Efficacy and Moderators of Virtual Reality for Cognitive Training in People with Dementia and Mild Cognitive Impairment: A Systematic Review and Meta-Analysis

Themis **Papaioannou**[a], Alexandra **Voinescu**[a], Karin **Petrini**[a,b], Danaë **Stanton Fraser**[a]

[a]*Department of Psychology, University of Bath, Claverton Down, Bath BA2 7AY, UK*

[b]*Centre for the Analysis of Motion, Entertainment Research and Applications, University of Bath, Claverton Down, Bath BA2 7AY, UK*

Running title*:* Virtual reality for MCI and dementia

Corresponding author Dr Alexandra Voinescu,
Department of Psychology, University of Bath
Claverton Down, Bath BA2 7AY, UK
01225 386582
a.voinescu@bath.ac.uk

ABSTRACT

Background. Mild cognitive impairment (MCI) and dementia result in cognitive decline which can negatively impact everyday functional abilities and quality of life. Virtual reality (VR) interventions could benefit the cognitive abilities of people with MCI and dementia, but evidence is inconclusive. ***Objective.*** To investigate the efficacy of VR training on global and domain-specific cognition, activities of daily living and quality of life. To explore the influence of priori moderators (e.g., immersion type, training type) on the effects of VR training. Adverse effects of VR training were also considered. ***Methods.*** A systematic literature search was conducted on all major databases for randomised control trial studies. Two separate meta-analyses were performed on studies with people with MCI and dementia. ***Results.*** Sixteen studies with people with MCI and four studies with people with dementia were included in each meta-analysis. Results showed moderate to large effects of VR training on global cognition, attention, memory, and construction and motor performance in people with MCI. Immersion and training type were found to be significant moderators of the effect of VR training on global cognition. For people with dementia, results showed moderate to large improvements after VR training on global cognition, memory, and executive function, but a subgroup analysis was not possible. ***Conclusion.*** Our findings suggest that VR training is an effective treatment for both people with MCI and dementia. These results contribute to the establishment of practical guidelines for VR interventions for patients with cognitive decline.

KEYWORDS

Virtual reality; Mild cognitive impairment; Cognitive training; Cognitive rehabilitation; Cognition; Dementia

INTRODUCTION

Dementia is caused by a range of neurological disorders and is characterised by progressive deterioration in cognitive and functional abilities. Current estimates of global prevalence of dementia are 50 million people worldwide while this number is projected to triple by 2050 [1]. MCI is considered the stage between normal cognition in older age and dementia and is also characterised by cognitive deficits, albeit without major impairments in functional abilities [2]. Current estimates of MCI prevalence are around 15-20% in people aged over 60 [3]. MCI is considered a risk factor for developing dementia while estimates of progression rates are around 5-10% [4]. Both conditions can have a negative impact on psychological well-being, activities of daily living (ADL) and quality of life (QOL) [5, 6].

Cognitive training aims to improve performance of specific-domain cognition, such as memory and attention, through repetition of tasks [7]. Cognitive training is typically administered with classic methods, such as paper and pencil or computerised tasks, although these methods have questionable or limited ecological validity [8]. Ecological validity refers to the extent the training is taking place in realistic settings and trains cognitive abilities needed in every-day life. For example, tasks such as performing arithmetic equations within a specified time have low ecological validity, as they are not tasks normally done daily [9]. VR on the other hand, allows interaction with realistic three-dimensional environments which can lead to more complex experiences and offer more ecological cognitive training compared to classic methods [8, 10]. Additionally, VR systems offer the option to train in a simulated environment without the risks that accompany the same task in a real-life setting (e.g., driving a car). Furthermore, with the advancements of the technology, VR systems have become more accessible and affordable, making them promising tools for cognitive training in people with MCI and dementia [11].

Complexity of the simulated environments of VR systems, might assist in slowing down cognitive decline in people with MCI and dementia [12]. Interaction with enriched environments is shown

to enhance neuroplasticity in Alzheimer's disease brains in transgenic mice models [13, 14]. Meta-analytical studies point towards the effectiveness of VR training on people with MCI and dementia [15-19]. Kim et al. [15] investigated the efficacy of VR training on global cognition in people with MCI and dementia and found that overall VR had a significant medium effect. Wu et al. [16] looked at the efficacy of VR training on specific domain cognitive functions such as executive function, short-term and long-term memory, as well as global cognition, when compared to more conventional cognitive training methods (e.g., computerised cognitive training, paper and pencil) in people with MCI. They reported superiority of VR-based interventions on global cognition and executive function, but not on short-term and long-term memory. A more recent meta-analysis by Zhong et al. [19] reported similar results. In contrast, a meta-analytical study by Hung et al. [17] found better effects of VR training on memory but not on global cognition and executive function, and benefits of VR training on language and visuospatial skills, which was not covered in the study by Wu et al. [16]. Effects on memory after VR training were also reported by Zhu et al. [18], who additionally reported significant effects of VR training on global cognition, attention and executive function.

These somehow contrasting results could be explained by differences in characteristics of the included studies. For example, some of the studies included in the meta-analyses used immersive technologies such as Head-Mounted Displays (HMD) and Connected Automated Virtual Environments (CAVE), whereas others have used less immersive technologies, such as desktop monitors. Evidence suggests that more immersive technologies might result in better cognitive performance in healthy individuals, for example in memory [20-22], and attention and executive function [23], although evidence is still inconclusive (e.g., [24]). Immersion refers to the technological capabilities of the VR system that increase its fidelity, e.g., stereoscopic view, motion tracking, field of view, so that the user experiences the virtual environment in a somewhat similar way to a real-world environment. Additionally, immersion is interlinked with

presence, the psychological experience of being transported into the virtual environment [25], in that more immersive systems tend to elicit more presence, although presence can be experienced also in non-immersive systems, albeit to a lesser extent [26]. Furthermore, presence appears to be the mediating factor for enhancing cognitive performance in healthy individuals [23, 27]. Therefore, differences of the effects found between the two meta-analyses could be attributed to the overall moderating effects of immersion or presence of the studies included in each meta-analysis.

Furthermore, VR studies are heterogenous not only in terms of the apparatus used (e.g., immersive, or non-immersive technology), but also in terms of the nature of the VR training programmes. For example, some applications involve activities that simulate real life scenarios, such as cooking, while others may involve exergames (i.e., video games that require body-movement interaction and physical activity) [28, 29]. Previous research has shown that different types of training can affect different cognitive domains as some platforms may stimulate one cognitive domain more than another [30]. Therefore, the type of training (e.g., exergame, simulation, game etc.) may also moderate the effects of VR training.

Other factors related to the type of VR system have also been shown to influence cognitive performance. For example, more embodied systems, such as those where users view the environment from a first-person perspective (compared to a third-person perspective) or include bodily cues (e.g., virtual hands that are visible when users move their hand in the field of view), seem to result in better memory performance [31, 32]. Bespoke VR training has also been shown to have better cognitive outcomes compared to commercial VR training [33]. Additionally, effects of VR training might differ when compared to different types of control group conditions. For example, cognitive training and other cognitive-based interventions seem to be more effective when compared to passive controls in which participants receive no treatment or their treatment is usual care, but not when compared to active controls where participants receive an

intervention that does not target cognition specifically, such as art discussions and physical activity [34, 35]. Furthermore, computerised cognitive training has been found to have greater effects on cognition when compared to other methods [36, 37], so perhaps VR training has different effects when compared to controls that used computerised cognitive training or more conventional methods (e.g., paper and pencil). Different types of VR intervention such as training or rehabilitation may also influence the effects, as they are based on different theoretical and conceptual assumptions [7]. Cognitive training refers to the repeated practice of standardised tasks that target specific cognitive domains [38], and is based on the assumption that regular practice of a specific cognitive skill will improve or maintain function of that domain which will then generalise to other functions [7]. On the other hand, cognitive rehabilitation is tailored to individual needs and aims to improve everyday cognitive function by building on the strengths of the individual to compensate for the cognitive deficits [7]. Furthermore, individual differences such as education is believed to play a role in the progression of cognitive decline in people with dementia. For example, although people with higher education are less prone to develop dementia and MCI, when they do develop these conditions, the cognitive decline is greater than for people with lower level of education [39, 40], and therefore cognitive outcomes after VR training may be influenced by education level. Finally, duration of the intervention may also increase the effects of cognitive training [41].

Although the aforementioned factors might influence the effects of VR training on cognitive performance, no study yet has investigated their role in people with MCI and dementia. Furthermore, the sole effects of VR training have not been investigated yet, as previous meta-analyses contained studies with mixed methods in the intervention group [16, 17]. For example, a study included in Hung et al. [17] and Wu et al. [16], used both auditory and VR training in the intervention group for 6 months [42]. In the initial phase, three auditory sessions were administered alternating with three VR sessions every 2 weeks and the cycle was repeated

every 2 weeks. After 3 months, one auditory and one VR session were administered every week. In the auditory training, participants were blindfolded and listened via headphones to stories accompanied by musical backgrounds. The same musical background was used in the VR training. Previous research has shown that listening to narratives improves memory performance [43], so auditory sessions could have benefitted participants' cognitive abilities, thus making it difficult to discriminate the sole effect of the VR training. Additionally, most meta-analyses [16, 18, 19] included studies that contained 2D graphics (they lack depth information, e.g., video games such as Tetris and Super Mario) (e.g., [44-46]), which does not fall under a strict definition of VR [8, 10].

Furthermore, because VR simulates real-life environments, it may offer greater potential for transfer of training to the ADL and QOL, compared to more traditional cognitive training methods [47]. Nevertheless, a recent meta-analysis on studies with people with MCI does not support this assumption [19]. It is possible people with MCI do not experience as severe impairments in the ADL as people with dementia [6], and therefore benefits on the ADL may be significant in people with dementia but not with MCI. However, no meta-analysis to date has examined the effects of VR training on the ADL of people with dementia. Additionally, no meta-analysis to date has investigated the effects of VR training on QOL, even though previous research suggests that VR training may improve QOL in people with MCI and dementia [48].

Finally, no study yet has pulled together the adverse effects of VR training. Exposure to immersive VR is associated with cybersickness, an adverse effect similar to motion sickness that results in discomfort and feeling unwell [49], which might result in less acceptability of VR training in people with MCI and dementia [50].

Thus, the aims of this study are to combine the evidence from randomised and quasi-randomised control trials in order to investigate a) the efficacy of sole VR training on global and domain-specific cognition in people with dementia and MCI compared to more conventional

methods such as computerised cognitive training and paper and pencil, b) the effects of different moderators such as immersion and presence on cognitive outcomes, c) the adverse effects and acceptability of VR training, and d) the efficacy of VR training on ADL and QOL in people with MCI and dementia.

METHOD

This meta-analysis was registered with PROSPERO (CRD42020178679) and follows the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines [51].

Information Sources and Study Selection

A systematic search was conducted on all major databases (PubMed, Embase, Web of Science, LILACS, ALOIS, Cochrane Central Register of Controlled Trials (CENTRAL), Scopus, PsycNET, ACM, IEEE, CINAHL, PEDro) and grey literature databases (Ethos Thesis, ProQuest dissertations and theses, Open access theses and dissertations and OpenGrey) for randomised control trials studying the effect of interventions using VR on one or more cognitive outcomes in people with dementia or MCI (the research strategy can be found in Appendix E of this article). There were no language and date restrictions. Furthermore, additional articles were obtained from a search on the reference list of other reviews. Two reviewers conducted an eligibility screening independently, initially based on title only, then based on title and abstract, and finally on full text. If the eligibility of a study was unclear, the authors of the study were contacted for clarification. Finally, the studies selected by each reviewer were combined in one list and any disagreement was resolved by a third reviewer.

Eligibility Criteria

Participants

Participants included adults of any age, background, and gender, diagnosed either with dementia (any subtype and severity) or MCI. For dementia, only studies that clearly reported use of internationally recognised criteria for diagnosis were included, (e.g., DSM 5 [52] or ICD 11

[53]). For MCI, studies were included if authors have reported a valid method for diagnosis (rigorous criteria) (e.g., Petersen criteria [54]), but also broader diagnosis criteria (broad criteria) (e.g., MoCA, ACE-R and MMSE scores cut-off scores [55]). Because we included both broad and rigorous criteria for MCI, we decided to include the MCI criteria (rigorous vs broad classification) as an extra moderator for the meta-analysis of studies on MCI. Studies with mixed cohorts were included only if data of relevant participants could be obtained for each group.

Interventions

Any 3D VR application which was used for cognitive training or cognitive rehabilitation was included. These could be custom or commercial VR applications, video games or exergames (video games that involve some form of physical activity). Both immersive (e.g., using HMD) and non-immersive (e.g., desktop screen) applications were included. Interventions that combined VR with some other method (e.g., music therapy or computerised cognitive training) were excluded.

Comparisons

We included active control groups that used any methods except VR (e.g., paper and pencil cognitive training, music therapy, computerised cognitive training, psychoeducation, multisensory stimuli) or passive control groups (no-contact or wait-list).

Outcomes

Primary outcomes were the change from baseline to post-test scores of global cognition and specific-domain cognition measures. For global cognition we included two outcomes: the combined scores of individual domain outcomes (i.e., combined measures of memory, attention, executive function, construction and motor performance, and verbal function and language), which we refer to as *combined cognitive functioning outcomes,* and measures of general cognition (e.g., MMSE and MoCA). Specific-domain cognition outcomes were memory, attention, executive function, verbal function and language, construction, and motor performance. Only

outcomes measured with a validated battery were included. Secondary outcomes were ADL (e.g., Assessment of Activities of Daily Living Scale [56]), self-efficacy (e.g., Self-efficacy Scale [57]), QOL (e.g., The Cornell-Brown Scale [58]), acceptability and treatment adherence (number of dropouts) and reports of adverse effects (e.g., cybersickness measures [59]).

Studies

We included randomised and quasi-randomised control trials following the Cochrane criteria [60].

Data Collection and Coding

Extracted continuous data were recorded as mean and standard deviations for each group (interventions and control) at baseline, post-test and follow up. If those data were not available, additional information was asked from the authors of the study. Coding of outcomes into cognitive domains was according to classification of assessment scales found in Lezak et al. [61] (the list of the scales and their categorisation into cognitive domains is included in Appendix B). If a study reported multiple measures for the same outcome (e.g., reporting both MoCA and MMSE for global cognition), a composite score was calculated for this outcome. To do that, we calculated the z scores for each measure and calculated the mean and standard deviation for each outcome [35]. For dichotomous outcomes we extracted the number of participants in each outcome group. Attrition rates were also included (dropouts in post-test, dropouts in follow-up tests).

Risk of Bias in Individual Studies and Publication bias

The Cochrane Collaboration's risk of bias tool was used to assess risk of bias in each study [60]. The studies were assessed independently by two authors as low risk, high risk or unclear risk against five domains. Disagreements were resolved by a discussion with a third reviewer. These domains are randomisation process, deviations from intended interventions, missing outcome

data, measurement of the outcome, and selection of the reported outcome. Funnel plots were used to assess publication bias [60].

Data Analysis

Analysis was performed separately for dementia and MCI, for both primary and secondary outcomes. When studies had more than one outcome for a single cognitive domain, the standardised mean difference and variance were combined into one estimate. Standardized mean differences were calculated as Hedges' *g*, as it gives a less biased effect size than Cohen's d [62], and 95% confidence interval of change between the change in outcome measures between VR and control groups from baseline to post-test and follow up. Hedges' *g* estimates of < 0.20 were considered very small, between 0.20 and 0.50 small, between 0.50 and 0.80 medium, and > 0.80 large [63]. For dichotomous outcomes (e.g., participant adherence), we expressed effects as risk ratios (RRs) along with 95% CIs. A random effects model was used to calculate all pooled effects. Statistical heterogeneity was assessed by visual inspection of forest plots and the I^2 statistic which described heterogeneity as low on 25%, moderate on 50% and large on 75% [64]. Sensitivity analysis was performed based on risk of bias ratings [65] to show any effect low-quality studies had on the outcome. We categorised the studies with low risk of bias as being of high quality, whereas studies with concerning or high risk of bias were considered low quality. To examine the effect of individual study characteristics on the outcomes a subgroup analysis of the following moderators was conducted, if the number of studies was above 10 [66, 67]: type of intervention (i.e., rehabilitation vs. training); type of application (immersive vs non-immersive); presence; first- or third-person perspective; type of control group (active, such as music therapy and psychoeducation, vs. passive, such as no-contact or wait-list); type of active control group (computerised tasks vs pen-and-paper), participant background (education) and duration of the intervention. The type of training was also considered in a

separate post hoc exploratory analysis, as it was not possible to predict it a priori. Data analysis was performed using Comprehensive Meta-analysis version 3 [68].

The GRADE framework was used to classify the certainty of evidence [69]. The certainty of evidence is downgraded by one or two levels across five domains: risk of bias, unexplained heterogeneity, indirectness of evidence, imprecision of results and publication bias (see below). Assessment is carried out per cognitive domain and each domain can be rated with one out of four possible outcomes: high level of certainty, which indicates that the true effect is most likely similar to the estimated effect, moderate level of certainty, which indicates the true effect is probably close to the estimated effect; low level of certainty, which indicates the true effect might be significantly different to the estimated effect; and very low level of certainty, which indicates the true effect is probably significantly different to the estimated effect. Downgrading of evidence was performed independently by two reviewers and was based on the GRADE handbook. The methodology for downgrading the evidence is explained in more detail on Appendix C.

