

Original Paper

Virtual Reality Gamification of Visual Search, Response Inhibition, and Visual Short-Term Memory Tasks for Cognitive Assessment: Experimental Study

Marios Hadjjaros^{1,2,3}, PhD; Andria Shimi^{1,3,4}, Prof Dr; Kleanthis Neokleous¹, PhD; Constantinos Pattichis^{1,2,3}, Prof Dr; Marios Avraamides^{1,4}, Prof Dr

¹CYENS—Centre of Excellence, Nicosia, Cyprus

²Department of Computer Science, University of Cyprus, Nicosia, Cyprus

³Biomedical Engineering Research Centre, University of Cyprus, Nicosia, Cyprus

⁴Department of Psychology, University of Cyprus, Nicosia, Cyprus

Corresponding Author:

Marios Hadjjaros, PhD

CYENS—Centre of Excellence

1 Dimarchou Lellou Demetriadi

Nicosia 1016

Cyprus

Phone: 357 99094970

Email: marios.hadjjaros@gmail.com

Abstract

Background: Cognitive tasks are foundational tools in psychology and neuroscience for studying attention, perception, and memory. However, they typically employ simple or artificial stimuli and require numerous repetitive trials, which can adversely affect participant engagement and ecological validity.

Objective: This study investigated whether gamified versions of 3 established cognitive tasks, namely, the Visual Search task (attention), the Whack-the-Mole task (response inhibition), and the Corsi block-tapping test (visual short-term memory), replicate the typical patterns of results reported for their traditional counterparts. It also examined whether the method of administration—in immersive virtual reality (VR) versus desktop computer, and in the laboratory versus at home—influences performance.

Methods: Seventy-five participants (male=24, female=51; age range 18-35 years; mean 23.15, SD 4.38 years) were randomly assigned to 1 of 3 administration conditions (n=25 each). In the VR-Lab condition, participants completed the tasks in immersive VR within the laboratory; in the Desktop-Lab condition, they completed the tasks on a 2D desktop screen in the laboratory; and in the Desktop-Remote condition, participants completed the tasks on their personal computers at home. All participants completed the same gamified tasks while seated, entering responses with either a mouse or a VR controller, depending on the condition.

Results: The results obtained from these gamified tasks across all 3 administration conditions replicated the typical performance patterns observed with their traditional counterparts, despite using more ecologically valid stimuli and fewer trials. However, administration modality did influence certain performance measures, particularly reaction times (RTs) and task efficiency. Specifically, in the Visual Search task, RTs were significantly faster in the VR-Lab condition (mean 1.24 seconds) than in the Desktop-Lab (mean 1.49 seconds; $P<.001$) and Desktop-Remote (mean 1.44 seconds; $P=.008$) conditions. In the Whack-the-Mole task, no significant group differences emerged in d' scores (VR-Lab: mean 3.79, Desktop-Remote: mean 3.75, Desktop-Lab: mean 3.62; $P=.49$), but RTs were slower in the Desktop-Remote condition (mean 0.64 seconds) than in the VR-Lab (mean 0.41 seconds; $P<.001$) and Desktop-Lab (mean 0.48 seconds; $P<.001$) conditions. For the Corsi block-tapping test, no significant group differences in span scores were found (VR-Lab: mean 5.48, Desktop-Lab: mean 5.68, and Desktop-Remote: mean 5.24; $P=.24$). Finally, a significant positive correlation was observed between RTs for Hits in the Whack-the-Mole task and feature search trials in the Visual Search task ($r=0.24$; $P=.04$).

Conclusions: Gamified cognitive tasks administered in VR replicated established behavioral patterns observed with their traditional versions while improving ecological validity and reducing task duration. Administration modality had limited effects on overall outcomes, although RTs were slower in remote settings. These findings support the feasibility of using

gamified VR tasks for scalable and ecologically valid cognitive assessment. Overall, the study underscores the potential of VR to increase participant engagement and enrich cognitive research through more immersive and motivating testing environments.

JMIR Form Res 2025;9:e65836; doi: [10.2196/65836](https://doi.org/10.2196/65836)

Keywords: attention; memory; virtual reality; visuospatial; serious games; cognitive assessment; Corsi Test; Visual Search; Go/No-Go; response inhibition

Introduction

Background

Cognitive tasks are commonly used in Psychology and Neuroscience to investigate various mental processes, such as attention, memory, and perception [1,2]. These tasks often involve simple stimuli, such as alphanumeric characters and shapes and a large number of trials, a deliberate design choice aimed at controlling for extraneous variables and ensuring the reliability of the results [3,4]. While this methodological approach is crucial for maintaining experimental rigor and validity, it may inadvertently lead to issues regarding participant engagement [5]. For example, the repetitive tasks and the simple stimuli may cause participant disinterest, boredom, and fatigue, which in turn may compromise data quality or introduce unintended biases [6,7]. Thus, while striving for methodological control, researchers must also consider strategies to enhance participant engagement to mitigate these potential challenges and maintain the integrity of their findings.

Virtual Reality as a Tool for Gamified Cognitive Assessment

Virtual reality (VR) technology presents a promising avenue for addressing the engagement challenges inherent in traditional cognitive tasks [8]. By leveraging VR, researchers can gamify tasks, transforming them into immersive and interactive experiences that are inherently more engaging for participants [9].

Numerous studies have indicated that gamification in cognitive tasks can significantly increase participant engagement and motivation [10,11]. Recent studies exploring VR gaming have emphasized the significant role of user engagement and motivation in enhancing the overall experience. One key factor driving VR game adoption is the concept of “flow,” which is a state of optimal immersion where players experience intense enjoyment and concentration. This “flow” state is essential for fostering deeper engagement with VR games, particularly as the player becomes more immersed in the virtual environment. The sense of spatial presence, or the feeling of being physically located within the VR world, is another major motivator, as it strengthens the emotional connection to the game and enhances relaxation or enjoyment [12,13].

Cognitive Tasks Under Investigation

The Visual Search task is a widely used method for investigating the mechanisms of attention involved in locating a target among distractors, mirroring real-world scenarios such

as finding a friend in a crowded cafeteria, searching for an item in the grocery store, or searching a target object among nontarget objects during security screening. In laboratory settings, a typical Visual Search task involves presenting a target object on the computer screen among distractors [14,15]. Stimuli often consist of alphanumeric characters or simple geometric shapes. Visual search tasks typically involve 2 key manipulations: similarity and display size. Similarity refers to the type of search that is determined by how visually distinct the target is from the distractors. In *feature search* trials, the target differs from distractors in terms of a single feature, such as its shape (eg, locating the letter “O” among several “Xs”) or its color (eg, locating a red “X” among white “Xs”). In *conjunction search* trials, the target is defined by a combination of features that are also present in the distractors (eg, locating a green “O” among green “Xs” and red “Os”). The second manipulation is the display size, which refers to the number of distractors presented in each trial alongside the target [16]. The typical pattern of results obtained across many studies is that (1) reaction times (RTs) (ie, the time it takes the participant to find the target from the onset of the display) are shorter for feature search trials than conjunction search trials, (2) RT increases significantly with display size, a phenomenon known as the “display size effect,” but only in conjunction search trials [15,17,18], and (3) RT remains relatively constant across display sizes in feature search trials, a phenomenon known as the “pop-out effect” [19].

The Whack-the-Mole task is a Go/No-Go task that provides a measure of response inhibition, that is, the ability to inhibit inappropriate prepotent responses [20]. Past research has shown that response inhibition is a key ability for cognitive [21] and psychosocial [22] development and is a reliable predictor of school readiness and academic achievement in domains such as math and reading [23]. Impairments in response inhibition are commonly observed in neurodevelopmental disorders such as autism [24] and attention-deficit/hyperactivity disorder [25,26] as well as in psychiatric and neurological disorders such as borderline personality disorder [27] and Parkinson disease [28]. In computerized Go/No-Go tasks, participants are required to execute an action rapidly (eg, press a button) upon detecting a specific stimulus (Go trials) and withhold a response when presented with a different stimulus (No-Go trials) [20,21]. To ensure task effectiveness, the majority of trials are typically Go trials (approximately 75%), priming the execution of a response [20,21]. Response inhibition is assessed by computing d' , a sensitivity metric that takes into account the proportion of hits (ie, correctly executed responses in Go trials) and false alarms (ie, incorrect execution of responses in No-Go trials) [20,29].

Although d' may vary depending on the manipulations of the experiment, such as the discriminability of the stimuli, its typical value, indicating moderate sensitivity, is between 3 and 4 [20,30]. In general, participants exhibit false alarms in 10% of the No-Go trials [20,21].

The Corsi Test is one of the most widely used tasks to assess visuospatial short-term memory. The original version [31] consists of 9 physical cubes placed on a board. Each trial involves the experimenter sequentially touching a series of blocks from one side of the board, and the participant, situated on the opposite side, is required to reproduce the sequence by touching the blocks in the same order. If the participant successfully reproduces the sequence, then the number of blocks touched in the sequence is increased by 1 [31-33]. The task continues until the participant fails to accurately reproduce the sequence [34]. The longest sequence that is reproduced correctly indexes the participant's visuospatial short-term memory span. Studies show that healthy adults have a typical span between 5 and 7 [35]. In more recent implementations, electronic versions of the task have been developed and delivered through computers or mobile devices, where targets light up in a specific sequence that participants reproduce by clicking or touching the targets. Even in these modern versions of the task, the typical span ranges between 5 and 7 [33].

Related Work

Immersive VR limits extraneous distractions from the surrounding environment [36,37]. This is because immersive VR setups entail wearing a head-mounted display (HMD) that effectively blocks out the external environment, thereby reducing potential environmental distractions that may interfere with task performance [38,39]. This feature may be particularly important in cognitive research, where maintaining participants' focus and attention is being assessed [40] and a measure of "true" ability under no external distraction is needed. It is, therefore, imperative to ascertain whether these gamified tasks, administered with or without immersive VR, produce the expected pattern of results as their traditional counterparts, an indication that they capture the same underlying cognitive processes. Recent studies have investigated the consistency and relative performance of computerized and VR-based cognitive tasks, revealing nuanced differences between the 2 modalities. These findings underscore the importance of task context and cognitive demands when evaluating their efficacy.

