# **Impact of Graphical Complexity on Action Performance in a VR Cooking Task**

Neha Kowtharapu

Luddy School of Informatics, Computing and Engineering, Indiana University Indianapolis

SP25: THESIS/PROJ Health Informatics (27933)

Professor Hee‑Tae Jung

May 8, 2025

# **1. Abstract**

Background – Graphical complexity, the sheer amount of visual detail on screen, can amplify cognitive load in immersive virtual‑reality (VR) tasks, yet its behavioural impact is poorly quantified.

Methods – Forty-six adults completed a balanced cross‑over VR cooking experiment in which they prepared **pie** or **stew** recipes presented in **simple** or **complex** scenes (four recipe‑types × recipe‑complexity conditions). Nine action‑performance metrics were extracted from time-stamped interaction logs. We tested condition effects with two-way ANOVAs (recipe type × recipe complexity) and assessed cognitive relevance with Spearman correlations against WAIS Block‑Design scores.

Results – Scene detail, not scripted recipe logic, was the primary driver of timing costs. **Average** and **75th‑percentile** inter‑action intervals showed main effects of recipe type (F = 4.65–4.69, p ≤ .033) and recipe complexity (F = 6.04–7.56, p ≤ .016) plus a strong interaction (F = 7.57–7.82, p ≤ .007), with the ostensibly Simple Stew condition slower than every other recipe in Tukey tests (Δ = 2.15–3.71 s, p < .016). Lower‑bound pacing (25th percentile) and timing variability (SD) also displayed significant interactions, again driven by Simple Stew. Only four metric‑condition pairs correlated with cognitive ability: faster median pace in Simple Pie (r = –.46, p =.035), steadier timing in Simple Stew (r = –.52, p =.010), more exploratory empty‑hand grabs in Complex Pie (r = .51, p =.016), and fewer recipe look‑ups in Complex Stew (r = –.44, p =.040).

Conclusions – A visually dense “simple” stew scene imposed greater timing and consistency demands than any “complex” pie, underscoring ingredient density as a hidden complexity lever. Median and SD latencies, plus targeted information‑seeking counts, emerged as sensitive digital markers of cognitive load in VR cooking and warrant inclusion in adaptive training and rehabilitation systems.

# **2. Introduction and Background**

Graphical complexity in immersive virtual reality (VR) environments can impose substantial cognitive demands on users, potentially affecting their action performance and underlying cognitive processes. In tasks requiring continuous interaction and decision-making, such as a VR cooking simulation, the variability and timing of user actions offer rich behavioural markers of cognitive load and executive control. Traditional cognitive assessments, like the Block Design subtest of the WAIS, capture spatial reasoning under controlled conditions but lack the ecological validity and fine‐grained temporal resolution needed to characterise real‐world task performance (Clay et al., 2019). In contrast, sensor‐driven metrics in VR such as the 25th, 50th, and 75th percentiles of inter‐action intervals, the average inter‐action interval, standard deviation of intervals, action repetition count, grab‑cookbook count, grab‑none count, and inter‑hand coordination ratio provide continuous, high‐resolution insights into how users manage perceptual, memory, and motor demands under varying levels of graphical detail.

Recent studies demonstrate the sensitivity of VR‐based measures to cognitive load and adaptability. Kwan et al. (2024) found that older adults with cognitive frailty exhibited significant changes in dual‐task performance and reaction times when switching between VR subtasks of differing complexity, highlighting the feasibility of using temporal action metrics to index attentional shifts. Similarly, Kang et al. (2021) showed that immersive VR training improved visuospatial skills and increased frontal–occipital connectivity, suggesting that movement and gaze metrics in VR can reflect neural adaptations to task demands. Tortora et al. (2023) further synthesized evidence that VR interventions offer superior ecological validity and task realism, making them well‐suited for probing higher‐order executive functions, such as cognitive flexibility, planning, and divided attention, through naturalistic activities.

Building on this foundation, our study examines how varying graphical complexity in a VR cooking task modulates the distribution (25th, 50th, 75th percentiles), variability (standard deviation), and frequency (action repetition, grab‑cookbook, grab‑none, inter-hand coordination ratios) of action intervals. By correlating these behavioural markers with standard cognitive measures (e.g., WAIS Block Design), we aim to develop an ecologically valid framework for assessing executive control and cognitive adaptability in healthy young adults. This approach promises scalable and sensitive assessment tools capable of detecting subtle changes in cognitive workload before clinical symptoms emerge, with potential applications in adaptive cognitive training and neurorehabilitation.