RESULTS

The search was conducted in August 2021, which covered all the years up to that point without date restriction and provided 20389 records. After removing duplicates, 16584 articles were screened based on titles and abstracts, of which 485 were assessed in full text. In total, 20 studies were eligible for inclusion, 16 of which concerned people with MCI and four people with dementia (Figure 1). Four articles [70-73] reported outcomes collected from two separate studies, and therefore, outcomes of those four articles were combined under two studies in our review. Two studies did not report means and standard deviations; however, we requested and received the data from the authors [74, 75]

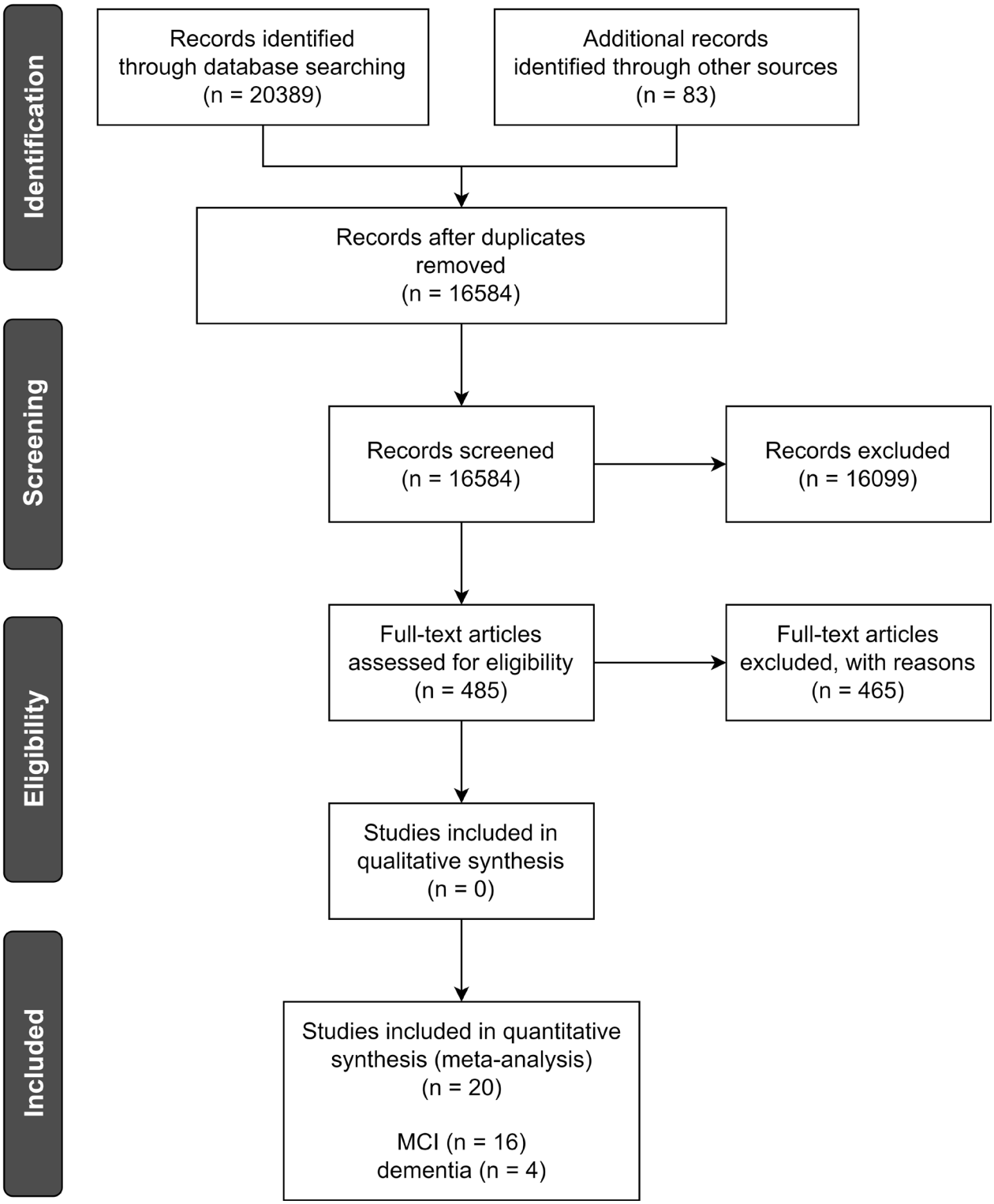

Figure 1. PRISMA flow diagram of first systematic search

Characteristics of Included Studies

Mild cognitive impairment

The 16 included studies consisted of 742 participants at baseline, with a mean age of 72.22 (*SD* = 4.97). The VR group consisted of 352 participants at baseline with a mean age of 72.01 (*SD* = 4.97) and a control group of 390 participants at baseline with mean age 72.43 (*SD* = 4.49). The active control group (i.e., groups that involved interventions other than VR, such as music therapy and psychoeducation) consisted of 294 participants at baseline with mean age 73.916 (*SD* = 5.31) and the passive control group (i.e., no-contact or wait-list groups) of 96 participants at baseline with mean age 70.08 (*SD* = 2.95). Eight studies reported rigorous criteria for the assessment of MCI (e.g., Petersen criteria [76]) [70, 71, 77-83], while the remaining eight used either broader criteria (e.g., MMSE scores) or did not report in detail all the criteria [74, 75, 84-89]. Four studies used immersive VR in the experimental condition [75, 80, 86, 89] whereas 10 used non-immersive VR [74, 77-79, 81-83, 85, 87, 88], one mixed (experimental condition included both immersive and non-immersive VR applications) [70, 71], and one did not report this information [84]. Ten studies used a first-person perspective in VR [70, 71, 75, 77-80, 86-89] and three a third-person perspective [74, 82, 85], while three did not report this information [81, 83, 84]. The applications used in the experimental condition were six bespoke [75, 78, 80, 81, 86, 87], eight commercial [74, 77, 79, 82, 83, 85, 88, 89] and one mixed (had both a commercial and a bespoke application) [70, 71], while one did not report this information [84]. All studies used cognitive training, and none reported cognitive rehabilitation. In the control group, 13 studies had an active control group and three had a passive control group (Table 1). In the active control group, four studies had computerised cognitive training [75, 77, 78, 86], eight had a non-computerised intervention, such as occupational therapy and cognitive stimulation therapy [70, 71, 74, 82, 84, 85, 87-89], and two did not report this information [81, 83]. In terms of the type of training, five studies contained an exergame [74, 77, 82, 83, 85], two contained a game [79, 81], four contained a simulation [78, 86-88], and four contained a mixed type (multiple types

of VR training) [70, 71, 75, 80, 89]. Duration of the interventions in the experimental and active control groups ranged between six and 24 weeks, and the total number of treatment sessions ranging between 10 and 40, while each session lasted between 25 and 100 minutes (Table 5, Appendix D).

Dementia

The four included studies consisted of 163 participants in total at baseline, with a mean age of 82.01 (*SD* = 5.90). The VR group consisted of 83 participants with mean age of 81.79 (*SD* = 5.65) and the active control of 79 participants with mean age of 82.32 (*SD* = 6.47). One study used a passive control [90] while the remaining three studies used an active control group [72, 91, 92]. All studies used cognitive training and non-immersive VR applications as intervention. Two studies used a first-person perspective in VR [72, 90], one used a third-person perspective [91], and one did not report this information [92]. Two studies used a bespoke VR application [90, 91], one study used commercial [72], and one did not report this information [92]. All the studies that had an active control group used non-computerised cognitive training [72, 91, 92]. In terms of the type of training, two studies contained an exergame [72, 92], one contained a game [91] and one a simulation [90]. Duration of the interventions in the experimental and active control groups ranged between six and 24 weeks, and the total number of treatment sessions ranging between 10 and 40, while each session lasted between 15 and 100 minutes (Table 5, Appendix D).

Risk of bias in included studies

One study reported that allocation sequence was concealed from participants or personnel before and until the assignment of participants to intervention or control groups [88]. Blinding of personnel delivering the intervention is typically not possible in VR training, as the sham group is not identical to the intervention group which makes it easy for personnel to guess the intervention group [93]. However, outcome assessors can be blinded to group allocation. Nine

studies reported assessors to be blinded to group allocation [70-72, 75, 77, 81, 83, 87, 91, 92]. Tarnanas et al. [78] reported double blinding but did not provide information of how this was achieved. The remaining studies [74, 80, 84-86, 88, 89] did not provide information, except Ramnath et al. [82] who reported that assessors were not blinded. Van de Weijer et al. [79] reported that, contrary to what was stated in their procedure protocol, assessors were not blinded. Only five studies mentioned intention-to-treat analysis [72, 75, 79, 82, 89]. The remaining fifteen studies did not report a methodology for dealing with missing data. Twelve studies reported dropout numbers for both intervention and control groups, ten of which gave reasons [70-72, 74, 75, 78, 79, 83, 89, 92], while the remaining two did not [88, 90]. From the description of the reasons that were given it was concluded that dropout rates did not influence the outcomes [94]. One study reported that there were no dropouts [77] and the remaining six did not report information about this [80, 84-87, 91]. Figure 2 shows the assessed risk of bias for each study.

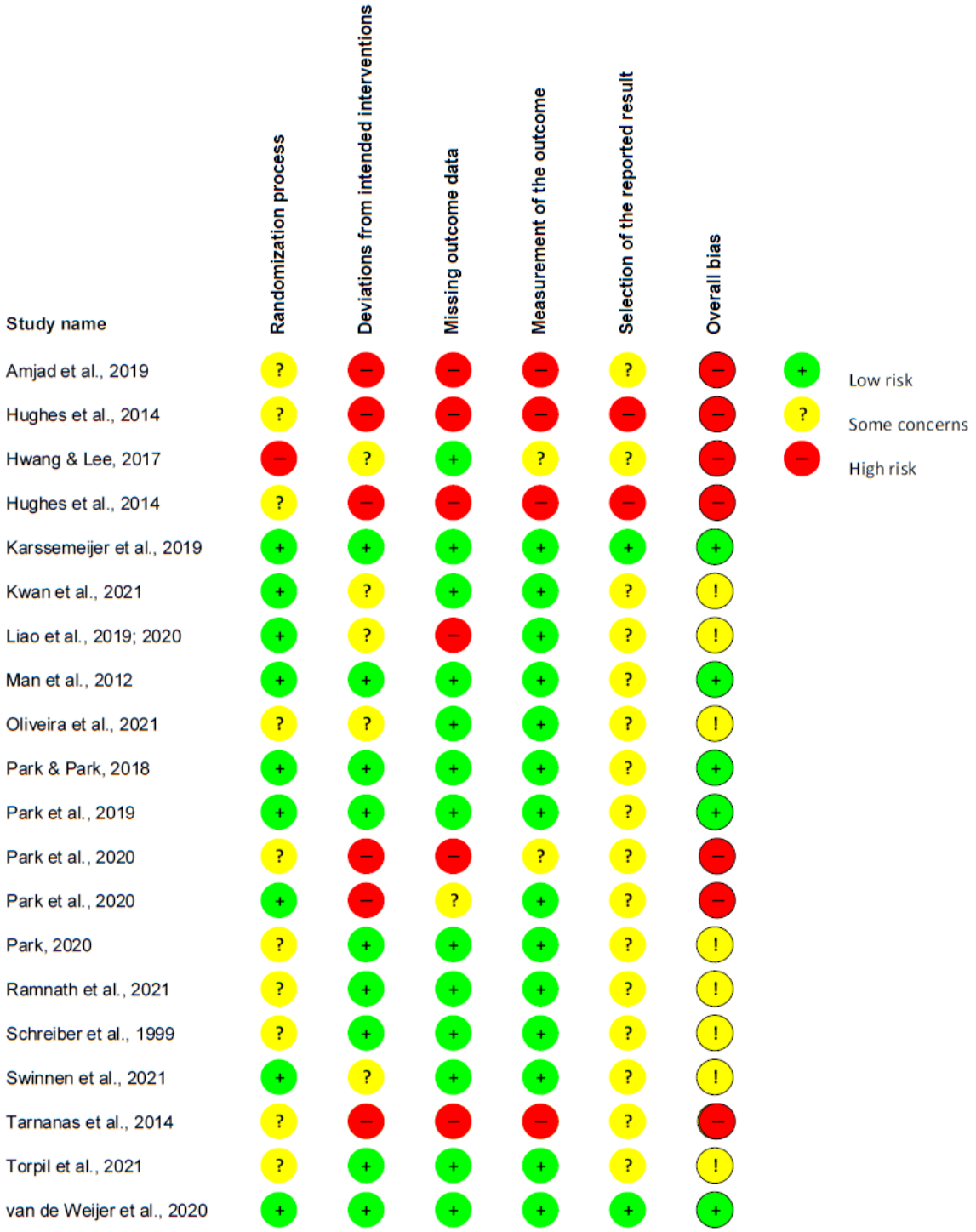

Figure 2. Assessed risk of bias for all studies.

Effects of VR in Mild Cognitive Impairment Outcomes (post-intervention)

Combined cognitive functioning outcomes

Based on the 16 studies included in the meta-analysis, the effect of VR training on the combined cognitive functioning outcomes (i.e., the combined scores of individual domain outcomes such as memory, attention, executive function, construction and motor performance, and verbal function and language), was large and statistically significant (g = 1.05, 95% confidence interval [CI] = [0.60, 1.51], p < .001). Heterogeneity across studies was considerable (I^2 = 87.96), but could be explained by immersion (Table 2). One study reported an extremely large SMD (g = 90.51, SE = 10.39) for the global cognition outcome, which was an outlier, therefore was removed from further analysis [74]. Having removed these data, the effect size was still large and statistically significant (g = 1.08, 95% CI = [0.71, 1.45], p < .001; Figure 3) while heterogeneity lowered slightly (I^2 = 79.77). The resulting funnel plot did not show significant asymmetry (Egger's intercept = 2.30, p = .29; Figure 12, Appendix A). Quality of evidence was moderate.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Amjad et al. 2019	2.305	0.414	0.171	1.494	3.116	5.573	0.000
Hughes et al., 2014	0.407	0.433	0.188	-0.441	1.256	0.941	0.347
Park & Park, 2018	0.401	0.229	0.053	-0.048	0.851	1.749	0.080
Ramnath et al., 2021	0.825	0.306	0.094	0.225	1.425	2.695	0.007
Torpil et al., 2021	0.997	0.268	0.072	0.471	1.523	3.713	0.000
van de Weijer et al., 2020	1.588	0.376	0.141	0.851	2.325	4.225	0.000
Park 2020	2.442	0.352	0.124	1.753	3.131	6.946	0.000
Thapa et al., 2021	0.481	0.248	0.061	-0.005	0.967	1.942	0.052
Liao et al., 2019; Liao et al., 2020	1.117	0.380	0.144	0.373	1.861	2.943	0.003
Park et al., 2020	0.238	0.423	0.179	-0.591	1.066	0.562	0.574
Kwan et al., 2021	0.047	0.506	0.256	-0.944	1.038	0.093	0.926
Hwang & Lee, 2017	3.270	0.615	0.378	2.065	4.475	5.317	0.000
Tarnanas et al., 2014	1.444	0.236	0.056	0.981	1.906	6.123	0.000
Man et al., 2012	0.507	0.302	0.091	-0.085	1.100	1.679	0.093
Park et al., 2019	0.245	0.422	0.178	-0.583	1.073	0.580	0.562
Park et al., 2020 (1)	1.498	0.391	0.153	0.732	2.264	3.834	0.000
	1.077	0.190	0.036	0.705	1.449	5.677	0.000

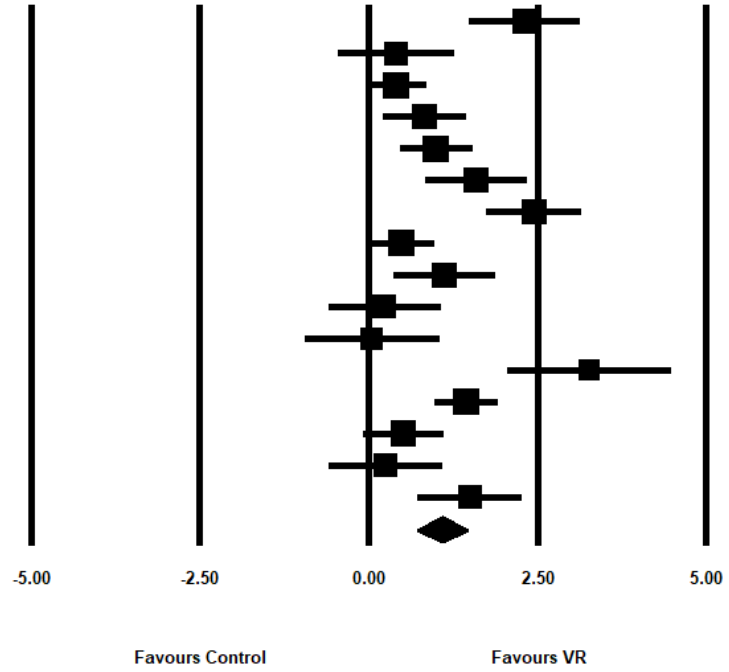

Figure 3. Effect on combined cognitive functioning outcomes (random-effects model).