For example, Barrett et al [41] compared VR, 3D desktop, and 2D versions of a category learning task in which participants classify visual stimuli into categories based on 2 relevant features while ignoring a third irrelevant one. Feedback is provided after each trial to support learning. RTs in the 2D condition were significantly faster than in the VR and 3D conditions, and fixation durations were shorter, indicating reduced cognitive load for the 2D environment. However, participants reported higher levels of immersion and engagement in the VR and 3D conditions, suggesting that while traditional computerized tasks may facilitate quicker responses, VR environments offer richer experiential benefits.

This trade-off between speed and immersion is also reflected in findings from Bhargava et al [42], who examined cognitive task performance across VR and mobile-based conditions using custom-designed games. Bhargava et al [42] validated 3 custom-designed games: 2 VR-based (Navigation and Hand-Eye Coordination) and 1 tablet-based (Memory) against the ACE-III (Addenbrooke's Cognitive Examination—III) cognitive assessment tool in a cohort of young adults. The ACE-III is a comprehensive tool used to assess cognitive functioning across 5 domains: attention, memory, language, verbal fluency, and visuospatial skills. The study found no significant differences in overall task performance between the VR and 3D mobile conditions; however, the VR tasks elicited significantly higher motivation and engagement scores. These findings suggest that VR can enhance user experience and engagement without compromising the validity or reliability of cognitive assessment.

This enhanced motivational effect of VR-based tasks is further reinforced by Faria et al [43], who explored VR in the context of cognitive rehabilitation. Similar findings are reported by Faria et al [43], who evaluated 2 approaches to personalized cognitive rehabilitation—1 VR-based and 1 traditional—with patients with chronic stroke. Both methods were equally effective in improving cognitive performance, but the VR approach led to higher participant satisfaction and engagement due to its interactive features. This suggests that VR can deliver comparable therapeutic outcomes with added motivational benefits.

In addition to motivation, several studies emphasize the ecological validity of VR, a point illustrated well in the work by Tan et al [44]. Tan et al [44] assessed a VR-based cognitive screening tool targeting 6 cognitive domains. Their results showed high accuracy in detecting cognitive impairments, comparable with traditional methods, with the added benefit of enhanced ecological validity in VR. Participants were able to perform naturalistic tasks, such as navigation and object interaction, which are difficult to replicate in conventional computerized tests.

This growing interest in ecological validity and realism of cognitive assessments in VR is echoed in the meta-analysis by Neğüt et al [45], which further consolidates the strengths and challenges of immersive assessments. Finally, Neğüt et al [45] conducted a meta-analysis comparing VR-based assessments with classical paper-and-pencil or computerized measures. They reported that VR tasks often felt more challenging due to their immersive and dynamic nature, which can increase perceived cognitive load. However, the ecological validity and potential for more realistic stimuli in VR environments often make them more representative of real-world cognitive challenges.

In summary, past research shows that VR-based tasks enhance immersion, engagement, and ecological validity, making them a compelling alternative to traditional assessments. While RTs may vary depending on task design and context, this trade-off is often outweighed by the benefits of more naturalistic and interactive environments. These advantages position VR as a valuable complement to

conventional methods, particularly in contexts that require high levels of interaction, naturalistic task performance, or strong participant motivation.

Building on this body of work, this study extends previous research by directly comparing attention and memory performance in immersive VR with desktop settings, both in-lab and remote, to assess the impact of presentation medium and testing environment on cognitive outcomes. Related to this question, several past studies using a variety of tasks have indicated that the quality of data collected on the web remains uncompromised [46-52]. However, web-based data frequently exhibit greater variability in contrast to data gathered in controlled laboratory environments [51]. This variability could either stem from the diminished control over experimental conditions, as previously mentioned, or it could be an intrinsic characteristic of studying more diverse participant cohorts, thereby becoming a variable of interest in its own right [53].

This issue of variability is particularly relevant in light of findings by Segen et al [54], whose online versus lab-based comparison provides an important context for interpreting remote testing results. Notably, Segen et al [54] found that the data collected in a conventional laboratory setting and those collected on the web produced very similar results, although the web-based data were more variable, with SEs being about 10% larger than those of the data collected in the lab. However, Segen et al [54] have used a scene recognition task that might be less prone to the environmental distraction that is typically present when carrying out a task remotely compared with the attention and memory tasks we used here.

Problem Statement and Objective

The heightened engagement resulting from the use of VR may come at a cost to experimental control. The dynamic and immersive nature of VR environments introduces additional variables that may influence participants' cognitive processes, potentially altering the patterns of results obtained compared with their traditional counterparts. Therefore, it becomes imperative for scientists to systematically evaluate whether gamifying cognitive tasks in VR environments yields findings that are consistent with those obtained with traditional methods [19]. Doing so would ensure that while enhancing engagement with gamification, researchers can still maintain the methodological rigor necessary for robust scientific conclusions. Therefore, the main objective of this study was to examine whether gamified versions of the Visual Search, the Whack-the-Mole, and the Corsi block tasks, each assessing a distinct cognitive process, produce similar patterns of results as their traditional versions. A second objective was to examine whether the way the gamified tasks are administered influences the pattern of results obtained. To this purpose, we compared results across 3 different conditions that involved the same gamified cognitive tasks: an immersive VR condition (VR-Lab), a desktop VR condition (Desktop-Lab; nonimmersive), and a desktop VR condition in which participants carried out the task at home (Desktop-Remote; nonimmersive). By conducting this comparison, we aimed to elucidate the influence of immersive VR on

task outcomes, providing insights into the effectiveness of different task formats.

Methods

Participants

The experimental study included 75 healthy young adults (male=24, female=51) with an overall mean age of 23.15 years (SD 4.38 years) recruited from the student community of the University of Cyprus and the CYENS—Centre of Excellence. Recruitment was conducted using a convenience sampling method through email announcements, web-based student forums and platforms, and direct invitations. All participants met the following inclusion criteria: (1) aged between 18 and 35 years, (2) normal or corrected-to-normal vision, and (3) no reported history of neurological, psychiatric, or cognitive impairments that could affect task performance. Eligibility was determined through a brief prescreening interview. Individuals with a self-reported history of neurological disorders (eg, epilepsy and traumatic brain injury), psychiatric conditions (eg, attention-deficit/hyperactivity disorder, anxiety, and depression), or uncorrected visual impairments were excluded from participation. Participants were randomly assigned to 1 of 3 between-subject conditions: VR-Lab, Desktop-Lab, or Desktop-Remote, with 25 participants in each group. No participants were excluded from the statistical analyses, and there were no dropouts. All laboratory-based sessions (VR-Lab and Desktop-Lab conditions) took place at the Experimental Psychology Laboratory at the University of Cyprus, a controlled environment designed to minimize external distractions.

Materials

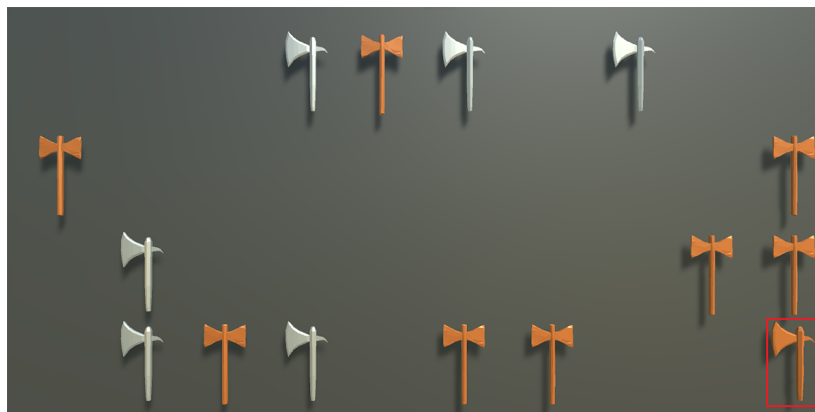
Visual Search Task

The Visual Search task used in this study involved the presentation of brown and gray axes (Multimedia Appendix 1), with the target object changing from trial to trial. In feature search trials, the target differed from the distractors in terms of color, for example, a gray axe presented among brown axes. In contrast, in conjunction search trials, the target was defined by a combination of features that were also present in the distractors. For example, the target could be a brown 1-sided axe presented among gray 1-sided axes and brown double-sided axes (Figure 1). To examine the influence of display size, we varied the number of axes in both trial types to 5, 15, 25, or 40. Participants were instructed to search the display for the target and provide a response as soon as they detected it. In every trial, the targets were presented for 1.5 seconds, followed by the search display. Overall, the task consisted of 96 trials (50% feature search trials and 50% conjunction search trials), with 8 trials in each possible combination of search type and display size. Before the experimental trials, participants performed 8 practice trials, 1 in each combination of search type (Feature vs Conjunction Search) and display size (5, 15, 25, or 40) to familiarize themselves with the task and procedure. Trials were presented in a different random order for each

participant. Participants in the VR-Lab condition searched for the target and selected it by directing a laser beam extended from the VR controller toward the target. The response was logged when the laser beam was on the target for 2 seconds; however, the RT was measured from the moment the beam first landed on the target. Participants in the Desktop-Lab and

Desktop-Remote conditions responded by moving the mouse cursor to the target and clicking on it. In all 3 conditions (VR-Lab, Desktop-Lab, and Desktop-Remote), participants had the whole board with the axes in their field of view and thus did not need to turn their heads to search.

Figure 1. Example of a conjunction search trial in the Visual Search task of the experimental study. In this trial, participants searched for a brown 1-sided axe (target) among gray 1-sided axes and brown double-sided axes (distractors). This is a trial with a display size of 15. The red outline is added for illustration purposes and was not present in the actual display viewed by participants.



Whack-the-Mole Task

In our version of the Whack-the-Mole task that was modeled after Casey et al [55], participants used a mallet to hit moles that popped up with random disguises from the hole in the center of the ground. In Go trials, a mole appeared with a varying disguise in each trial, and participants were instructed to hit it as fast as they could by either pulling the trigger on the handheld controller with the VR-Lab condition or clicking the mouse in the Desktop-Lab and Desktop-Remote conditions. Participants in the VR-Lab condition held a virtual mallet mapped to the controller in their dominant hand and hit the mole by pulling the trigger with their index finger. Similarly, participants in the Desktop-Lab and Desktop-Remote conditions clicked the mouse with the index

finger of their dominant hand to hit the mole. In No-Go trials, mushrooms appeared instead of the mole, and participants were instructed to withhold a response (Figure 2). Before the experimental task, participants were shown the 7 disguises of the mole and the 3 types of mushrooms and carried out 10 practice trials. As shown in Figure 2, the task included 7 different mole disguises and 3 different types of mushrooms. Each participant completed 100 trials, with 75 being the Go trials and 25 being the No-Go trials. The 2 trial types were presented randomly within the task. In all 3 conditions (VR-Lab, Desktop-Lab, and Desktop-Remote), participants had a clear view of the ground area from which the mole emerged, without the need to turn their heads.

