virtual worlds

MDPI

Perspective

Beyond the Testing Room: Virtual Reality as a Paradigmatic Solution to Ecological Validity Deficits in Neuropsychological Memory Assessment

Ninette Simonian and Nicco Reggente *

Institute for Advanced Consciousness Studies, 2811 Wilshire Blvd., Suite 510, Santa Monica, CA 90403, USA
* Correspondence: nicco@advancedconsciousness.org

Abstract

Traditional neuropsychological memory assessments lack ecological validity and often fail to capture how memory functions in everyday life. This limits early detection of cognitive decline and reduces correspondence with patient complaints and caregiver observations. We argue that virtual reality (VR) mostly addresses these limitations. VR-based assessments immerse individuals in naturalistic environments that engage authentic cognitive processing while maintaining experimental control. We review empirical evidence demonstrating that VR assessments show superior diagnostic sensitivity for distinguishing healthy aging from mild cognitive impairment (MCI) and Alzheimer's disease (AD), particularly through tasks that integrate memory with spatial navigation and executive function. VR-derived performance metrics also correlate more strongly with subjective experiences and caregiver reports than traditional tests. We propose that VR represents a fundamental reconceptualization of memory assessment, though challenges regarding standardization and accessibility must be addressed.

Keywords: virtual reality; ecological validity; memory assessment; neuropsychology; MCI; Alzheimer's Disease; spatial navigation; cognitive screening

1. Introduction

Memory assessment is an integral part of clinical neuropsychology, informing diagnosis, treatment planning, and prognosis across a wide range of neurological and psychiatric conditions. Clinicians rely on memory evaluation to detect cognitive decline, monitor disease progression, and assess functional capacity in patients with subjective memory complaints, suspected dementia, traumatic brain injury, stroke, and other conditions affecting cognition [1,2]. The public importance of accurate memory assessment is underscored by widespread concern about age-related cognitive decline: approximately 31% of US adults identify Alzheimer's disease (AD) as their most feared diagnosis [3]. Early detection of neurocognitive disorders fundamentally shapes treatment efficacy and patient outcomes, making reliable and valid memory assessment essential for clinical practice [1].

Despite decades of development, traditional neuropsychological memory assessments have been criticized for limitations in standardization, cultural adaptation, and alignment with both subjective experience and modern understanding of brain function [4]. This review focuses on ecological validity as a generative constraint—the methodological gap from which these alignment failures derive. We argue that ecological validity deficits propagate through a causal chain: assessments conducted in artificial, distraction-free environments

Check for updates

Academic Editors: Kenneth Y. T. Lim and Michael Vallance

Received: 20 October 2025
Revised: 20 January 2026
Accepted: 26 January 2026
Published: 2 February 2026

fail to capture real-world cognitive demands, which limits their scope and capacity to differentiate clinical populations, which in turn produces poor correspondence with subjective complaints and caregiver observations, while simultaneously remaining vulnerable to compensatory strategies that mask underlying deficits. These four limitations—ecological validity, diagnostic differentiation, subjective alignment, and compensatory susceptibility—structure our analysis.

We review virtual reality technology as a paradigmatic solution to this constellation of constraints. VR-based assessments immerse individuals in interactive environments simulating real-world cognitive demands—such as navigating a store, preparing a meal, and managing a schedule—while maintaining experimental control necessary for standardized measurement. This approach is particularly consequential for early detection: traditional diagnostic criteria define MCI by memory deficits with preserved Activities of Daily Living [5,6], yet VR studies indicate that functional impairment in complex Instrumental ADLs may characterize predementia when assessed under ecologically valid conditions [5]. We synthesize empirical evidence demonstrating VR's superior diagnostic sensitivity for distinguishing healthy aging from MCI and AD, stronger correlation with patient and caregiver reports, and resistance to compensatory masking—advantages detailed in Sections 2–4, respectively.

2. Shortcomings of Traditional Tests

2.1. Limited Ecological Validity

Common screening instruments (e.g., Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), and Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-COG)) create a disconnect between evaluation methods and real-world cognitive demands. Similarly, standardized instruments such as the Auditory Verbal Learning Test (AVLT; [4]) and the California Verbal Learning Test (CVLT; [7]) assess isolated cognitive functions in controlled, distraction-free environments [8]. These conditions do not reflect the complexity of real-world scenarios. These measures show deficiencies in verisimilitude (how well tasks mirror real-world contexts) and veridicality (how well performance predicts real-life outcomes) [9–11]. Both are central to ecological validity.

Word-list assessments exemplify this disconnect. They require memorizing unrelated word sequences, a task rarely encountered in daily life [12]. Deficits in this isolated skill may have a limited impact on everyday functioning. Traditional assessments overlook practical tasks such as cooking, shopping, and medication management [13–15]. These skills are often disrupted early in cognitive decline, particularly among older adults.

By removing contextual support, artificial testing environments create incomplete representations of functional memory [16]. This generates discrepancies between clinical and real-world performance. While these measures can detect impairment [17], their limited ecological validity reduces insight into how deficits manifest in real-world contexts. Traditional test scores show modest and inconsistent correlations with everyday functioning measures [12,18–21].

2.2. Bounded Scope and Diagnostic Differentiation

Traditional memory assessments are conducted in artificial settings and capture only a narrow range of memory functioning, failing to reflect the complexity of everyday cognition and thereby limiting their clinical utility. This restricted scope is problematic for differentiating memory problems. For example, distinguishing MCI from dementia or identifying subtle decline along the aging-to-pathology continuum requires detecting gradations that traditional tests often miss [6]. These distinctions determine whether

https://doi.org/10.3390/virtualworlds5010007

patients receive watchful waiting, cognitive training, pharmacological intervention, or comprehensive care planning.

Standard neuropsychological tools fail to capture how memory actually operates in complex, real-world environments. These assessments tend to isolate cognitive skills under ideal conditions, overlooking distractions, multitasking, and contextual cues that influence everyday memory performance. Consequently, they offer an incomplete and sometimes misleading picture of an individual's functional capacity, especially in the early stages of cognitive decline. Traditional assessments also remain susceptible to compensatory strategies, a limitation addressed in Section 2.4. In practice, patients' adaptive behaviors, like using reminders or routines, are clinically meaningful yet unmeasured.

Efforts like the Rivermead Behavioral Memory Test (RBMT) have attempted to bridge this gap by integrating real-world tasks such as remembering routes, messages, and appointments [22]. However, even these advances fall short by lacking sufficient complexity or breadth, particularly in areas like prospective memory [23], which are critical for daily functioning. Many assessments also struggle with diagnostic accuracy, either failing to identify individuals with MCI or misclassifying healthy adults as impaired [24]. This diagnostic uncertainty stems largely from the constrained scope of traditional assessments, which fail to capture the episodic, contextual, and adaptive nature of real-world memory functioning.

2.3. Disconnect from Subjective and Caregiver Reports

Subjective complaints and caregiver observations theoretically represent the ecologically valid endpoints that memory assessment should ultimately predict. Traditional memory assessments align poorly with patients' subjective experiences of cognitive decline [25]. Despite producing "objective" results, conventional tests often fail to capture the memory failures that concern patients in daily life.

2.3.1. Subjective Memory Complaints

Pearman and Storandt [25] found little correspondence between standard memory test scores and subjective memory complaints among healthy older adults, suggesting that what individuals notice and report—forgetting appointments, misplacing everyday items, losing conversational threads—is often unrelated to what conventional assessments actually measure [25]. Chaytor and Schmitter-Edgecombe [12] and Reid and MacLullich [26] argue that subjective complaints are frequently dismissed in clinical settings because they are not reflected in test performance.

This disconnect is problematic because subjective complaints may be among the earliest indicators of cognitive decline, particularly in prodromal AD. Luo and Craik [27] found that detectable differences between healthy aging and AD often emerge only in later stages. Current tools may therefore miss early deficits that patients perceive. Clinicians face considerable difficulty distinguishing normal aging from early pathology, or even mood-related memory lapses, during initial evaluation [28]. This challenge is compounded by the fact that memory, executive function, and functional capacity are complementary domains, key to diagnosing neurocognitive disorders [29]. Hodges [30] identified episodic memory and spatial orientation as early markers of AD, yet traditional assessments often fail to capture deficits in these domains until later stages.

2.3.2. Caregiver Reports

Caregiver reports are a staple of cognitive assessment, particularly in dementia evaluations where patient self-report becomes unreliable, offering valuable insight into daily functioning, symptom trajectory, and the broader impact of cognitive decline on family systems [31,32]. However, caregiver reports present notable limitations when used as a proxy for neuropsychological assessment, stemming from four distinct mechanisms. First, caregiver burden and psychological state: emotional distress, depression, and caregiving burden significantly bias symptom recognition and reporting [32], with greater negative affect predicting more reported prospective memory failures [33] and caregiver anxiety producing discrepancies with clinician impressions [34]. Second, caregiver cognitive status: caregivers' own executive functioning predicts reporting accuracy, with lower cognitive scores yielding less accurate ADL ratings—particularly as "role reversal" dynamics mean caregivers infer rather than observe patient capabilities [32]. Spousal caregivers themselves show reduced cognitive functioning compared to matched controls, revealing bidirectional influences within the caregiving dyad [35]. Third, compensatory behavioral dynamics: caregivers who perform tasks for rather than with patients lose visibility into actual functional limitations, creating systematic reporting errors in both directions [36]. Fourth, disease staging effects: caregiver-clinician disagreements peak in mild dementia, precisely when accurate detection matters most, as caregivers often misattribute early neuropsychiatric symptoms to normal aging [34].

Methodological challenges compound these issues: caregiver reports emphasize visible behavioral changes while missing subtle cognitive symptoms [37,38], vary widely in standardization [39], and rely on subjective observational measures that may not reflect true capabilities [40]. Notably, patients with AD report lapses in both prospective and retrospective memory, whereas caregivers weigh their observations toward prospective failures—forgetting appliances, medications, and appointments—given their disruptive and hazardous nature [35]. These are precisely the everyday memory demands that word-list recall tasks fail to assess.

Rather than dismissing patient-caregiver discrepancies as merely reporter unreliability, we suggest these inconsistencies are also revelatory of fundamental limitations in what traditional assessments capture. As Díaz-Orueta et al. [38] argue, contrasting formal results with subjective impressions from caregivers and patients highlights the growing gap between laboratory measures and real-world cognition.

The patient-caregiver relationship itself may be affected by these assessment discrepancies. When patients perceive their memory as intact while caregivers observe daily failures, conflicts may arise regarding autonomy, safety, and care needs [31,41]. Torlaschi et al. [39] found that caregiver burden was significantly predicted by patient neuropsychiatric symptoms and reduced functional autonomy, and was negatively associated with quality of healthcare communication. These findings highlight how cognitive decline is a family-system challenge, with assessment discrepancies potentially contributing to relational tension and suboptimal care coordination.

2.4. Susceptibility to Compensatory Mechanisms

Traditional assessments remain susceptible to compensatory mechanisms that can mask underlying cognitive deficits. This susceptibility represents perhaps the most concerning limitation, as it can lead to false reassurance and delayed diagnosis precisely when early intervention would be most beneficial.