General cognition

Ten studies reported outcomes for general cognition (e.g., MMSE or MoCA). The pooled effect size was medium and statistically significant (g = 0.53, 95% CI = [0.26, 0.80], p < .001). However, the CIs were wide which increases uncertainty in the effect estimate (Figure 4). Heterogeneity across studies was moderate (I^2 = 45.22), but could be explained by immersion (Appendix F, Table 6). The resulting funnel plot did not show significant asymmetry (Egger's intercept = 0.22, p = .90; Figure 13, Appendix A). Quality of evidence was moderate; downgraded by one level for imprecision, due to CIs including both a small and a large effect.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Hughes et al., 2014	0.407	0.433	0.188	-0.441	1.256	0.941	0.347
Ramnath et al., 2021	0.644	0.301	0.090	0.055	1.234	2.143	0.032
Torpil et al., 2021	0.997	0.268	0.072	0.471	1.523	3.713	0.000
van de Weijer et al., 2020	1.588	0.376	0.141	0.851	2.325	4.225	0.000
Thapa et al., 2021	0.222	0.244	0.060	-0.257	0.700	0.909	0.364
Liao et al., 2019; Liao et al., 2020	0.263	0.337	0.114	-0.397	0.924	0.781	0.435
Park et al., 2020	0.043	0.420	0.176	-0.780	0.865	0.102	0.919
Kwan et al., 2021	0.047	0.506	0.256	-0.944	1.038	0.093	0.926
Tarnanas et al., 2014	0.349	0.212	0.045	-0.067	0.764	1.646	0.100
Park et al., 2020 (1)	0.604	0.338	0.114	-0.059	1.267	1.786	0.074
	0.531	0.138	0.019	0.260	0.802	3.845	0.000

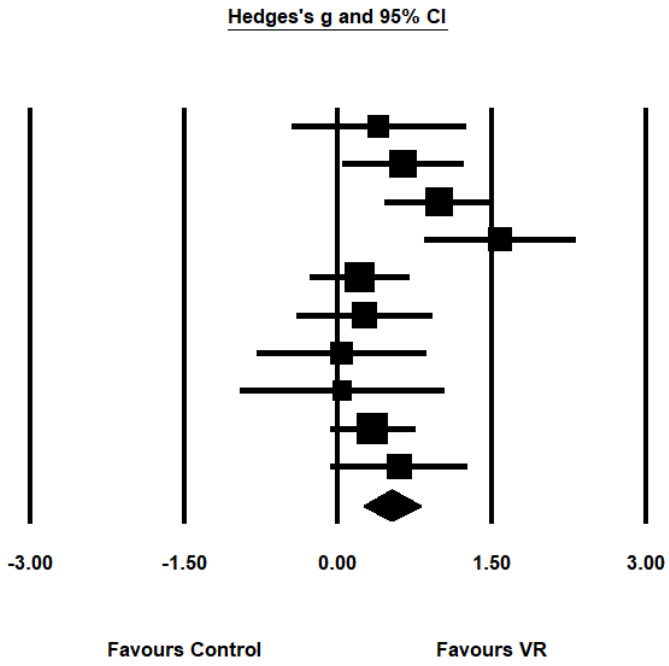

Figure 4. Effect on measures of general cognition outcomes (random-effects model).

Memory

Six studies reported outcomes for memory. The combined effect size was large and statistically significant (g = 1.01, 95% CI = [0.27, 1.75], p = .01). However, the CIs were very wide which increases uncertainty in the effect estimate (Figure 5). Heterogeneity across studies was considerable (I^2 = 89.40). The funnel plot did not reveal asymmetry (Figure 14, Appendix A), but formal testing was not conducted because of the small number of studies. Quality of evidence was very low; downgraded by two levels for inconsistency, due to considerable and not explained heterogeneity, and two levels for imprecision, due to small sample size (< 400) and CI including both a small and a large effect.

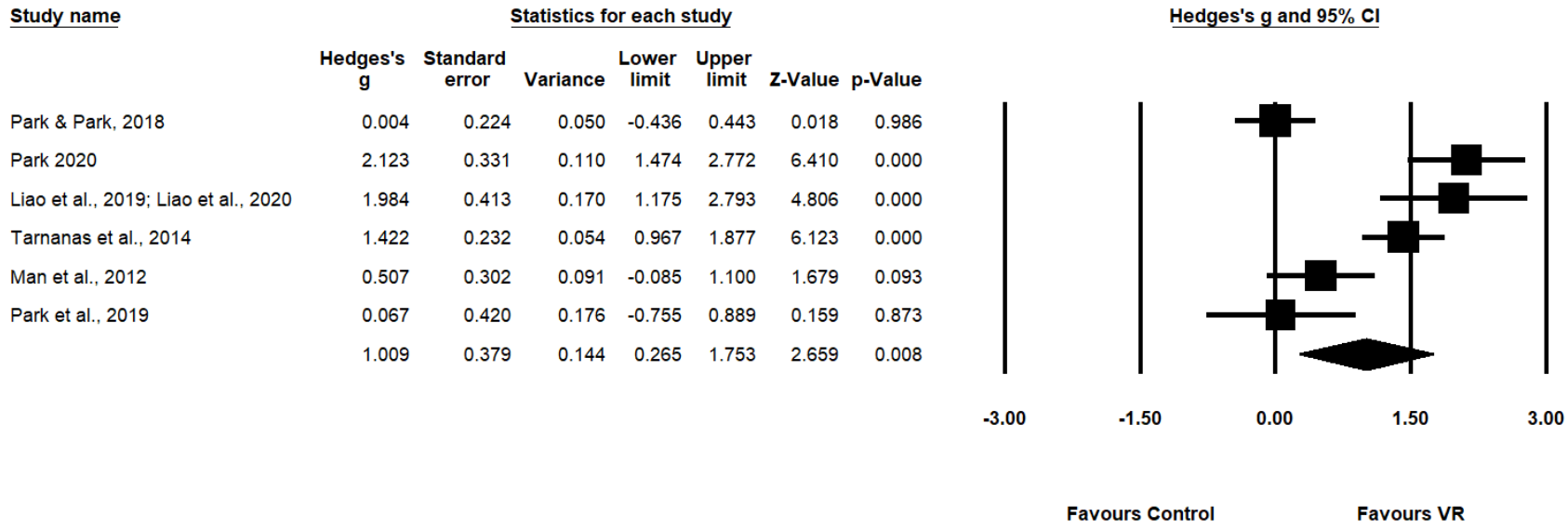

Figure 5. Effect on memory outcomes (random-effects model).

Attention, processing speed and working memory

Ten studies reported outcomes for attention, processing and working memory. The combined effect size was large and statistically significant (g = 1.42, 95% CI = [0.92, 1.92], p < .001). However, the CIs were very wide which increases uncertainty in the effect estimate (Figure 6). Heterogeneity across studies was considerable (I^2 = 81.72), but could be explained by the type of control group (i.e., active control group which involved a type of intervention other than VR, such as music therapy and psychoeducation, or passive control group, which typically refers to no-contact or wait-lists) (Appendix F, Table 8). The resulting funnel plot did not show significant asymmetry (Egger's intercept = 3.04, p = .26; Figure 15, Appendix A). Quality of evidence was high.

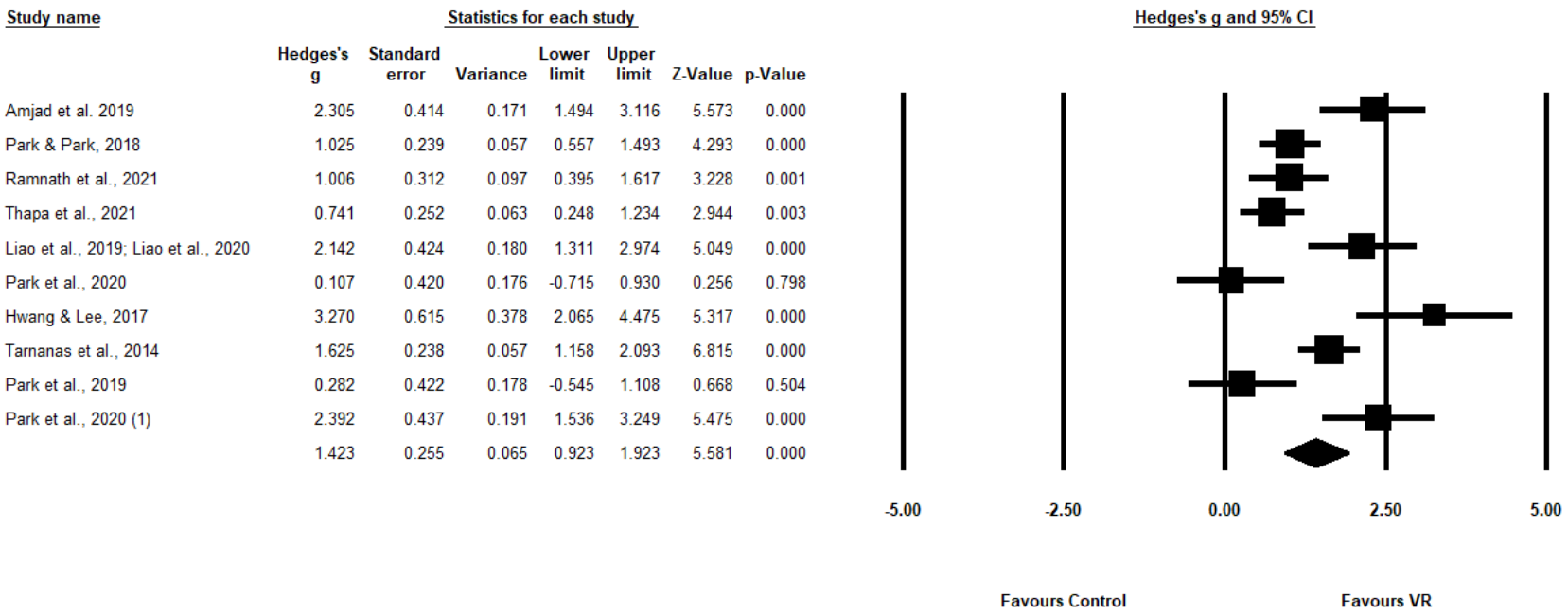

Figure 6. Effect on attention, processing speed and working memory outcomes (random-effects model).

Executive function

Four studies reported outcomes for executive function. The combined effect size was large and not statistically significant (g = 0.87, 95% CI = [-0.17, 1.92], p = .10). The CIs were very wide which increases uncertainty in the effect estimate (Figure 7). Heterogeneity across studies was considerable (I^2 = 89.08). The funnel plot revealed asymmetry (Figure 16, Appendix A), but formal testing was not conducted because of the small number of studies. Quality of evidence was very low; downgraded by two levels for inconsistency, due to considerable and not explained heterogeneity, and two levels for imprecision, due to small sample size (< 400) and CI including both a negative and a large positive effect.

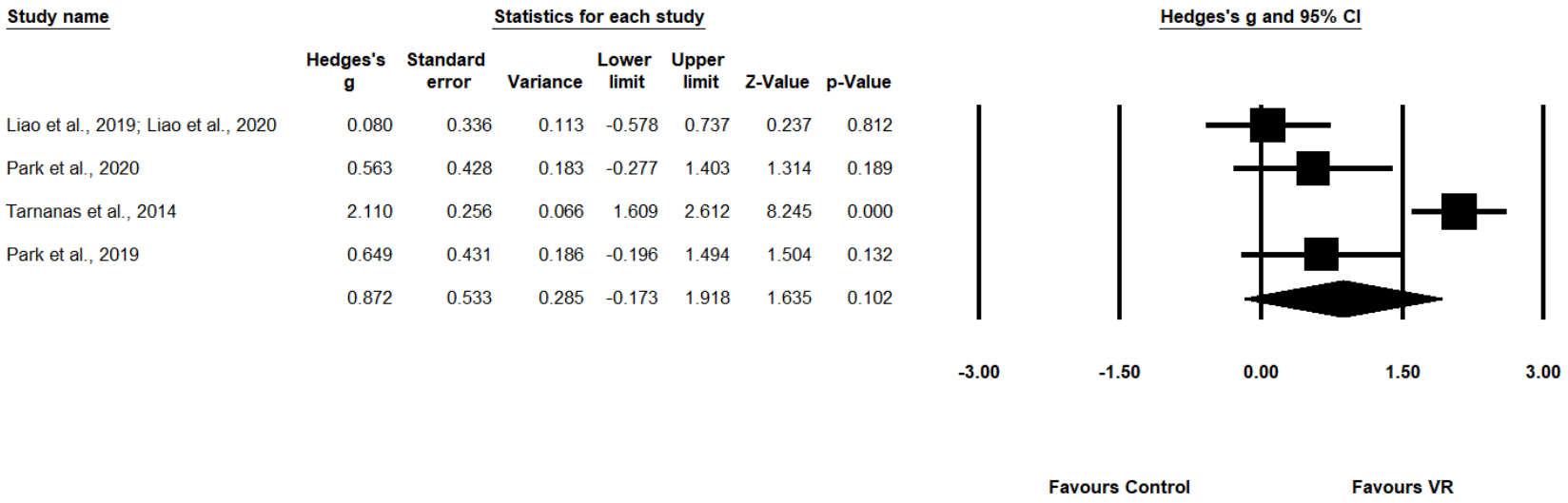

Figure 7. Effect on executive function outcomes (random-effects model).

Construction and Motor Performance

Four studies reported outcomes for construction and motor performance. The combined effect size was large and statistically significant (g = 1.16, 95% CI = [0.03, 2.29], p = .04). The CIs were very wide which increases uncertainty in the effect estimate (Figure 8). Heterogeneity across studies was considerable (I^2 = 93.45). The funnel plot did not show asymmetry (Figure 17, Appendix A), but formal testing was not conducted because of the small number of studies. Quality of evidence was very low; downgraded by two levels for inconsistency, due to considerable and not explained heterogeneity, and two levels for imprecision, due to small sample size (< 400) and CI including both a small and a large effect.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Park & Park, 2018	0.174	0.225	0.050	-0.266	0.615	0.776	0.438
Park 2020	2.762	0.371	0.138	2.035	3.488	7.446	0.000
Tarnanas et al., 2014	1.572	0.237	0.056	1.108	2.037	6.641	0.000
Park et al., 2019	0.160	0.420	0.177	-0.663	0.984	0.382	0.703
	1.164	0.577	0.333	0.033	2.295	2.017	0.044

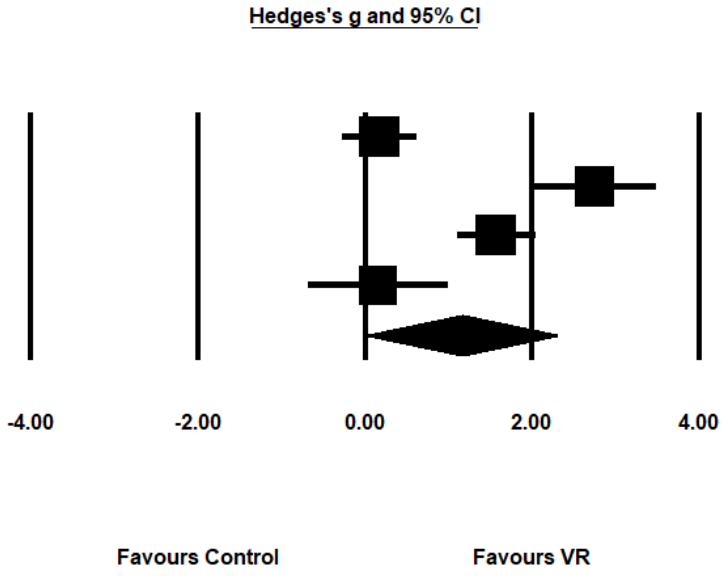

Figure 8. Effect on construction and motor performance outcomes (random-effects model).

Verbal Function and Language

Two studies have reported outcomes for verbal function and language. The combined effect size was large and not statistically significant (g = 0.86, 95% CI = [-0.62, 2.35], p = .25). The CIs were very wide which increases uncertainty in the effect estimate (Figure 9). Heterogeneity across studies was considerable (I^2 = 89.88). A funnel plot was not possible to be generated as the number of studies was less than three. Quality of evidence was very low; downgraded by two levels for inconsistency, due to considerable and not explained heterogeneity, and two levels for imprecision, due to small sample size (< 400) and CI including both a negative and a large positive effect.

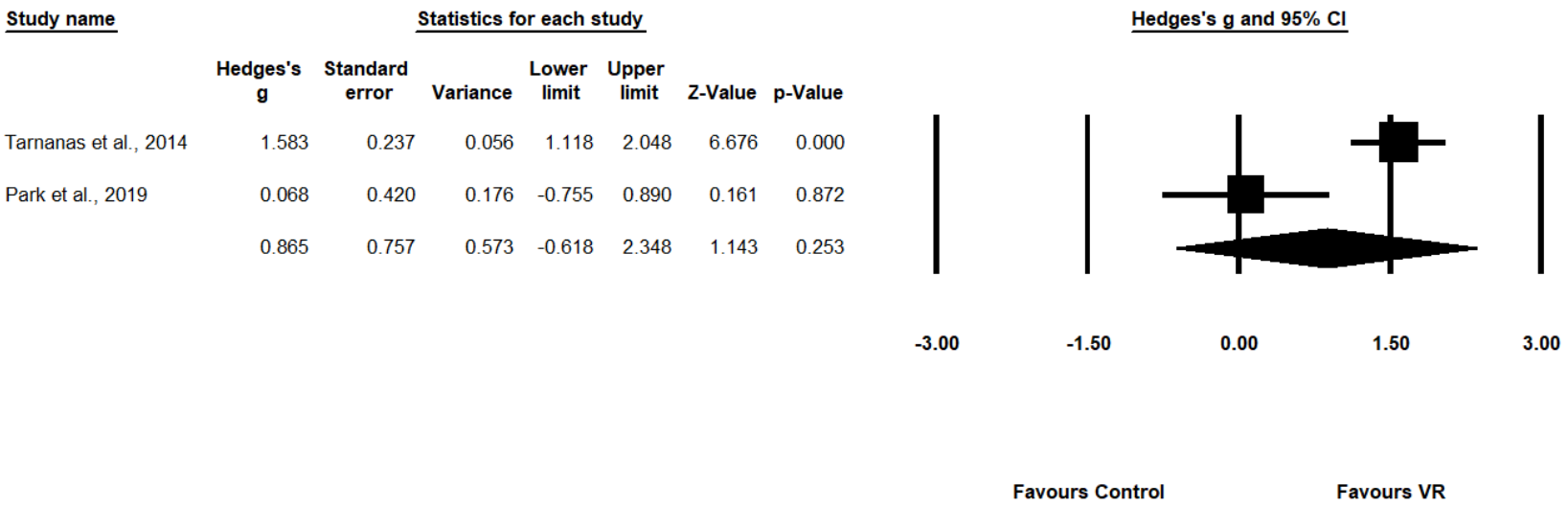

Figure 9. Effect on verbal function and language outcomes (random-effects model).