Figure 2. An illustration of the Whack-the-Mole task as used in the study. On the left, the 3 different types of mushrooms (top row) and the 7 different mole disguises (middle and bottom rows) are shown. On the right, an example Go trial is shown. In the VR-Lab condition, participants had to pull the trigger on the handheld controller to respond. In the Desktop-Lab and Desktop-Remote conditions, they had to click the mouse. In No-Go trials, showing different types of mushrooms, participants had to refrain from any response.



Corsi Test

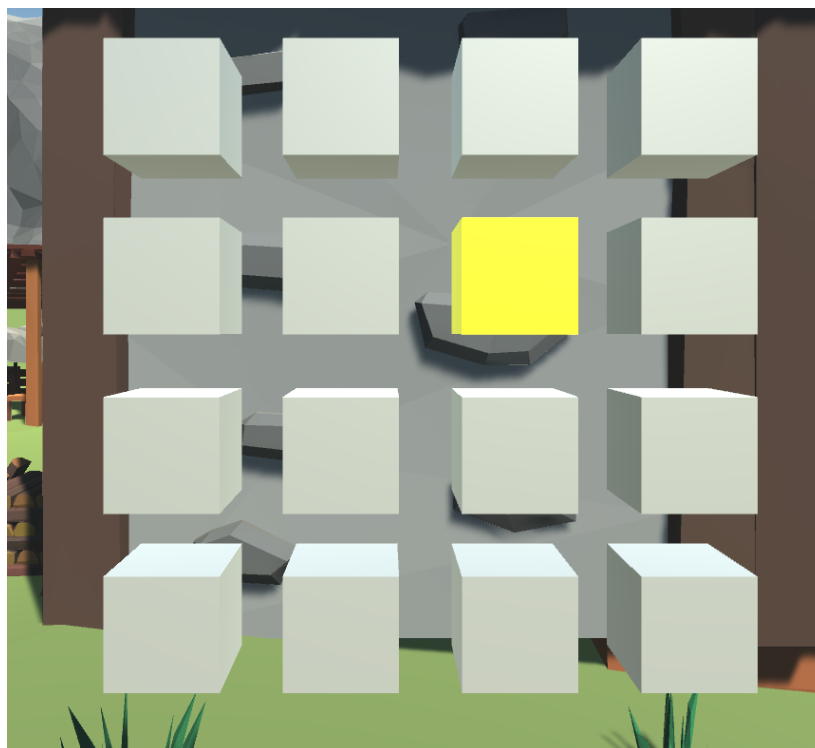
In this task, participants were presented with a 4×4 matrix of gray cubes (Figure 3). Each trial involved a unique sequence

of cubes changing color from gray to yellow sequentially, and participants were required to replicate the sequence by tapping the blocks that changed color in the same order.

Participants in the VR-Lab condition used the VR controllers to touch the blocks, whereas those in the Desktop-Lab and Desktop-Remote conditions clicked the blocks with the mouse. The task began with a single cube and the length of the sequence increased as participants responded correctly. Participants were given 3 attempts at each span length and advanced to the next length if they responded correctly in at least 2 attempts. The block span for each participant was

measured as the length of the last sequence that was repeated correctly in at least 2 out of 3 attempts. After completing the tasks, participants in the Desktop-Remote condition were asked to click on a save button that sent their data to a cloud database. In all 3 conditions (VR-Lab, Desktop-Lab, and Desktop-Remote), the whole 4×4 matrix was within the participants' field of view, so no head turning was required.

Figure 3. An illustration of the matrix in the Corsi Test as used in this study. Participants viewed a sequence of cubes turning yellow and then recreated it by either touching the cubes with the handheld controllers (VR-Lab condition) or clicking on each cube with the mouse (Desktop-Lab and Desktop-Remote participants).



Procedure

Participants were initially contacted via email with details about the study and available time slots for participation. Those interested registered electronically by selecting a suitable time slot. Upon arrival at the lab, participants read a detailed description of the study and signed a consent form if they agreed to take part. No participants withdrew from the experimental procedure. Participants in the Desktop-Remote condition followed the same process, with the only difference being that they completed the study from home through a secure hyperlink and communicated with the experimenter (first author) via video call. They completed the tasks on their personal computers, and they were instructed to select a quiet, distraction-free space and to use a desktop or laptop computer with a stable internet connection. During the experiment, all participants were seated and instructed to use only their hand to operate either the mouse or the VR controller, ensuring that movement was limited primarily to the wrist and minor elbow motion. This protocol was applied consistently across all 3 tasks, regardless of condition. Before starting the experiment, participants in all conditions underwent a brief orientation

session where they were familiarized with the tasks and response methods.

Participants in the VR-Lab condition carried out the tasks in an HTC Vive Pro HMD and used the Vive handheld controllers to respond. Participants in the Desktop-Lab condition carried out the task on the same desktop computer that was used for the VR-Lab condition but viewed the tasks on a 24" LCD display. The order of the 3 tasks was counter-balanced across conditions using a 3×3 Latin square design. For all tasks, data in the VR-Lab and Desktop-Lab conditions were automatically recorded by the computer running the experiment. In the 2 lab conditions, data were collected locally on the computer of the lab, while in the Desktop-Remote condition, data were encrypted and stored in a secure database.

Statistical Design

Separate mixed-design ANOVAs were conducted on mean RTs and d' , depending on the cognitive task (see "Results" section for task-specific details). The within-subject variables for each task are detailed in the "Results" section, while the between-subject variable across all tasks was the mode

of administration. In addition to the ANOVAs, Pearson correlations were computed between task scores to explore individual differences. All analyses were carried out using the Jamovi software package [56].

Ethical Considerations

The study conformed to European and national legislation and fundamental ethical principles, including those reflected in the Charter of Fundamental Rights of the European Union, the European Convention on Human Rights and its Supplementary Protocols, and the World Medical Association Declaration of Helsinki. The protocol of the study was given ethics clearance by the Cyprus National Bioethics Committee (approval number: EEBK EΠ 2021.01.99). All participants read and signed informed consent forms either in person or digitally. Data were collected and maintained without any participant identification information.

Results

Overview

To investigate whether the 3 gamified tasks replicated the pattern of results reported in the literature for their traditional counterparts and to compare performance across modes of administration, we computed inferential statistics separately for each task. We report the findings in the next subsections.

Visual Search

For Visual Search, we carried out a repeated-measures ANOVA on RT with search type and display size as the within-subject variables and the administration condition as the between-subject variable.

Results revealed significant main effects for search type, $F_{1,72}=787.12$, $P<.001$, $\eta^2=0.38$; the display size, $F_{3,216}=150.49$, $P<.001$, $\eta^2=0.17$; and the administration condition, $F_{2,72}=8.49$, $P=.001$, $\eta^2=0.02$. There was also a significant interaction between search type and display size, $F_{3,216}=145.42$, $P<.001$, $\eta^2=0.16$. As seen in Figure 4, the interaction was driven by the presence of a display size effect (ie, RT increasing from 5 to 15 and then to 25 and further to 40 objects) in conjunction search trials but not in feature search trials. The significant effect for the administration condition was due to shorter RT in the VR-Lab (mean 1.24 seconds) condition than in the Desktop-Lab (mean 1.49 seconds; $P<.001$) and Desktop-Remote (mean 1.44 seconds; $P=.008$) conditions (Figure 5). The difference between the Desktop-Lab and Desktop-Remote conditions was not significant ($P=1.00$). Importantly, no significant interaction involving the administration condition was found. The effect size for search type ($\eta^2=0.38$) is large, suggesting a robust difference in RTs between feature and conjunction searches, consistent with prior literature. The effect size for display size ($\eta^2=0.17$) is also large, highlighting the substantial impact of increasing display size on RTs. In contrast, the main effect of the administration condition ($\eta^2=0.02$) has a small effect size, indicating that the mode of task administration had a statistically significant but relatively small influence on RTs. Finally, the interaction effect between search type and display size was associated with a large effect size ($\eta^2=0.16$), reinforcing the well-documented finding that increasing display size disproportionately affects conjunction searches compared with feature searches.

Figure 4. The average reaction time in the Visual Search task as a function of search type (feature search vs conjunction search) and display size (5, 15, 25, and 40 objects) for the VR-Lab, Desktop-Lab, and Desktop-Remote conditions. Results are shown in distinct panels for the VR-Lab, the Desktop-Lab, and the Desktop-Remote conditions. Error bars represent 95% confidence intervals.