A useful analogy emerges from genetics: just as identical phenotypes can arise from distinct genotypes, equivalent behavioral performance can emerge from fundamentally different underlying neural architectures. Conventional tests often fail to account for this phenotypic equivalence problem. Banville et al. [40] illustrate that individuals with traumatic brain injury may succeed on prospective memory tasks while relying on compensatory mechanisms—slower task execution and reduced precision—to manage increased cognitive load. Similarly, those in preclinical AD stages shift from hippocampal-dependent to striatal-based navigation strategies [42]—adaptations that preserve the behavioral phenotype while fundamentally altering the cognitive substrate being employed.

Neuroimaging offers a window into this genotype-phenotype dissociation. Agosta et al. [43], commenting on Skouras et al. [44], demonstrated that neural compensation occurs in individuals who perform normally on traditional tests: alternative neural regions are recruited when primary circuits begin to fail, maintaining cognitive output despite underlying degradation. The preclinical phase involves complex network reorganization whose fragility remains invisible when examining only performance scores on simplified, domain-specific assessments—assessments whose focused demands actively facilitate compensatory recruitment that would fail under the distributed demands of real-world memory functioning.

Research on spatial navigation exemplifies this phenomenon: Kunz et al. [45] demonstrated that early degradation in grid-cell-like activity in the entorhinal cortex—a region particularly vulnerable to AD—could be behaviorally masked through reliance on alternative neural systems, helping explain why navigation impairments often emerge late in the disease trajectory despite early pathological changes. The neural genotype deteriorates while the behavioral phenotype remains deceptively intact.

These masking effects carry profound clinical implications. Mancuso et al.'s [46] systematic review demonstrated that traditional assessment performance fails to predict functional outcomes such as informant memory diaries or clinician ratings. Furthermore, when success relies on executive functioning or linguistic strategies rather than memory per se [47], assessments inadvertently privilege individuals with higher cognitive reserve or education, exacerbating diagnostic inequities.

Beyond performance metrics lies a fundamental mismatch in cognitive demands: while everyday remembering requires binding what, where, and when details into integrated episodic experiences [48], traditional tests fragment these elements into isolated verbal recall tasks. This artificial decomposition not only misrepresents the phenomenology of memory but also creates assessment conditions optimized for compensatory strategies that would fail under the distributed demands of naturalistic contexts.

3. Addressing Critical Limitations in Traditional Assessment with Virtual Reality (VR)

3.1. Establishing Validity Through Traditional Correlations

VR-based memory assessments show strong correlations with traditional measures, confirming they capture similar cognitive constructs [46]. This creates a paradox: we validate VR against the very tests we critique. Ideally, the benchmark would be real-world criteria like subjective experiences and caregiver reports. However, demonstrating alignment with established measures is a necessary first step before arguing that VR captures memory functioning more accurately.

Barclay et al. [49] conducted a focused review of VR environments developed specifically for memory assessment, analyzing 30 studies across 22 unique virtual environments. They identified object memory, spatial memory, and feature binding as the most frequently assessed components, the same core memory domains targeted by traditional neuropsychological batteries. This systematic analysis confirmed that VR assessments effectively measure established memory constructs while highlighting how methodological variations in navigation style, encoding instructions, and environmental design influence cognitive load and performance. Similarly, Initial validation studies establish VR's convergent validity with established measures while extending beyond conventional approaches. Mancuso et al. [46] analyzed 24 studies exploring the convergent and divergent validity of VR memory tasks compared to traditional measures, finding strong correlations between VR and conventional memory tests.

https://doi.org/10.3390/virtualworlds5010007

Da Costa et al. [13] introduced two immersive VR tasks (the SOIVET Maze and SOIVET Route) to assess spatial orientation in older adults, finding that both tasks correlated with traditional visuoperceptual and executive measures while maintaining high feasibility in older adult participants. This correlation with established measures provides validation that VR assessments capture similar cognitive constructs as traditional methods. Similarly, Kourtesis and MacPherson [50] used the immersive Virtual Reality Everyday Assessment Lab (VR-EAL) to explore how attention, planning, and delayed recognition impact prospective memory. Their findings confirmed the established theoretical frameworks by showing that event-based prospective memory relies on delayed recognition and visuospatial attention, while time-based prospective memory depends more on planning, patterns consistent with traditional assessment findings.

Corriveau Lecavalier et al. [51] validated the "Virtual Shop" for assessing episodic memory in older adults, where participants memorized and retrieved items in a realistic virtual convenience store with background noise and semantic distractors, demonstrating strong construct validity with VR performance significantly correlating with traditional memory and executive tasks. Bottiroli et al. [52] used the Smart Aging Platform, a VR-based "serious game" that demonstrated the ability to distinguish between healthy controls and various clinical groups while correlating with traditional neuropsychological measures. Likewise, Gottlieb et al. [53] adapted the Rey Auditory Verbal Learning Test (RAVLT) into an immersive VR format, maintaining the original test's structure while embedding it in a simulated real-life scenario, with results showing that the VR-RAVLT preserved key features such as serial position effects and demonstrated strong construct and discriminant validity relative to the standard RAVLT.

A growing body of research supports VR as a promising tool for memory assessment, with early studies demonstrating that VR replicates findings from traditional pen-and-paper tests. While these conventional measures account for some variance in cognitive performance and serve as our established benchmark, they are not necessarily the "gold standard" and may only capture a portion of real-world cognitive functioning. By building on this foundation, VR extends assessment beyond the limitations of traditional tools, offering the ability to probe memory and executive functions in ecologically valid, interactive environments that more closely reflect everyday cognitive demands and capture what we truly care about in understanding human memory performance [16,54].

3.1.1. Domain-Specific Efficacy: VR Benefits by Memory Type

Contrary to the view that VR uniformly enhances assessment, evidence also suggests its utility varies significantly depending on the specific memory domain and task demands [46,55–58]. Understanding these distinctions is essential for interpreting VR assessment outcomes and designing targeted diagnostic protocols.

3.1.2. Spatial Memory: Allocentric vs. Egocentric Navigation

VR shows particular advantages for assessing allocentric navigation (world-centered spatial processing), which relies on the medial temporal lobe and hippocampus [59,60]. Standard tests often fail to capture allocentric processing because they lack large-scale environmental continuity. VR tasks have demonstrated that deficits in allocentric navigation, finding a path based on environmental landmarks, often precede deficits in egocentric navigation (body-centered turns) in amnestic MCI [60]. However, as the disease progresses to AD, both systems become impaired [59]. The benefit of VR is specific to tasks requiring path integration, continuously updating one's position based on movement cues, which cannot be replicated by static paper-and-pencil tasks [45].

3.1.3. Episodic Memory: Feature Binding

Traditional word-list tests assess item memory ("what"). VR outperforms these tests specifically in assessing feature binding, the integration of what, where, and when [61]. Plancher et al. [58] found that while older adults might perform comparably to younger adults on simple item recall in VR, they show significant deficits in binding items to spatial and temporal contexts, deficits that standard tests often miss. However, active versus passive exploration affects outcomes. Active navigation generally improves spatial memory through motor integration [56,62]. Yet for older adults or those with executive dysfunction, active navigation can act as dual-task interference, lowering episodic recall compared to passive viewing [53,62].

3.1.4. Prospective Memory: Time- vs. Event-Based Tasks

VR shows differential sensitivity depending on the type of prospective memory trigger, appearing particularly superior for time-based prospective memory (e.g., "check the clock every 5 min") compared to event-based memory [63]. Time-based tasks in VR require self-initiated strategic monitoring while multitasking, such as shopping while remembering to perform a future action, creates cognitive load that mimics real life. Standard clinical tests often lack this multitasking interference, allowing patients to compensate with focused attention and masking deficits that would manifest in everyday settings [22,61].

3.2. Addressing Ecological Validity Limitations

VR addresses ecological validity concerns by immersing individuals in interactive, life-like environments, mediated through head-mounted displays (HMDs) and motion-tracking sensors, eliciting cognitive processing while maintaining strict psychometric rigor [64]. As such, VR assessment tools simulate the complex, multisensory demands of everyday life [38,64]. For instance, in the Virtual Environment Grocery Store (VEGS), participants do not merely recall word lists; they must navigate 3D aisles to locate specific products, manage a budget, and inhibit attention to irrelevant auditory announcements, thereby engaging the complex executive-memory binding required for independent living [10,46]. Similarly, virtual kitchen environments require patients to execute multi-step tasks, such as preparing a cup of coffee or cooking a meal, allowing clinicians to assess sequencing, safety awareness, and object usage in a safe, clutter-free simulation of domestic life [46].

Crucially, VR distinguishes itself from unstructured naturalistic observation through experimental control, the capacity to manipulate environmental variables with a precision impossible in the physical world [64]. In VR, researchers can standardize the timing, intensity, and location of dynamic stimuli, ensuring that every participant encounters the exact same environmental challenges [55]. This control extends to the physical laws of the environment; researchers can enforce invisible boundaries to restrict movement without altering visual cues, or instantaneously teleport participants between contexts to test context-dependent memory, manipulations that are infeasible in physical laboratory settings [65]. By fostering presence, the subjective sensation of "being there", VR elicits naturalistic behaviors and motoric responses that correlate more strongly with real-world functioning than abstract psychometric scores [64]. Consequently, VR captures the fine-grained data on timing, precision, and strategy necessary to detect subtle deficits in contextual and spatiotemporal memory often obscured by compensatory mechanisms in standard assessments [64].

Empirical Evidence: Immersive Features and Experimental Controls

Early validation studies have demonstrated VR's superiority in creating ecologically valid assessment conditions and identifying the specific immersive features and experimental controls that drive diagnostic sensitivity. Wallet et al. [62] demonstrated that the visual fidelity of the environment, specifically the inclusion of realistic textures and colors in a 3D reproduction of the Bordeaux district, offers functional purpose. High-fidelity cues were found to be essential for constructing the egocentric representations (landmark knowledge) required for successful wayfinding and route reproduction in the real world [62]. Conversely, impoverished vertical environments (wireframe models) forced reliance on abstract survey knowledge, which did not transfer as effectively to daily life tasks [62].

The level of immersion and stereoscopic rendering further modulates these cognitive gains. As noted by Smith, stereoscopic displays provide binocular depth cues that, while not always altering spatial configuration memory, can increase "remember" judgments (recollection) for context-consistent objects [55]. This granular impact of immersion extends to the sensory field; Ragan et al. found that increasing the Field of View (FoV) and Field of Regard (FoR) significantly reduced memory errors, suggesting that the sensory envelopment provided by VR aids in binding spatial sequences [66]. Furthermore, Kim et al. [67] found that stereoscopic cues facilitate feedforward motor control, resulting in smoother movement trajectories during interaction with virtual objects. However, they notably observed that this motor advantage was absent in older adults, attributing this to age-related declines in stereopsis [67]. These findings support the argument that while features like stereoscopic rendering directly impact perceptual encoding and motor planning, their diagnostic utility may be mitigated by the sensory capabilities of the aging cohort [67].