Sensitivity analysis

Sensitivity analysis indicated that the difference between lower quality studies (k = 12, g = 1.22, 95% CI = [0.77, 1.67], p < .001, I^2 = 80.55) and higher quality studies k = 4, g = 0.66, 95% CI = [0.13, 1.20] p = .014, I^2= 64.55) was not significant (p = .12).

Adverse effects

None of the studies reported detailed data for both intervention and control groups, so a meta-analysis was not conducted. One study [80] reported dizziness (4.2%) and fatigue (8.3%), but the data were not reported. Another study [75] reported that most participants did not report cybersickness, and the few who did experienced mild symptoms (*Mdn* = 4.63, *IQR* = 18.33), while one participant dropped out due to cybersickness [75]. Five studies [70, 71, 77, 81, 82, 86] reported that participants did not experience any adverse effects during the session. The remaining nine studies [74, 78, 79, 83-85, 87-89] did not report if participants experienced adverse effects.

Activities of daily living

Three studies reported outcomes for ADL. The combined effect size was small and statistically non-significant (g = 0.26, 95% CI = [-0.13, 0.65], p = .18). The CIs were relatively wide which increases uncertainty in the effect estimate (Figure 18, Appendix A). Evidence of heterogeneity was not present (I^2 = 0.00). The funnel plot did not reveal asymmetry (Figure 19, Appendix A), but formal testing was not conducted because of the small number of studies. Quality of evidence was low; downgraded by two levels for imprecision due to small sample size (< 400) and CI including both a negative and a large positive effect.

Quality of life

Only one study measured QOL [77]. The measure used was the Short Form 36 Health Survey Questionnaire (SF-36) which consists of 36 items and eight subscales: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional and mental health

[95]. The VR group showed significant (p < .05) improvement in vitality (VR: M_{change} = 9.05, SD_{change} = 1.17; Control: M_{change} = 2.69, SD_{change} = 1.6), role-emotional (VR: M_{change} = 8.31, SD_{change} = 1.20; Control: M_{change} = 4.15, SD_{change} = 0.71), mental health (VR: M_{change} = 11.62, SD_{change} = 1.63; Control: M_{change} = 6.95, SD_{change} = 1.75), and bodily pain (VR: M_{change} = 4.21, SD_{change} = 2.17; Control: M_{change} = 0.10, SD_{change} = 0.38).

Acceptability and treatment adherence

Nine studies reported dropout numbers. Dropout was more likely to occur in the VR group but failed to reach significance (OR = 1.15, CI = [0.50, 2.64], p = .75). The CIs were very wide which increases uncertainty in the effect estimate (Figure 20, Appendix A). Heterogeneity across studies was low (I^2 = 30.84). The funnel plot did not show asymmetry (Figure 21, Appendix A), but formal testing was not conducted because of the small number of studies. Quality of evidence was high.

Moderators of VR Efficacy

Due to the small number of studies in some cognitive domains a subgroup analysis was conducted for the following outcomes: combined cognitive functioning , global cognition, attention, processing speed and working memory [66, 67]. Type of VR intervention (i.e., training vs rehabilitation) could not be examined, as all studies included only cognitive training and none cognitive rehabilitation. Subgroup analysis on the combined cognitive functioning outcomes revealed that immersion could explain the high heterogeneity found between the studies (Table 2 and Table 3). Subgroup analysis on the global cognition outcomes revealed that immersion could also explain the high heterogeneity found between the studies (Appendix F, Table 6 and Table 7). Subgroup analysis on the attention, processing speed and working memory outcomes revealed that the type of control group (i.e., (i.e., active control group which involved a type of intervention other than VR, such as music therapy and psychoeducation, or passive control

group, which typically refers to no-contact or wait-lists) could explain the high heterogeneity found between the studies (Appendix F, Table 8 and Table 9).

Additionally, we conducted an exploratory content analysis on the types of training included in the studies, as we could not predict them a priori. The types of training were categorised based on the elements of the application used in each of the 16 studies. The categories that emerged were:

Video games. An electronic game that is comprised of a set of rules and has variable and quantifiable outcomes that the player tries to influence by exerting effort [96]. Examples of studies that used this type of training are [79, 81].

Exergames. These are video games that promote (either via using or requiring) players' physical movements (exertion) that is generally more than sedentary and includes strength, balance, and flexibility activities [97]. Examples of studies that used this type of training are [74, 82, 83].

Simulations. These are applications that simulate real life scenarios and do not involve specific goals[96]. Examples of studies that used this type of training are [78, 87].

Mixed. This category includes studies that used more than one type of application described above. Examples of studies that used this type of training are [70, 71, 75].

Subgroup analysis on the combined cognitive functioning outcomes (i.e., the combined scores of individual domain outcomes) revealed that the type of training could explain the high heterogeneity found between the studies (Table 2). Similarly, a subgroup analysis on the general cognition outcomes (e.g., MoCA and MMSE tests) showed that the type of training could explain the heterogeneity found between the studies (Table 6). Type of training does not seem to moderate the effects of VR training on attention, processing speed, and working memory (Table 8).

Effects of VR in Mild Cognitive Impairment Outcomes (follow-up)

Only one study [85] performed a follow-up assessment at one year post-test. Results showed that the VR group did not differ significantly in global cognition (M = 28.41, SD = 4.12) from the control group (M = 28.88, SD = 6.83). Similarly, the VR group did not differ significantly in ADL measures (M = 180.60, SD = 99.21) from the control group (M = 182.46, SD = 86.63).

Effects of VR in Dementia Outcomes (post-intervention)

Combined cognitive functioning outcomes

Based on four studies, the effect of VR training on the combined cognitive functioning outcomes (i.e., the combined scores of individual domain outcomes such as memory, attention, executive function, construction and motor performance, and verbal function and language), was large and statistically significant (g = 1.14, 95% CI = [0.41, 1.87], p = .002; Figure 10). Heterogeneity across studies was substantial (I^2 = 73.85). The resulting funnel plot did not show asymmetry (Figure 22, Appendix A), but formal testing was not conducted because of the small number of studies. Quality of evidence was very low; downgraded by two levels for inconsistency, as heterogeneity was considerable and could not be explained, and one level for imprecision, due to small sample size (< 400).

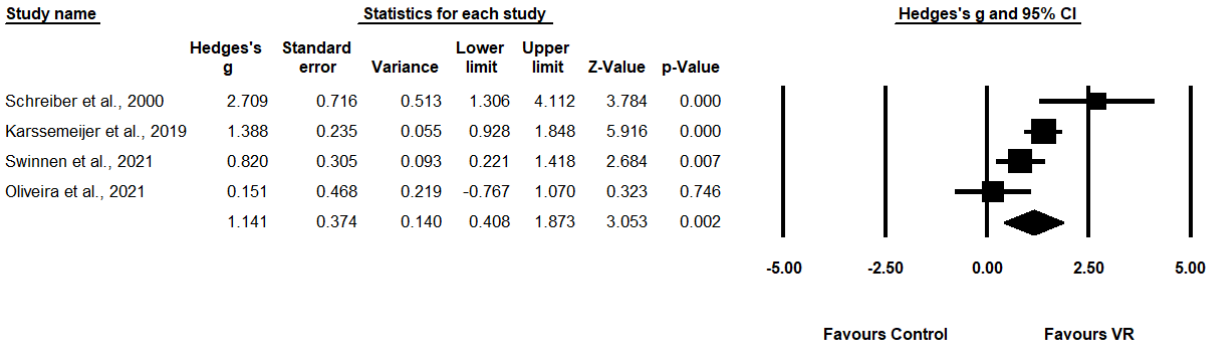

Figure 10. Effect on combined cognitive functioning outcomes (random-effects model).

General cognition

Two studies reported outcomes for general cognition (e.g., MMSE or MoCA). The pooled effect size was medium and statistically not significant (g = 0.59, 95% CI = [- 0.02, 1.20], p = .06). However, the CIs were wide which increases uncertainty in the effect estimate (Figure 11). Heterogeneity across studies was low (I^2 = 26.38). A funnel plot was not possible to be generated as the number of studies was less than three. Quality of evidence was low; downgraded two levels for imprecision, due to small sample size (< 400) and CI including both a negative and a positive effect.

Memory

Two studies reported outcomes for memory. The pooled effect size was medium and statistically significant (g = 2.14, 95% CI = [1.67, 2.61], p < .001). However, the CIs were wide which increases uncertainty in the effect estimate (Figure 11). Evidence of heterogeneity was not present (I^2 = 0.00). A funnel plot was not possible to generate as the number of studies was less than three. Quality of evidence was moderate; downgraded by one level for imprecision, due to small sample size (< 400).

Executive function

Two studies reported outcomes for executive functioning. The pooled effect size was medium and statistically significant (g = 0.51, 95% CI = [0.13, 0.89], p = .01. However, the CIs were wide which increases uncertainty in the effect estimate (Figure 11). Evidence of heterogeneity was not present (I^2 = 0.00). A funnel plot was not possible to generate as the number of studies was less than three. Quality of evidence was low; downgraded by two levels for imprecision, due to small sample size (< 400) and CI including both a small and a large effect.

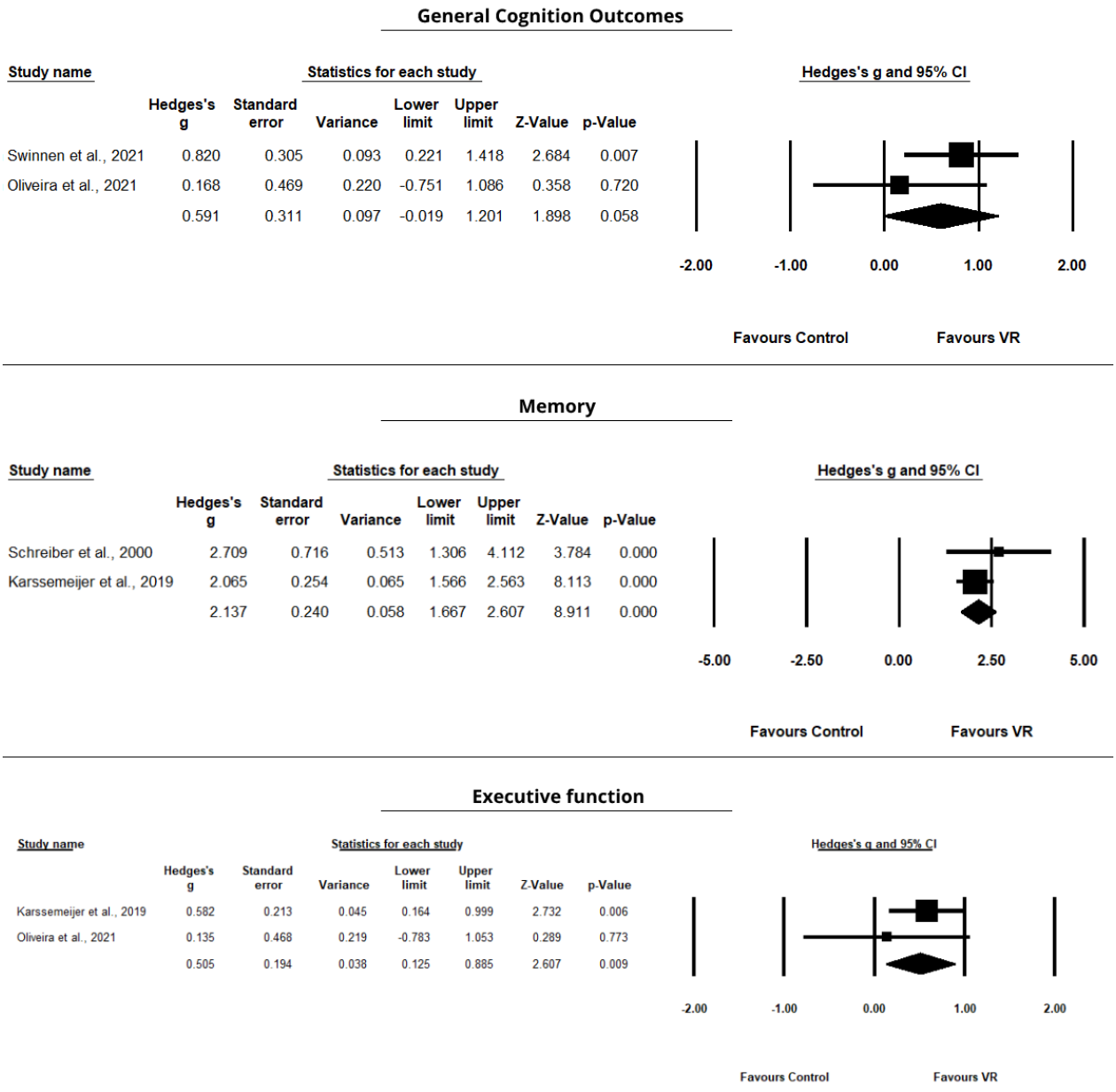

Figure 11. Effect on general cognition outcomes, memory, and executive function (random-effects model).

Attention, processing speed and working memory

Only one study included measures of attention, processing speed, and working memory [72], so a meta-analysis was not performed. In this study people with dementia trained either in a VR, or by doing aerobic or relaxation and flexibility exercises. Results from this study showed benefits on psychomotor speed after VR training (M^1 = 0.02, SD^1 = 0.87) and aerobic exercise (M^1 = 0.32, SD^1 = 0.64), but not after relaxation and flexibility exercises (M^1 = -0.25, SD^1 = 1.04). Additionally, no effects were found on working memory after VR training (M^1 = -0.25, SD^1 = 0.79), aerobic (M^1 = 0.04, SD^1 = 0.80), and relaxation and flexibility exercises (M^1 = -0.12, SD^1 = 1.02).

Sensitivity analysis

Sensitivity analysis indicated that the difference between the studies with lower quality (k = 3, g = 1.09, 95% CI = [-0.05, 2.23], p = .06, I^2 = 77.72) and the studies with higher quality k = 1, g = 1.39, 95% CI = [0.93, 1.85] p < .001, I^2= 0.00) was not significant (p = .63).

Adverse effects

None of the studies reported detailed data for both intervention and control groups, so a meta-analysis was not conducted. One study reported that participants did not experience any adverse effects during the session [92]. The remaining three studies [72, 90, 91] did not report if participants experienced adverse effects.

Activities of daily living

Three studies reported outcomes for the ADL. The combined effect size was small and statistically non-significant (g = 0.23, 95% CI = [-0.24, 0.74], p = .15). The CIs were wide which increases uncertainty in the effect estimate (Figure 23, Appendix A). Evidence of heterogeneity was not present (I^2 = 0.00). The resulting funnel plot did not show asymmetry (Figure 24, Appendix A), but formal testing was not conducted because of the small number of studies.

[1] Means and *SD*s were computed based on *Z* scores as reported by the study authors.

Quality of evidence was low; downgraded by two levels for imprecision due to small sample size (< 400) and CI including both a negative and a large positive effect.

Quality of life

Only one study measured QOL [92]. The measure used was the Dementia Quality of Life which consists of 29 items, measuring five domains: self-esteem, positive affect, negative affect, feelings of belonging, and sense of aesthetics [98]. There was no significant improvement of the QOL after the intervention (p = .16) (VR: M_{change} = 0.80, SD_{change} = 1.14; Control: M_{change} = 0.00, SD_{change} = 1.27). SDs change were imputed according to the formula suggested by the Cochrane handbook [60].

Acceptability and treatment adherence

Two studies reported dropout numbers. There was no difference of dropouts between the interventions and the control group (OR = 0.85, CI = [0.32, 2.31], p = .76). The CIs were very wide which increases uncertainty in the effect estimate (Figure 25, Appendix A). Evidence of heterogeneity was not present (I^2 = 0.00). A funnel plot was not possible to generate as the number of studies was less than three. Quality of evidence was low; downgraded by two levels for imprecision, due to small sample size (< 400) and CI including both a small and a large effect.

Moderators of VR Efficacy

A subgroup analysis was not conducted, as the number of studies in the meta-analysis was below 10 [60].