# **2.1. Significance:**

 **Guides the design of VR rehabilitation environments with appropriate visual complexity to avoid cognitive overload.**  
By identifying how graphical detail influences action‑timing and movement patterns, our findings will inform developers on optimal scene complexity that maximizes engagement without overwhelming users’ attentional resources (Hamad & Jia, 2022).

 **Validates novel digital markers of executive function action timing metrics that may complement traditional tests.**  
Demonstrating robust correlations between VR-derived metrics (e.g., inter-action percentiles, variability measures, and object interaction counts) and WAIS Block Design performance will support their integration as complementary tools in cognitive assessment and early detection of executive dysfunction.

# **3. Problem Statement**

Traditional neuropsychological tests, such as the WAIS Block Design, offer valuable snapshots of visuospatial reasoning but reveal little about how people flexibly allocate attention and coordinate actions in real-time, everyday contexts. Virtual‑reality (VR) cooking tasks can capture continuous streams of gaze and motor data, yet no study has systematically manipulated scene complexity to determine which micro‑level action metrics (25th, 50th, 75th inter‑action percentiles; average and SD of inter‑action intervals; action‑repetition, grab‑cookbook, grab‑none counts; inter‑hand‑coordination ratio) are most sensitive to cognitive load. Clinicians and researchers, therefore, lack clear evidence linking these temporal features to executive‑function benchmarks such as Block Design performance (Tan et al., 2024).

# **3.1. Specific Aims:**

**Aim 1 – Task-level effects:**  
Quantify how increasing graphical complexity (Simple vs. Complex scenes for Pie and Stew recipes) changes nine predefined action metrics 25th, 50th, and 75th percentiles of inter‑action intervals; average inter‑action interval; action‑repetition count; grab‑cookbook count; grab‑none count; interval standard deviation; and inter‑hand‑coordination ratio during VR meal preparation.

**Aim 2 – Metric sensitivity:**  
Identify which of the nine action metrics most reliably discriminate between low‑ and high‑complexity conditions, thereby pinpointing the digital markers most sensitive to visual‑load–induced cognitive stress.

**Aim 3 – Cognitive linkage:**  
Examine relationships between each action metric and WAIS Block Design scores to determine whether timing-based markers of VR task performance predict established measures of visuospatial reasoning and cognitive flexibility.

# **4. Literature Review of Action‑Timing Metrics in VR Cooking Tasks**

1. **25th‑Percentile Inter-Action Interval (Lower‑Bound Pace)**

The 25th‑percentile represents the fastest quartile of response latencies and is often interpreted as the point at which action sequences approach automaticity. Shorter lower-bound intervals typically emerge when cognitive load is low and sensorimotor mappings are well learned. Experimental work in immersive VR shows that elevated cognitive demands elongate these rapid responses (Catania et al., 2023), reflecting reduced automatic control under load. Conversely, VR motor‑cognitive training programs that strengthen divided‑attention skills have been reported to compress early‑quartile latencies over time, indicating increased processing efficiency.

1. **Median (50th‑Percentile) Inter-Action Interval**

The median interval provides a robust estimate of an individual’s typical pacing, less sensitive to extreme pauses than the mean. Meta-analytic evidence demonstrates moderate reductions in median reaction times after repeated VR practice, suggesting that immersive environments can foster more efficient central processing of routine actions. Because the measure captures the central tendency of action speed, it serves as a stable comparator across tasks of differing complexity.

1. **75th‑Percentile Inter-Action Interval (Upper‑Bound Pause)**

The 75th‑percentile captures slower responses that frequently correspond to deliberation, visual search, or error correction. Prolonged upper‑quartile latencies have been linked to heightened cognitive load and context switching demands in VR adaptation studies. Accordingly, reductions in this metric following adaptive training are interpreted as evidence of improved executive control over intermittent burst‑pause dynamics (Makmee & Wongupparaj, 2025).