Furthermore, VR allows for the isolation of cognitive mechanisms through "matched-pairs" control designs [55]. To control for the confounding variable of motor planning during navigation, studies by Plancher et al. [68] and Jang et al. [69] employed protocols where "active" participants navigated a route freely, while "passive" participants viewed a video recording of that exact route. This ensured that both groups received identical visual stimulation, isolating the cognitive effect of volitional control from perceptual input.

Additionally, VR permits the standardization of dynamic variables that are uncontrollable in the real world. For instance, the "New Zealand shopping mall" environment allowed researchers to systematically divide the street into "low-distraction" and "high-distraction" zones (manipulating visual and auditory noise like traffic and footsteps) to precisely measure the impact of environmental stressors on prospective memory in head-injury patients [38]. This capacity for control extends to complex functional assessments. Díaz-Orueta et al. describe the "Memo Test," a virtual furniture store where the patient acts as a manager [38]. This environment introduces controlled interruptions, such as customers ringing a bell or approaching a counter, requiring the subject to prioritize tasks (e.g., entering codes, preparing orders [38]. This design reveals "frontal" memory deficits, such as difficulties in monitoring and organizing stored content, which are often obscured in conventional testing [38]. Similarly, the VRCPAT (Virtual Reality Cognitive Performance Assessment Test) utilizes a virtual city populated with standardized distractors (animals, parked cars) to measure learning and memory over a fixed 15-min interval, distinguishing clinical populations through strictly controlled exposure to stimuli [38].

Finally, the diagnostic superiority of VR in distinguishing MCI and AD from healthy aging appears driven by the system's ability to decouple memory systems through specific environmental manipulations. Jonson et al. highlight that VR adaptations of tasks like the Morris Water Maze can isolate allocentric processing (landmark-independent navigation) from egocentric strategies [70]. By varying starting positions within a standardized virtual pool or park, these environments force users to rely on distal cues [70]. Since allocentric deficits are often the earliest marker of AD, these controlled environmental manipulations offer diagnostic granularity that static paper-and-pencil tests cannot replicate [70].

https://doi.org/10.3390/virtualworlds5010007

3.3. Expanding Scope and Diagnostic Differentiation

VR immerses participants in realistic, dynamic environments that capture the complexity of everyday cognitive tasks, naturally bridging the gap between laboratory performance metrics and real-world functioning. By simulating lifelike scenarios, such as navigating a virtual grocery store, shopping for target items while ignoring distractors, or responding to time-sensitive announcements, VR mirrors the multitasking, navigation, and time management demands that patients and caregivers identify as challenging in daily life [71].

Beyond replicating real-world tasks, VR offers unparalleled flexibility in creating diverse, tailored environments, from cities and classrooms to supermarkets and homes, designed to target specific cognitive challenges [71–73]. Researchers can manipulate environmental features, introduce distractors, or adjust sensory load to probe subtle cognitive differences that conventional assessments might overlook [57].

Context-dependent memory assessment exemplifies VR's expanded capability. Building on Godden and Baddeley's [74] findings that recall improves when learning and retrieval environments match, VR enables systematic manipulation of contextual cues without the logistical constraints of real-world manipulations. Essoe et al. [75] demonstrated that immersive virtual environments enhance context-dependent memory through both physical and mental reinstatement of learning contexts, with fMRI findings showing that the fidelity of neural contextual reinstatement correlates with performance. VR also offers unique methodological affordances that allow researchers to manipulate environments beyond physical limitations, for example, teleporting participants between settings, imposing invisible barriers, altering visual cues in real time, or shifting perspectives [48]. Such flexibility is especially powerful in spatial navigation research, where infinitely large, novel environments can be generated to distinguish between place-based and sequence-based encoding strategies [48,76].

VR also allows precise assessment of spatial orientation and navigational processes through maze navigation, route learning, and wayfinding tasks [46,71]. Furthermore, VR enables sophisticated assessment of egocentric-to-allocentric spatial transformation, a cognitive process fundamental to memory decline [77–81], yet methodologically challenging to evaluate through conventional means. By presenting participants with surprise memory tests from an allocentric "bird's eye" perspective after they have navigated virtual environments from an egocentric viewpoint (e.g., [82]), researchers can directly probe this critical spatial translation mechanism. This capability reveals a paradoxical advantage: VR surpasses even real-world testing for certain memory processes, as constructing analogous ego-to-allo conversion tasks in physical environments would require prohibitive logistical complexity while offering less experimental control, similar to the difficulties of conducting real-world autobiographical memory studies (e.g., [83]).

Such ease and affordances are particularly valuable for detecting early deficits in pathological aging, as impairments in spatial navigation often precede other cognitive symptoms in Alzheimer's disease (AD) or mild cognitive impairment (MCI) [45]. Studies demonstrate that VR can detect subtle navigation difficulties, capturing both impaired and preserved processes with ecological relevance [56].

3.3.1. Superior Sensitivity and Diagnostic Capabilities

VR-based assessments demonstrate enhanced sensitivity in detecting cognitive impairments compared to traditional measures [70]. Nolin et al. [84] evaluated VR's effectiveness in assessing prospective memory in elderly individuals with MCI compared to cognitively healthy individuals. While traditional prospective memory tasks showed no significant differences between groups, the MCI group performed worse on VR tasks, revealing VR assessments' superior sensitivity to cognitive impairments associated with aging. Further-

https://doi.org/10.3390/virtualworlds5010007

more, VR task performance correlated positively with Montreal Cognitive Assessment scores, while traditional tasks showed no such correlation.

Neguț et al. [85] conducted a meta-analytic review examining the sensitivity of VR-based neuropsychological tools in distinguishing cognitive impairment. Analyzing 18 studies, they found a large overall effect size favoring healthy controls over clinical groups across executive function, memory, and visuospatial abilities, suggesting that VR assessments are highly effective in detecting cognitive deficits. Clay et al. [86] evaluated five studies investigating immersive VR for assessment in AD and MCI, finding that VR assessment is most valuable when grounded in theoretical models of neurodegeneration and designed to test specific neural systems affected early in AD pathology.

These contemporary meta-analytic findings build upon foundational work establishing VR's capacity to dissociate preservation from impairment across spatial processing systems. Prospective memory strength assessed through lifelike VR errand scenarios differentiates older adults from those with mild AD while simultaneously ascertaining functional capacity for real-world task completion [87,88], and VR-based memory tests sensitive to population-specific behaviors have facilitated clinical classification that permits early intervention [61,89]. Seminal case studies demonstrated that an MCI patient exhibited recall deficits when required to change viewpoints despite normal recognition of topographical scenes [90,91]—a dissociation revealing that simpler recognition processes can mask transformation-dependent memory failures. Subsequent work confirmed that intact egocentric navigation alongside impaired allocentric strategies in virtual Morris Water Maze paradigms distinguishes amnestic MCI from AD [92] and differentiates amnesics from controls [93]. This allo-to-egocentric transformation deficit is particularly pronounced in AD, where patients show reduced ability to leverage bird's-eye perspective solutions when subsequently navigating first-person mazes [94], alongside spatial and non-verbal episodic memory impairment during temporally ordered goal navigation [95].

Tarnanas et al. [5] demonstrated that the Virtual Reality Day-Out Task (VR-DOT), which requires executive prioritization during a fire evacuation, predicted conversion from MCI to AD with greater accuracy than standard cognitive measures. The task quantified subtle errors in executive planning and psychomotor speed that standard ADL questionnaires fail to capture. Similarly, Zygouris et al. [96] found that a Virtual Supermarket task achieved 87.3% classification accuracy between healthy older adults and MCI using variables reflecting executive inhibitory control (e.g., purchasing unlisted items) and processing speed.

3.3.2. Neurocognitive Mechanisms: Why VR Engages Vulnerable Neural Circuits

VR's diagnostic sensitivity also stems from its ability to engage neural circuits compromised early in AD, circuits that remain largely untested during static neuropsychological testing [44,45,59].

The Entorhinal Cortex and Grid-Cell Network

VR navigation tasks directly engage the entorhinal cortex (EC), the site of earliest tau pathology in AD [45]. The EC contains grid cells essential for path integration—updating position based on self-motion cues [45]. Kunz et al. [45] demonstrated that reduced grid-cell-like representations in the EC during VR navigation could be detected in young adults at genetic risk for AD (APOE-ε4 carriers) long before performance deficits appeared on standard episodic memory tests. Because paper-and-pencil tests do not require translation of self-motion into a cognitive map, they fail to stress-test the entorhinal grid-cell network, missing the earliest functional markers of disease. The most compelling validation of this mechanism came from Howett et al. [97], who demonstrated that an immersive VR path integration task specifically designed to test entorhinal cortex function could differentiate MCI patients with positive AD biomarkers from those with negative biomarkers and healthy controls with high sensitivity and specificity.

The Retrosplenial Cortex and Mental Frame Syncing

VR tasks challenge the retrosplenial cortex (RSC) by requiring continuous translation between egocentric (body-centered) and allocentric (world-centered) frames of reference [59,77]. This "mental frame syncing" is critical for navigation and is supported by the RSC and posterior cingulate cortex, regions showing hypometabolism early in MCI [59,77]. While standard tests often assess these frames in isolation (e.g., drawing a map from a fixed perspective), VR navigation forces rapid, real-time switching and integration [77]. Consequently, VR detects "heading disorientation" and translation deficits that characterize early AD but are masked in static testing formats [59]. Park [98] provided empirical support for this mechanism, reporting that a Spatial Cognitive Task using VR (SCT-VR) outperformed the Montreal Cognitive Assessment in distinguishing MCI from healthy aging, achieving 94.4% sensitivity and 96.4% specificity, diagnostic accuracy attributed specifically to the task's engagement of allocentric spatial representation, a hippocampal-dependent function compromised early in MCI.

Hippocampal Feature Binding

VR environments impose cognitive load that taxes feature-binding mechanisms in the hippocampus. Unlike standard tests assessing item memory in isolation, realistic VR scenarios require simultaneous binding of "what" (objects), "where (spatial location), and "when" (temporal sequence) [58,61]. Plancher et al. found that while older adults may preserve item memory in VR, they show specific deficits in this multimodal binding process [58]. By simulating these complex binding demands, VR exposes hippocampal inefficiencies, specifically the inability to downregulate hyperactivity during complex task performance, that correlate with early amyloid and tau deposition [44]. Lecouvey et al. [88] further demonstrated this sensitivity using a VR town-navigation task to examine prospective memory in individuals with mild AD. Participants completed both time-based and event-based intentions with varying cue-action link strengths; AD patients showed pronounced deficits across all intention types, especially in retrospective components of prospective memory, the memory for what one intended to do, demonstrating VR's effectiveness in capturing the nuanced binding deficits that characterize hippocampal dysfunction.

3.3.3. Beyond Memory Assessment

The potential of VR as a diagnostic platform extends beyond memory assessment to other domains [9,99–101]. This technology has enabled the development of "serious games," immersive simulations that support not only diagnosis but also prognosis and treatment of cognitive impairments, particularly through tasks that resemble everyday activities [102]. For example, simulated meal preparation tasks or spatial navigation challenges allow for multidimensional assessment of memory, executive function, and attention in realistic contexts, while simultaneously offering therapeutic potential [59,77,103]. Similarly, virtual renditions of classic mnemonics (e.g., virtual memory palaces [91]) could be adapted to incidentally assist individuals with remembering non-spatial information by arranging it on the scaffolding of spatial memory, which benefits from increased evolutionary prowess [25].

