Effects of Visual-Haptic Weight Congruence on Perceived Realism and Performance in VR Object Manipulation

BOON Yi Jun, Younes Trichine

Abstract

This study investigates how visuo-haptic weight congruence affects VR object manipulation. While congruent conditions received higher realism ratings ($p < .001$, $d = 2.71$), they showed no objective performance advantage over incongruent conditions. The key finding was an asymmetric mismatch effect: when objects looked heavy but felt light, thirty participants made 2.8× more corrections ($t(29) = 6.12$, $p < .001$, $d = 1.12$) and exhibited 26% lower path efficiency than the reverse mismatch. This demonstrates that mismatch direction matters more than congruence itself, with underestimation causing severe performance degradation while overestimation has minimal cost. Results suggest VR designers should prioritize avoiding visual overestimation of haptic weight to prevent user performance deficits.

1 Introduction

The simulation of weight sensation in Virtual Reality (VR) is still a challenge to be tackled for users. The integration of visual and haptic cues can still produce a similar experience for the user's perception. Our study investigates how the visuo-haptic weight simulation affects performance, mainly when manipulating objects. This can be via examining the mismatch between these two types of cues, and whether the direction of this mismatch has an impact on performance. While the research discussed in the following section suggests different ways to integrate the two cues in discussion, we will be hypothesizing the following: Congruent visual-haptic pairings will give a better performance than incongruent pairings (H1). We also expect that the underestimation mismatch (objects look heavy but feel light) would have more impact than the overestimation one (objects look light but feel heavy) (H2). Lastly, we expect to have higher realism ratings on congruent conditions than on incongruent conditions (H3).

2 Lit. Review

Weight simulation in VR still remains a challenge among researchers of the field. As virtual objects lack a proper representation of inertia, it becomes hard to reflect their real weights and thus results in interactions that are physically inaccurate and less consistent with the real world. The papers reviewed below counter such limitations through the use of different types of cues (visual, haptic, or proprioceptive) and learned cognitive associations. [1] investigates how haptic illusions can improve motion precision in tasks. The visual perturbations placed on the heavy objects can affect the movement patterns of the user, creating an anticipated course of actions. Producing these subtle manipulations can improve the accuracy levels in such motion-controlled tasks. In addition to having physics simulations, users can also perceive weight through patterns of vibrations [2]. Although there is no established link between vibration intensities and simulated weight, users are more sensitive when different haptic weight cues are combined. Indeed, such a combination can create the illusion of a dynamic mass, as proposed in [3], where two haptic illusions could create a stronger perception of mass to the user than either technique alone. However, research suggests that the detection of such illusions can differ based on individual differences in sensory activity [4]. As some users rely more on vision or haptics than the other one, illusions can easily break if there is a slight mismatch between the two cues. Therefore, it is essential to sync them in time to ensure congruence for the participants. A further approach to simulate physical forces would be through the use of physical proxies [5]. The user can see the change of the visual object's shape and movement, then the distortion level is recorded, where they finally realize it mismatches the physical one. Similar to the approaches discussed, we seek to identify how visuo-haptic integration can affect perception and performance. While various weight simulation techniques exist, the combined effects on experience and performance remain unexplored. Specifically, the mismatch direction may impact the user's performance during the VR simulation. When users are more sensitive to upward mismatches (looking heavy but is light) than to downward mismatches, it could lead to an impairment in successfully performing tasks. Our study explores whether underestimating weight would lead to a worse performance than overestimating it (through expecting upward and downward mismatches).

3 Prototype

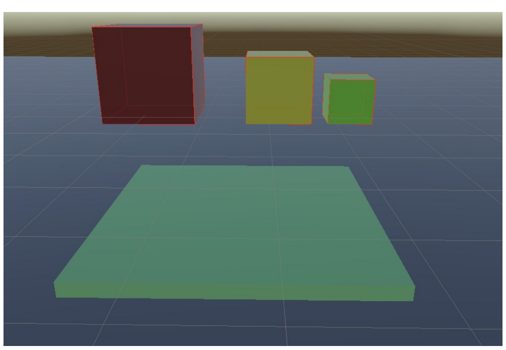

Figure 1: VR experimental scene of the target zone and different cube sizes and colors.