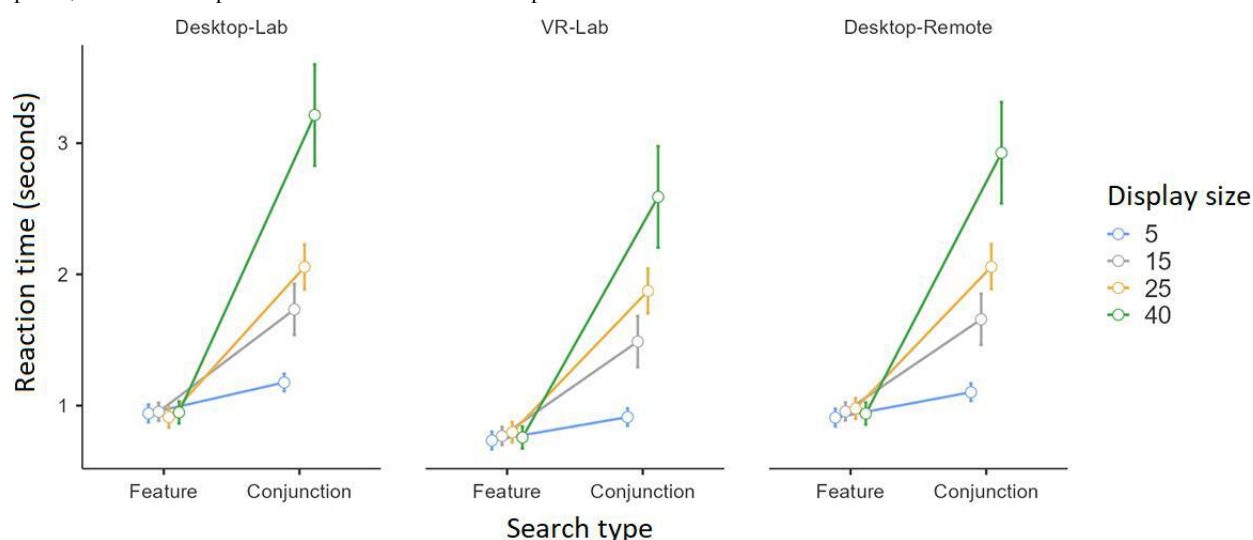
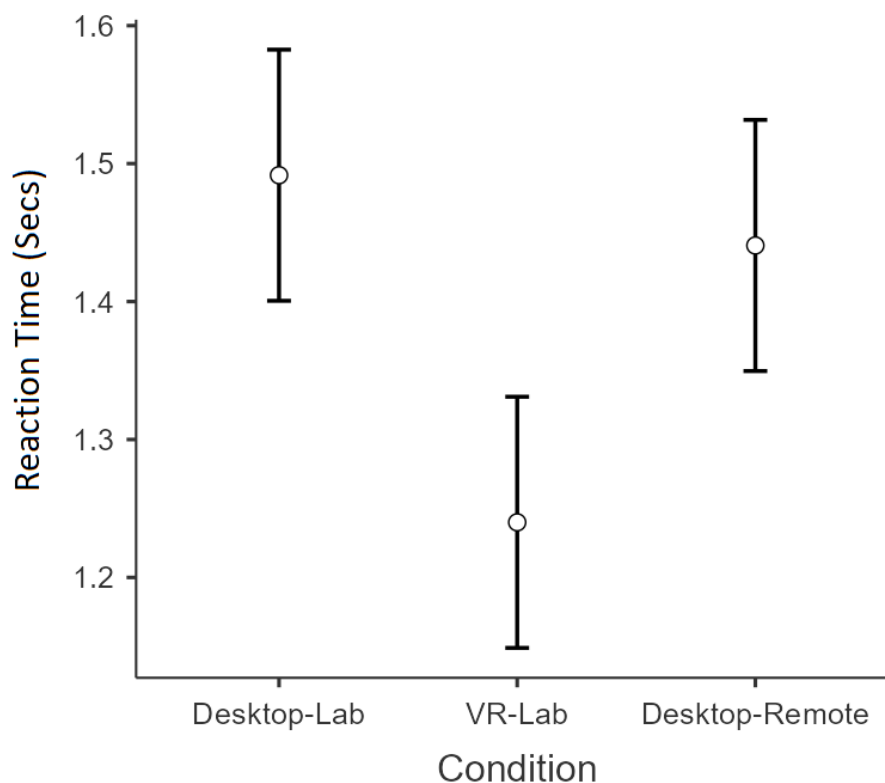


Figure 5. Mean reaction times in the Visual Search task, aggregated across search type and display size, for each of the 3 administration conditions: VR-Lab, Desktop-Lab, and Desktop-Remote. Overall, reaction time was shorter in the VR-Lab condition than in the Desktop-Lab and Desktop-Remote conditions. Error bars represent 95% confidence intervals.



Whack-the-Mole

To analyze the accuracy data for the Whack-the-Mole task, we first computed for each participant the percentages of hits and false alarms. Based on the percentages, we computed the d' score for each participant. The d' score is a sensitive discrimination measure that reflects the degree to which participants accurately report the presence or absence of the target in the display. The d' score was calculated using the formula: $d' = z(\text{hit rate}) - z(\text{false alarm rate})$.

The average d' scores were 3.79 for the VR-Lab, 3.75 for the Desktop-Remote, and 3.62 for the Desktop-Lab

condition (Figure 6), in line with the values typically reported in the literature [20,30]. A one-way ANOVA showed that the difference in d' across the VR-Lab, Desktop-Lab, and Desktop-Remote conditions was not significant, $F_{2,72}=0.71$, $P=.49$, $\eta^2=0.02$. In contrast, the analysis on mean RTs for hits only (ie, correct responses in Go trials) showed that participants were significantly slower in the Desktop-Remote condition (mean 0.64 second) than in the Desktop-Lab (mean 0.48 second) and VR-Lab (mean 0.41 second) conditions, both P values $<.001$ (Figure 7). The difference in RTs between the VR-Lab and the Desktop-Lab conditions was marginally significant ($P=.07$).

Figure 6. The average d' values in the Whack-the-Mole task across the 3 administration conditions. The d' is computed, according to Signal Detection Theory, by taking into account the percentage of hits (ie, correct responses to targets) and false alarms (ie, incorrect responses to distractors). Error bars depict 95% confidence intervals.

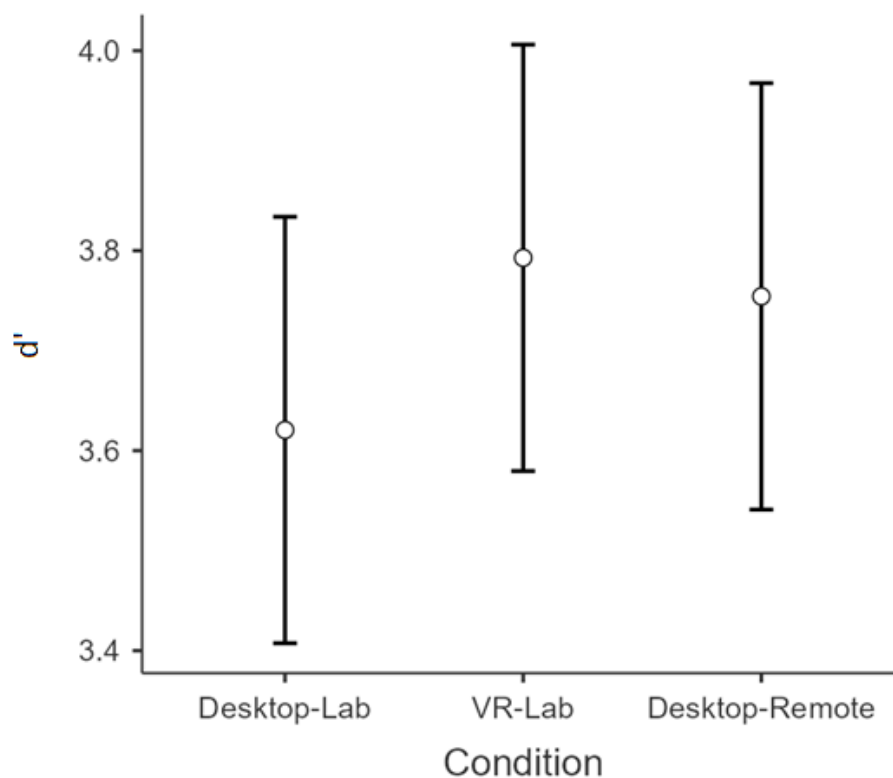
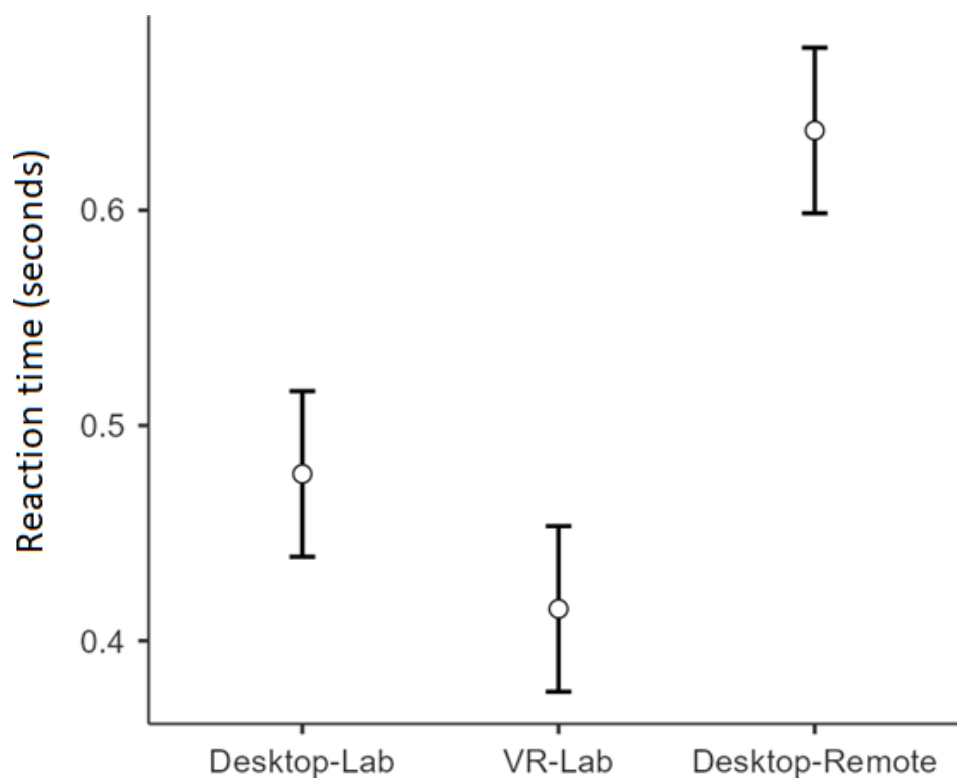


Figure 7. Average reaction time for Hits (ie, correct responses to targets) in the Whack-the-Mole task across the 3 administration conditions. Error bars depict 95% confidence intervals.