Recent advances in AI-integrated VR assessments offer a glimpse into the future of cognitive diagnostics, demonstrating the potential to go beyond static measurements and toward dynamic, predictive modeling. For example, Altozano et al. [104] developed a system that uses VR and deep-learning algorithms to detect autism spectrum disorder (ASD) in young children with over 85% accuracy, outperforming traditional assessment methods. This system analyzes motor movements and gaze patterns within immersive virtual environments that mirror everyday situations, capturing naturalistic responses rather than the artificial behaviors often observed in controlled laboratory settings. This model of embedding AI-driven analysis within ecologically valid virtual tasks represents a powerful blueprint for what memory prognosis could become: a future in which early, affordable, and precise detection of cognitive decline, such as in AD or MCI, is possible through continuous monitoring of behavior in lifelike scenarios, rather than isolated test scores. As with ASD, leveraging biomarkers like motor behavior, attentional shifts, and spatial navigation within VR could unlock richer, more holistic understandings of cognitive trajectories.

VR also holds the potential to offer a more personalized, nuanced picture of cognitive health, capturing subtle deficits while generating rich behavioral data that connect with how people feel and function in daily life [25]. This shift represents a critical step toward closing the ecological gap in memory assessment and making cognitive evaluation more meaningful and accurate in clinical practice.

3.4. Connecting to Subjective and Caregiver Reports

VR addresses the limitations of subjective and informant reporting by providing an objective, observable platform that captures the contextual complexity individuals encounter in daily life.

3.4.1. VR as Objective Benchmark for Functional Capacity

VR-generated data can complement caregiver reports by providing objective, performance-based assessments in simulated everyday environments. Allain et al. [8] developed a virtual kitchen task in which performance on a coffee-making simulation was significantly correlated with caregiver ratings on the Lawton-Brody IADL scale. Notably, the VR task outperformed traditional cognitive assessments in predicting functional status, offering a more nuanced and ecologically valid picture of everyday functioning. By immersing patients in realistic scenarios, virtual kitchens, grocery stores, and household environments, clinicians can directly observe specific IADL deficits that are often subjects of dispute between patients and caregivers. This objective benchmarking validates or challenges caregiver reports, helping resolve discrepancies caused by the "do for" dynamic or caregiver burden [36].

Pieri et al. [105] highlight that VR-based tools can simulate complex IADL tasks such as cooking, managing a household, or grocery shopping, allowing clinicians to observe behavior in real time. These immersive scenarios not only enhance ecological validity but can also serve to objectively validate or challenge caregiver reports, uncovering preserved abilities or undetected impairments that paper-based tools may overlook. Knight and Titov [87] demonstrate how VR creates environments where patients must autonomously coordinate multiple cognitive processes to complete everyday memory tasks that mirror real-world challenges reported by patients and caregivers. Their analysis shows how VR platforms assess prospective memory-relevant domains through tasks like remembering to check doors, responding to environmental cues, and following medication regimens.

3.4.2. Superior Alignment with Subjective Experiences

Research demonstrates that VR captures subjective cognitive experiences more effectively than traditional assessments. Cushman et al. [106] showed VR tasks' sensitivity to progressive navigation deficits across the spectrum from normal aging through MCI to early AD, concluding that VR enables assessment of cognitive functions in ways directly relevant to the lived experiences of patients with pathological aging. Plancher et al. [58] demonstrated that VR memory tasks successfully detected subjective memory complaints and age-related cognitive changes in healthy older adults that traditional tests missed.

In a study examining episodic memory across healthy older adults, patients with amnestic MCI, and individuals with AD, Plancher et al. [68] found that AD patients performed significantly worse than both MCI and healthy participants, with allocentric spatial memory being especially effective in distinguishing MCI from controls. Critically, performance on the VR task correlated more strongly with participants' everyday memory complaints than conventional tests in both healthy older adults and individuals with MCI, highlighting superior ecological relevance. This suggests VR captures the multifaceted nature of memory failures, such as binding contextual features of what, where, and when, that patients experience and report, but standardized testing misses.

Rizzo et al. [107] highlighted VR's strength in evaluating memory within complex, realistic scenarios that correspond to the functional difficulties patients and families encounter in daily life, providing greater ecological validity than paper-based assessments and uncontrolled naturalistic observations. Pieri et al. [105] concluded from their analysis of 287 studies that VR effectively assesses not only memory but also spatial navigation, executive functions, attention, and daily living skills. Rather than replacing traditional tests, VR assessments were recommended as complementary tools that provide deeper insights into both impairments and preserved abilities.

3.4.3. Bridging Objective Metrics and Lived Experience

VR tasks are often correlated with executive function and global cognition, suggesting VR assessments more accurately reflect the interdependent nature of cognitive processes in real-world settings, a "functional perspective" [46] that captures memory as it naturally operates alongside other cognitive domains. Spiers et al. [91] assessed topographical and episodic memory in a patient with hippocampal damage, finding that the patient's impaired performance in VR-based navigation and episodic recall mirrored their real-world memory deficits, supporting VR as an effective method for detecting hippocampus-related impairments.

Together, these findings highlight VR's unique capacity to bridge objective performance metrics and subjective experiences, enhancing the interpretability and allowing researchers and clinicians to connect task performance with real-world outcomes for both patients and caregivers.

3.4.4. Gaps in Current Research

Despite these advances, the literature reveals specific gaps regarding VR and the caregiving dyad. First, while caregiver executive function impacts reporting accuracy, there is no research examining the "observer effect"—whether witnessing a patient's objective VR performance changes caregiver perception or reduces anxiety about patient capabilities. Second, longitudinal studies are needed to examine the temporal relationship between caregiver burden and patient quality-of-life discrepancies, and whether regular VR-based functional assessments can align these perspectives over the disease trajectory. Third, while digital storytelling shows promise for relationship building, empirical data supporting its use as a standard rehabilitative tool remains limited and requires standardized validation protocols.

Additionally, no studies have directly examined whether VR-based assessments improve patient-caregiver concordance, facilitate shared clinical decision-making, or reduce conflicts regarding patient autonomy and care needs. Research should investigate whether ecologically valid assessments produce results that both patients and caregivers recognize as reflecting daily functioning—potentially improving diagnostic acceptance and care planning. Finally, studies examining whether VR assessment feedback affects caregiver burden or the quality of patient-caregiver communication would help establish whether addressing the ecological validity gap translates into meaningful improvements in clinical relationships and care outcomes.

Future work must investigate not only how VR assesses patients, but how it might serve as a clinical tool to harmonize the discordant realities often experienced by patients and their caregivers.

3.5. VR and Compensatory Mechanisms

The dynamic, immersive nature of VR environments places greater demands on genuine cognitive processing, limiting opportunities for compensatory strategies that may mask underlying deficits in traditional assessments. By simulating complex, real-world tasks and providing controlled yet attention-demanding conditions, VR reduces reliance on learned strategies or external aids. The strong sense of presence elicits naturalistic behavior, minimizing the likelihood that participants can "game" the assessment. This enhanced complexity and immersiveness allow VR to provide more authentic measures of underlying cognitive capacity, revealing subtle impairments that traditional neuropsychological tests may miss [48]. Belger et al. [64] developed the immersive Virtual Memory Task (imVMT), which employs controller-free natural interaction to sustain deep immersion while preventing reliance on learned test-taking strategies. Validated across 35 neurological patients, including those with sensorimotor impairments, the imVMT maintained high participant engagement and ecological validity across varying difficulty levels, demonstrating that immersive environments can elicit authentic memory performance rather than strategic compensation.

Virtual environments are able to capture nuanced behaviors in realistic settings, enabling researchers to observe both cognitive deficits and the adaptive strategies individuals use to compensate [105,108]. For example, in VR tasks set in apartments or supermarkets, participants may use spatial landmarks or categorical grouping, strategies that are often invisible in standard list-learning tests. Barnett et al. [57] found that the Virtual Environment Grocery Store (VEGS) was particularly sensitive to memory impairment in older adults, who recalled fewer items in the VR task compared to traditional testing, suggesting that VR's immersive, multimodal demands may reveal deficits that compensatory strategies can mask in conventional assessments.

Research on spatial navigation supports VR's utility. Campbell et al. [109] argued that conventional tools like the Tower of London lack "neuroanatomical specificity," while VR tasks offer context-specific insights into functional decline. Kunz et al. [45] showed that early degradation in grid-cell-like activity in the entorhinal cortex, a region vulnerable to AD, could be behaviorally masked by reliance on alternative neural systems. This helps explain why navigation impairments often emerge late in the disease.

Antonova et al. [110] demonstrated flexible use of memory systems by combining VR with neuroimaging, showing that individuals switched between hippocampal and striatal strategies depending on task demands. Similarly, Allison et al. [42] found that people in preclinical AD stages exhibited impairments in hippocampal-dependent wayfinding but not in caudate-dependent route learning, suggesting an early shift toward compensatory strategies that VR can detect before standard assessments do.

A systematic review by Mancuso et al. [46] found that VR-based tasks often correlate with both memory and executive functioning measures, highlighting how these systems interact during compensatory strategy use. Similarly, and as discussed above, Agosta et al. [43] used fMRI with VR to reveal neural compensation even in individuals who performed at normal levels on traditional tests, showing how VR combined with neuroimaging can provide unique insights into brain network changes that might not be detectable with traditional assessments.

Cogné et al. [56] highlighted VR's value in detecting spatial navigation difficulties as prodromal markers of AD, probing both impaired and preserved spatial processes with ecological relevance. Consistent with this, da Costa et al. [111] found that tasks requiring active navigation within virtual environments were more effective in distinguishing MCI patients from healthy controls than traditional or passive tests, suggesting that VR's immersive, dynamic nature limits reliance on compensatory strategies that might otherwise mask deficits.

4. Discussion

4.1. Summary of Key Findings

Ecological validity. VR's capacity to foster presence—the percept of inhabiting virtual space—elicits naturalistic, self-directed behavior where participants often forget they are being assessed [56,85], fundamentally altering the psychological context of evaluation. This immersive quality yields measurable consequences: VR-based prospective memory tasks incorporating event, time, and activity-based cues outperformed the RBMT in stroke survivors [38], while high-fidelity virtual environments demonstrated successful transfer to real-world wayfinding and spatial mapping [62]. Critically, these advantages extend to clinical populations with motor limitations; the immersive Virtual Memory Task maintained engagement among neurological patients, including those with sensorimotor impairments, through controller-free natural interaction [64]. A paradoxical methodological advantage emerges: VR surpasses even real-world testing for practical assessments of certain memory processes (e.g., ego-to-allocentric conversion, context-dependent memory manipulation) because constructing analogous tasks in physical environments would require prohibitive logistical complexity while offering less experimental control.

