roiser JP, Riedel WJ, Blackwell AD (2014) Cognitive impairment in depression: a systematic review and meta-analysis. *Psychol Med* 44:2029–2040. <https://doi.org/10.1017/S0033291713002535>
- Toulouse E, Piéron, H (1986) *Toulouse-Piéron. Test of Perception and Attention*. TEA, Madrid
- Wechsler D (1997a) *Wechsler Adult Intelligence Scale (WAIS)-III administration and scoring manual*. San Antonio, TX, The Psychological Corporation
- Wechsler D (1997b) *Wechsler Memory Scale (WMS)-III administration and scoring manual*. San Antonio, TX, The Psychological Corporation
- Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, Olesen J, Allgulander C, Alonso J, Faravelli C, Fratiglioni L, Jennun P, Lieb R, Maercker A, van Os J, Preisig M, Salvador-Carulla L, Simon R, Steinhausen HC (2011) The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* 21:655–679. <https://doi.org/10.1016/j.euroneuro.2011.07.018>
- World Health Organization (WHO) (2010) *International statistical classification of diseases and related health problems (10th revision)*. WHO Press, Switzerland
- Yang S, Chun MH, Son YR (2014) Effect of virtual reality on cognitive dysfunction in patients with brain tumor. *Ann Rehabil Med* 38:726. <https://doi.org/10.5535/arm.2014.38.6.726>
- Zajac-Lamparska L, Wilkość-Dębczyńska M, Wojciechowski A, Podhorecka M, Polak-Szabela A, Warchoń Ł, Kędziora-Kornatowska K, Araszkievicz A, Izdebski P (2019) Effects of virtual reality-based cognitive training in older adults living without and with mild dementia: a pretest–posttest design pilot study. *BMC Res Notes* 12:776. <https://doi.org/10.1186/s13104-019-4810-2>

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