1. **Average Inter-Action Interval**

The arithmetic mean aggregates all latencies, integrating both rapid and slow responses. Systematic reviews indicate that average inter‑action times in VR tasks decline when training incorporates progressive difficulty and real-time feedback, highlighting the metric’s sensitivity to general learning effects (Moulaei et al., 2024). However, its susceptibility to skew from occasional long pauses underscores the importance of evaluating it in concert with percentile and variability measures.

1. **Standard Deviation of Inter-Action Intervals**

The standard deviation (SD) quantifies temporal consistency. Elevated SD values signify erratic pacing and have been associated with increased cognitive load or uncertainty in immersive environments (Juliano et al., 2022). Successful VR interventions aimed at executive‑function enhancement routinely report decreases in latency variability, suggesting tighter cognitive regulation of action sequences.

1. **Action Repetition Count**

Action repetition count tallies consecutive interactions with the same object and is interpreted as an indicator of hesitancy, procedural uncertainty, or inefficient motor planning. Reviews of adaptive VR game mechanics emphasise the need to monitor and minimise unnecessary repetitions to sustain engagement and reduce cognitive fatigue (Guzmán et al., 2024; Wang et al., 2021). Interventions that provide contextual prompts or adaptive difficulty curves have been shown to lower repetition rates, reflecting smoother task execution.

1. **Grab‑Cookbook Count**

Grab‑cookbook events capture the frequency with which users reference an in-game instructional aid. High counts suggest reliance on external memory supports and, by extension, elevated task complexity or incomplete schema internalisation. Studies on authentic VR learning spaces note that strategically placed reference objects can scaffold performance but should diminish in use as proficiency grows. Sustained high grab‑cookbook counts, therefore, flag lingering cognitive load in recipe planning and sequencing (Zhu et al., 2021).

1. **Grab‑None Count**

Grab‑none events represent attempted interactions that fail to register an object and are treated as execution errors or mis‑perceptions. In immersive cognitive‑screening paradigms, error-to-action ratios correlate with broader indices of cognitive impairment, underscoring that grab‑none counts as a sensitive marker of visuomotor accuracy. Reductions in this metric after practice signal improved object recognition and spatial mapping.

1. **Inter-Hand Coordination Ratio**

The inter-hand coordination ratio reflects the balance of left‑ versus right-hand actions. Ratios approaching unity denote symmetrical engagement, whereas large deviations indicate unilateral dominance. VR motor‑cognitive protocols that demand bimanual coordination have demonstrated shifts toward more balanced ratios alongside gains in dual‑task performance, linking the measure to executive control over distributed motor resources. Excessive reliance on a single hand, especially under high‑load conditions, may also signify compensatory strategies when cognitive capacity is strained.

# **5. Methodology**

# **5.1. Participants**

Forty-six adult volunteers were recruited for a VR cooking study that employed a balanced cross‑over design. Upon arrival, each participant completed the **Block Design** subtest of the *Wechsler Adult Intelligence Scale* (WAIS), providing a standardized visuo-spatial cognitive score (BDT) that served as the reference measure for subsequent correlation analyses.

# **5.2. Study Design and Counter‑Balancing**

Participants were randomly assigned to one of two counter-balanced sequences so that everyone completed one **simple** and one **complex** recipe while alternating recipe type:  
  • **Sequence A:** *Simple Pie → Complex Stew*  
  • **Sequence B:** *Complex Pie → Simple Stew*

This arrangement controlled for potential order effects and ensured that recipe type (pie vs stew) and complexity (simple vs complex) were independently manipulated.

# **A diagram of a process flow AI-generated content may be incorrect.**

*Figure 1. Methodology*

# **5.3. VR Apparatus and Calibration**

Experiments were conducted with a head‑mounted display, a pair of handheld controllers, and five inertial motion trackers affixed to the upper arms, forearms, and chest. Before the cooking tasks, headset fit, tracker placement, and controller functionality were verified. Participants then completed a short orientation tutorial that covered object grabbing, utensil manipulation, and ingredient placement, giving them time to adjust in either a seated or standing posture according to individual comfort and safety.

# **5.4. Task Execution**

Each participant carried out two recipe-based simulations. Step-by-step instructions were embedded in a virtual recipe book that advanced only after a correct action, providing immediate visual confirmation. Everyday cooking behaviours such as reaching, selecting, stirring, and activating appliances were incorporated. Task complexity was systematically varied through ingredient count, action dependencies, and decision-making requirements, creating the four experimental conditions: Simple Pie, Simple Stew, Complex Pie, and Complex Stew.