Diagnostic differentiation. The most compelling empirical support concerns VR's superior sensitivity in distinguishing clinical populations. Park's Spatial Cognitive Task [98] achieved 94.4% sensitivity and 96.4% specificity in differentiating MCI from healthy aging—diagnostic precision attributed to the necessary engagement of allocentric spatial representation, which was compromised early in disease progression. Zygouris et al. [96] reported 87.3% classification accuracy using variables extracted from the Virtual Supermarket task reflective of executive inhibitory control and processing speed. Perhaps most consequentially, Howett et al. [97] demonstrated that immersive path integration tasks could differentiate MCI patients with positive AD biomarkers from those with negative biomarkers, linking behavioral performance to underlying neuropathology rather than clinical staging alone. Tarnanas et al. [5] extended this prognostic capacity, showing that VR-DOT predicted MCI-to-AD conversion with greater accuracy than standard cognitive measures by quantifying subtle executive planning errors that ADL questionnaires systematically miss. The starkest demonstration of differential sensitivity comes from Nolin et al. [84]: traditional prospective memory tasks revealed no significant differences between MCI and healthy groups, while VR tasks administered to the same participants exposed clear group differentiation. This sensitivity further permits differential staging along the MCI-to-AD continuum: intact egocentric but impaired allocentric navigation characterizes amnestic MCI, whereas both spatial processing systems degrade in AD [92]—a clinically consequential dissociation inaccessible to assessments lacking large-scale spatial demands.

Subjective and caregiver alignment. VR-derived performance metrics demonstrate stronger correspondence with lived experience than conventional assessments. Plancher et al. [68] found that VR task performance correlated more robustly with participants' everyday memory complaints than traditional tests in both healthy older adults and individuals with MCI. Allain et al. [8] reported that virtual kitchen performance not only correlated with caregiver Lawton-Brody IADL ratings but outperformed traditional cognitive assessments in predicting functional status, thus providing objective benchmarking against which caregiver observations can be validated. This alignment emerges because VR captures the multifaceted phenomenology of everyday memory failure (i.e., the binding of what, where, and when into integrated episodic experiences) that patients notice and report, but word-list recall structurally cannot assess.

Compensatory mechanism resistance. VR's immersive complexity creates assessment conditions where compensatory strategies that succeed on simplified tasks become insufficient or, when examined alongside neuroimaging, turn compensatory recruitment into diagnostic metrics. For example, Allison et al. [42] detected preserved caudate-dependent route learning alongside impaired hippocampal-dependent wayfinding in preclinical AD, revealing compensatory shifts to striatal strategies invisible to conventional assessment. Kunz et al. [45] demonstrated reduced grid-cell-like representations in the entorhinal cortex among young APOE-ε4 carriers during VR navigation—deficits detectable long before performance decrements appeared on standard episodic memory tests. Agosta et al. [43] used fMRI combined with VR to reveal neural compensation in individuals performing normally on traditional measures, exposing brain network reorganization masked by surface-level scores. Case-level evidence corroborates this pattern: MCI patients demonstrate normal topographical scene recognition yet exhibit pronounced deficits when recall requires viewpoint transformation [90,91], confirming that preserved performance on simpler recognition processes can mask failures in transformation-dependent spatial memory. One potential mechanistic explanation emerges: VR tasks engage distributed neural systems simultaneously (e.g., hippocampal binding, entorhinal path integration, retrosplenial frame translation), overwhelming the compensatory recruitment that succeeds when these circuits are tested in isolation.

4.2. Limitations

Despite this substantial evidence base, VR-based assessment has not yet achieved widespread clinical implementation. The transition from experimental validation to routine diagnostic use confronts pragmatic barriers that warrant systematic consideration—and evidence is not uniformly positive. In some cases, added cognitive load and interface complexity may introduce confounds that obscure rather than reveal cognitive status. These challenges extend beyond hardware accessibility to encompass patient usability, integration with clinical workflows, and methodological standardization [65].

4.2.1. Patient Usability and Safety

A primary concern in administering VR to older adults is the incidence of cybersickness, characterized by nausea and disorientation caused by a mismatch between the visual and vestibular systems [70,107]. Susceptibility to cybersickness increases with age and exposure duration, potentially confounding assessment results; poor performance may reflect physical discomfort rather than cognitive deficit [55]. While prior research indicates that adverse effects are typically mild and temporary, the application of immersive virtual environments in clinical populations raises important ethical and safety concerns, particularly regarding symptom susceptibility. Notably, a subset of users experience short-term health effects when using immersive HMDs, though they tend to dissipate rapidly [107].

Furthermore, the "digital divide" remains a barrier. Older adults with limited technology literacy may experience anxiety or increased cognitive load when navigating virtual interfaces [8,35,70], making it difficult to distinguish between genuine memory impairment and difficulty utilizing the interface [8,35,70,84]. Additionally, physical comorbidities common in neurodegenerative populations, such as tremors or hemiparesis, may limit a patient's ability to use standard handheld controllers, necessitating the development of more accessible, gesture-based, or passive navigation interfaces [8,64,70]. This necessitates the development of more accessible, gesture-based, or passive navigation interfaces to ensure assessments measure cognitive ability rather than motor limitation [64].

4.2.2. Integration with Clinical Workflows

Practical constraints within healthcare systems also pose significant hurdles. Rizzo et al. [101] note that clinical adoption has historically been hamstrung not only by cost but by the complexity of the systems and clinician unfamiliarity with VR equipment. Standard screening tools (e.g., the MMSE or MoCA) are favored for their brevity (5–10 min) and minimal resource requirements [36]. In contrast, VR assessments often require setup time, patient calibration, and tutorial phases to ensure the participant understands the controls, which may disrupt the tight scheduling of primary care or outpatient memory clinics [35]. Furthermore, the successful deployment of VR requires clinician training to manage hardware and troubleshoot software issues, demanding time and human resources that are often scarce in clinical settings [8,101]. Physical space is also a constraint; immersive VR often requires dedicated, obstruction-free areas to ensure patient safety during navigation, which may not be available in standard examination rooms [65].

4.2.3. Standardization and Normative Data

Finally, the lack of standardization across VR platforms limits their diagnostic utility. Unlike traditional neuropsychological tests, which have rigid administration protocols, VR environments vary widely in design, complexity, and interactivity [70,107]. For VR to serve as a reliable diagnostic tool, the field must progress from feasibility studies and pilot tests to the establishment of standardized protocols that can be shared across the research community [107,112]. Currently, many VR studies rely on small, homogenous samples [113], preventing the establishment of reliable cut-off scores required for individual clinical diagnosis. Furthermore, current research often lacks "dismantle studies" necessary to specify the "active ingredients" of VR interventions—determining which specific components (e.g., immersion, interactivity, or content) drive clinical outcomes [101].

5. Conclusions

VR assessment is at a critical point where technology and theory converge. By presenting multiple simultaneous cognitive demands, VR environments reveal distributed neural degradation that determines real-world functional capacity [42,114]. These demands overwhelm compensatory mechanisms that might succeed on simpler, isolated tasks. This capability becomes particularly crucial for differential diagnosis along the MCI-to-dementia continuum, where treatment pathways diverge dramatically based on subtle cognitive gradations that traditional assessments systematically obscure (Table 1).

The path forward demands methodological rigor: standardized protocols targeting specific neuropathological signatures, validation against ecological outcomes rather than merely traditional tests, and integration into clinical workflows that preserve the technology's diagnostic advantages while ensuring accessibility across diverse populations. As VR transitions from experimental to clinical use, its capacity to reveal neural compensation through ecologically valid yet controlled assessment may change how we detect, monitor, and intervene in neurodegenerative disease.

Table 1. Methodological constraints of traditional neuropsychological assessment and corresponding virtual reality solutions for ecologically valid memory evaluation.

Assessment Domain	Conventional Neuropsychological Restraints	VR-Enhanced Methodological Innovations
A) Ecological Validity Deficits	Tests occur in artificial, distraction-free environments that fail to reflect everyday cognitive demands. Tasks like word lists or isolated memory drills rarely resemble real-world activities.	Immersive VR environments simulate realistic everyday tasks while maintaining experimental control, fostering presence and naturalistic behavior.
B) Limited Scope and Diagnostic Differentiation	Traditional assessments measure a narrow range of memory skills and often cannot distinguish between normal aging, MCI, and early AD, missing subtle, context-dependent impairments.	VR allows complex, multi-domain tasks in dynamic environments to assess prospective memory, spatial memory, navigation, and executive function simultaneously.
C) Disconnect from Subjective and Caregiver Reports	Test scores often do not align with what patients notice or what caregivers observe in daily life, limiting clinical insight into functional impairments.	VR simulates everyday tasks that participants struggle with, producing performance data that correlates better with self-reports and caregiver ratings.
D) Susceptibility to Compensatory Mechanisms	Patients can perform normally on tests using alternative strategies, masking true cognitive deficits and delaying early detection.	Immersive VR tasks create complex, multi-step, attention-demanding environments that limit reliance on compensatory strategies.

Author Contributions: Conceptualization, N.R.; Writing—Original Draft Preparation, N.S. and N.R.; Writing—Review and Editing, N.S. and N.R.; Supervision, N.R.; Funding Acquisition, N.R. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported indirectly through funding from the Pacific Institute of Medical Research awarded to N.R (IACS001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study.

Acknowledgments: We extend our gratitude to Jesse Rissman, Joey K. Y. Essoe, and Felix Schoeller for their thoughtful discussions over the past several years, which have substantively enriched the conceptual development and intellectual foundations of this manuscript.

Conflicts of Interest: The authors declare no competing financial or non-financial interests that could have influenced this research.

References

1. Akpan, A.; Tabue-Teguo, M.; Fougère, B. Neurocognitive Disorders: Importance of Early/Timely Detection in Daily Clinical Practice. *J. Alzheimers Dis.* **2019**, *70*, 317–322. [CrossRef]
2. Torregrossa, W.; Torrisi, M.; De Luca, R.; Casella, C.; Rifici, C.; Bonanno, M.; Calabrò, R.S. Neuropsychological Assessment in Patients with Traumatic Brain Injury: A Comprehensive Review with Clinical Recommendations. *Biomedicines* **2023**, *11*, 1991. [CrossRef]
3. Roberts, J.R.; Maxfield, M. A 2-Study Psychometric Evaluation of the Modified Dementia Worry Scale. *Am. J. Alzheimers Dis. Other Demen.* **2021**, *36*, 1533317521995322. [CrossRef]
4. Van der Elst, W.; van Boxtel, M.P.J.; van Breukelen, G.J.P.; Jolles, J. Rey's Verbal Learning Test: Normative Data for 1855 Healthy Participants Aged 24-81 Years and the Influence of Age, Sex, Education, and Mode of Presentation. *J. Int. Neuropsychol. Soc.* **2005**, *11*, 290–302. [CrossRef]
5. Tarnanas, I.; Schlee, W.; Tsolaki, M.; Müri, R.; Mosimann, U.; Nef, T. Ecological Validity of Virtual Reality Daily Living Activities Screening for Early Dementia: Longitudinal Study. *JMIR Serious Games* **2013**, *1*, e1. [CrossRef]
6. Tarnanas, I.; Tsolaki, M.; Nef, T.; Müri, R.M.; Mosimann, U.P. Can a Novel Computerized Cognitive Screening Test Provide Additional Information for Early Detection of Alzheimer's Disease? *Alzheimers Dement.* **2014**, *10*, 790–798. [CrossRef] [PubMed]