Effects of VR in Dementia Outcomes (at follow-up)

Only one study [72] performed a follow-up assessment. Results showed that the effect of VR training on attention that was found at post-test (12 weeks) was maintained at the 24 weeks

follow-up in the VR group (M^2 = -0.13, SD^2 = 0.98) and the aerobic control group (M^2 = 0.35, SD^2 = 0.73), but not in the relaxation and flexibility exercises control group (M^2 = - 0.39, SD^2 = 1.37).

DISCUSSION

The first aim of this review was to compare the effects of VR training with more conventional methods, such as computerised cognitive training and paper and pencil tasks, on general and domain-specific cognition in people with MCI and dementia. The second aim was to investigate the role of different moderators on the effects of VR training. The third aim was to evaluate current evidence of VR training in terms of safety and acceptability when compared to more conventional methods in people with MCI and dementia. Finally, the fourth aim was to evaluate the efficacy of VR training on ADL and QOL measures when compared to more conventional methods in people with MCI and dementia. Comparisons between VR training and control groups were carried out separately for MCI and dementia. Additionally, we assessed quality of evidence for each outcome to provide a more comprehensive evaluation of the use of VR training in clinical practice.

Mild Cognitive Impairment

Global and domain-specific cognition

Results from 16 randomised control trials showed that people with MCI receiving VR training improved significantly on overall global cognitive functioning outcomes (Hedge's g = 1.08) and general cognition measures (e.g., MMSE) (Hedge's g = 0.53) compared to those receiving more conventional methods such as computerised cognitive training or paper and pencil tasks, or to those not receiving an intervention. This suggests that VR training is more effective for global cognition than more conventional methods. Additionally, confidence in the effects of VR training on global cognition is moderate according to the GRADE assessment of quality of evidence.

Further, sensitivity analysis showed no significant differences between the studies with lower and medium quality, which increases confidence of results.

Regarding individual cognitive domains, VR training significantly improved memory, attention, and construction and motor performance, but not executive function or verbal function and language. However, quality of evidence ranged from low to very low due to unexplained heterogeneity, small sample sizes, and wide CIs. Our findings suggest that VR training benefits global cognition and possibly specific-domain cognition such as memory, attention, and constructive and motor performance in people with MCI, but more studies are needed to draw a strong conclusion about the effects on domain-specific cognition. Furthermore, one study included in our review suggests no retention of the cognitive effects over long term for both the VR training and control group [85], although a meta-analysis was not possible as no other studies reported follow-up measures. Future studies should include measures of cognitive outcomes over time.

Our review supports previous findings that suggest efficacy of VR training compared to more conventional methods in people with MCI [16-19]. Furthermore, our results are in line with previous reviews on the benefits of VR training in people with MCI regarding global cognition [15, 16, 18, 19], but not regarding the outcomes on individual cognitive domains [16, 17, 19]. For example, previous reviews report significant benefits of VR training on executive function [16, 19] and language [17], whereas we did not identify a significant improvement in those domains. A possible explanation for these conflicting findings is the fact that these reviews used different categorisation for individual-domain outcomes, that was not specified, compared to our coding, which was based on Lezak et al. [61]. For example, in our review Trail Makin Test A and B was coded as an *attention, processing speed, and working memory* measurement whereas in other reviews it was coded as *executive function* measurements [16, 17, 19]. On the other hand, we found a significant and large effect of VR training on memory (g = 1.01), whereas previous

reviews show either smaller [17] or no effects [16, 19]. One explanation for these differences could be linked to different characteristics between studies included in each review [28], which may influence the pooled effects on domain-specific cognitive outcomes. High heterogeneity, in terms of apparatus and training programme, between VR training studies has been reported previously in other systematic reviews [28, 29], which could influence the cognitive outcome. This is also corroborated by the high heterogeneity reported for most outcomes in our study. Another explanation could be that other meta-analyses included studies that contained an intervention that was not strictly VR according to our definition (e.g., [42, 44-46]). In our meta-analysis, we included interventions that used solely VR training.

Moderators of VR

In this review we tested several factors that may moderate the effects of VR training on cognitive function. These factors were immersion, user's point of view (first- or third-person perspective), type of control (active control group which involved a type of intervention other than VR, such as music therapy and psychoeducation, or passive control group, which typically refers to no-contact or wait-lists; computerised task or paper and pencil), diagnostic criteria of MCI, education, and duration of training. In addition, we performed an exploratory content analysis in order to identify the different types of training that were used in the studies. The types that emerged were video games, exergames, simulation and mixed.

A subgroup analysis revealed that immersion and training type moderated the effects of VR training on the combined cognitive functioning outcomes (i.e., the combined scores of individual domain outcomes such as memory, attention, executive function, construction and motor performance, and verbal function and language) and general cognition (e.g., MoCA and MMSE tests), and could explain the high heterogeneity found between the studies. Non-immersive applications seem to have greater effects on the combined cognitive functioning outcomes and on general cognitive outcomes than immersive applications. A possible explanation for this

finding is that participants were less familiar with immersive technologies, which may have hindered the effects of VR training. Indeed, research suggests that older people, such as the people with MCI included in our meta-analysis (mean age was 72.22, *SD* = 4.97), have decreased performance when using more immersive technologies, compared to young people who do not show a difference between immersive and non-immersive technologies [24]. This is corroborated by other research that shows immersive applications to be more difficult than more conventional methods and with hindered performance, especially in clinical populations [99]. Using more advanced technologies, such as VR, may be hindered in older and clinical populations due to barriers, such as cognitive (e.g., attention and working memory), motivational (e.g., trust in own abilities), physical (e.g., flexibility and hand-eye coordination), and perceptual abilities (e.g., sensory capabilities such as visual and auditory acuity) [100]. As immersive technologies could have great potential in clinical applications as they provide close-to-real world experiences, future research should focus on the usability of such technology in clinical populations, e.g., people with MCI. Indeed, a recent review on usability of VR clinical systems in older people found that most usability studies involved healthy or heterogeneous clinical populations, or VR physiotherapy training programmes [101].

In terms of the type of training, video games were found to have a much larger effect (g = 2.05) on the combined cognitive functioning outcomes, compared to exergames, simulations, or mixed training (e.g., containing simulation and exergame components). The same results were found in the subgroup analysis on the general cognition outcomes. A possible explanation for this finding is that video games may be perceived as more enjoyable than other types of intervention and thus more motivating, which has been shown to increase adherence to the intervention and therefore its effectiveness [102]. A study that compared types of motivation when playing video games or exergames, found that exergames scored lower in the Need for Competence compared to video games, which according to the self-determination theory concerns our desire

to experience a sense of challenge in a competitive environment [103]. Considering this, it might be possible that more effective VR training may be linked to greater Need for Competence and that low challenge could negatively impact motivation [104]. Simulations, on the other hand, have no goals and therefore no challenges by definition which might negatively impact motivation [96].

The type of control group (passive, such as no-contact or wait-list, or active, such as music therapy and psychoeducation) seemed to also moderate the effects of VR training on attention, processing speed, and working memory. VR training showed bigger effects when compared to active control groups, such as paper and pencil cognitive training, music therapy, and computerised cognitive training, or passive control groups (i.e., no-contact or wait-list). The significance of the difference between passive and active controls should be taken with caution, considering that there was only one study in the passive group, however, the finding that VR training has a significant effect when compared to active controls is encouraging.

The remaining moderators were not found to significantly influence the effects of VR training. Some of these factors may have no effect on the benefits of VR training in people with MCI, due to deterioration of specific cognitive abilities in this population. For example, embodied cognition, which would be required to differentiate the effects of first-person and third-person perspective on VR training, may be more impaired in people with MCI or other neurodegenerative diseases [105]. Alternatively, specific moderators such as embodiment may affect very specific components of cognitive domains. For example, first-person and third-person perspective, may be more relevant in cognition that involves egocentric spatial memory and near-space representations [105]. It should also be noted that for some moderators the number of studies between subgroups were sometimes imbalanced and/or contained a small number of studies in one subgroup (e.g., studies with first-person versus a third-person perspective), which reduces the power of the subgroup analysis [106]. Additionally, caution should be exercised when

interpreting subgroup analyses, as evidence is not direct and other covariates might influence the results [107].

Safety and acceptability

From the seven studies that reported adverse effects, one mentioned dizziness and fatigue for a small number of participants [80], and another one reported mild symptoms for a few participants and one dropout due to cybersickness [75]. Both of these studies contained immersive VR, while from the remaining five studies that reported no adverse effects, two contained immersive VR [70, 71, 86] and three contained non-immersive VR [77, 81, 82]. Nevertheless, we could not conduct a meta-analysis, as there were no data for the active control groups. Future studies should document incidents of adverse effects for both intervention and control groups, in order to make an informed decision about the safety of VR training systems in people with MCI.

VR training was well-accepted in people with MCI with no significant dropout rates. The low number of studies reporting adverse effects and the acceptability of VR training, indicates good feasibility of VR training in people with MCI. However, formal documentation of adverse effects and studies with better quality is needed to allow an informed decision.

Our findings suggest feasibility of VR training in people with MCI, which supports previous research on feasibility of VR training in people with MCI [102] and other neurological disorders, such as stroke and traumatic brain injury [33]. However, one should bear in mind that most studies included in the review used a non-immersive system, and only four of the studies contained an immersive VR training [70, 71, 75, 80, 86]. Furthermore, two of the four studies that used immersive VR training reported mild adverse effects such as dizziness and fatigue [75, 80], which could indicate the infeasibility of immersive VR training for some. However, more studies with immersive VR training in people with MCI are needed in order to draw conclusions regarding the feasibility of immersive VR training, as data in the present review is limited.

Activities of daily living and quality of life

Our findings suggest that VR training does not have a significant effect on ADL in people with MCI, although quality of evidence was low. This is in line with a previous meta-analysis which reported similar findings [19]. One possible explanation is that cognitive training delivered via VR targets mostly specific cognitive functions and not ADL, and so cognitive training is not able to target and improve ADL functions. Similarly, it might be the case that the effect of VR on specific cognitive domains shows limited transfer to real life settings including ADL [108]. It could also be that because the ADL are not so severely impaired in people with MCI, the scales for measuring ADL may not be so sensitive to capture any change in MCI [6].

This was the first meta-analysis that included measures of QOL. Nevertheless, only one study included QOL measures and therefore a meta-analysis was not conducted [77]. In this study, participants trained for 10 weeks by playing on a Nintendo Wii in a virtual environment (e.g., playing table tennis with an opponent similarly to a real-life scenario) by using motion controller, or with a computerised cognitive training programme (CoTras). Results revealed significant improvements of QOL domains such as vitality (energy and fatigue), emotional role limitations, mental health, and bodily pain, after playing an exergame on Nintendo Wii, but not in others such as physical functioning, physical role limitations, general health perceptions and social functioning. Based on the authors' conclusions, the VR training did not contain elements that may increase social functioning, (e.g., social interaction during play) or physical function (e.g., lower extremity activities). On the other hand, previous studies have shown that VR exergames can benefit mental health [109], emotional and psychological problems [110], and bodily pain [111] in older adults. Additionally, the movement component of exergames may promote physical activity similar to traditional physical activities [112], which in turn may increase vitality [113], as well as other mental and physical domains of QOL measures [114]. Nevertheless, more studies are needed to make an informed decision on the QOL benefits of VR training.

Dementia

Global and domain-specific cognition

Results from four randomised control trials indicate that people with dementia receiving VR training improved significantly on overall global cognitive functioning outcomes (Hedge's g = 1.14), compared to those receiving more conventional methods such as computerised cognitive training or paper and pencil tasks, or to those not receiving an intervention. This is in line with previous research that report similar findings [15, 18]. A sensitivity analysis did not reveal any significant differences between the studies with lower and medium quality, which increases confidence in the results. On the other hand, the effect of VR training on general cognition measures (e.g., MMSE, MoCA) does not corroborate this finding, as it was found to be non-significant. Nevertheless, quality of evidence ranged from low to very low, so more studies are needed to examine these effects.

In terms of the effects of VR training on individual cognitive domains, our findings suggest that VR training is beneficial to memory and executive functioning. This is interesting, as we did not find the same significant effect of VR training on executive function in people with MCI. Furthermore, the effect of VR training on the memory of people with dementia was twice as large compared to that found in people with MCI (g = 2.14 compared to g = 1.01, respectively). Considering that people with dementia experience greater impairments in memory and executive functions than people with MCI [54, 115], it could be assumed that VR training has greater benefits for those with greater cognitive impairment. Additionally, the effects on memory and executive function have clinical significance, as both of these two functions have been linked to impaired ADL in dementia [116].

On the other hand, only one study included measures of attention, processing speed, and working memory, and reported positive effects of VR training on processing speed which were retained after 12 weeks without training [72]. Furthermore, none of the included studies

measured verbal function and construction and motor performance. Considering the lack of other cognitive measures, the small number of studies, and that quality of evidence was low for most cognitive domains, except for memory, more research in this area is needed to draw firm conclusions of the effects of VR training on dementia.

Safety and acceptability

One study [92] provided information regarding adverse effects, and reported that none of the participants experienced any adverse effect [92]. A meta-analysis was not possible as no other study reported adverse effects. Nevertheless, we performed a meta-analysis on the numbers of dropout to investigate adherence to the intervention compared to the control group. The analysis showed no significant differences of dropout numbers between the intervention and control groups, which suggests good feasibility of VR training. This is in line with previous research that shows feasibility of both immersive and less immersive VR applications in people with dementia [102, 117, 118]. It should be noted however, that all the studies included in our meta-analysis for dementia used a non-immersive system. Future studies should consider including immersive VR training as well as adverse effects measures in order to draw conclusions regarding the feasibility of immersive VR training, as data in the present review is limited.

Activities of daily living and quality of life

This was the first meta-analysis that investigated the effects of VR training on the ADL of people with dementia. Our findings suggest that VR training does not have a significant effect on ADL, contrary to the effects reported on other clinical populations, such as Parkinson's disease and stroke patients [119-121]. This contradiction could be attributed to the fact that different impairments are involved in dementia and Parkinson's disease, and that the VR training may be able to target the deficits in Parkinson's disease but not those in dementia [122]. Alternatively, this inconsistency with previous findings could be explained by the fact that the studies included in our meta-analysis used less immersive systems, whereas more immersive systems could

allow more successful transference of skills, as users are able to perform actions in a more ecological manner [33]. Furthermore, our results do not support previous research [123], and suggest significant benefits on executive function after VR training, but not on the ADL. This might indicate that other cognitive domains may contribute to the decline of ADL, such as language and attention [116], or maybe the participants were less affected on the ADL and the effect was negligible. Finally, it should also be considering that the findings on ADL were based on a small number of studies. Future research should focus on examining the link between VR cognitive training and ADL.

Only one study included QOL measures and therefore a meta-analysis was not conducted [92]. In this study, people with dementia trained for 8 weeks with a “Dividat Senso” device which consists of a step training platform which is sensitive to pressure changes. Participants played various exergames by using their lower limbs. Results from this study did not return a significant benefit of VR training on QOL measures. It may be that VR training must incorporate other aspects in order to benefit QOL. For example, VR training in a group setting has been shown to promote a positive and encouraging environment that allows opportunities for social interaction [124], which has been shown to benefit QOL [125]. As there is a close link between QOL and depression among older adults, improving mood could benefit QOL as well [126]. Nevertheless, more studies are needed to make an informed decision on the QOL benefits of VR training as research on this is on its infancy.

How could the effects of VR training be explained?

The benefits of VR training on cognition compared to more conventional methods may be explained by multiple aspects. First, VR offers more complex environments compared to conventional methods, which might slow down the progression of MCI [12]. Studies on transgenic mice has shown that interaction with enriched environments increases neuroplasticity in AD brains [13, 14]. Second, presence, which is inherently linked to VR systems might act as

an underlying mechanism for the positive effects of VR training. Studies on healthy young and older adults that investigated the mediating effect of presence in immersive VR systems reported a positive relationship with cognitive performance [23, 27]. Presence is thought to share similar features with selective attention which allows users to focus on the stimuli of virtual environment and exclude unrelated real-world stimuli [127]. Third, VR experiences increase motivation and engagement which may result in higher adhesion to training [102]. Nevertheless, none of the studies included in the present review reported measurements of presence, motivation, and engagement. To study the role of these factors as an underlying mechanism explaining the positive effects of VR, future studies should include these measures.

Implications for clinical practice

A key issue is to identify the underlying clinical impact and to see if gains in cognition observed in VR can translate to real life improvements. Most effects that were expressed via standardized mean differences were of moderate and large magnitude, which suggests that VR-based interventions have clinical significance. Major clinical improvements based on large effects were reported for people with MCI, on global cognition, memory, attention, and constructive and motor performance. No significant improvement was reported on ADL after VR training, but large effects were reported for QOL. Video games and non-immersive applications seem to have greater effects on VR training.

In terms of benefits in people with dementia, VR training shows moderate improvements in global cognition, memory, executive function, and processing speed. Global cognition and memory showed large cognitive improvement after VR training whereas executive function and processing speed showed moderate improvements. Improvements on ADL and QOL are not significant.