# **5.5. Data Collection and Management**

Four synchronized CSV logs were written for every participant–task pairing:

1. **Action Data** – Time-stamped records of all interaction events.
2. **Completed‑Step Log** – Chronological task‑progress entries.
3. **View Log** – Head‑gaze time stamps of on-screen objects.
4. **Movement Data** – Six-axis position and rotation samples from each tracker.

Files were labelled only with anonymized participant IDs and task codes. The present analysis focused exclusively on the **Action Data** files to address the research question:

“How does graphical complexity in a VR cooking environment affect action performance during pie and stew preparation?”

# **5.6. Feature Extraction**

For each Action Data file, the following metrics were computed:

* 25th, 50th, and 75th percentiles of inter‑action intervals
* Mean inter-action interval
* Standard deviation of inter-action intervals
* Action‑repetition count
* Grab‑None count
* Grab‑Cookbook count
* Inter‑Hand Coordination Index

# **5.7. Data Analysis**

**(a) Metric–BDT Associations**

* Spearman’s rank correlation coefficient (r) was calculated between each extracted metric and the participant’s WAIS Block Design Test score, assessing the monotonic relationship between visuo-spatial cognitive ability and action performance.

**(b) Group Comparisons Across Conditions**

* For every metric, a two-way ANOVA was conducted with Recipe Type (Pie vs Stew) and Complexity (Simple vs Complex) as fixed factors, using data from all four conditions.
* Tukey HSD post‑hoc tests followed each significant ANOVA to localize pairwise differences among Simple Pie, Complex Pie, Simple Stew, and Complex Stew.

# **5.8. Interpretation and Reporting**

Correlation coefficients, ANOVA F‑statistics, and Tukey HSD mean‑difference estimates were synthesized to determine (i) which action‑performance metrics were most sensitive to increases in graphical complexity, (ii) how recipe type moderated these effects, and (iii) the extent to which each metric captured underlying cognitive ability as indexed by WAIS‑BDT.

# **6. Results**

# **6.1. Links between cognitive score and action metrics**

Only four metric–condition pairs were meaningfully related to Block‑Design performance. Faster pacing in **Simple Pie** (median inter-action interval) and steadier timing in **Simple Stew** (SD of inter-action interval) rose with visuo-spatial skill, whereas high scorers made more exploratory empty‑hand grabs in **Complex Pie** and consulted the cookbook less in **Complex Stew**. Every other r value hovered near zero, confirming that most timing and count measures were not cognition‑sensitive.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Metric** | **Simple Pie** | **Complex Pie** | **Simple Stew** | **Complex Stew** |
| 25th Percentile | r = -0.42, p = 0.058 | r = -0.09, p = 0.68 | r = 0.06, p = 0.790 | r = 0.05, p = 0.840 |
| Median | **r = -0.46, p = 0.035** | r = -0.11, p = 0.61 | r = -0.11, p = 0.620 | r = 0.08, p = 0.740 |
| 75th Percentile | r = -0.27, p = 0.23 | r = -0.21, p = 0.32 | r = -0.06, p = 0.760 | r = -0.05, p = 0.810 |
| Average inter-action interval | r = -0.23, p = 0.31 | r = -0.17, p = 0.44 | r = -0.14, p = 0.560 | r = 0.06, p = 0.810 |
| Standard Deviation | r = -0.02, p = 0.91 | r = -0.01, p = 0.96 | **r = -0.52, p = 0.010** | r = 0.14, p = 0.510 |
| Inter-hand Coordination Index | r = -0.09, p = 0.70 | r = -0.30, p = 0.18 | r = -0.14, p = 0.530 | r = -0.26, p = 0.260 |
| Action repetition count | r = -0.25, p = 0.28 | r = 0.12, p = 0.60 | r = -0.09 , p = 0.690 | r = -0.23, p = 0.320 |
| Grab-None count | r = -0.09, p = 0.70 | **r = 0.51, p = 0.016** | r = -0.07, p = 0.750 | r = 0.09, p = 0.710 |
| Grab-Cookbook count | r = 0.01, p = 0.96 | r = -0.01, p = 0.97 | r = -0.01, p = 0.950 | **r = -0.44, p = 0.040** |

*Table 1: Spearman correlations between Block‑Design score and action‑performance metrics*

A graph with blue dots and red text

AI-generated content may be incorrect.A graph with blue dots and red text

AI-generated content may be incorrect.