7. Norman, M.A.; Evans, J.D.; Miller, W.S.; Heaton, R.K. Demographically Corrected Norms for the California Verbal Learning Test. *J. Clin. Exp. Neuropsychol.* **2000**, *22*, 80–94. [CrossRef] [PubMed]
8. Allain, P.; Foloppe, D.A.; Besnard, J.; Yamaguchi, T.; Etcharry-Bouyx, F.; Gall, D.L.; Nolin, P.; Richard, P. Detecting Everyday Action Deficits in Alzheimer's Disease Using a Nonimmersive Virtual Reality Kitchen. *J. Int. Neuropsychol. Soc.* **2014**, *20*, 468–477. [CrossRef] [PubMed]
9. Horan, W.P.; Moore, R.C.; Belanger, H.G.; Harvey, P.D. Utilizing Technology to Enhance the Ecological Validity of Cognitive and Functional Assessments in Schizophrenia: An Overview of the State-of-the-Art. *Schizophr. Bull. Open* **2024**, *5*, sgae025. [CrossRef]
10. Parsons, T.D.; McPherson, S.; Interrante, V. Enhancing Neurocognitive Assessment Using Immersive Virtual Reality. In *IEEE AutoTestcon, Proceedings of the 2013 1st Workshop on Virtual and Augmented Assistive Technology (VAAT), Lake Buena Vista, FL, USA, 17–17 March 2013*; IEEE: Lake Buena Vista, FL, USA, 2013; pp. 27–34.
11. Sbordone, R.J.; Long, C. *Ecological Validity of Neuropsychological Testing*; CRC Press: Boca Raton, FL, USA, 1996.
12. Chaytor, N.; Schmitter-Edgecombe, M. The Ecological Validity of Neuropsychological Tests: A Review of the Literature on Everyday Cognitive Skills. *Neuropsychol. Rev.* **2003**, *13*, 181–197. [CrossRef]
13. da Costa, R.Q.M.; Pompeu, J.E.; Moretto, E.; Silva, J.M.; Dos Santos, M.D.; Nitrini, R.; Brucki, S.M.D. Two Immersive Virtual Reality Tasks for the Assessment of Spatial Orientation in Older Adults with and Without Cognitive Impairment: Concurrent Validity, Group Comparison, and Accuracy Results. *J. Int. Neuropsychol. Soc.* **2022**, *28*, 460–472. [CrossRef]
14. Oliveira, J.; Gamito, P.; Rosa, B.; Bértolo, D.; Ribeiro, J.; Sousa, T.; Morais, D.; Ferreira, F.; Lopes, P. Ecologically-Oriented Approach for Cognitive Assessment in the Elderly. In *REHAB 16, Proceedings of the 4th Workshop on ICTs for Improving Patients Rehabilitation Research Techniques, Lisbon, Portugal, 13–14 October 2016*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 32–35.
15. Snyder, P.J.; Jackson, C.E.; Petersen, R.C.; Khachaturian, A.S.; Kaye, J.; Albert, M.S.; Weintraub, S. Assessment of Cognition in Mild Cognitive Impairment: A Comparative Study. *Alzheimers Dement. J. Alzheimers Assoc.* **2011**, *7*, 338–355. [CrossRef]
16. Negut, A. Cognitive Assessment and Rehabilitation in Virtual Reality: Theoretical Review and Practical Implications. *J. Appl. Psychol.* **2014**, *16*, 1–7.
17. Parsons, T.D.; Rizzo, A.A. Neuropsychological Assessment Using the Virtual Reality Cognitive Performance Assessment Test. In Proceedings of the 7th ICDVRAT with Art-Abilitation, Maia and Porto, Portugal, 8–10 September 2008.
18. Farias, S.T.; Mungas, D.; Reed, B.R.; Cahn-Weiner, D.; Jagust, W.; Baynes, K.; DeCarli, C. The Measurement of Everyday Cognition (ECog): Scale Development and Psychometric Properties. *Neuropsychology* **2008**, *22*, 531–544. [CrossRef] [PubMed]
19. Odhuba, R.A.; van den Broek, M.D.; Johns, L.C. Ecological Validity of Measures of Executive Functioning. *Br. J. Clin. Psychol.* **2005**, *44*, 269–278. [CrossRef]
20. Ouellet, É.; Boller, B.; Corriveau-Lecavalier, N.; Cloutier, S.; Belleville, S. The Virtual Shop: A New Immersive Virtual Reality Environment and Scenario for the Assessment of Everyday Memory. *J. Neurosci. Methods* **2018**, *303*, 126–135. [CrossRef] [PubMed]
21. Schultheis, M.T.; Himelstein, J.; Rizzo, A.A. Virtual Reality and Neuropsychology: Upgrading the Current Tools. *J. Head Trauma Rehabil.* **2002**, *17*, 378–394. [CrossRef] [PubMed]
22. Steibel, N.M.; Olchik, M.R.; Yassuda, M.S.; Finger, G.; Gomes, I. Influence of Age and Education on the Rivermead Behavioral Memory Test (RBMT) among Healthy Elderly. *Dement. Neuropsychol.* **2016**, *10*, 26–30. [CrossRef]
23. Fish, J.; Wilson, B.A.; Manly, T. The Assessment and Rehabilitation of Prospective Memory Problems in People with Neurological Disorders: A Review. *Neuropsychol. Rehabil.* **2010**, *20*, 161–179. [CrossRef]
24. De Roeck, E.E.; De Deyn, P.P.; Dierckx, E.; Engelborghs, S. Brief Cognitive Screening Instruments for Early Detection of Alzheimer's Disease: A Systematic Review. *Alzheimers Res. Ther.* **2019**, *11*, 21. [CrossRef]
25. Pearman, A.; Storandt, M. Predictors of Subjective Memory in Older Adults. *J. Gerontol. B Psychol. Sci. Soc. Sci.* **2004**, *59*, P4–P6. [CrossRef]
26. Reid, L.M.; MacLullich, A.M.J. Subjective Memory Complaints and Cognitive Impairment in Older People. *Dement. Geriatr. Cogn. Disord.* **2006**, *22*, 471–485. [CrossRef]
27. Luo, L.; Craik, F.I. Aging and Memory: A Cognitive Approach. *Can. J. Psychiatry* **2008**, *53*, 346–353. [CrossRef] [PubMed]
28. Grön, G.; Bittner, D.; Schmitz, B.; Wunderlich, A.P.; Riepe, M.W. Subjective Memory Complaints: Objective Neural Markers in Patients with Alzheimer's Disease and Major Depressive Disorder. *Ann. Neurol.* **2002**, *51*, 491–498. [CrossRef]
29. Psychiatry Online. Diagnostic and Statistical Manual of Mental Disorders. Available online: https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425787 (accessed on 16 October 2025).
30. Hodges, J.R. Alzheimer's Centennial Legacy: Origins, Landmarks and the Current Status of Knowledge Concerning Cognitive Aspects. *Brain J. Neurol.* **2006**, *129*, 2811–2822. [CrossRef] [PubMed]
31. Stasolla, F.; Di Gioia, M.; Messina, I.; Treglia, F.; Passaro, A.; Zullo, A.; Dragone, M. Assessing and Recovering Alzheimer's Disease: A Comparative Analysis of Standard Neuropsychological Approaches and Virtual Reality Interventions with the Use of Digital Storytelling. *Front. Psychol.* **2024**, *15*, 1406167. [CrossRef]

32. Aldaco, J.P.H.; Olmos, W.; Baez, A.; O'Brien, T.; Kozuki, J.; Alving, L.; Lent, D.; Woo, E. A-105 The Utility of Subjective Reports in Predicting Objective Prospective Memory Outcomes in Amnestic and Non-Amnestic Mild Cognitive Impairment. *Arch. Clin. Neuropsychol.* **2022**, *37*, 1257. [CrossRef]
33. Ojea Ortega, T.; González Álvarez de Sotomayor, M.M.; Pérez González, O.; Fernández Fernández, O. A new assessment for episodic memory. Episodic memory test and caregiver's episodic memory test. *Neurología* **2013**, *28*, 488–496. [CrossRef]
34. Cuoco, S.; Blundo, C.; Ricci, M.; Cappiello, A.; Bisogno, R.; Carotenuto, I.; Avallone, A.R.; Erro, R.; Pellecchia, M.T.; Amboni, M.; et al. Psychometric Properties of the Caregiver's Inventory Neuropsychological Diagnosis Dementia (CINDD) in Mild Cognitive Impairment and Dementia. *J. Neural Transm.* **2024**, *131*, 173–180. [CrossRef] [PubMed]
35. Chua, S.I.L.; Tan, N.C.; Wong, W.T.; Allen, J.C.; Quah, J.H.M.; Malhotra, R.; Østbye, T. Virtual Reality for Screening of Cognitive Function in Older Persons: Comparative Study. *J. Med. Internet Res.* **2019**, *21*, e14821. [CrossRef]
36. Bressan, L.A.; Vale, F.d.A.C.; Speciali, J.G. The Daily Life of Patients with Dementia: A Comparative Study between the Information Provided by the Caregiver and Direct Patient Assessment. *Dement. Neuropsychol.* **2007**, *1*, 288–295. [CrossRef]
37. Stella, F.; Forlenza, O.V.; Laks, J.; de Andrade, L.P.; de Castilho Cação, J.; Govone, J.S.; de Medeiros, K.; Lyketsos, C.G. Caregiver Report versus Clinician Impression: Disagreements in Rating Neuropsychiatric Symptoms in Alzheimer's Disease Patients. *Int. J. Geriatr. Psychiatry* **2015**, *30*, 1230–1237. [CrossRef]
38. Díaz-Orueta, U.; Climent, G.; Cardas-Ibáñez, J.; Alonso, L.; Olmo-Osa, J.; Tirapu-Ustárroz, J. Memory Assessment by Means of Virtual Reality: Its Present and Future. *Rev. Neurol.* **2016**, *62*, 75–84. [CrossRef]
39. Torlaschi, V.; Maffoni, M.; Maltauro, G.; Pierobon, A.; Vigorè, M.; Maestri, R.; Chimento, P.; Buonocore, M.; Mancardi, G.; Fundarò, C. The Patient–Caregiver Dyad: The Impact of Cognitive and Functional Impairment. *Neurol. Sci.* **2022**, *43*, 2481–2490. [CrossRef]
40. Banville, F.; Nolin, P.; Lalonde, S.; Henry, M.; Dery, M.-P.; Villemure, R. Multitasking and Prospective Memory: Can Virtual Reality Be Useful for Diagnosis? *Behav. Neurol.* **2010**, *23*, 209–211. [CrossRef]
41. Morganti, F.; Gattuso, M.; Singh Solorzano, C.; Bonomini, C.; Rosini, S.; Ferrari, C.; Pievani, M.; Festari, C. Virtual Reality-Based Psychoeducation for Dementia Caregivers: The Link between Caregivers' Characteristics and Their Sense of Presence. *Brain Sci.* **2024**, *14*, 852. [CrossRef] [PubMed]
42. Allison, S.L.; Fagan, A.M.; Morris, J.C.; Head, D. Spatial Navigation in Preclinical Alzheimer's Disease. *J. Alzheimers Dis.* **2016**, *52*, 77–90. [CrossRef] [PubMed]
43. Agosta, F.; Canu, E.; Filippi, M. Virtual Reality and Real-Time Neurofeedback Functional MRI: A Breakthrough in Foreseeing Alzheimer's Disease? *Brain* **2020**, *143*, 722–726. [CrossRef]
44. Skouras, S.; Torner, J.; Andersson, P.; Koush, Y.; Falcon, C.; Minguillon, C.; Fauria, K.; Alpiste, F.; Blenow, K.; Zetterberg, H.; et al. Earliest Amyloid and Tau Deposition Modulate the Influence of Limbic Networks during Closed-Loop Hippocampal Downregulation. *Brain J. Neurol.* **2020**, *143*, 976–992. [CrossRef]
45. Kunz, L.; Schröder, T.N.; Lee, H.; Montag, C.; Lachmann, B.; Sariyska, R.; Reuter, M.; Stirnberg, R.; Stöcker, T.; Messing-Floeter, P.C.; et al. Reduced Grid-Cell-like Representations in Adults at Genetic Risk for Alzheimer's Disease. *Science* **2015**, *350*, 430–433. [CrossRef]
46. Mancuso, V.; Sarcinella, E.D.; Bruni, F.; Arlati, S.; Di Santo, S.G.; Cavallo, M.; Cipresso, P.; Pedroli, E. Systematic Review of Memory Assessment in Virtual Reality: Evaluating Convergent and Divergent Validity with Traditional Neuropsychological Measures. *Front. Hum. Neurosci.* **2024**, *18*, 1380575. [CrossRef]
47. Toplak, M.E.; West, R.F.; Stanovich, K.E. Practitioner Review: Do Performance-Based Measures and Ratings of Executive Function Assess the Same Construct? *J. Child Psychol. Psychiatry* **2013**, *54*, 131–143. [CrossRef] [PubMed]
48. Reggente, N. VR for Cognition and Memory. In *Virtual Reality in Behavioral Neuroscience: New Insights and Methods*; Maymon, C., Grimshaw, G., Wu, Y.C., Eds.; Current Topics in Behavioral Neurosciences; Springer International Publishing: Cham, Switzerland, 2023; Volume 65, pp. 189–232.
49. Barclay, P.A.; Parker, J.; Sims, V. Scoping Review of Virtual Environments for Assessing Episodic Memory. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2018**, *62*, 1494. [CrossRef]
50. Kourtesis, P.; MacPherson, S.E. An Ecologically Valid Examination of Event-Based and Time-Based Prospective Memory Using Immersive Virtual Reality: The Influence of Attention, Memory, and Executive Function Processes on Real-World Prospective Memory. *Neuropsychol. Rehabil.* **2023**, *33*, 255–280. [CrossRef]
51. Corriveau Lecavalier, N.; Ouellet, É.; Boller, B.; Belleville, S. Use of Immersive Virtual Reality to Assess Episodic Memory: A Validation Study in Older Adults. *Neuropsychol. Rehabil.* **2020**, *30*, 462–480. [CrossRef]
52. Bottiroli, S.; Bernini, S.; Cavallini, E.; Sinforiani, E.; Zucchella, C.; Pazzi, S.; Cristiani, P.; Vecchi, T.; Tost, D.; Sandrini, G.; et al. The Smart Aging Platform for Assessing Early Phases of Cognitive Impairment in Patients with Neurodegenerative Diseases. *Front. Psychol.* **2021**, *12*, 635410. [CrossRef]
53. Gottlieb, A.; Doniger, G.M.; Kimel-Naor, S.; Ben-Gal, O.; Cohen, M.; Iny, H.; Beeri, M.S.; Plotnik, M. Development and Validation of Virtual Reality-Based Rey Auditory Verbal Learning Test. *Front. Aging Neurosci.* **2022**, *14*, 980093. [CrossRef]