Strengths and limitations

Our review has several strengths. First, it is the first review to investigate the effects of VR training separately in people with MCI and dementia. Previous reviews have treated dementia and MCI as one group [15, 18], however, research suggests that there are differences between the two conditions in terms of impairment and neurorehabilitation mechanisms [128, 129]. Second, we included interventions that contained solely VR training in comparison to previous reviews that included mixed interventions or non-VR interventions. Third, we conducted a broad search on several big databases, including grey literature databases, in order to capture all relevant studies. Furthermore, we did not exclude any languages from our search, and we only included studies that used a randomised or quasi randomised control trial design, which increases quality of evidence. Additionally, our review is the first to explore the effect of different moderators on the effects of VR training and the first to identify immersion and type of training as significant moderators for the effects on global cognition. However, there were some limitations that we could not control for, which must be considered when interpreting the results. First, the limited number of studies should be considered, especially in the meta-analysis of VR training effects on dementia, and therefore caution must be exercised when interpreting the results. Second, the sample size for some individual cognitive domains as well as for the ADL and QOL was small, which lowers confidence in results [130]. Third, a subgroup analysis was not possible in all cognitive domains, as the number of studies for these domains was less than the recommended (10) [60]. Fourth, for most outcomes we found high heterogeneity that could not be explained by the priori moderators, albeit based on small and group-imbalanced sample sizes. Other factors such as presence, motivation, engagement, and complexity of environment may moderate these effects of VR training. Fifth, only one study with people with MCI and one study with people with dementia examined retention of cognitive benefits and QOL measures, so more studies with these measures are needed in order to make an informed decision. Additionally, none of the studies have included data of adverse effects for the control group, so a

comparison between intervention and control groups was not possible. Finally, most studies used a non-immersive system, especially those that investigated the effects of VR training on dementia, so more studies are needed that include immersive VR training.

CONCLUSIONS

In people with MCI, VR training may benefit global cognition, memory, attention, constructive and motor performance, and QOL, but it may not benefit executive function, verbal function and language, or ADL. Furthermore, VR training seems to be safe and well-accepted by people with MCI. However, more studies with immersive VR training are needed in order to be confident of these effects. Less immersive systems showed greatest improvements on cognitive outcomes. Additionally, video games were most effective, followed by exergames and simulations, and then by training that combined two or more of these VR training methods. Presence, motivation, engagement, and complexity of virtual environments was not measured in any of the included studies, so future studies must examine the effect of these moderators. Finally, more studies must include short-term and long-term follow-ups and QOL measures.

In people with dementia, VR training may benefit global cognition, memory, executive function, and processing speed, but not ADL or QOL, although more randomised control trials are needed to make an informed decision. Additionally, more studies employing immersive VR training are needed to better understand whether more immersive methods may have greater benefit in these individuals.

ACKNOWLEDGMENTS

This research was supported by the University of Bath as part of a +3 PhD studentship held by TP. KP's research is partly funded by the Centre for the Analysis of Motion, Entertainment Research and Applications (CAMERA 2.0; EP/T022523/1).

CONFLICT OF INTEREST/DISCLOSURE STATEMENT

The authors have no conflict of interest to report.

REFERENCES

[1] World Health Organization (2017) *Global action plan on the public health response to dementia 2017-2025*, World Health Organization, Geneva.

[2] Burton CL, Strauss E, Bunce D, Hunter MA, Hultsch DF (2009) Functional Abilities in Older Adults with Mild Cognitive Impairment. *Gerontology* **55**, 570-581, 10.1159/000228918.

[3] Roberts R, Knopman DS (2013) Classification and Epidemiology of MCI. *Clinics in Geriatric Medicine* **29**, 753-772, 10.1016/j.cger.2013.07.003.

[4] Mitchell AJ, Shiri-Feshki M (2009) Rate of progression of mild cognitive impairment to dementia – meta-analysis of 41 robust inception cohort studies. *Acta Psychiatrica Scandinavica* **119**, 252-265, 10.1111/j.1600-0447.2008.01326.x.

[5] Bárrios H, Narciso S, Guerreiro M, Maroco J, Logsdon R, de Mendonça A (2013) Quality of life in patients with mild cognitive impairment. *Aging & Mental Health* **17**, 287-292, 10.1080/13607863.2012.747083.

[6] Rodríguez-Bailón M, Montoro-Membila N, Garcia-Morán T, Arnedo-Montoro ML, Funes Molina MJ (2015) Preliminary cognitive scale of basic and instrumental activities of daily living for dementia and mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology* **37**, 339-353, 10.1080/13803395.2015.1013022.

[7] Clare L, Woods RT (2004) Cognitive training and cognitive rehabilitation for people with early-stage Alzheimer's disease: A review. *Neuropsychological Rehabilitation* **14**, 385-401, 10.1080/09602010443000074.

[8] Parsons TD (2015) Virtual Reality for Enhanced Ecological Validity and Experimental Control in the Clinical, Affective and Social Neurosciences. *Frontiers in Human Neuroscience* **9**, 660, 10.3389/fnhum.2015.00660.

[9] Finn M, McDonald S (2011) Computerised cognitive training for older persons with mild cognitive impairment: A pilot study using a randomised controlled trial design. *Brain Impairment* **12**, 187-199, 10.1375/brim.12.3.187.

[10] Rizzo AA, Buckwalter JG (1997) Virtual reality and cognitive assessment and rehabilitation: the state of the art. *Studies in Health Technology and Informatics*, 123-146, 10.3233/978-1-60750-888-5-123.

[11] Freeman D, Reeve S, Robinson A, Ehlers A, Clark D, Spanlang B, Slater M (2017) Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine* **47**, 2393-2400, 10.1017/S003329171700040X.

[12] Garcia L, Kartolo A, Méthot-Curtis E (2012) A discussion of the use of virtual reality in dementia. . *Virtual Reality in Psychological, Medical and Pedagogical Applications*, 123-136, 10.5772/46412.

[13] Lazarov O, Robinson J, Tang Y-P, Hairston IS, Korade-Mirnics Z, Lee VMY, Hersh LB, Sapolsky RM, Mirnics K, Sisodia SS (2005) Environmental Enrichment Reduces Aβ Levels and Amyloid Deposition in Transgenic Mice. *Cell* **120**, 701-713, 10.1016/j.cell.2005.01.015.

[14] Jankowsky JL, Melnikova T, Fadale DJ, Xu GM, Slunt HH, Gonzales V, Younkin LH, Younkin SG, Borchelt DR, Savonenko AV (2005) Environmental Enrichment Mitigates Cognitive Deficits in a Mouse Model of Alzheimer's Disease. *The Journal of Neuroscience* **25**, 5217, 10.1523/JNEUROSCI.5080-04.2005.

[15] Kim O, Pang Y, Kim J-H (2019) The effectiveness of virtual reality for people with mild cognitive impairment or dementia: a meta-analysis. *BMC Psychiatry* **19**, 219, 10.1186/s12888-019-2180-x.

[16] Wu J, Ma Y, Ren Z (2020) Rehabilitative Effects of Virtual Reality Technology for Mild Cognitive Impairment: A Systematic Review With Meta-Analysis. *Frontiers in Psychology* **11**, 1811, 10.3389/fpsyg.2020.01811.

[17] Hung SC, Ho AY, Lai IH, Lee CS, Pong AS, Lai FH (2020) Meta-Analysis on the Effectiveness of Virtual Reality Cognitive Training (VRCT) and Computer-Based Cognitive Training (CBCT) for Individuals with Mild Cognitive Impairment (MCI). *Electronics* **9**, 10.3390/electronics9122185.

[18] Zhu S, Sui Y, Shen Y, Zhu Y, Ali N, Guo C, Wang T (2021) Effects of Virtual Reality Intervention on Cognition and Motor Function in Older Adults With Mild Cognitive Impairment or Dementia: A Systematic Review and Meta-Analysis. *Frontiers in Aging Neuroscience* **13**, 10.3389/fnagi.2021.586999.

[19] Zhong D, Chen L, Feng Y, Song R, Huang L, Liu J, Zhang L (2021) Effects of virtual reality cognitive training in individuals with mild cognitive impairment: A systematic review and meta-analysis. *International Journal of Geriatric Psychiatry* **36**, 1829-1847, 10.1002/gps.5603.

[20] Ijaz K, Ahmadpour N, Naismith SL, Calvo RA (2019) An Immersive Virtual Reality Platform for Assessing Spatial Navigation Memory in Predementia Screening: Feasibility and Usability Study. *JMIR Ment Health* **6**, 10.2196/13887.

[21] Murcia-López M, Steed A (2016) The Effect of Environmental Features, Self-Avatar, and Immersion on Object Location Memory in Virtual Environments. *Frontiers in ICT* **3**, 24, 10.3389/fict.2016.00024.

[22] Schöne B, Wessels M, Gruber T (2019) Experiences in Virtual Reality: a Window to Autobiographical Memory. *Current Psychology* **38**, 715-719, 10.1007/s12144-017-9648-y.

[23] Huang K-T (2019) Exergaming Executive Functions: An Immersive Virtual Reality-Based Cognitive Training for Adults Aged 50 and Older. *Cyberpsychology, Behavior, and Social Networking* **23**, 143-149, 10.1089/cyber.2019.0269.

[24] Plechatá A, Sahula V, Fayette D, Fajnerová I (2019) Age-Related Differences With Immersive and Non-immersive Virtual Reality in Memory Assessment. *Frontiers in Psychology* **10**, 10.3389/fpsyg.2019.01330.

[25] Slater M, Wilbur S (1997) A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* **6**, 603-616, 10.1162/pres.1997.6.6.603.

[26] Kober SE, Kurzmann J, Neuper C (2012) Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: An EEG study. *International Journal of Psychophysiology* **83**, 365-374, 10.1016/j.ijpsycho.2011.12.003.

[27] Cadet LB, Chainay H (2020) Memory of virtual experiences: Role of immersion, emotion and sense of presence. *International Journal of Human-Computer Studies* **144**, 10.1016/j.ijhcs.2020.102506.

[28] Sayma M, Tuijt R, Cooper C, Walters K (2020) Are We There Yet? Immersive Virtual Reality to Improve Cognitive Function in Dementia and Mild Cognitive Impairment. *The Gerontologist* **60**, 502-512, 10.1093/geront/gnz132.

[29] Sefcik J, Petrovsky D, Cacchione P, Oh S, Demiris G (2020) Virtual Reality Use Among Persons With Dementia: An Integrative Review. *Innovation in Aging* **4**, 822-823, 10.1093/geroni/igaa057.3000.

[30] Rahe J, Petrelli A, Kaesberg S, Fink GR, Kessler J, Kalbe E (2015) Effects of cognitive training with additional physical activity compared to pure cognitive training in healthy older adults. *Clinical interventions in aging* **10**, 297-310, 10.2147/CIA.S74071.

[31] Bréchet L, Mange R, Herbelin B, Theillaud Q, Gauthier B, Serino A, Blanke O (2019) First-person view of one's body in immersive virtual reality: Influence on episodic memory. *PLOS ONE* **14**, 10.1371/journal.pone.0197763.

[32] Freton M, Lemogne C, Bergouignan L, Delaveau P, Lehéricy S, Fossati P (2014) The eye of the self: precuneus volume and visual perspective during autobiographical memory retrieval. *Brain Structure and Function* **219**, 959-968, 10.1007/s00429-013-0546-2.

[33] Voinescu A, Sui J, Stanton Fraser D (2021) Virtual Reality in Neurorehabilitation: An Umbrella Review of Meta-Analyses. *Journal of Clinical Medicine* **10**, 10.3390/jcm10071478.

[34] Martin M, Clare L, Altgassen AM, Cameron MH, Zehnder F (2011) Cognition-based interventions for healthy older people and people with mild cognitive impairment. *Cochrane Database of Systematic Reviews*, 10.1002/14651858.CD006220.pub2.

[35] Bahar-Fuchs A, Martyr A, Goh AMY, Sabates J, Clare L (2019) Cognitive training for people with mild to moderate dementia. *Cochrane Database of Systematic Reviews*, 10.1002/14651858.CD013069.pub2.

[36] Hill NTM, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A (2016) Computerized Cognitive Training in Older Adults With Mild Cognitive Impairment or Dementia: A

Systematic Review and Meta-Analysis. *American Journal of Psychiatry* **174**, 329-340, 10.1176/appi.ajp.2016.16030360.

[37] Kueider AM, Parisi JM, Gross AL, Rebok GW (2012) Computerized Cognitive Training with Older Adults: A Systematic Review. *PLOS ONE* **7**, 10.1371/journal.pone.0040588.

[38] Gates N, Valenzuela M (2010) Cognitive exercise and its role in cognitive function in older adults. *Current psychiatry reports* **12**, 20-27, 10.1007/s11920-009-0085-y.

[39] Meng X, D'Arcy C (2012) Education and Dementia in the Context of the Cognitive Reserve Hypothesis: A Systematic Review with Meta-Analyses and Qualitative Analyses. *PLOS ONE* **7**, 10.1371/journal.pone.0038268.

[40] Scarmeas N, Albert SM, Manly JJ, Stern Y (2006) Education and rates of cognitive decline in incident Alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry* **77**, 308, 10.1136/jnnp.2005.072306.

[41] Bamidis PD, Fissler P, Papageorgiou SG, Zilidou V, Konstantinidis EI, Billis AS, Romanopoulou E, Karagianni M, Beratis I, Tsapanou A, Tsilikopoulou G, Grigoriadou E, Ladas A, Kyrillidou A, Tsolaki A, Frantzidis C, Sidiropoulos E, Siountas A, Matsi S, Papatriantafyllou J, Margioti E, Nika A, Schlee W, Elbert T, Tsolaki M, Vivas AB, Kolassa I-T (2015) Gains in cognition through combined cognitive and physical training: the role of training dosage and severity of neurocognitive disorder. *Frontiers in Aging Neuroscience* **7**, 152, 10.3389/fnagi.2015.00152

[42] Optale G, Urgesi C, Busato V, Marin S, Piron L, Priftis K, Gamberini L, Capodieci S, Bordin A (2009) Controlling Memory Impairment in Elderly Adults Using Virtual Reality Memory Training: A Randomized Controlled Pilot Study. *Neurorehabilitation and Neural Repair* **24**, 348-357, 10.1177/1545968309353328.

[43] Batini F, Toti G, Bartolucci M (2016) Neuropsychological benefits of a narrative cognitive training program for people living with dementia: A pilot study. *Dementia & Neuropsychologia* **10**, 127-133, 10.1590/S1980-5764-2016DN1002008.

[44] Schwenk M, Sabbagh M, Lin I, Morgan P, Grewal GS, Mohler J, Coon DW, Najafi B (2016) Sensor-based balance training with motion feedback in people with mild cognitive impairment. *Journal of rehabilitation research and development* **53**, 945-958, 10.1682/JRRD.2015.05.0089.

[45] Delbroek T, Vermeylen W, Spildooren J (2017) The effect of cognitive-motor dual task training with the biorescue force platform on cognition, balance and dual task performance in institutionalized older adults: a randomized controlled trial. *Journal of Physical Therapy Science* **29**, 1137-1143, 10.1589/jpts.29.1137.

[46] Yang H-L, Chu H, Kao C-C, Chiu H-L, Tseng I-J, Tseng P, Chou K-R (2019) Development and effectiveness of virtual interactive working memory training for older people with mild cognitive impairment: a single-blind randomised controlled trial. *Age and Ageing* **48**, 519-525, 10.1093/ageing/afz029.

[47] Rose FD, Attree EA, Brooks BM, Andrews TK (2001) Learning and Memory in Virtual Environments: A Role in Neurorehabilitation? Questions (and Occasional Answers) from the University of East London. *Presence: Teleoperators and Virtual Environments* **10**, 345-358, 10.1162/1054746011470208.

[48] D'Cunha NM, Nguyen D, Naumovski N, McKune AJ, Kellett J, Georgousopoulou EN, Frost J, Isbel S (2019) A Mini-Review of Virtual Reality-Based Interventions to Promote Well-Being for People Living with Dementia and Mild Cognitive Impairment. *Gerontology* **65**, 430-440, 10.1159/000500040.

[49] Davis S, Nesbitt K, Nalivaiko E (2014) in *Proceedings of the 2014 Conference on Interactive Entertainment* Association for Computing Machinery, Newcastle, NSW, Australia, pp. 1–9, 10.1145/2677758.2677780.

[50] da Costa RQM, Pompeu JE, de Mello DD, Moretto E, Rodrigues FZ, dos Santos MD, Nitrini R, Morganti F, Brucki SMD (2018) Two new virtual reality tasks for the assessment of spatial orientation: Preliminary results of tolerability, sense of presence and usability. *Dementia & Neuropsychologia* **12**, 196-204, 10.1590/1980-57642018dn12-020013

[51] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J, Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology* **62**, 1-34, 10.1016/j.jclinepi.2009.06.006.

[52] American Psychiatric Association (2013) *Diagnostic and statistical manual of mental disorders (DSM-5®)*, American Psychiatric Pub.

[53] World Health Organization (2019) *International classification of diseases for mortality and morbidity statistics (11th Revision).* .

[54] Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E (1999) Mild Cognitive Impairment: Clinical Characterization and Outcome. *Archives of Neurology* **56**, 303-308, 10.1001/archneur.56.3.303.

[55] Lonie JA, Tierney KM, Ebmeier KP (2009) Screening for mild cognitive impairment: a systematic review. *International Journal of Geriatric Psychiatry* **24**, 902-915, 10.1002/gps.2208.

[56] Bucks RS, Ashworth DL, Wilcock GK, Siegfried K (1996) Assessment of Activities of Daily Living in Dementia: Development of the Bristol Activities of Daily Living Scale. *Age and Ageing* **25**, 113-120, 10.1093/ageing/25.2.113.

[57] Sherer M, Maddux JE, Mercandante B, Prentice-Dunn S, Jacobs B, Rogers RW (1982) The Self-Efficacy Scale: Construction and Validation. *Psychological Reports* **51**, 663-671, 10.2466/pr0.1982.51.2.663.