A graph of a diagram

AI-generated content may be incorrect.A graph with blue dots and red text

AI-generated content may be incorrect.

***Figure 1. Relationship between visuo-spatial ability and action‑performance metrics across VR cooking conditions*** *Composite scatter plot displaying the four metric‑condition pairs that showed significant Spearman correlations with WAIS Block‑Design scores: (a) shorter median inter‑action intervals in Simple Pie (r = –0.46, p =.035), (b) higher Grab‑None counts in Complex Pie (r = 0.51, p =.016), (c) lower timing variability in Simple Stew (SD; r = –0.52, p =.010), and (d) fewer cookbook consultations in Complex Stew (Grab‑Cookbook; r = –0.44, p =.040). Trend lines illustrate the direction and strength of each relationship.*

# **6.2.  Recipe‑type × Complexity effects**

Two-way ANOVAs pinpointed timing, not counts, as the workload bottleneck:

* **Average inter-action interval** and the **75th‑percentile latency** showed recipe‑type, complexity and interaction effects (F = 4.65–7.82, p ≤ .033). **Stew‑Simple** was significantly slower than every other condition (Tukey p < .016).
* **25th‑percentile latency** and **timing variance (SD)** exhibited interaction-only effects (F = 4.97 and 5.54, p ≤ .028), driven by a Stew‑Simple > Pie‑Simple contrast.
* **Grab‑Cookbook** counts were consistently higher for pies than stews (recipe‑type F = 7.98, p =.006).
* Median timing, Grab‑None, Action repetition, and Inter‑hand coordination showed no reliable main or interaction effects (all p ≥ .085).

|  |  |  |  |
| --- | --- | --- | --- |
| **Metric** | **Factor** | **F-value** | **p-value** |
| 25th percentile | Recipe Type x Recipe Complexity | F = 4.97 | p = 0.028 |
| Average inter-action interval | Recipe Type | F = 4.65 | p = 0.033 |
|  | Recipe Complexity | F = 6.04 | p = 0.016 |
|  | Recipe Type x Recipe Complexity | F = 7.82 | p = 0.006 |
| 75th percentile | Recipe Type | F = 4.69 | p = 0.033 |
|  | Recipe Complexity | F = 7.56 | p = 0.007 |
|  | Recipe Type x Recipe Complexity | F = 7.57 | p = 0.007 |
| Standard Deviation | Recipe Type x Recipe Complexity | F = 5.54 | p = 0.021 |
| Grab-Cookbook count | Recipe Type | F = 7.98 | p = 0.006 |

*Table 2: Significant two‑way ANOVA effects*

# **6.3. Pairwise comparisons (Tukey HSD)**

Tukey's honest significant difference tests were run for every metric to pinpoint which specific recipe type × recipe complexity combinations drove the significant ANOVA effects. Only contrasts that survived the family-wise α = .05 threshold are discussed below:

* **Timing speed** — For both **Average** and **75th‑percentile** inter‑action intervals, the **Stew ‑ Simple** condition was slower than every other recipe (Pie ‑ Complex, Pie ‑ Simple, Stew ‑ Complex), confirming that even a “simple” stew imposes a heavier timing burden than any pie.
* **Lower‑tail latency (25th percentile)** — The only meaningful contrast was **Stew ‑ Simple > Pie ‑ Simple**, showing that the quickest actions were already delayed in Stew before adding recipe complexity.
* **Timing variability (SD of inter-action interval)** — **Stew ‑ Simple** again led the field, displaying greater variability than **Pie ‑ Simple** and **Stew ‑ Complex**, underscoring how the stew format inflates pace and consistency demands.