Corsi Test

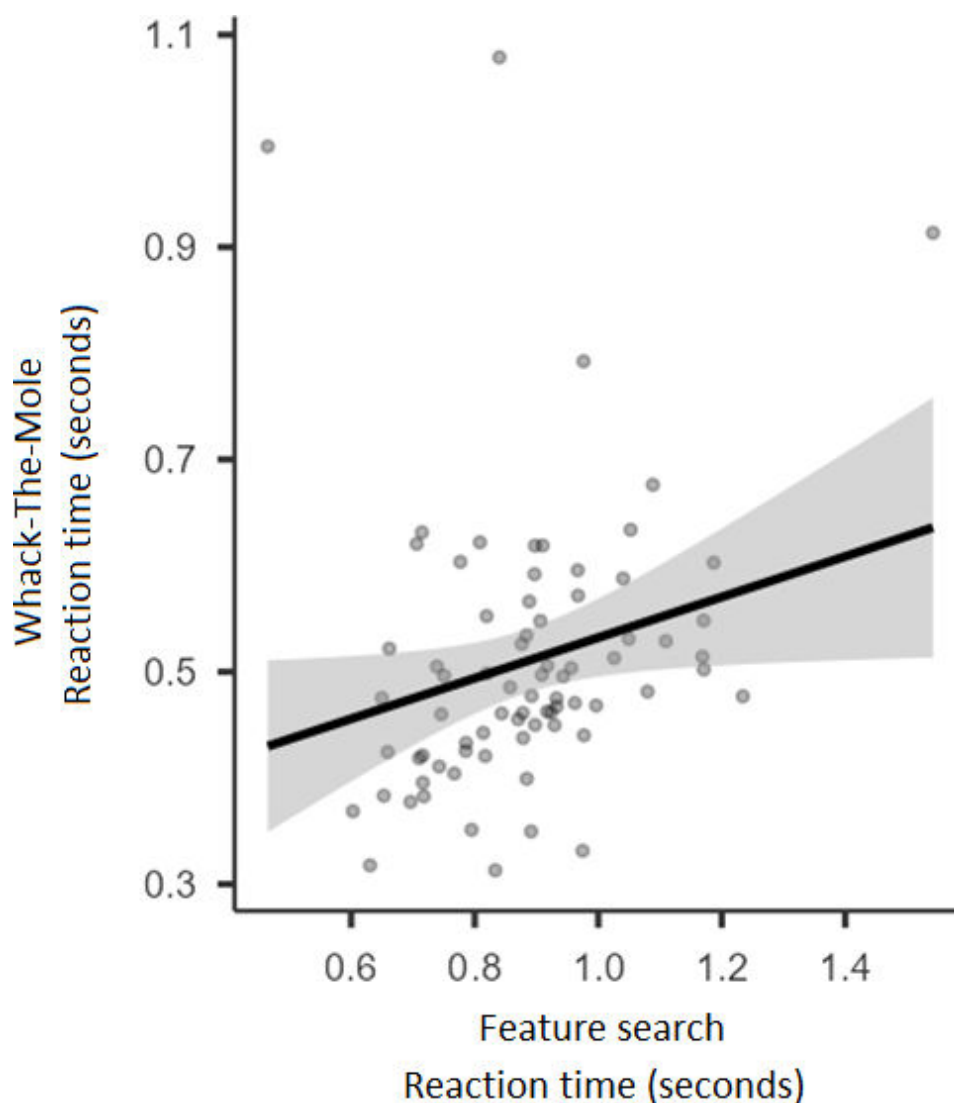
As shown in [Multimedia Appendix 2](#), the average span scores in the Corsi Test were 5.48 in the VR-Lab condition, 5.68 in the Desktop-Lab condition, and 5.24 in the Desktop-Remote condition. Despite the numeric difference across the 3 conditions, a one-way ANOVA showed that the effect of the administration condition on the Corsi span score was not significant, $F_{2,72}=1.45$, $P=.24$, $\eta^2=0.04$. These findings indicate that the administration condition had no impact on spatial working memory performance, as measured by the Corsi Test.

Individual Differences Within Attentional Tasks

To examine a possible relation between the 2 distinct attentional processes, that is, Visual Search and response

inhibition, we carried out Pearson correlation analyses between the mean RT in feature search and conjunction search trials of the Visual Search task separately and the mean RT of Hits in the Whack-the-Mole task. For these, we averaged the RT across displays for each search type and we combined the data from the 3 administration conditions per task. We found a significant positive correlation between RT for Hits in the Whack-the-Mole task and RT in feature search trials of the Visual Search task, $r_{73}=0.24$; $P=.04$ ([Figure 8](#)). The correlation between RT for Hits in the Whack-the-Mole task and RT for the conjunction search trials of the Visual Search task was not significant, $r_{73}=-0.04$; $P=.74$.

Figure 8. Scatter plot depicting the correlation of Whack-the-Mole reaction time for Hits and Visual Search reaction time for Feature Search trials. The significant positive correlation indicates that participants who were faster to respond in Go trials of the Whack-the-Mole task tended to respond faster in the Feature Search trials of the Visual Search task.



Discussion

Principal Results

The main objective of this study was to investigate whether the gamified cognitive tasks we developed yield the same patterns of results as their traditional versions. A secondary aim was to determine whether the mode of administration (immersive VR vs desktop VR) and the task administered remotely or in the lab affect performance outcomes.

With regard to the main objective, we found that our gamified tasks replicated the traditional patterns observed in previous studies. Specifically, in the Visual Search task, RTs were shorter for feature search trials than for conjunction search trials, aligning with the established finding that feature searches are overall easier [15]. Additionally, in conjunction search trials, RT increased significantly with display size, a finding known as the display size effect [15,17,18]. Conversely, RTs remained relatively constant across display sizes in feature search trials, demonstrating the well-established pop-out effect [15]. In the Whack-the-Mole task, the overall response inhibition performance aligns with past studies indicating a typical d' score between 3 and 4, which reflects moderate sensitivity [20,49]. Our participants exhibited false alarms in about 10% of No-Go trials, consistent with the rate reported in the literature [20-22]. Furthermore, in assessing visual short-term memory, our findings showed that our participants had a span between 5 and 7, which is comparable with that reported by studies using the original version of the task [35], as well as those using modern electronic implementations delivered via computers or mobile devices [33].

While replicating the expected pattern of results across the 3 administration conditions, we also found some differences across these conditions. For example, performance in the Visual Search task was faster in the VR-Lab condition by 250 milliseconds than in the Desktop-VR condition. One possibility is that this advantage for immersive VR is due to diminished distraction resulting from blocking out extraneous environmental information when wearing an HMD. However, it could also be due to methodological differences between the 2 administration conditions. Specifically, participants in the Desktop-Lab condition responded by moving and clicking the mouse on the target object. Thus, RT encompassed the time to locate the target, the time to move the mouse, and the time to click. In contrast, in the VR-Lab condition, participants pointed the VR controller to the target object without requiring them to press any button. Thus, RT in this condition involved the time to locate the target and the time to move the controller in 3D space. Perhaps moving the controller in 3D space is performed faster than moving a mouse cursor on a screen. Or perhaps, the difference can be attributed to the need to click the mouse in the desktop condition but not in the VR condition. Notably, we found no difference between the Desktop-Lab and Desktop-Remote conditions, suggesting that differences in the environmental settings (eg, varying screen sizes and distances from the screen at participant homes) and environmental distractors that may be greater at home than

the laboratory do not interfere significantly with the assessment of Visual Search abilities.

In the Whack-the-Mole task, performance in the VR-Lab condition was marginally faster by 70 milliseconds than in the Desktop-Lab condition. In contrast to the Visual Search task, in the Whack-the-Mole task, the response requirements were rather similar across the 2 administration conditions. That is, in the VR-Lab condition, participants pulled the trigger of the controller to hit the mole, while in the desktop versions, they clicked the mouse. Thus, taking the results of the 2 tasks together, it seems more likely that in the Whack-the-Mole task, the advantage of immersive VR is due to blocking out access to distracting information from the external environment. Nevertheless, we cannot safely exclude the possibility that the difference is due to subtle differences in the response mode. Perhaps the time needed to pull the trigger on the VR controller is shorter than clicking the mouse button. However, it should be noted that the performance in the Desktop-Lab condition was faster by 160 milliseconds than in the Desktop-Remote condition, despite the fact that the response mode in these 2 conditions was identical. This finding suggests that the differences in overall performance across administration conditions in the Whack-the-Mole task are not methodological.

A possible explanation for the RT differences across the administration conditions of the Whack-the-Mole task is offered by the Perceptual Load Theory. This theory was proposed by Lavie and Tsal [57,58] in the mid-1990s as a potential resolution to the early versus late selection debate in attention and perception. Lavie and Tsal argued that when a task is perceptually demanding (referred to as high load), people dedicate all processing resources to the relevant stimuli early on in the information-processing stream. This results in no interference from distracting irrelevant stimuli, as no spare resources are available to process them. In contrast, when the task is not perceptually demanding (ie, a low-load task), spare resources are available after processing the relevant stimuli. These resources, according to the theory, automatically spill into the processing of irrelevant stimuli. Thus, in low-load tasks, selection takes place late with irrelevant stimuli exerting an influence on performance [59-62]. Notably, in our own research using empirical methods and computational modeling [63], we have shown that inducing a narrow focus of attention may mediate whether distractor effects are present. That is, when the tasks force participants to zoom in their attention to stimuli on the screen, rather than having a more relaxed focus of attention, distracting information can exert no influence even if the task is of low load. Here, the Whack-the-Mole can be more easily classified as low load than high load. More importantly, it can be executed efficiently by maintaining a broader focus of attention, that is, by just focusing attention on a large part in the center of the screen rather than zooming in on a small stimulus. Carrying out the task this way allows for external distractors to slow down performance. If participants' homes include more environmental distractors than a laboratory, this would translate to more extraneous influence on performance and slowing down of responses. This could also account

for the advantage of the VR-Lab condition as donning an HMD serves to effectively block environmental distractors that could influence performance.

A question that arises, though, is why there was no difference between the Desktop-Lab and Desktop-Remote conditions in the Visual Search task? On one hand, the conjunction search trials of the Visual Search task can be thought of as a high-low task that requires a narrow focus of attention, limiting resources to the processing of the relevant stimuli. This would explain why there were no differences across administration conditions. On the other hand, this cannot apply to the feature search trials. Feature search trials are more likely a low load task and, as such, would allow for the processing of distracting information. Yet, no differences were found across administration conditions. Therefore, it seems that further research, in which response modes and other methodological differences are controlled for, is needed to draw more definitive conclusions.

Comparison With Prior Work

This study examined the impact of administration mode (VR-Lab, Desktop-Lab, and Desktop-Remote) on cognitive performance across 3 tasks: Visual Search, Whack-the-Mole, and Corsi Test. Overall, the results demonstrate that while VR tasks provide a highly immersive and engaging platform for cognitive assessment, the performance differences between VR and traditional desktop-based conditions are task-dependent and vary in magnitude.