We developed a VR experimental system using Unity 2022.3 and a Meta Quest 3 headset to investigate how visual and haptic weight cues affect object manipulation and perceived realism. The system uses a fully automated 3×3 within-subjects design, from trial randomization to data logging.

The system runs on different components. The experiment is managed by the Unity software (figure 1), where the random 36 trials are run to spawn a cube for each trial. Using the VR controllers, objects have mass and drag, allowing users to affect their movements. Heavier objects resist acceleration and produce stronger haptic vibration during grabbing, and vice versa. The data logger then records movement data, including completion time, path efficiency, overshoots, and movement corrections. The trials start recording data once the cube is grabbed, and end once they hit the target zone after 500 ms.

4 Experimental Plan

Based on what we noticed from previous research and our ongoing approach, we propose the following three hypotheses:

1. **Congruence Performance (H1):** Congruent visual-haptic weight conditions will have better performance than incongruent conditions
2. **Asymmetric Mismatch (H2):** The impact on performance will depend on the mismatch direction, with *underestimation* of weight (looks heavy/ feels light) causing more errors than the case with *overestimation* (looks light/feels heavy).
3. **Congruence Realism (H3):** Subjectively, congruent conditions will have higher realism ratings than incongruent conditions.

We use a 3x3 within-subjects design, with thirty students as participants under the same physical conditions. The object manipulation task will go under all possible conditions of visual and haptic weight as follows:

- **Visual Cues:** Three cubes vary in scale (0.7, 1.0, 1.5) and color (Green, Yellow, Red).

- **Haptic Cues:** Three cubes vary in the following physics parameters:
  - *Light:* 0.3 kg mass, low drag, minimal vibration.
  - *Medium:* 4.0 kg mass, moderate drag and vibration.
  - *Heavy:* 15.0 kg mass, high drag, strong vibration.

Participants are required to complete 36 randomized trials, in which they move a cube to the target zone in each trial. The following metrics are recorded:

1. **Objective Performance Measures:** Completion time, Overshoot count (target re-entries), Movement corrections (change in direction), and Path efficiency. *(Used to test H1 and H2).*
2. **Subjective Realism Measure:** Participants rate *Perceived Realism* on a 1–7 Likert scale after each trial, and are required to answer yes/no on visual/haptic congruence. *(Used to test H3).*

4.1 Analysis Plan

For all analyses, participant-level averaging will be used to account for repeated measures. Data will be analyzed using separate statistical approaches for the objective and subjective hypotheses:

- **H1 (Congruence Performance):** Analyzed using a 3×3 Repeated-Measures ANOVA on participant performance metrics (completion time, corrections, overshoots, path efficiency), comparing congruent and incongruent conditions. Paired t-tests will be used to consider specific condition differences.
- **H2 (Asymmetric Mismatch):** Analyzing incongruent conditions: Comparing the two mismatch directions: *Looks-Heavier* (visually heavy/haptically light) versus *Feels-Heavier* (visually light/haptically heavy). Participant-level paired t-tests should reveal any asymmetric effects across all performance metrics. Effect sizes (Cohen's d and eta squared, η) will be reported.
- **H3 (Congruence Realism):** Analyzed using the subjective realism ratings data. Paired t-tests will compare mean realism ratings between congruent and incongruent conditions.

5 Results

Thirty participants completed 1,080 valid trials and realism ratings across nine conditions (Mean = 36 trials/participant). Repeated-measures ANOVAs revealed significant effects for corrections ($F(8, 29) = 15.60$, $p < .001$, $\eta^2 = 0.32$), overshoots ($F(8, 29) = 7.10$, $p < .001$, $\eta^2 = 0.18$), and path efficiency ($F(8, 29) = 34.12$, $p < .001$, $\eta^2 = 0.51$), but not completion time ($F(8, 29) = 0.78$, $p = .621$).