No other pairwise differences reached significance, reinforcing that the **Stew ‑ Simple** scenario uniquely taxes action timing across the board.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | **Metric** |  |  | | --- | |  | | **Contrast (Condition A > Condition B)** | **Mean difference (Δ)** | |  | | --- | | **p‑value** | |
| 25th Percentile | Stew – Simple > Pie – Simple | Δ = 0.390 s | p = 0.0217 |
| Average inter-action interval | Stew – Simple > Pie – Complex | Δ = 2.154 s | p = 0.0155 |
|  | Stew – Simple > Pie – Simple | Δ = 2.490 s | p = 0.0042 |
|  | Stew – Simple > Stew – Complex | Δ = 2.649 s | p = 0.0020 |
| 75th Percentile | Stew – Simple > Pie – Complex | Δ = 3.091 s | p = 0.0077 |
|  | Stew – Simple > Pie – Simple | Δ = 3.298 s | p = 0.0046 |
|  | Stew – Simple > Stew – Complex | Δ = 3.712 s | p = 0.0011 |
| Standard deviation | Stew – Simple > Pie – Simple | Δ = 3.192 s | p = 0.0185 |
|  | Stew – Simple > Stew – Complex | Δ = 3.088 s | p = 0.0243 |

*Table 3: Tukey HSD – significant recipe type × recipe complexity contrasts*

# **7. Discussion**

# **7.1. Principal findings**

Graphical load shaped how participants executed cooking actions: even the ostensibly “simple” stew scene stretched average and 75th‑percentile inter-action intervals beyond every other recipe, and it also widened lower‑tail latency and timing variability. Those inflation effects point to a baseline cost imposed by the stew’s denser ingredient palette and colour palette before additional recipe‑complexity logic is introduced. At the same time, only a handful of timing‑variability and information‑seeking behaviours median pace in Simple Pie, SD of timing in Simple Stew, empty‑hand grabs in Complex Pie and cookbook look‑ups in Complex Stew mirrored baseline visuo‑spatial skill, suggesting that most action metrics are driven more by on‑screen clutter than by individual cognitive differences.

7**.2. Relation to prior work**

These results echo lab-based findings that upper‑tail latency rises with scene clutter yet add nuance by showing that “simple” object-rich scenes can be more taxing than “complex” but visually cleaner ones. The mixed cognitive correlations align with studies reporting that fine-grained VR timing metrics capture executive or visuo-spatial capacity only when tasks leave room for self-directed strategy (Kang et al., 2021); here, exploratory grabs and skipped look‑ups emerged exactly in the two visually busiest contexts, supporting that view.

### **7.3. Interpretation and mechanisms**

We argue that stew layouts impose a larger perceptual‑search burden. Participants sift through many small, similarly coloured items before committing to an action, inflating both mean and extreme latencies. For high-ability users, those costs shrink to some degree, seen in steadier rhythms and reduced recipe checks, yet exploratory empty grabs rise in the most cluttered pie, hinting at a trade-off between strategic probing and decisiveness. Thus, action‑metric sensitivity appears to hinge on where the “bottleneck” sits, stimulus identification (stews) versus plan selection (complex pies).

### **7.4. Strengths and limitations**

Strengths include a balanced cross-over design, unified headset‑tracker logging, and the use of nine complementary timing/count metrics. Limitations stem from relying on a single cognitive measure (Block‑Design), analysing only action logs (view and movement streams remain untapped), and studying one culinary theme; findings may not generalise to radically different VR tasks or populations with motor impairments. Sample size (n = 46) also caps power for detecting small interaction effects.

### **8. Conclusion**

Graphical complexity in VR cooking exerts a stronger pull on action timing than scripted recipe logic, with visually saturated “simple” stews proving more taxing than cleaner “complex” pies. Only timing‑variability and targeted information-seeking behaviours reliably signalled underlying visuo-spatial skill, positioning them as candidate metrics for adaptive VR systems that need to sense cognitive load on the fly.

### **9. Implications and future work**

Designers of therapeutic or training VR kitchens should treat ingredient density as its complexity lever: even basic recipes can overload users if the visual field is busy. Median and SD latencies, plus context-specific information‑seeking counts, emerge as practical markers for adaptive difficulty scaling or cognitive‑load feedback. Future studies should pair these action metrics with gaze and movement features, broaden cognitive batteries, and test whether real-time difficulty adjustments based on the median of inter-action interval improve learning or rehabilitation outcomes. Expanding to larger, clinically diverse cohorts will clarify how motor or cognitive deficits modulate the stew‑versus‑pie cost gradient.