Our results from the Visual Search task replicate the well-documented effects of search type and display size [14,15,19], with longer RT observed for conjunction searches than for feature searches and RT increasing with display size in conjunction search trials. Notably, participants in the VR-Lab condition exhibited significantly shorter RTs than those in the Desktop-Lab and Desktop-Remote conditions, suggesting that the immersive nature of VR may enhance visual attention and processing speed in this task. This finding aligns with previous research indicating that VR environments can improve cognitive task performance through increased engagement and spatial immersion [41]. Barrett et al [41] reported longer RT and higher fixation durations in VR compared with 2D environments, which they attributed to the added cognitive demands of immersion. Our results diverge from those of Barrett et al, as the VR-Lab condition here resulted in faster RT, potentially due to the optimized laboratory setup and controlled task parameters. Notably, these patterns of results were obtained with much fewer trials than typically included in the traditional versions of the tasks. For instance, while we ran our gamified Visual Search task with only 96 trials, we produced the same pattern of results as Huang and Pashler [64] and Bravo and Nakayama [65], who ran 1000 and 240 trials, respectively.

The Whack-the-Mole task revealed no significant differences in d' scores across administration conditions, indicating that accuracy in response inhibition tasks is consistent across VR and desktop modalities. These findings are consistent with those by Bhargava et al [42], who found that VR tasks achieved comparable accuracy

to traditional assessments while enhancing engagement. However, RTs for Hits were significantly slower in the Desktop-Remote condition than in both the VR-Lab and Desktop-Lab conditions. This delay may reflect a lack of environmental standardization in remote settings, which has been noted as a limitation of remote cognitive testing [45]. The marginally faster RT in the VR-Lab condition than in the Desktop-Lab condition ($P=.07$) further highlights the potential for VR to enhance cognitive-motor coordination, as observed in other VR-based cognitive tasks [43]. In addition, while Shimi et al [20] and Casey et al [55] had 220 trials and 300 trials, respectively, in the Whack-the-Mole task they administered, we obtained the same results with only 100 trials.

The Corsi Test revealed no significant differences in average span scores across conditions, supporting the conclusion that spatial working memory is not influenced by the mode of administration. This aligns with prior findings that traditional and VR-based versions of the Corsi Test produce comparable results [31,33]. While the observed numeric differences were less and nonsignificant, the slightly lower scores in the Desktop-Remote condition may reflect environmental distractions or variability in participant equipment, which can affect remote testing reliability [45].

The positive correlation between RT for Hits in the Whack-the-Mole task and feature search trials in the Visual Search task suggests an underlying link between response inhibition and attentional processes, particularly in tasks involving rapid response execution. However, the absence of a significant correlation between RT in conjunction search trials and response inhibition suggests that these processes may operate independently under higher cognitive demands. These findings add to the growing literature on individual differences in attentional control [44] and further validate the use of gamified tasks in exploring these relationships.

Our findings emphasize the importance of task design and administration context in cognitive testing. While VR can enhance engagement and, in some cases, improve task performance, the added cognitive load of immersive environments may not universally benefit all cognitive processes [41,45]. The only minor differences in accuracy and spatial memory performance across administration modes highlight the robustness of these tasks to different delivery methods, including remote testing. However, slower RT in remote conditions underscores the need for careful standardization and participant instructions when deploying tasks outside laboratory settings.

Future studies should explore the impact of task complexity and individual differences on performance across VR and traditional modalities. Additionally, longitudinal research is needed to assess whether the observed benefits of VR in certain cognitive domains translate into practical advantages in real-world applications, such as rehabilitation and training.

Limitations

In this study, we did not assess participants' prior experience with VR. Previous exposure to VR could have influenced RTs, with participants who have used VR systems before potentially showing faster RTs due to greater familiarity with the technology. This factor should be considered in future studies, and including a questionnaire to assess past VR experience would be a valuable addition to control for this variable and better understand its impact on performance.

While this study focuses on the potential engagement benefits of VR-based cognitive tasks, it is important to acknowledge that VR can also introduce challenges. One key limitation is the potential for cybersickness, eye strain, and general fatigue associated with prolonged VR use. These factors can negatively impact participant performance [66,67] and may have contributed to variability in the results, particularly among participants who were less familiar with VR technology. Although we did not directly assess these issues by asking participants, no participants in the VR group reported any difficulties or discomfort with VR during or after their participation. While this suggests that cybersickness or fatigue was not a major confound, future studies could benefit from incorporating standardized self-report questionnaires or physiological measures to better isolate the impact of VR engagement on cognitive performance.

Notably, we did not include a paper-and-pencil comparison group, as our primary aim was to contrast the gamified VR tasks with traditional desktop-based digital assessments. Including a paper-and-pencil comparison could provide insights to cross-modality generalizability, but it would also introduce additional variables unrelated to our core research objective, such as differences in response modalities and

environmental influences. Nevertheless, we acknowledge that such a comparison could be valuable in future research to examine broader ecological validity, recognizing that certain tasks in our study (eg, Whack-the-Mole) are not feasible in a paper-based format.

Conclusions

This study demonstrated that gamified VR tasks can replicate well-established cognitive performance patterns using fewer trials, highlighting their potential as efficient and engaging tools for assessing attention, response inhibition, and visual short-term memory. By comparing performance across immersive VR, desktop-based, and remote administration conditions, we found that VR-based tasks maintained ecological validity and produced results consistent with traditional methods, suggesting that increased immersion does not compromise data quality. These advantages are particularly relevant for reducing participant fatigue and enhancing motivation, factors essential for high-quality cognitive assessment.

At the same time, VR introduces methodological variables that require further attention. In this study, we did not directly assess cybersickness, fatigue, or prior VR experience, which may influence individual performance. In addition, comparisons to paper-and-pencil tasks were not included, as the tasks examined here are typically administered in computerized form. Future research should explore these aspects systematically and investigate how task-specific features and environmental settings interact with immersion to affect cognitive outcomes. Overall, our findings support the use of immersive VR as a viable and scalable platform for cognitive assessment, especially in applications that benefit from high engagement and ecological realism.

Acknowledgments

This project has received partial funding from the European Union's Horizon 2020 Research and Innovation Programme (grant agreement number 739578) and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation, and Digital Policy.

Data Availability

The aggregated data file on which all analyses were carried out is freely available on the Open Science Framework at <https://osf.io/e9ry2>

Conflicts of Interest

None declared.

Multimedia Appendix 1

Example of a feature search trial used in the Visual Search task of this experimental study. In this trial, participants searched for a gray 2-sided axe (target) among 24 brown 2-sided axes (distractors). This is a trial with a display size of 25. The red outline is added for illustration purposes and was not present in the actual display viewed by participants.

[PNG File (Portable Network Graphics File), 428 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Mean span scores in the Corsi block-tapping task across the 3 administration conditions. Span scores reflect the longest sequence of targets correctly recalled by participants. Error bars represent 95% confidence intervals.

[PNG File (Portable Network Graphics File), 22 KB-Multimedia Appendix 2]