54. Parsons, T.D. Virtual Reality for Enhanced Ecological Validity and Experimental Control in the Clinical, Affective and Social Neurosciences. *Front. Hum. Neurosci.* **2015**, *9*, 660. [CrossRef] [PubMed]
55. Smith, S.A. Virtual Reality in Episodic Memory Research: A Review. *Psychon. Bull. Rev.* **2019**, *26*, 1213–1237. [CrossRef] [PubMed]
56. Cogné, M.; Taillade, M.; N'Kaoua, B.; Tarruella, A.; Klinger, E.; Larrue, F.; Sauzéon, H.; Joseph, P.-A.; Sorita, E. The Contribution of Virtual Reality to the Diagnosis of Spatial Navigation Disorders and to the Study of the Role of Navigational Aids: A Systematic Literature Review. *Ann. Phys. Rehabil. Med.* **2017**, *60*, 164–176. [CrossRef]
57. Barnett, M.D.; Chek, C.J.W.; Shorter, S.S.; Parsons, T.D. Comparison of Traditional and Virtual Reality-Based Episodic Memory Performance in Clinical and Non-Clinical Cohorts. *Brain Sci.* **2022**, *12*, 1019. [CrossRef]
58. Plancher, G.; Gyselinck, V.; Nicolas, S.; Piolino, P. Age Effect on Components of Episodic Memory and Feature Binding: A Virtual Reality Study. *Neuropsychology* **2010**, *24*, 379–390. [CrossRef]
59. Vlček, K.; Laczó, J. Neural Correlates of Spatial Navigation Changes in Mild Cognitive Impairment and Alzheimer's Disease. *Front. Behav. Neurosci.* **2014**, *8*, 89. [CrossRef]
60. Weniger, G.; Ruhleder, M.; Lange, C.; Wolf, S.; Irle, E. Egocentric and Allocentric Memory as Assessed by Virtual Reality in Individuals with Amnestic Mild Cognitive Impairment. *Neuropsychologia* **2011**, *49*, 518–527. [CrossRef]
61. La Corte, V.; Sperduti, M.; Abichou, K.; Piolino, P. Episodic Memory Assessment and Remediation in Normal and Pathological Aging Using Virtual Reality: A Mini Review. *Front. Psychol.* **2019**, *10*, 173. [CrossRef]
62. Wallet, G.; Sauzéon, H.; Pala, P.A.; Florian, L.; Zheng, X.; N'Kaoua, B. Virtual/Real Transfer of Spatial Knowledge: Benefit from Visual Fidelity Provided in a Virtual Environment and Impact of Active Navigation. *Cyberpsychol. Behav. Soc. Netw.* **2011**, *14*, 417–423. [CrossRef] [PubMed]
63. Hogan, C.; Cornwell, P.; Fleming, J.; Man, D.W.K.; Shum, D.H.K. Assessment of Prospective Memory after Stroke Utilizing Virtual Reality. *Virtual Real.* **2023**, *27*, 333–346. [CrossRef]
64. Belger, J.; Blume, M.; Akbal, M.; Chojecki, P.; de Mooij, J.; Gaebler, M.; Klotzsche, F.; Krohn, S.; Lafci, M.T.; Quinque, E.; et al. The Immersive Virtual Memory Task: Assessing Object-Location Memory in Neurological Patients Using Immersive Virtual Reality. *Neuropsychol. Rehabil.* **2024**, *34*, 870–898. [CrossRef] [PubMed]
65. Reggente, N.; Essoe, J.K.-Y.; Aghajan, Z.M.; Tavakoli, A.V.; McGuire, J.F.; Suthana, N.A.; Rissman, J. Enhancing the Ecological Validity of fMRI Memory Research Using Virtual Reality. *Front. Neurosci.* **2018**, *12*, 408. [CrossRef]
66. Ragan, E.D.; Sowndararajan, A.; Kopper, R.; Bowman, D.A. The Effects of Higher Levels of Immersion on Procedure Memorization Performance and Implications for Educational Virtual Environments. *Presence Teleoperators Virtual Environ.* **2010**, *19*, 527–543. [CrossRef]
67. Kim, H.; Kim, Y.; Lee, J.; Kim, J. Stereoscopic Objects Affect Reaching Performance in Virtual Reality Environments: Influence of Age on Motor Control. *Front. Virtual Real.* **2024**, *5*, 1475482. [CrossRef]
68. Plancher, G.; Tirard, A.; Gyselinck, V.; Nicolas, S.; Piolino, P. Using Virtual Reality to Characterize Episodic Memory Profiles in Amnestic Mild Cognitive Impairment and Alzheimer's Disease: Influence of Active and Passive Encoding. *Neuropsychologia* **2012**, *50*, 592–602. [CrossRef]
69. Jang, S.; Vitale, J.; Jyung, R.; Black, J. Direct Manipulation Is Better than Passive Viewing for Learning Anatomy in a Three-Dimensional Virtual Reality Environment. *Comput. Educ.* **2016**, *106*, 150–165. [CrossRef]
70. Jonson, M.; Avramescu, S.; Chen, D.; Alam, F. The Role of Virtual Reality in Screening, Diagnosing, and Rehabilitating Spatial Memory Deficits. *Front. Hum. Neurosci.* **2021**, *15*, 628818. [CrossRef] [PubMed]
71. Parsons, T.D.; Barnett, M. Validity of a Newly Developed Measure of Memory: Feasibility Study of the Virtual Environment Grocery Store. *J. Alzheimers Dis.* **2017**, *59*, 1227–1235. [CrossRef]
72. Parsons, T.; Duffield, T. Paradigm Shift Toward Digital Neuropsychology and High-Dimensional Neuropsychological Assessments: Review. *J. Med. Internet Res.* **2020**, *22*, e23777. [CrossRef]
73. Plancher, G.; Nicolas, S.; Piolino, P. Virtual Reality as a Tool for Assessing Episodic Memory. In Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology, Bordeaux, France, 27–29 October 2008; Association for Computing Machinery: New York, NY, USA, 2008; pp. 179–182.
74. Godden, D.R.; Baddeley, A.D. Context-Dependent Memory in Two Natural Environments: On Land and Underwater. *Br. J. Psychol.* **1975**, *66*, 325–331. [CrossRef]
75. Essoe, J.K.-Y.; Reggente, N.; Ohno, A.A.; Baek, Y.H.; Dell'Italia, J.; Rissman, J. Enhancing Learning and Retention with Distinctive Virtual Reality Environments and Mental Context Reinstatement. *npj Sci. Learn.* **2022**, *7*, 31. [CrossRef]
76. Iglói, K.; Doeller, C.F.; Paradis, A.-L.; Benchenane, K.; Berthoz, A.; Burgess, N.; Rondi-Reig, L. Interaction Between Hippocampus and Cerebellum Crus I in Sequence-Based but Not Place-Based Navigation. *Cereb. Cortex* **2015**, *25*, 4146–4154. [CrossRef]
77. Colombo, D.; Serino, S.; Tuena, C.; Pedroli, E.; Dakanalis, A.; Cipresso, P.; Riva, G. Egocentric and Allocentric Spatial Reference Frames in Aging: A Systematic Review. *Neurosci. Biobehav. Rev.* **2017**, *80*, 605–621. [CrossRef]
78. Chen, Y.; Byrne, P.; Crawford, J.D. Time Course of Allocentric Decay, Egocentric Decay, and Allocentric-to-Egocentric Conversion in Memory-Guided Reach. *Neuropsychologia* **2011**, *49*, 49–60. [CrossRef]