### **10. References**

Catania, V., Rundo, F., Panerai, S., & Ferri, R. (2023). Virtual reality for the rehabilitation of acquired cognitive disorders: A narrative review. *Bioengineering*, *11*(1), 35. https://doi.org/10.3390/bioengineering11010035

Clay, V., König, P., & König, S. U. (2019). Eye tracking in virtual reality. *Journal of Eye*

*Movement Research, 12*(1). https://doi.org/10.16910/jemr.12.1.3

Guzmán, D. E., Rengifo, C. F., Guzmán, J. D., & Cena, C. E. G. (2024). Virtual reality games for cognitive rehabilitation of older adults: A review of adaptive games, domains and techniques. *Virtual Reality*, *28*(2). https://doi.org/10.1007/s10055-024-00968-3

Hamad, A., & Jia, B. (2022). How virtual reality technology has changed our lives: An overview of the current and potential applications and limitations. *International Journal of Environmental Research and Public Health*, *19*(18), 11278. https://doi.org/10.3390/ijerph191811278

Juliano, J. M., Schweighofer, N., & Liew, S. (2022). Increased cognitive load in immersive virtual reality during visuomotor adaptation is associated with decreased long-term retention and context transfer. *Journal of NeuroEngineering and Rehabilitation*, *19*(1). https://doi.org/10.1186/s12984-022-01084-6

Kang, J. M., Kim, N., Lee, S. Y., Woo, S. K., Park, G., Yeon, B. K., Park, J. W., Youn, J., Ryu, S., Lee, J., & Cho, S. (2021). Effect of cognitive training in fully immersive virtual reality on visuospatial function and Frontal-Occipital functional connectivity in Predementia: Randomized controlled trial. *Journal of Medical Internet Research*, *23*(5), e24526. https://doi.org/10.2196/24526

Kwan, R. Y. C., Liu, J., Sin, O. S. K., Fong, K. N. K., Qin, J., Wong, J. C. Y., & Lai, C. (2024). Effects of virtual reality motor-cognitive training for older people with cognitive frailty: Multicentered randomized controlled trial. *Journal of Medical Internet Research*, *26*, e57809. https://doi.org/10.2196/57809

Makmee, P., & Wongupparaj, P. (2025). VR cognitive-based intervention for enhancing cognitive functions and well-being in older adults with mild cognitive impairment: Behavioural and EEG evidence. *Psychosocial Intervention*, *34*(1), 37–51. https://doi.org/10.5093/pi2025a4

Moulaei, K., Sharifi, H., Bahaadinbeigy, K., & Dinari, F. (2024). Efficacy of virtual reality-based training programs and games on the improvement of cognitive disorders in patients: a systematic review and meta-analysis. *BMC Psychiatry*, *24*(1). https://doi.org/10.1186/s12888-024-05563-z

Tan, N. C., Lim, J. E., Sultana, R., Quah, J. H. M., & Wong, W. T. (2024). A virtual reality cognitive screening tool based on the six cognitive domains. *Alzheimer S & Dementia Diagnosis Assessment & Disease Monitoring*, *16*(4). https://doi.org/10.1002/dad2.70030

Tortora, C., Di Crosta, A., La Malva, P., Prete, G., Ceccato, I., Mammarella, N., Di Domenico, A., & Palumbo, R. (2023). Virtual reality and cognitive rehabilitation for older adults with mild cognitive impairment: A systematic review. *Ageing Research Reviews*, *93*, 102146. https://doi.org/10.1016/j.arr.2023.102146

Wang, A., Thompson, M., Uz-Bilgin, C., & Klopfer, E. (2021). Authenticity, interactivity, and collaboration in virtual reality games: Best practices and lessons learned. *Frontiers in Virtual Reality*, *2*. https://doi.org/10.3389/frvir.2021.734083

Zhu, S., Sui, Y., Shen, Y., Zhu, Y., Ali, N., Guo, C., & Wang, T. (2021). Effects of virtual reality intervention on cognition and motor function in older adults with mild cognitive impairment or dementia: A Systematic review and meta-analysis. *Frontiers in Aging Neuroscience*, *13*. https://doi.org/10.3389/fnagi.2021.586999

### **Appendix**

### **Appendix A:**

**Spearman Correlation Matrix for All Metrics and Conditions: Simple Pie**

A graph with blue dots and red text

AI-generated content may be incorrect.A graph with blue dots and red text

AI-generated content may be incorrect.

A graph with blue dots and red text

AI-generated content may be incorrect.A graph with blue dots and red text

AI-generated content may be incorrect.