References

1. Smith EE, Jonides J. Storage and executive processes in the frontal lobes. *Science*. Mar 12, 1999;283(5408):1657-1661. [doi: [10.1126/science.283.5408.1657](https://doi.org/10.1126/science.283.5408.1657)] [Medline: [10073923](https://pubmed.ncbi.nlm.nih.gov/10073923/)]
2. Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci*. 2001;24(1):167-202. [doi: [10.1146/annurev.neuro.24.1.167](https://doi.org/10.1146/annurev.neuro.24.1.167)] [Medline: [11283309](https://pubmed.ncbi.nlm.nih.gov/11283309/)]
3. Eriksen BA, Eriksen CW. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept Psychophys*. Jan 1974;16(1):143-149. [doi: [10.3758/BF03203267](https://doi.org/10.3758/BF03203267)]
4. Posner MI, Petersen SE. The attention system of the human brain. *Annu Rev Neurosci*. 1990;13:25-42. [doi: [10.1146/annurev.ne.13.030190.000325](https://doi.org/10.1146/annurev.ne.13.030190.000325)] [Medline: [2183676](https://pubmed.ncbi.nlm.nih.gov/2183676/)]
5. Mekler ED, Brühlmann F, Opwis K, Tuch AN. Disassembling gamification: the effects of points and meaning on user motivation and performance. Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; Apr 27 to May 2, 2013; Paris, France.
6. Lang A, Bradley SD, Park B, Shin M, Chung Y. Parsing the resource pie: using STRTs to measure attention to mediated messages. *Media Psychol*. Nov 2006;8(4):369-394. [doi: [10.1207/s1532785xmep0804_3](https://doi.org/10.1207/s1532785xmep0804_3)]
7. Warm JS, Parasuraman R, Matthews G. Vigilance requires hard mental work and is stressful. *Hum Factors*. Jun 2008;50(3):433-441. [doi: [10.1518/001872008X312152](https://doi.org/10.1518/001872008X312152)] [Medline: [18689050](https://pubmed.ncbi.nlm.nih.gov/18689050/)]
8. Parsons TD. Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Front Hum Neurosci*. 2015;9:660. [doi: [10.3389/fnhum.2015.00660](https://doi.org/10.3389/fnhum.2015.00660)] [Medline: [26696869](https://pubmed.ncbi.nlm.nih.gov/26696869/)]
9. Slater M, Sanchez-Vives MV. Enhancing Our Lives with Immersive Virtual Reality. *Front Robot AI*. Dec 19, 2016;3:74. [doi: [10.3389/frobt.2016.00074](https://doi.org/10.3389/frobt.2016.00074)]
10. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". Presented at: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; Sep 28-30, 2011; Tampere, Finland.
11. Hamari J, Koivisto J, Sarsa H. Does gamification work?—A literature review of empirical studies on gamification. Presented at: 2014 47th Hawaii International Conference on System Sciences; Jan 6-9, 2014; Waikoloa, HI, USA.
12. Lee CC, Hsiao KL, Chen CC. Exploring the benefit and sacrifice factors of virtual reality gameplay. *Front Psychol*. 2020;11:251. [doi: [10.3389/fpsyg.2020.00251](https://doi.org/10.3389/fpsyg.2020.00251)] [Medline: [32210870](https://pubmed.ncbi.nlm.nih.gov/32210870/)]
13. Huang W, Roscoe RD, Johnson-Glenberg MC, Craig SD. Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion. *J Comput Assist Learn*. Jun 2021;37(3):745-758. [doi: [10.1111/jcal.12520](https://doi.org/10.1111/jcal.12520)]
14. Mittelstaedt JM, Wacker J, Stelling D. VR aftereffect and the relation of cybersickness and cognitive performance. *Virtual Real*. Jun 2019;23(2):143-154. [doi: [10.1007/s10055-018-0370-3](https://doi.org/10.1007/s10055-018-0370-3)]
15. Wolfe JM. Major issues in the study of visual search: part 2 of "40 years of feature integration: special issue in memory of Anne Treisman". *Atten Percept Psychophys*. Feb 2020;82(2):383-393. [doi: [10.3758/s13414-020-02022-1](https://doi.org/10.3758/s13414-020-02022-1)] [Medline: [32291612](https://pubmed.ncbi.nlm.nih.gov/32291612/)]
16. Morar SS, Macredie RD, Cribbin T. An investigation of visual cues used to create and support frames of reference and visual search tasks in desktop virtual environments. *Virtual Real*. Oct 1, 2002;6(3):140-150. [doi: [10.1007/s100550200015](https://doi.org/10.1007/s100550200015)]
17. Smilek D, Enns JT, Eastwood JD, Merikle PM. Relax! Cognitive strategy influences visual search. *Vis Cogn*. Aug 2006;14(4-8):543-564. [doi: [10.1080/13506280500193487](https://doi.org/10.1080/13506280500193487)]
18. Mueller ST. A partial implementation of the Bica Cognitive Decathlon Using the psychology experiment building language (PEBL). *Int J Mach Conscious*. Dec 2010;02(2):273-288. [doi: [10.1142/S1793843010000497](https://doi.org/10.1142/S1793843010000497)]
19. Lumsden J, Edwards EA, Lawrence NS, Coyle D, Munafò MR. Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. *JMIR Serious Games*. Jul 15, 2016;4(2):e11. [doi: [10.2196/games.5888](https://doi.org/10.2196/games.5888)] [Medline: [27421244](https://pubmed.ncbi.nlm.nih.gov/27421244/)]
20. Shimi A, Tsestou V, Hadjjaros M, Neokleous K, Avraamides M. Attentional skills in soccer: evaluating the involvement of attention in executing a goalkeeping task in virtual reality. *Appl Sci (Basel)*. Jan 2021;11(19):9341. [doi: [10.3390/app11199341](https://doi.org/10.3390/app11199341)]
21. Cragg L, Nation K. Go or no-go? Developmental improvements in the efficiency of response inhibition in mid-childhood. *Dev Sci*. Nov 2008;11(6):819-827. [doi: [10.1111/j.1467-7687.2008.00730.x](https://doi.org/10.1111/j.1467-7687.2008.00730.x)] [Medline: [19046150](https://pubmed.ncbi.nlm.nih.gov/19046150/)]
22. Nigg J, Nikolas M, Friderici K, Park L, Zucker RA. Genotype and neuropsychological response inhibition as resilience promoters for attention-deficit/hyperactivity disorder, oppositional defiant disorder, and conduct disorder under conditions of psychosocial adversity. *Dev Psychopathol*. 2007;19(3):767-786. [doi: [10.1017/S0954579407000387](https://doi.org/10.1017/S0954579407000387)] [Medline: [17705902](https://pubmed.ncbi.nlm.nih.gov/17705902/)]

23. Allan NP, Hume LE, Allan DM, Farrington AL, Lonigan CJ. Relations between inhibitory control and the development of academic skills in preschool and kindergarten: a meta-analysis. *Dev Psychol.* Oct 2014;50(10):2368-2379. [doi: [10.1037/a0037493](https://doi.org/10.1037/a0037493)] [Medline: [25069051](https://pubmed.ncbi.nlm.nih.gov/25069051/)]
24. Schmitt LM, White SP, Cook EH, Sweeney JA, Mosconi MW. Cognitive mechanisms of inhibitory control deficits in autism spectrum disorder. *J Child Psychol Psychiatry.* May 2018;59(5):586-595. [doi: [10.1111/jcpp.12837](https://doi.org/10.1111/jcpp.12837)] [Medline: [29052841](https://pubmed.ncbi.nlm.nih.gov/29052841/)]
25. Aron AR, Poldrack RA. The cognitive neuroscience of response inhibition: relevance for genetic research in attention-deficit/hyperactivity disorder. *Biol Psychiatry.* Jun 1, 2005;57(11):1285-1292. [doi: [10.1016/j.biopsych.2004.10.026](https://doi.org/10.1016/j.biopsych.2004.10.026)] [Medline: [15950000](https://pubmed.ncbi.nlm.nih.gov/15950000/)]
26. van Hulst BM, de Zeeuw P, Vlaskamp C, Rijks Y, Zandbelt BB, Durston S. Children with ADHD symptoms show deficits in reactive but not proactive inhibition, irrespective of their formal diagnosis. *Psychol Med.* Nov 2018;48(15):2515-2521. [doi: [10.1017/S0033291718000107](https://doi.org/10.1017/S0033291718000107)] [Medline: [29415788](https://pubmed.ncbi.nlm.nih.gov/29415788/)]
27. Barker V, Romaniuk L, Cardinal RN, Pope M, Nicol K, Hall J. Impulsivity in borderline personality disorder. *Psychol Med.* Jul 2015;45(9):1955-1964. [doi: [10.1017/S0033291714003079](https://doi.org/10.1017/S0033291714003079)] [Medline: [25600066](https://pubmed.ncbi.nlm.nih.gov/25600066/)]
28. Evans AH, Strafella AP, Weintraub D, Stacy M. Impulsive and compulsive behaviors in Parkinson's disease. *Mov Disord.* Aug 15, 2009;24(11):1561-1570. [doi: [10.1002/mds.22505](https://doi.org/10.1002/mds.22505)] [Medline: [19526584](https://pubmed.ncbi.nlm.nih.gov/19526584/)]
29. Hirose S, Chikazoe J, Watanabe T, et al. Efficiency of go/no-go task performance implemented in the left hemisphere. *J Neurosci.* Jun 27, 2012;32(26):9059-9065. [doi: [10.1523/JNEUROSCI.0540-12.2012](https://doi.org/10.1523/JNEUROSCI.0540-12.2012)] [Medline: [22745505](https://pubmed.ncbi.nlm.nih.gov/22745505/)]
30. Gutiérrez-Cobo MJ, Cabello R, Fernández-Berrocal P. The three models of emotional intelligence and performance in a hot and cool go/no-go task in undergraduate students. *Front Behav Neurosci.* 2017;11:33. [doi: [10.3389/fnbeh.2017.00033](https://doi.org/10.3389/fnbeh.2017.00033)] [Medline: [28275343](https://pubmed.ncbi.nlm.nih.gov/28275343/)]
31. Corsi PM. Human Memory and the Medial Temporal Region of the Brain [PhD Thesis]. ProQuest Information & Learning; 1973.
32. Claessen MHG, van der Ham IJM, van Zandvoort MJE. Computerization of the standard Corsi block-tapping task affects its underlying cognitive concepts: a pilot study. *Appl Neuropsychol Adult.* 2015;22(3):180-188. [doi: [10.1080/23279095.2014.892488](https://doi.org/10.1080/23279095.2014.892488)] [Medline: [25258029](https://pubmed.ncbi.nlm.nih.gov/25258029/)]
33. Siddi S, Preti A, Lara E, et al. Comparison of the touch-screen and traditional versions of the Corsi block-tapping test in patients with psychosis and healthy controls. *BMC Psychiatry.* Jun 23, 2020;20(1):329. [doi: [10.1186/s12888-020-02716-8](https://doi.org/10.1186/s12888-020-02716-8)] [Medline: [32576254](https://pubmed.ncbi.nlm.nih.gov/32576254/)]
34. Hadjjaros M, Neokleous K, Schiza E, Matsangidou M, Avraamides MN, Pattichis CS. A game-based cognitive assessment for visuospatial tasks: evaluation in healthy adults. Presented at: 2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE); Oct 25-27, 2021:1-5; Kragujevac, Serbia. [doi: [10.1109/BIBE52308.2021.9635507](https://doi.org/10.1109/BIBE52308.2021.9635507)]
35. Kessels RP, van Zandvoort MJ, Postma A, Kappelle LJ, de Haan EH. The Corsi Block-Tapping Task: standardization and normative data. *Appl Neuropsychol.* 2000;7(4):252-258. [doi: [10.1207/S15324826AN0704_8](https://doi.org/10.1207/S15324826AN0704_8)] [Medline: [11296689](https://pubmed.ncbi.nlm.nih.gov/11296689/)]
36. Dahlstrom-Hakki I, Alstad Z, Asbell-Clarke J, Edwards T. The impact of visual and auditory distractions on the performance of neurodiverse students in virtual reality (VR) environments. *Virtual Real.* Mar 2024;28(1). [doi: [10.1007/s10055-023-00933-6](https://doi.org/10.1007/s10055-023-00933-6)]
37. Hung SC, Ho AY, Lai IH, Lee CS, Pong AS, Lai FH. Meta-analysis on the effectiveness of virtual reality cognitive training (VRCT) and computer-based cognitive training (CBCT) for individuals with mild cognitive impairment (MCI). *Electronics (Basel).* Dec 18, 2020;9(12):2185. [doi: [10.3390/electronics9122185](https://doi.org/10.3390/electronics9122185)]
38. Slater M, Wilbur S. A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence Teleoperators Virtual Environ.* Dec 1997;6(6):603-616. [doi: [10.1162/pres.1997.6.6.603](https://doi.org/10.1162/pres.1997.6.6.603)]
39. Bailenson JN, Beall AC, Loomis J, Blascovich J, Turk M. Transformed social interaction: decoupling representation from behavior and form in collaborative virtual environments. *Presence Teleoperators Virtual Environ.* Aug 1, 2004;13(4):428-441. [doi: [10.1162/1054746041944803](https://doi.org/10.1162/1054746041944803)]
40. Bailenson J, Patel K, Nielsen A, Bajscy R, Jung SH, Kurillo G. The effect of interactivity on learning physical actions in virtual reality. *Media Psychol.* Sep 17, 2008;11(3):354-376. [doi: [10.1080/15213260802285214](https://doi.org/10.1080/15213260802285214)]
41. Barrett RCA, Poe R, O'Camb JW, et al. Comparing virtual reality, desktop-based 3D, and 2D versions of a category learning experiment. *PLoS One.* 2022;17(10):e0275119. [doi: [10.1371/journal.pone.0275119](https://doi.org/10.1371/journal.pone.0275119)] [Medline: [36201546](https://pubmed.ncbi.nlm.nih.gov/36201546/)]
42. Bhargava Y, Kottapalli A, Baths V. Validation and comparison of virtual reality and 3D mobile games for cognitive assessment against ACE-III in 82 young participants. *Sci Rep.* Oct 13, 2024;14(1):23918. [doi: [10.1038/s41598-024-75065-1](https://doi.org/10.1038/s41598-024-75065-1)] [Medline: [39397120](https://pubmed.ncbi.nlm.nih.gov/39397120/)]
43. Faria AL, Pinho MS, Bermúdez I Badia S. A comparison of two personalization and adaptive cognitive rehabilitation approaches: a randomized controlled trial with chronic stroke patients. *J Neuroeng Rehabil.* Jun 16, 2020;17(1):78. [doi: [10.1186/s12984-020-00691-5](https://doi.org/10.1186/s12984-020-00691-5)] [Medline: [32546251](https://pubmed.ncbi.nlm.nih.gov/32546251/)]