79. Tu, S.; Spiers, H.J.; Hodges, J.R.; Piguet, O.; Hornberger, M. Egocentric versus Allocentric Spatial Memory in Behavioral Variant Frontotemporal Dementia and Alzheimer's Disease. *J. Alzheimer's Dis.* **2017**, *59*, 883–892. [CrossRef] [PubMed]
80. Serino, S.; Cipresso, P.; Morganti, F.; Riva, G. The Role of Egocentric and Allocentric Abilities in Alzheimer's Disease: A Systematic Review. *Ageing Res. Rev.* **2014**, *16*, 32–44. [CrossRef]
81. Tuena, C.; Mancuso, V.; Stramba-Badiale, C.; Pedroli, E.; Stramba-Badiale, M.; Riva, G.; Repetto, C. Egocentric and Allocentric Spatial Memory in Mild Cognitive Impairment with Real-World and Virtual Navigation Tasks: A Systematic Review. *J. Alzheimer's Dis.* **2021**, *79*, 95–116. [CrossRef]
82. Reggente, N.; Essoe, J.K.Y.; Baek, H.Y.; Rissman, J. The Method of Loci in Virtual Reality: Explicit Binding of Objects to Spatial Contexts Enhances Subsequent Memory Recall. *J. Cogn. Enhanc.* **2020**, *4*, 12–30. [CrossRef]
83. Rissman, J.; Chow, T.E.; Reggente, N.; Wagner, A.D. Decoding fMRI Signatures of Real-World Autobiographical Memory Retrieval. *J. Cogn. Neurosci.* **2016**, *28*, 604–620. [CrossRef]
84. Nolin, P.; Banville, F.; Cloutier, J.; Allain, P. Virtual Reality as a New Approach to Assess Cognitive Decline in the Elderly. *Acad. J. Interdiscip. Stud.* **2013**, *2*, 612. [CrossRef]
85. Neguț, A.; Matu, S.-A.; Sava, F.A.; David, D. Virtual Reality Measures in Neuropsychological Assessment: A Meta-Analytic Review. *Clin. Neuropsychol.* **2016**, *30*, 165–184. [CrossRef]
86. Clay, F.; Howett, D.; FitzGerald, J.; Fletcher, P.; Chan, D.; Price, A. Use of Immersive Virtual Reality in the Assessment and Treatment of Alzheimer's Disease: A Systematic Review. *J. Alzheimers Dis.* **2020**, *75*, 23–43. [CrossRef] [PubMed]
87. Knight, R.G.; Titov, N. Use of Virtual Reality Tasks to Assess Prospective Memory: Applicability and Evidence. *Brain Impair.* **2009**, *10*, 3–13. [CrossRef]
88. Lecouvey, G.; Morand, A.; Gonneaud, J.; Piolino, P.; Orriols, E.; Pélerin, A.; Ferreira Da Silva, L.; de La Sayette, V.; Eustache, F.; Desgranges, B. An Impairment of Prospective Memory in Mild Alzheimer's Disease: A Ride in a Virtual Town. *Front. Psychol.* **2019**, *10*, 241. [CrossRef]
89. Déjos, M.; Sauzéon, H.; N'Kaoua, B. La Réalité Virtuelle Au Service de l'évaluation Clinique de La Personne Âgée: Le Dépistage Précoce de La Démence. *Rev. Neurol.* **2012**, *168*, 404–414. [CrossRef]
90. King, J.A.; Burgess, N.; Hartley, T.; Vargha-Khadem, F.; O'Keefe, J. Human Hippocampus and Viewpoint Dependence in Spatial Memory. *Neurosci. J.* **2002**, *12*, 811–820. [CrossRef]
91. Spiers, H.J.; Burgess, N.; Hartley, T.; Vargha-Khadem, F.; O'Keefe, J. Bilateral Hippocampal Pathology Impairs Topographical and Episodic Memory but Not Visual Pattern Matching. *Hippocampus* **2001**, *11*, 715–725. [CrossRef]
92. Laczó, J.; Andel, R.; Vyhnalek, M.; Vlcek, K.; Magerova, H.; Varjassyova, A.; Tolar, M.; Hort, J. Human Analogue of the Morris Water Maze for Testing Subjects at Risk of Alzheimer's Disease. *Neurodegener. Dis.* **2010**, *7*, 148–152. [CrossRef]
93. Goodrich-Hunsaker, N.J.; Livingstone, S.A.; Skelton, R.W.; Hopkins, R.O. Spatial Deficits in a Virtual Water Maze in Amnesic Participants with Hippocampal Damage. *Hippocampus* **2009**, *20*, 481–491. [CrossRef]
94. Morganti, F.; Stefanini, S.; Riva, G. From Allo- to Egocentric Spatial Ability in Early Alzheimer's Disease: A Study with Virtual Reality Spatial Tasks. *Cogn. Neurosci. Alzheimer's Dis.* **2013**, *4*, 171–180. [CrossRef]
95. Kalova, E.; Vlček, K.; Jarolimova, E.; Bures, J. Allothetic Orientation and Sequential Ordering of Places Is Impaired in Early Stages of Alzheimer's Disease: Corresponding Results in Real Space Tests and Computer Tests. *Behav. Brain Res.* **2005**, *159*, 175–186. [CrossRef]
96. Zygouris, S.; Giakoumis, D.; Votis, K.; Doumpoulakis, S.; Ntovas, K.; Segkouli, S.; Karagiannidis, C.; Tzovaras, D.; Tsolaki, M. Can a Virtual Reality Cognitive Training Application Fulfill a Dual Role? Using the Virtual Supermarket Cognitive Training Application as a Screening Tool for Mild Cognitive Impairment. *J. Alzheimers Dis.* **2015**, *44*, 1333–1347. [CrossRef]
97. Howett, D.; Castegnaro, A.; Krzywicka, K.; Hagman, J.; Marchment, D.; Henson, R.; Rio, M.; King, J.A.; Burgess, N.; Chan, D. Differentiation of Mild Cognitive Impairment Using an Entorhinal Cortex-Based Test of Virtual Reality Navigation. *Brain J. Neurol.* **2019**, *142*, 1751–1766. [CrossRef]
98. Park, J.-H. Can the Virtual Reality-Based Spatial Memory Test Better Discriminate Mild Cognitive Impairment than Neuropsychological Assessment? *Int. J. Environ. Res. Public Health* **2022**, *19*, 9950. [CrossRef]
99. Bohil, C.J.; Alicea, B.; Biocca, F.A. Virtual Reality in Neuroscience Research and Therapy. *Nat. Rev. Neurosci.* **2011**, *12*, 752–762. [CrossRef]
100. Brooks, B.M.; Rose, F.D. The Use of Virtual Reality in Memory Rehabilitation: Current Findings and Future Directions. *NeuroRehabilitation* **2003**, *18*, 147–157. [CrossRef]
101. Rizzo, A.S.; Koenig, S.T.; Talbot, T.B. Clinical Virtual Reality: Emerging Opportunities for Psychiatry. *Focus Am. Psychiatr. Publ.* **2018**, *16*, 266–278. [CrossRef]
102. Zucchella, C.; Sinforiani, E.; Tamburin, S.; Federico, A.; Mantovani, E.; Bernini, S.; Casale, R.; Bartolo, M. The Multidisciplinary Approach to Alzheimer's Disease and Dementia. A Narrative Review of Non-Pharmacological Treatment. *Front. Neurol.* **2018**, *9*, 1058. [CrossRef]

103. Maguire, E.A.; Woollett, K.; Spiers, H.J. London Taxi Drivers and Bus Drivers: A Structural MRI and Neuropsychological Analysis. *Hippocampus* **2006**, *16*, 1091–1101. [CrossRef]
104. Altozano, A.; Minissi, M.E.; Alcañiz, M.; Marín-Morales, J. Introducing 3DCNN ResNets for ASD Full-Body Kinematic Assessment: A Comparison with Hand-Crafted Features. *Expert Syst. Appl.* **2025**, *270*, 126295. [CrossRef]
105. Pieri, L.; Tosi, G.; Romano, D. Virtual Reality Technology in Neuropsychological Testing: A Systematic Review. *J. Neuropsychol.* **2023**, *17*, 382–399. [CrossRef]
106. Cushman, L.A.; Stein, K.; Duffy, C.J. Detecting Navigational Deficits in Cognitive Aging and Alzheimer Disease Using Virtual Reality. *Neurology* **2008**, *71*, 888–895. [CrossRef]
107. Rizzo, A.; Gambino, G.; Sardo, P.; Rizzo, V. Being in the Past and Perform the Future in a Virtual World: VR Applications to Assess and Enhance Episodic and Prospective Memory in Normal and Pathological Aging. *Front. Hum. Neurosci.* **2020**, *14*, 297. [CrossRef]
108. Gaggioli, A. *Advanced Technologies in Rehabilitation: Empowering Cognitive, Physical, Social and Communicative Skills Through Virtual Reality, Robots, Wearable Systems and Brain-Computer Interfaces*; IOS Press: Amsterdam, The Netherlands, 2009.
109. Campbell, Z.; Zakzanis, K.K.; Jovanovski, D.; Joordens, S.; Mraz, R.; Graham, S.J. Utilizing Virtual Reality to Improve the Ecological Validity of Clinical Neuropsychology: An fMRI Case Study Elucidating the Neural Basis of Planning by Comparing the Tower of London with a Three-Dimensional Navigation Task. *Appl. Neuropsychol.* **2009**, *16*, 295–306. [CrossRef]
110. Antonova, E.; Parslow, D.; Brammer, M.; Simmons, A.; Williams, S.; Dawson, G.R.; Morris, R. Scopolamine Disrupts Hippocampal Activity during Allocentric Spatial Memory in Humans: An fMRI Study Using a Virtual Reality Analogue of the Morris Water Maze. *J. Psychopharmacol.* **2011**, *25*, 1256–1265. [CrossRef]
111. da Costa, R.Q.M.; Pompeu, J.E.; de Viveiro, L.A.P.; Brucki, S.M.D. Spatial Orientation Tasks Show Moderate to High Accuracy for the Diagnosis of Mild Cognitive Impairment: A Systematic Literature Review. *Arq. Neuropsiquiatr.* **2020**, *78*, 713–723. [CrossRef]
112. Pieri, L.; Serino, S.; Cipresso, P.; Mancuso, V.; Riva, G.; Pedroli, E. The ObReco-360°: A New Ecological Tool to Memory Assessment Using 360° Immersive Technology. *Virtual Real.* **2022**, *26*, 639–648. [CrossRef]
113. Dassel, K.B.; Schmitt, F.A. The Impact of Caregiver Executive Skills on Reports of Patient Functioning. *Gerontologist* **2008**, *48*, 781–792. [CrossRef] [PubMed]
114. Gazova, I.; Vlcek, K.; Laczo, J.; Nedelska, Z.; Hyncicova, E.; Mokrisova, I.; Sheardova, K.; Hort, J. Spatial Navigation—A Unique Window into Physiological and Pathological Aging. *Front. Aging Neurosci.* **2012**, *4*, 26794. [CrossRef] [PubMed]