A graph with blue dots and a line

AI-generated content may be incorrect.A graph with blue dots and red text

AI-generated content may be incorrect.

A graph with blue dots and numbers

AI-generated content may be incorrect.A graph with blue dots and numbers

AI-generated content may be incorrect.

A graph with blue dots and red text

AI-generated content may be incorrect.

**Spearman Correlation Matrix for All Metrics and Conditions: Complex Pie**

A graph with blue dots and a line

AI-generated content may be incorrect.A graph with blue dots and red text

AI-generated content may be incorrect.

A graph of a number of numbers

AI-generated content may be incorrect.A graph with blue dots and red text

AI-generated content may be incorrect.

A graph with blue dots and a line

AI-generated content may be incorrect.A graph with blue dots and red text

AI-generated content may be incorrect.

A graph with blue dots and red text

AI-generated content may be incorrect.A graph with blue dots and a line

AI-generated content may be incorrect.

A graph with blue dots and red text

AI-generated content may be incorrect.

**Spearman Correlation Matrix for All Metrics and Conditions: Simple Stew**

**A graph with blue dots and numbers

AI-generated content may be incorrect.A graph with blue dots and numbers

AI-generated content may be incorrect.**

A graph with blue dots and a line

AI-generated content may be incorrect.A graph with blue dots and numbers

AI-generated content may be incorrect.

A graph with blue dots and numbers

AI-generated content may be incorrect.A graph with blue dots and red text

AI-generated content may be incorrect.

A graph with blue dots and red text

AI-generated content may be incorrect.A graph of a diagram

AI-generated content may be incorrect.

A graph with blue dots and red text

AI-generated content may be incorrect.

**Spearman Correlation Matrix for All Metrics and Conditions: Complex Stew**

A graph with blue dots and a line

AI-generated content may be incorrect.A graph with blue dots and a line

AI-generated content may be incorrect.

A graph of a person's hand

AI-generated content may be incorrect.A graph with blue dots and numbers

AI-generated content may be incorrect.

A graph with blue dots and numbers

AI-generated content may be incorrect.A graph with blue dots and numbers

AI-generated content may be incorrect.

A graph with blue dots and red text

AI-generated content may be incorrect.A graph with blue dots and numbers

AI-generated content may be incorrect.

A graph with blue dots and red text

AI-generated content may be incorrect.

### **Appendix B:**

### **2-Way ANOVA Results**

**Median of inter-action interval**

A graph showing a diagram

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

**25th Percentile**

**A graph of different colored squares

AI-generated content may be incorrect.**

**A black and white screen with numbers and symbols

AI-generated content may be incorrect.**

**75th Percentile**

**A graph with a chart and a diagram

AI-generated content may be incorrect.**

**A black background with white text

AI-generated content may be incorrect.**

**Grab-None Count**

**A graph showing a diagram

AI-generated content may be incorrect.**

**A black background with white text

AI-generated content may be incorrect.**

**Action Repetition Count**

**A graph with colored squares and lines

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Average inter-action interval**

**A graph with a diagram

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Inter-hand coordination index**

**A diagram of a graph

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Standard Deviation**

**A chart with different colored squares

AI-generated content may be incorrect.**

**A black background with white text

AI-generated content may be incorrect.**

**Grab-Cookbook count**

**A graph with colorful boxes and text

AI-generated content may be incorrect.**

**A black background with white text

AI-generated content may be incorrect.**

### **Appendix C:**

### **Pairwise comparisons (Tukey HSD)**

**Median of inter-action interval**

**A screen shot of a computer

AI-generated content may be incorrect.**

**25th Percentile**

**A screenshot of a computer

AI-generated content may be incorrect.**

**75th Percentile**

**A screenshot of a computer screen

AI-generated content may be incorrect.**

**Grab-None Count**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Action Repetition Count**

**A screenshot of a computer screen

AI-generated content may be incorrect.**

**Average inter-action interval**

**A screenshot of a computer screen

AI-generated content may be incorrect.**

**Inter-hand coordination index**

**A screenshot of a computer screen

AI-generated content may be incorrect.**

**Standard Deviation**

**A screenshot of a computer program

AI-generated content may be incorrect.**

**Grab-Cookbook count**

**A screenshot of a computer screen

AI-generated content may be incorrect.**