[58] Ready RE, Ott BR, Grace J, Fernandez I (2002) The Cornell-Brown Scale for Quality of Life in Dementia. *Alzheimer Disease & Associated Disorders* **16**, 10.1097/00002093-200204000-00008

[59] Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG (1993) Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* **3**, 203-220, 10.1207/s15327108ijap0303_3.

[60] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (2019) *Cochrane Handbook for Systematic Reviews of Interventions*, John Wiley & Sons.

[61] Lezak MD, Howieson DB, Bigler ED, Tranel D (2012) *Neuropsychological assessment, 5th ed*, Oxford University Press, New York, NY, US.

[62] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2021) *Introduction to Meta-analysis*, John Wiley & Sons.

[63] Cohen J (2013) *Statistical Power Analysis for the Behavioral Sciences*, Academic press.

[64] Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. *BMJ* **327**, 557, 10.1136/bmj.327.7414.557.

[65] Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savović J, Schulz KF, Weeks L, Sterne JAC (2011) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **343**, 10.1136/bmj.d5928.

[66] Deeks JJ, Higgins JP, Altman DG, Group obotCSM (2019) Analysing data and undertaking meta-analyses In *Cochrane Handbook for Systematic Reviews of Interventions*, pp. 241-284, 10.1002/9781119536604.ch10.

[67] Burke JF, Sussman JB, Kent DM, Hayward RA (2015) Three simple rules to ensure reasonably credible subgroup analyses. *BMJ : British Medical Journal* **351**, 10.1136/bmj.h5651.

[68] Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2013) Biostat.

[69] Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, deBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schünemann HJ (2011) GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* **64**, 383-394, 10.1016/j.jclinepi.2010.04.026.

[70] Liao Y-Y, Chen IH, Lin Y-J, Chen Y, Hsu W-C (2019) Effects of Virtual Reality-Based Physical and Cognitive Training on Executive Function and Dual-Task Gait Performance in Older Adults With Mild Cognitive Impairment: A Randomized Control Trial. *Frontiers in Aging Neuroscience* **11**, 162, 10.3389/fnagi.2019.00162.

[71] Liao Y-Y, Tseng H-Y, LIN Y-J, Wang C-J, Hsu W-C (2020) Using virtual reality-based training to improve cognitive function, instrumental activities of daily living and neural efficiency in older adults with mild cognitive impairment. *European Journal of Physical and Rehabilitation Medicine*, 47-57, 10.23736/s1973-9087.19.05899-4.

[72] Karssemeijer EGA, Aaronson JA, Bossers WJR, Donders R, Olde Rikkert MGM, Kessels RPC (2019) The quest for synergy between physical exercise and cognitive stimulation via exergaming in people with dementia: a randomized controlled trial. *Alzheimer's Research & Therapy* **11**, 3, 10.1186/s13195-018-0454-z.

[73] Karssemeijer EGA, Bossers WJR, Aaronson JA, Sanders LMJ, Kessels RPC, Olde Rikkert MGM (2019) Exergaming as a Physical Exercise Strategy Reduces Frailty in People With Dementia: A Randomized Controlled Trial. *Journal of the American Medical Directors Association* **20**, 1502-1508.e1501, https://doi.org/10.1016/j.jamda.2019.06.026.

[74] Amjad I, Toor H, Niazi IK, Pervaiz S, Jochumsen M, Shafique M, Haavik H, Ahmed T (2018) Xbox 360 Kinect Cognitive Games Improve Slowness, Complexity of EEG, and Cognitive Functions in Subjects with Mild Cognitive Impairment: A Randomized Control Trial. *Games for Health Journal* **8**, 144-152, 10.1089/g4h.2018.0029.

[75] Kwan RYC, Liu JYW, Fong KNK, Qin J, Leung PK-Y, Sin OSK, Hon PY, Suen LW, Tse M-K, Lai CKY (2021) Feasibility and Effects of Virtual Reality Motor-Cognitive Training in Community-Dwelling Older People With Cognitive Frailty: Pilot Randomized Controlled Trial. *JMIR Serious Games* **9**, e28400, 10.2196/28400.

[76] Petersen RC (2004) Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine* **256**, 183-194, 10.1111/j.1365-2796.2004.01388.x.

[77] Park J-H, Park J-H (2017) Does cognition-specific computer training have better clinical outcomes than non-specific computer training? A single-blind, randomized controlled trial. *Clinical Rehabilitation* **32**, 213-222, 10.1177/0269215517719951.

[78] Tarnanas I, Tsolakis A, Tsolaki M (2014) Assessing Virtual Reality Environments as Cognitive Stimulation Method for Patients with MCI In *Technologies of Inclusive Well-Being:*

Serious Games, Alternative Realities, and Play Therapy, Brooks AL, Brahnam S, Jain LC, eds. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 39-74, 10.1007/978-3-642-45432-5_4.

[79] van de Weijer SCF, Duits AA, Bloem BR, de Vries NM, Kessels RPC, Köhler S, Tissingh G, Kuijf ML (2020) Feasibility of a Cognitive Training Game in Parkinson's Disease: The Randomized Parkin'Play Study. *European Neurology* **83**, 426-432, 10.1159/000509685.

[80] Park J-H, Liao Y, Kim D-R, Song S, Lim JH, Park H, Lee Y, Park KW (2020) Feasibility and Tolerability of a Culture-Based Virtual Reality (VR) Training Program in Patients with Mild Cognitive Impairment: A Randomized Controlled Pilot Study. *International Journal of Environmental Research and Public Health* **17**, 10.3390/ijerph17093030.

[81] Park J-H (2020) Effects of virtual reality-based spatial cognitive training on hippocampal function of older adults with mild cognitive impairment. *International Psychogeriatrics*, 1-7, 10.1017/S1041610220001131.

[82] Ramnath U, Rauch L, Lambert EV, Kolbe-Alexander T (2021) Efficacy of interactive video gaming in older adults with memory complaints: A cluster-randomized exercise intervention. *PLOS ONE* **16**, e0252016, 10.1371/journal.pone.0252016.

[83] Torpil B, Şahin S, Pekçetin S, Uyanık M (2020) The Effectiveness of a Virtual Reality-Based Intervention on Cognitive Functions in Older Adults with Mild Cognitive Impairment: A Single-Blind, Randomized Controlled Trial. *Games for Health Journal* **10**, 109-114, 10.1089/g4h.2020.0086.

[84] Hwang J, Lee S (2017) The effect of virtual reality program on the cognitive function and balance of the people with mild cognitive impairment. *Journal of Physical Therapy Science* **29**, 1283-1286, 10.1589/jpts.29.1283.

[85] Hughes TF, Flatt JD, Fu B, Butters MA, Chang C-CH, Ganguli M (2014) Interactive video gaming compared with health education in older adults with mild cognitive impairment: a feasibility study. *International Journal of Geriatric Psychiatry* **29**, 890-898, 10.1002/gps.4075.

[86] Park E, Yun B-J, Min Y-S, Lee Y-S, Moon S-J, Huh J-W, Cha H, Chang Y, Jung T-D (2019) Effects of a Mixed Reality-based Cognitive Training System Compared to a Conventional Computer-assisted Cognitive Training System on Mild Cognitive Impairment: A Pilot Study. *Cognitive and Behavioral Neurology* **32**, 10.1097/WNN.0000000000000197.

[87] Man DWK, Chung JCC, Lee GYY (2012) Evaluation of a virtual reality-based memory training programme for Hong Kong Chinese older adults with questionable dementia: a pilot study. *International Journal of Geriatric Psychiatry* **27**, 513-520, 10.1002/gps.2746.

[88] Park J-S, Jung Y-J, Lee G (2020) Virtual Reality-Based Cognitive–Motor Rehabilitation in Older Adults with Mild Cognitive Impairment: A Randomized Controlled Study on Motivation and Cognitive Function. *Healthcare* **8**, 10.3390/healthcare8030335.

[89] Thapa N, Park HJ, Yang J-G, Son H, Jang M, Lee J, Kang SW, Park KW, Park H (2020) The Effect of a Virtual Reality-Based Intervention Program on Cognition in Older Adults with Mild Cognitive Impairment: A Randomized Control Trial. *Journal of Clinical Medicine* **9**, 10.3390/jcm9051283.

[90] Oliveira J, Gamito P, Souto T, Conde R, Ferreira M, Corotnean T, Fernandes A, Silva H, Neto T (2021) Virtual Reality-Based Cognitive Stimulation on People with Mild to Moderate Dementia due to Alzheimer's Disease: A Pilot Randomized Controlled Trial. *International Journal of Environmental Research and Public Health* **18**, 10.3390/ijerph18105290.

[91] Schreiber M (1999) Potential of an Interactive Computer-based Training in the Rehabilitation of Dementia: An Initial Study. *Neuropsychological Rehabilitation* **9**, 155-167, 10.1080/713755596.

[92] Swinnen N, Vandenbulcke M, de Bruin ED, Akkerman R, Stubbs B, Firth J, Vancampfort D (2021) The efficacy of exergaming in people with major neurocognitive disorder residing in long-term care facilities: a pilot randomized controlled trial. *Alzheimer's Research & Therapy* **13**, 70, 10.1186/s13195-021-00806-7.

[93] Boutron I, Tubach F, Giraudeau B, Ravaud P (2004) Blinding was judged more difficult to achieve and maintain in nonpharmacologic than pharmacologic trials. *Journal of Clinical Epidemiology* **57**, 543-550, 10.1016/j.jclinepi.2003.12.010.

[94] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng H-Y, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT (2019) RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, 10.1136/bmj.l4898.

[95] Ware J, John E., Brook RH, Davies-Avery A, Williams KN, Stewart AL, Rogers WH, Donald CA, Johnston SA (1980) Conceptualization and measurement of health for adults in the health insurance study: model of health and methodology. .

[96] Juul J (2005) *Half-real: video games between real rules and fictional worlds*, MIT Press, Cambridge, Mass. London.

[97] Voinescu A, Papaioannou T, Petrini K, Stanton Fraser D (2021) Exergaming for dementia and mild cognitive impairment. *Cochrane Database of Systematic Reviews*, 10.1002/14651858.CD013853.

[98] Brod M, Stewart AL, Sands L, Walton P (1999) Conceptualization and Measurement of Quality of Life in Dementia: The Dementia Quality of Life Instrument (DQoL). *The Gerontologist* **39**, 25-36, 10.1093/geront/39.1.25.

[99] Neguţ A, Matu S-A, Sava FA, David D (2016) Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures: A meta-analytic approach. *Computers in Human Behavior* **54**, 414-424, https://doi.org/10.1016/j.chb.2015.08.029.

[100] Wildenbos GA, Peute L, Jaspers M (2018) Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). *International Journal of Medical Informatics* **114**, 66-75, https://doi.org/10.1016/j.ijmedinf.2018.03.012.

[101] Tuena C, Pedroli E, Trimarchi PD, Gallucci A, Chiappini M, Goulene K, Gaggioli A, Riva G, Lattanzio F, Giunco F, Stramba-Badiale M (2020) Usability Issues of Clinical and Research Applications of Virtual Reality in Older People: A Systematic Review. *Frontiers in Human Neuroscience* **14**, 10.3389/fnhum.2020.00093.

[102] Manera V, Chapoulie E, Bourgeois J, Guerchouche R, David R, Ondrej J, Drettakis G, Robert P (2016) A Feasibility Study with Image-Based Rendered Virtual Reality in Patients with Mild Cognitive Impairment and Dementia. *PLOS ONE* **11**, 10.1371/journal.pone.0151487.

[103] Osorio G, Moffat DC, Sykes J (2012) Exergaming, Exercise, and Gaming: Sharing Motivations. *Games for Health Journal* **1**, 205-210, 10.1089/g4h.2011.0025.

[104] Rheinberg F, Engeser S (2018) Intrinsic Motivation and Flow In *Motivation and Action*, Heckhausen J, Heckhausen H, eds. Springer International Publishing, Cham, pp. 579-622, 10.1007/978-3-319-65094-4_14.

[105] Tuena C, Serino S, Pedroli E, Stramba-Badiale M, Riva G, Repetto C (2021) Building Embodied Spaces for Spatial Memory Neurorehabilitation with Virtual Reality in Normal and Pathological Aging. *Brain Sciences* **11**, 1067.

[106] Pigott TD (2020) Power of Statistical Tests for Subgroup Analysis in Meta-Analysis In *Design and Analysis of Subgroups with Biopharmaceutical Applications*, Ting N, Cappelleri JC, Ho S, Chen D-G, eds. Springer International Publishing, Cham, pp. 347-368, 10.1007/978-3-030-40105-4_17.

[107] Richardson M, Garner P, Donegan S (2019) Interpretation of subgroup analyses in systematic reviews: A tutorial. *Clinical Epidemiology and Global Health* **7**, 192-198, 10.1016/j.cegh.2018.05.005.

[108] Levac DE, Huber ME, Sternad D (2019) Learning and transfer of complex motor skills in virtual reality: a perspective review. *Journal of NeuroEngineering and Rehabilitation* **16**, 121, 10.1186/s12984-019-0587-8.

[109] Studenski S, Perera S, Hile E, Keller V, Spadola-Bogard J, Garcia J (2010) Interactive video dance games for healthy older adults. *The journal of nutrition, health & aging* **14**, 850-852, 10.1007/s12603-010-0119-5.

[110] Jahouh M, González-Bernal JJ, González-Santos J, Fernández-Lázaro D, Soto-Cámara R, Mielgo-Ayuso J (2021) Impact of an Intervention with Wii Video Games on the Autonomy of Activities of Daily Living and Psychological–Cognitive Components in the Institutionalized Elderly. *International Journal of Environmental Research and Public Health* **18**, 10.3390/ijerph18041570.

[111] Brauner P, Ziefle M (2015) in *Internet of Things. User-Centric IoT*, eds. Giaffreda R, Vieriu R-L, Pasher E, Bendersky G, Jara AJ, Rodrigues JJPC, Dekel E, Mandler B Springer International Publishing, Cham, pp. 145-150, 10.1007/978-3-319-19656-5_21.

[112] Peng W, Lin J-H, Crouse J (2011) Is Playing Exergames Really Exercising? A Meta-Analysis of Energy Expenditure in Active Video Games. *Cyberpsychology, Behavior, and Social Networking* **14**, 681-688, 10.1089/cyber.2010.0578.

[113] Witard OC, McGlory C, Hamilton DL, Phillips SM (2016) Growing older with health and vitality: a nexus of physical activity, exercise and nutrition. *Biogerontology* **17**, 529-546, 10.1007/s10522-016-9637-9.

[114] Acree LS, Longfors J, Fjeldstad AS, Fjeldstad C, Schank B, Nickel KJ, Montgomery PS, Gardner AW (2006) Physical activity is related to quality of life in older adults. *Health and Quality of Life Outcomes* **4**, 37, 10.1186/1477-7525-4-37.

[115] Wajman JR, Cecchini MA, Bertolucci PHF, Mansur LL (2019) Quanti-qualitative components of the semantic verbal fluency test in cognitively healthy controls, mild cognitive impairment, and dementia subtypes. *Applied Neuropsychology: Adult* **26**, 533-542, 10.1080/23279095.2018.1465426.

[116] Hall JR, Vo HT, Johnson LA, Barber RC, O'Bryant SE (2011) The Link between Cognitive Measures and ADLs and IADL Functioning in Mild Alzheimer's: What Has Gender Got to Do with It? *International Journal of Alzheimer’s Disease* **2011**, 276734, 10.4061/2011/276734.

[117] Flynn D, van Schaik P, Blackman T, Femcott C, Hobbs B, Calderon C (2003) Developing a Virtual Reality–Based Methodology for People with Dementia: A Feasibility Study. *CyberPsychology & Behavior* **6**, 591-611, 10.1089/109493103322725379.

[118] Yun SJ, Kang M-G, Yang D, Choi Y, Kim H, Oh B-M, Seo HG (2020) Cognitive Training Using Fully Immersive, Enriched Environment Virtual Reality for Patients With Mild Cognitive Impairment and Mild Dementia: Feasibility and Usability Study. *JMIR Serious Games* **8**, e18127, 10.2196/18127.

[119] Triegaardt J, Han TS, Sada C, Sharma S, Sharma P (2020) The role of virtual reality on outcomes in rehabilitation of Parkinson's disease: meta-analysis and systematic review in 1031 participants. *Neurological Sciences* **41**, 529-536, 10.1007/s10072-019-04144-3.

[120] Lina C, Guoen C, Huidan W, Yingqing W, Ying C, Xiaochun C, Qinyong Y (2020) The Effect of Virtual Reality on the Ability to Perform Activities of Daily Living, Balance During Gait, and Motor Function in Parkinson Disease Patients: A Systematic Review and Meta-Analysis. *American Journal of Physical Medicine & Rehabilitation* **99**, 10.1097/PHM.0000000000001447.

[121] Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M (2018) Virtual Reality for Stroke Rehabilitation. *Stroke* **49**, 10.1161/STROKEAHA.117.020275.

[122] Rasovska H, Rektorova I (2011) Instrumental activities of daily living in Parkinson's disease dementia as compared with Alzheimer's disease: Relationship to motor disability and cognitive deficits: A pilot study. *Journal of the Neurological Sciences* **310**, 279-282, https://doi.org/10.1016/j.jns.2011.07.048.