44. Tan NC, Lim JE, Sultana R, Quah JHM, Wong WT. A virtual reality cognitive screening tool based on the six cognitive domains. *Alzheimers Dement (Amst)*. 2024;16(4):e70030. [doi: [10.1002/dad2.70030](https://doi.org/10.1002/dad2.70030)] [Medline: [39713248](https://pubmed.ncbi.nlm.nih.gov/39713248/)]
45. Neğu A, Matu SA, Sava FA, David D. Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures: a meta-analytic approach. *Comput Human Behav*. Jan 2016;54:414-424. [doi: [10.1016/j.chb.2015.08.029](https://doi.org/10.1016/j.chb.2015.08.029)]
46. Dandurand F, Shultz TR, Onishi KH. Comparing online and lab methods in a problem-solving experiment. *Behav Res Methods*. May 2008;40(2):428-434. [doi: [10.3758/brm.40.2.428](https://doi.org/10.3758/brm.40.2.428)] [Medline: [18522052](https://pubmed.ncbi.nlm.nih.gov/18522052/)]
47. Gould SJJ, Cox AL, Brumby DP, Wiseman S. Home is where the lab is: a comparison of online and lab data from a time-sensitive study of interruption. *Hum Comput*. 2015;2(1):45-67. [doi: [10.15346/hc.v2i1.4](https://doi.org/10.15346/hc.v2i1.4)]
48. de Leeuw JR, Motz BA. Psychophysics in a web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behav Res*. Mar 2016;48(1):1-12. [doi: [10.3758/s13428-015-0567-2](https://doi.org/10.3758/s13428-015-0567-2)]
49. Armitage J, Eerola T. Reaction time data in music cognition: comparison of pilot data from lab, crowdsourced, and convenience web samples. *Front Psychol*. 2019;10:2883. [doi: [10.3389/fpsyg.2019.02883](https://doi.org/10.3389/fpsyg.2019.02883)] [Medline: [31969849](https://pubmed.ncbi.nlm.nih.gov/31969849/)]
50. Bartneck C, Duenser A, Moltchanova E, Zawieska K. Comparing the similarity of responses received from studies in Amazon's Mechanical Turk to studies conducted online and with direct recruitment. *PLoS One*. 2015;10(4):e0121595. [doi: [10.1371/journal.pone.0121595](https://doi.org/10.1371/journal.pone.0121595)] [Medline: [25876027](https://pubmed.ncbi.nlm.nih.gov/25876027/)]
51. Hilbig BE. Reaction time effects in lab- versus web-based research: experimental evidence. *Behav Res Methods*. Dec 2016;48(4):1718-1724. [doi: [10.3758/s13428-015-0678-9](https://doi.org/10.3758/s13428-015-0678-9)] [Medline: [26542972](https://pubmed.ncbi.nlm.nih.gov/26542972/)]
52. Saunders DR, Bex PJ, Woods RL. Crowdsourcing a normative natural language dataset: a comparison of Amazon Mechanical Turk and in-lab data collection. *J Med Internet Res*. May 20, 2013;15(5):e100. [doi: [10.2196/jmir.2620](https://doi.org/10.2196/jmir.2620)] [Medline: [23689038](https://pubmed.ncbi.nlm.nih.gov/23689038/)]
53. Sauce B, Matzel LD. The causes of variation in learning and behavior: why individual differences matter. *Front Psychol*. 2013;4:395. [doi: [10.3389/fpsyg.2013.00395](https://doi.org/10.3389/fpsyg.2013.00395)] [Medline: [23847569](https://pubmed.ncbi.nlm.nih.gov/23847569/)]
54. Segen V, Avraamides M, Slattery T, Colombo G, Wiener JM. Comparable performance on a spatial memory task in data collected in the lab and online. *PLoS One*. 2021;16(11):e0259367. [doi: [10.1371/journal.pone.0259367](https://doi.org/10.1371/journal.pone.0259367)] [Medline: [34843521](https://pubmed.ncbi.nlm.nih.gov/34843521/)]
55. Casey BJ, Trainor RJ, Orendi JL, et al. A developmental functional MRI study of prefrontal activation during performance of a go-no-go task. *J Cogn Neurosci*. Nov 1997;9(6):835-847. [doi: [10.1162/jocn.1997.9.6.835](https://doi.org/10.1162/jocn.1997.9.6.835)] [Medline: [23964603](https://pubmed.ncbi.nlm.nih.gov/23964603/)]
56. The Jamovi Project. Jamovi (version 23) [computer software]. 2022. URL: <https://www.jamovi.org> [Accessed 2025-06-30]
57. Lavie N, Tsai Y. Perceptual load as a major determinant of the locus of selection in visual attention. *Percept Psychophys*. Mar 1994;56(2):183-197. [doi: [10.3758/BF03213897](https://doi.org/10.3758/BF03213897)]
58. Lavie N. Perceptual load as a necessary condition for selective attention. *J Exp Psychol Hum Percept Perform*. 1995;21(3):451-468. [doi: [10.1037/0096-1523.21.3.451](https://doi.org/10.1037/0096-1523.21.3.451)]
59. Cartwright-Finch U, Lavie N. The role of perceptual load in inattention blindness. *Cognition*. Mar 2007;102(3):321-340. [doi: [10.1016/j.cognition.2006.01.002](https://doi.org/10.1016/j.cognition.2006.01.002)] [Medline: [16480973](https://pubmed.ncbi.nlm.nih.gov/16480973/)]
60. Lavie N, Hirst A, de Fockert JW, Viding E. Load theory of selective attention and cognitive control. *J Exp Psychol Gen*. Sep 2004;133(3):339-354. [doi: [10.1037/0096-3445.133.3.339](https://doi.org/10.1037/0096-3445.133.3.339)] [Medline: [15355143](https://pubmed.ncbi.nlm.nih.gov/15355143/)]
61. Lavie N. Distracted and confused?: Selective attention under load. *Trends Cogn Sci (Regul Ed)*. Feb 2005;9(2):75-82. [doi: [10.1016/j.tics.2004.12.004](https://doi.org/10.1016/j.tics.2004.12.004)]
62. Rees G, Frith CD, Lavie N. Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science*. Nov 28, 1997;278(5343):1616-1619. [doi: [10.1126/science.278.5343.1616](https://doi.org/10.1126/science.278.5343.1616)] [Medline: [9374459](https://pubmed.ncbi.nlm.nih.gov/9374459/)]
63. Neokleous K, Shimi A, Avraamides MN. Modeling the effects of perceptual load: saliency, competitive interactions, and top-down biases. *Front Psychol*. 2016;7:1. [doi: [10.3389/fpsyg.2016.00001](https://doi.org/10.3389/fpsyg.2016.00001)] [Medline: [26858668](https://pubmed.ncbi.nlm.nih.gov/26858668/)]
64. Huang L, Pashler H. Attention capacity and task difficulty in visual search. *Cognition*. Jan 2005;94(3):B101-B111. [doi: [10.1016/j.cognition.2004.06.006](https://doi.org/10.1016/j.cognition.2004.06.006)] [Medline: [15617666](https://pubmed.ncbi.nlm.nih.gov/15617666/)]
65. Bravo MJ, Nakayama K. The role of attention in different visual-search tasks. *Percept Psychophys*. Sep 1992;51(5):465-472. [doi: [10.3758/BF03211642](https://doi.org/10.3758/BF03211642)]
66. Alaeifard M, Safaei M. Head movement patterns as predictors of cybersickness in virtual reality games. *Int J Adv Hum Comput Interact*. Jul 19, 2024;2(2):1. [doi: [10.1016/j.cognition.2004.06.006](https://doi.org/10.1016/j.cognition.2004.06.006)]
67. Weech S, Kenny S, Lenizky M, Barnett-Cowan M. Narrative and gaming experience interact to affect presence and cybersickness in virtual reality. *Int J Hum Comput Stud*. Jun 2020;138:102398. [doi: [10.1016/j.ijhcs.2020.102398](https://doi.org/10.1016/j.ijhcs.2020.102398)]

Abbreviations

ACE-III: Addenbrooke's Cognitive Examination—III

HMD: head-mounted display

RT: reaction time

VR: virtual reality

Edited by Amaryllis Mavragani, Lars Riedemann; peer-reviewed by Brendan Concannon, Jose Ferrer Costa, Liliana Mendes; submitted 27.08.2024; final revised version received 05.05.2025; accepted 25.05.2025; published 29.07.2025

Please cite as:

Hadjiaros M, Shimi A, Neokleous K, Pattichis C, Avraamides M

Virtual Reality Gamification of Visual Search, Response Inhibition, and Visual Short-Term Memory Tasks for Cognitive Assessment: Experimental Study

JMIR Form Res 2025;9:e65836

URL: <https://formative.jmir.org/2025/1/e65836>

doi: [10.2196/65836](https://doi.org/10.2196/65836)

© Marios Hadjiaros, Andria Shimi, Kleanthis Neokleous, Constantinos Pattichis, Marios Avraamides. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 29.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.