[123] Martyr A, Clare L (2012) Executive Function and Activities of Daily Living in Alzheimer's Disease: A Correlational Meta-Analysis. *Dementia and Geriatric Cognitive Disorders* **33**, 189-203, 10.1159/000338233.

[124] Dove E, Astell A (2017) The Kinect Project: Group motion-based gaming for people living with dementia. *Dementia* **18**, 2189-2205, 10.1177/1471301217743575.

[125] Jing W, Willis R, Feng Z (2016) Factors influencing quality of life of elderly people with dementia and care implications: A systematic review. *Archives of Gerontology and Geriatrics* **66**, 23-41, https://doi.org/10.1016/j.archger.2016.04.009.

[126] Sivertsen H, Bjørkløf GH, Engedal K, Selbæk G, Helvik AS (2015) Depression and Quality of Life in Older Persons: A Review. *Dementia and Geriatric Cognitive Disorders* **40**, 311-339, 10.1159/000437299.

[127] Witmer BG, Singer MJ (1998) Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* **7**, 225-240, 10.1162/105474698565686.

[128] Rajji TK (2019) Impaired brain plasticity as a potential therapeutic target for treatment and prevention of dementia. *Expert Opinion on Therapeutic Targets* **23**, 21-28, 10.1080/14728222.2019.1550074.

[129] Barban F, Mancini M, Cercignani M, Adriano F, Perri R, Annicchiarico R, Carlesimo GA, Ricci C, Lombardi MG, Teodonno V, Serra L, Giulietti G, Fadda L, Federici A, Caltagirone C, Bozzali M (2017) A Pilot Study on Brain Plasticity of Functional Connectivity Modulated by Cognitive Training in Mild Alzheimer's Disease and Mild Cognitive Impairment. *Brain Sciences* **7**, 10.3390/brainsci7050050.

[130] Jackson D, Turner R (2017) Power analysis for random-effects meta-analysis. *Research Synthesis Methods* **8**, 290-302, 10.1002/jrsm.1240.

[131] Wettenbank Overheid.nl, Leerplichtwet 1969, https://wetten.overheid.nl/BWBR0002628/2020-04-01#Paragraaf2

[132] Baek L, Parc J (2010) Korean Education and Foreign Direct Investment: Focusing on Entry Modes. *Journal of International Business and Economy* **11**, 89-111, 10.51240/jibe.2010.1.5.

[133] World Education News and Reviews W, Education in South Africa, https://wenr.wes.org/2017/05/education-south-africa.

[134] Council of Higher Education SiT, Turkish Higher Education System, https://studyinturkey.gov.tr/StudyinTurkey/ShowDetail?rID=Ec/rgHEN8Zg=&&cId=PE4Nr0mMoY4=.

Table 1. Summary characteristics of included studies

Study	Group	N[a]	Age, Mean (*SD*)	Sex (f:m)	Education (years) Mean (*SD*)	MMSE[a], Mean (*SD*)	Intervention	Duration (total minutes)[d]	Cognitive domains (measures)	Setting	Time of follow up
						Mild Cognitive Impairment					
Amjad et al. [74]	VR	22	62.80 (5.08)	N/R	N/R	22.68 (1.34)	Body and Brain Exercises, by Dr. Kawashima (XBOX Kinect)	27,000	Global cognition (MMSE, MoCA)	Hospital	N/A
	AC	22	65.56 (5.00)	N/R	N/R	22.83 (2.09)	Motion exercise: Joint range of motion and stretching exercises of upper and lower limbs	27,000	Attention (TMT- A+B)	Hospital	
Hwang & Lee [84]	VR	12	74.10 (6.00)	8:4	N/R	22.4 (0.7)	Not specified	12,000	Attention (VST, WST)	N/R	N/A
	AC	12	70.10 (5.30)	9:3	N/R	22.3 (0.7)	Traditional occupational therapy	12,000		N/R	
Hughes et al. [85]	VR	10	78.50 (7.10)	8:2	13.8 (2.4)	27.2 (1.9)	Boom Blox, Wii Play & Sports Resort (Nintendo Wii)	51,840	Global cognition (CAMCI)	Church located within the study area	N/A

	AC	10	76.20 (4.30)	6:4	13.1 (1.9)	27.1 (1.8)	Cognitive stimulation matching Wii cognitive simulation (Healthy aging education program)	51,840	ADL (TIAODL)	Church located within the study area	
Liao et al. [70, 71]	VR	21	75.50 (5.20)	11:7	9.3 (3.8)	27.2 (1.9)	Custom application & Job Simulator by Owlchemy labs (VIVE HTC, Microsoft Kinect)	77,760	Attention (TMT-A+B, D-TMT, SCWT)	N/R	N/A
	AC	21	73.10 (6.80)	12:4	9.9 (2.1	27.2 (1.6)	Combined physical and cognitive training	77,760	Memory (VLT) Global cognition (MoCA) Executive function (ExInt-25) ADL (AODLS)	N/R	
Park et al. [80]	VR	12	71.80 (6.61)	7:3	7.2 (3.61)	25.3 (2.41)	Custom application, (VIVE HTC, Microsoft Kinect)	17,280	Global cognition (K-MMSE)	N/R	N/A
	PC	12	69.45 (7.45)	6:4	8 (2.9)	26.18 (1.78)	N/A	N/A	Attention (DS-FB, SCWT	N/R	

									Executive function (WFT)		
Park & Park [77]	VR	39	66.95 (4.10)	19:20	8.54 (4.25)	26.41 (1.94)	Wii Sports Resort, (Nintendo Wii)	27,000	Attention (DS-FB, TMT-B, SCWT)	N/R	N/A
	AC	39	67.64 (4.55)	17:22	8.74 (4.51)	26.67 (1.68)	Computerised Cognitive Training (CCT) (CoTras)	27,000	Memory (RAVLT) Construction and motor performance (ROCFT, WAIS-RBDT) QOL (HR-QOL)	N/R	
Tarnanas et al. [78]	VR	39	70.50 (4.30)	20:12	Not specified	26.8 (3.6)	Virtual Museum	144,000	Memory (RAVLT)	N/R	N/A
	AC	39	69.70 (4.50)	23:16	Not specified	26.2 (3.6)	computerised cognitive training, DVD-based educational programs on history, art and literature, or participated at puzzle solving exercises	144,000	Construction and motor performance (ROCFT) Attention (DS-FB, SCWT, TMT-B)	N/R	
	PC	36	70.90	21:13	Not	26.2	N/A	N/A		N/A	

			(4.40)		specified	(3.1)			Global cognition (MMSE) Language (BNT) Executive function (CLF)		
Man et al. [87]	VR	20	80.30 (1.21)	17:3	Not specified[b]	21.5 (3.79)	Custom application- home and convenience store for memory training	3,000	Memory (FOME, MMQ)	N/R	N/A
	AC	24	80.28 (1.31)	22:2	Not specified[b]	23.00 (3.96)	Psycho-educational programme similar to VR	3,000		N/R	
van de Weijer et al. [79]	VR	21	64.65 (7.40)	N/R	N/R	N/R	AquaSnap	N/R	Global cognition (MyCQ, z-scores of SCWT, CLF, RAVLT, ROCFT, BNT, LLT, JLO)	Participants' home	N/A
	PC	20	64.01 (7.41)	N/R	N/R	N/R	N/A	N/R		N/A	
Park et al. [86]	VR	10	70.60 (4.29)	8:2	7.09 (3.36)	26.60 (1.35)	Mixed Reality System for Health	9,720	Executive (VF)	N/R	N/A

	AC	11	73.36 (5.50)	9:2	7.09 (3.36)	26.73 (1.49)	Computerised cognitive training (CCT) (Comcog, by Maxmedica)	9,720	Language (BNT) Memory (WL-LRR) Construction and motor performance (CPR) Attention (TMT-A+B)	N/R	
Park [81]	VR	28	71.93 (3.11)	16:12	8.42 (4.23)	26.71 (1.23)	Bespoke games where participants had to find gems	1,080	Construction and motor performance (WAIS-Block design) Memory (VLT)	N/R	N/A
	PC	28	72.04 (2.42)	17:11	8.78 (4.13)	26.43 (1.52)	N/A	N/A		N/R	
Park et al. [88]	VR	20	75.80 (8.50)	8:10	6.00[e] (N/R)	N/R	MOTOCOG systems with which participants performed activities such as driving, bathing, cooking, and shopping.	900	Global cognition (MoCA) Attention (TMT-A+B, DS-FB)	N/R	N/A
	AC	20	77.20 (7.20)	10:7	6.35[e] (N/R)	N/R	Tabletop activities (puzzles, wood blocks, card play, stick	900		N/R	

							construction activity, maze, and pencil-paper activities)				
Ramnath et al. [82]	VR	23	70.80 (4.52)	N/R	11.52[f] (N/R)	24.60 (2.69)	XBOX Kinect Sports	1,440	Attention (SCWT, NB 0, NB 1, NB 2)	Retirement home	N/A
	AC	22	74.14 (5.80)	N/R	10.82[f] (N/R)	25.00 (2.67)	Group-based low intensity multimodal supervised exercise	1,440	Global cognition (MMSE)	Retirement home	
Torpil et al. [83]	VR	32	70.12 (2.57)	19:11	13.47[g] (N/R)	N/R	Exergame in XBOX Kinect	1,080	Global cognition (LOTCA-G)	N/R	N/A
	AC	32	70.30 (2.73)	17:14	13.55[g] (N/R)	N/R	N/R	1,080		N/R	
Kwan et al.[75]	VR	9	73.44 (4.77)	8:1	6.67[e] (N/R)	N/R	Bespoke application with multiple mini games and exergames	480	Global cognition (MoCA)	Elderly community centre	N/A
	AC	8	77.25 (8.40)	7:1	6.65[e] (N/R)	N/R	Under-desk ergometer (DeskCycle 2) and cognitive games (Card Pairs, Mind Game Double Memory, Flashcard Maths, and Mind Game Double	480		Elderly community centre	

							Connect the dots				
Thapa et al. [89]	VR	34	72.60 (5.40)	28:6	9.30 (4.00)	26.00 (1.80)	Bespoke application that consisted of four mini-games and exergames: Juice making, Crow Shooting, Fireworks and Love house	2,400	Global cognition (MMSE) Attention (TMT-A+B, SDST)	N/R	N/A
	AC	34	72.70 (5.60)	24:10	8.40 (3.50)	26.30 (3.30)	Educational program on general health care	720-1,200		N/R	
							Dementia				
Karssemeijer et al. [72, 73]	VR	38	79.0 (6.9)	18:20	13.22[c] (N/R)	22.9 (3.4)	Combined cognitive -aerobic bicycle training (by Bike Labyrinth)	38,880-64,800	Attention (TMT-A+B, SCWT, WAIS-DS, WMS-SS)	Community centres	24 weeks
	AC 1	38	80.9 (6.1)	17:21	13.09[c] (N/R)	22.5 (3.1)	Aerobic training	38,880	Executive function (LF, RSCT)	Community centres	
	AC 2	39	79.8 (6.5)	18:21	13.43[c] (N/R)	21.9 (3.1)	Relaxation and flexibility exercises	38,880	Memory (LLT) ADL (KADL)	Community centres	

Schreiber et al. [91]	VR	7	80.86 (4.60)	5:2	N/R	22.14 (2.97)	Custom application “MultiTask” in which participants had to find specific targets and rooms	300	Memory (NAI, RBMT)	N/R	N/A
	AC	7	78.86 (6.72)	6:1	N/R	19.86 (4.56)	Social stimulation	300		N/R	
Swinnen et al. [92]	VR	28	84.70 (5.60)	18:5	N/R	18.00 (4.4)	Exergames played in the “Dividat Senso”	375	Global cognition (MoCA)	N/R	N/A
	AC	27	85.30 (6.50)	17:5	N/R	17.00 (4.2)	Music videos on TV screen	375	QOL (DQoL) ADL (KADL)	N/R	
Oliveira et al. [90]	VR	10	82.6 (5.42)	7:3	Not specified[b]	18.60 (6.48)	Systemic Lisbon Battery	540	Executive function (FAB)	Residential care homes	N/A
	PC	7	84.14 (6.30)	5:2	Not specified[b]	13.00 (7.53)	N/A	N/A	Global cognition (MMSE) ADL (IADL)	Care units	

Notes: VR virtual reality, AC active control, PC passive control, N/R not reported, N/A not applicable, SD standard deviation, f female, m male, MMSE Mini-Mental State Examination, MoCA Montreal cognitive assessment, CCT computerised cognitive training, TMT- A+B Trail making test A

+ B, VST Visual span test, WST word-colour test, CAMCI Computer Assessment of Memory and Cognitive Impairment, TIAODL Timed instrumental activities of daily living, D-TMT Delta trail making test, SCWT Stroop colour-word test, VLT Verbal learning test, ExInt-25 Executive interview-25, AODLS Activities of daily living scale, DS-FB Digit span forward and backword, WFT Word fluency test, RAVLT Rey auditory verbal learning test, HR-QOL Health related quality of life test, ROCFT Rey-Osterrieth complex figure test, WAIS-RBDT Wechsler adult intelligence scale – Revised block design test, K-MMSE Korean Mini-Mental State Examination, BNT Boston Naming Test, CLF Category and letter fluency, FOME Fuld object memory evaluation, MMQ Multifactorial memory questionnaire, LLT Location leaning test, JLO Judgment of line orientation, MyCQ My cognition quotient, VF Verbal fluency, WL-LRR Word list- learning-recall-recognition, CPR Constructional praxis and recall, LF Letter fluency, WAIS-DS Wechsler adult intelligence scale – Digit span, WMS-SS Wechsler memory scale – Spatial span, RSCT Rule shift card test, NAI Nuremberg aging inventory, RBMT Rivermead behavioural memory test, NB 0 N-Back task 0, NB 1 N-Back task 2, NB 2 N-Back task 2, LOTCA-G LOTCA-G Loewenstein Occupational Therapy Cognitive Assessment-Geriatric, SDST Symbol digit substitution test, FAB Frontal Assessment Battery, DQoL Dementia Quality of Life questionnaire, KADL Katz ADL Index, IADL Lawton–Brody Instrumental Activities of Daily Living Scale.

[a] Sample size N and MMSE are at baseline.

[b] Education was provided by authors but in a format that was not possible to extract.

[c] Years of education were imputed based on an approximation of the educational years in Netherlands (8 years of primary school, 4-6 years of middle school, 3 years Bachelor, 1-2 years master's degree and 4 years PhD) provided by Wettenbank Overheid.nl [131].

[d] More details can be found in Appendix D.

[e] Years of education were imputed based on an approximation of the educational years in Korea (6 years of primary school, 6 years of middle and high school, 4 years of university) [132].

[f] Years of education were imputed based on an approximation of the educational years in South Africa (9 years of primary school, 5 years of secondary school, 3 years of university) [133].

[g] Years of education were imputed based on an approximation of the educational years in Turkey (4 years of primary school, 4 years of middle school, 4 years of high school, 4 years of university) [134].

Table 2. Subgroup analysis for combined cognitive functioning outcomes with categorical variables for types of intervention, types of controls and MCI diagnosis criteria (mixed effects model).

Outcome	Moderator	*K*	*g*	*P*	I^2	*95%* CI	Q_b	*p*
Cognitive functioning	Active controls/	12	0.98	<.001	75.10	[0.54, 1.33]	0.57	.44
	Passive controls	3	1.44	.02	87.57	[0.21, 2.67]		
	Immersive VR/	5	0.48	.004	4.50	[0.15, 0.81]	7.07	.01*
	Non-immersive VR	10	1.22	<.001	79.78	[0.78, 1.65]		
	Broad MCI diagnosis criteria/	8	1.04	.002	82.38	[0.39, 1.68]	0.05	.83
	Rigorous MCI diagnosis criteria	8	1.13	<.001	78.63	[0.67, 1.59]		
	First-person perspective/	10	0.78	<.001	66.73	[0.42, 1.13]	0.49	.49
	Third-person perspective	3	1.17	.03	83.11	[0.12, 2.22]		
	Computerised applications in control group /	4	0.60	.09	79.06	[-0.09, 1.28]	1.83	.18
	Non-computerised applications in control group	8	1.21	<.001	79.72	[0.65, 1.77]		
	Commercial	8	1.02	<.001	73.36	[0.59, 1.44]	0.13	.71
	Bespoke	6	0.86	.02	84.82	[0.13, 1.59]		
	Exergame	5	0.95	.001	77.20	[0.36, 1.53]	10.38	.016*
	Game	2	2.03	<.001	63.69	[1.19, 2.86]		

	Simulation	4	0.95	.002	72.60	[0.34, 1.57]
	Mixed	4	0.52	.01	21.47	[0.12, 0.92]

Notes. k, Number of studies included in the analysis; *g,* Hedge's *g*; 95% CI, 95% Confidence interval around the weighted mean effect size; I^2, Heterogeneity within study, Q_b, Heterogeneity between studies, * indicates significance.

Table 3. Meta-regression analysis for combined cognitive functioning outcomes with numeric variables for education and duration (mixed effects model).

Outcome	**Moderator**	***k***	β	***SE***	**95% CI**	***Z***	***p***
Cognitive functioning	Education	11	0.006	0.088	[-0.17, 0.18]	0.07	.95
	Duration of intervention sessions (in total minutes)	15	0.00003	0.0002	[-0.0004, 0.0005]	0.13	.90

Notes. k, number of studies included in the analysis; 95% CI, 95% confidence interval around the weighted mean effect size; *SE*, standard error; β, meta-regression coefficient; *Z*, value for testing statistical significance for one coefficient.