H1: Congruence Effect

Paired t-tests showed no significant differences between congruent and incongruent conditions for completion time ($t(29) = 0.68$, $p = .505$), corrections ($t(29) = -0.71$, $p = .482$), or path efficiency ($t(29) = 0.02$, $p = .986$). Only overshoots differed ($t(29) = 2.11$, $p = .044$, $d = 0.39$), opposite to predictions. As displayed in Figure 2, **H1 was not supported.** There was no significant difference in movement corrections ($p = .482$).

H2: Asymmetric Mismatch Direction

Focusing on incongruent conditions, underestimation produced dramatically worse performance:

- **Corrections:** 2.8× more (M = 102.40 vs 36.49), $t(29) = 6.12$, $p < .001$, $d = 1.12$
- **Overshoots:** 5.7× more (M = 0.86 vs 0.15), $t(29) = 3.75$, $p < .001$, $d = 0.68$
- **Path Efficiency:** 26% lower (M = 0.52 vs 0.70), $t(29) = -12.38$, $p < .001$, $d = 1.26$

Completion time showed no difference ($p = .350$). Corroborating on Figure 3, **H2 was strongly supported.**

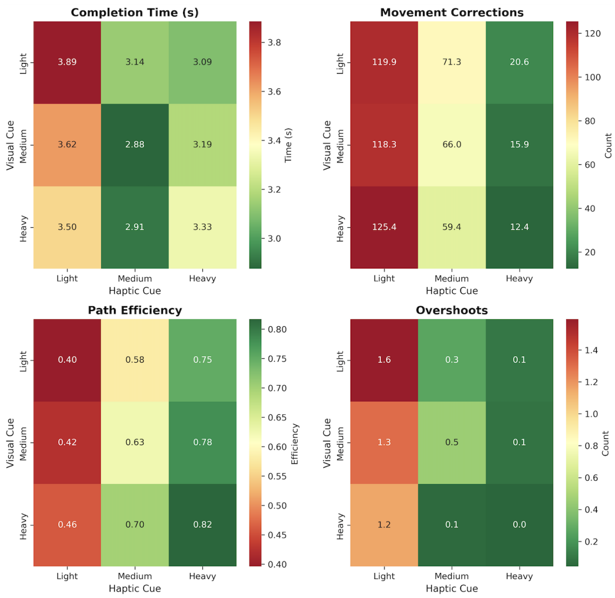

Figure 2: H1: Performance metrics across 3×3 condition matrix. (A) Completion time (s), (B) Movement corrections, (C) Path efficiency, (D) Overshoots. Diagonal cells (LL, MM, HH) = congruent conditions. Green = better performance, red = worse.

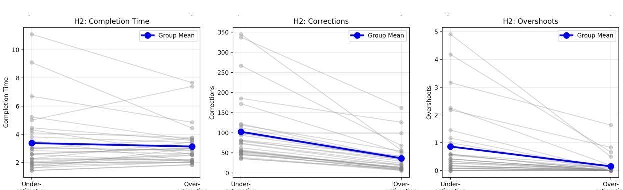

Figure 3: H2: Asymmetric mismatch effects. Each gray line = individual participant. Red = group means. (A) Completion time: no difference. (B) Corrections: dramatically higher for underestimation ($p < .001$, $d = 1.12$). (C) Path efficiency: substantially lower for underestimation ($p < .001$, $d = 1.26$).

H3: Congruence and Perceived Realism

Congruent conditions had higher ratings than incongruent, $p < .001$, $d = 2.71$. Detection accuracy was 91.4% (congruent: 95.8%, incongruent: 89.2%). According to these results and Figures 4 and 5, **H3 was strongly supported.** Mismatch magnitude apparently affected realism more than performance, where underestimation and overestimation received the same realism ratings (M = 3.47 vs 3.35, $p = .231$), despite differences in performance.

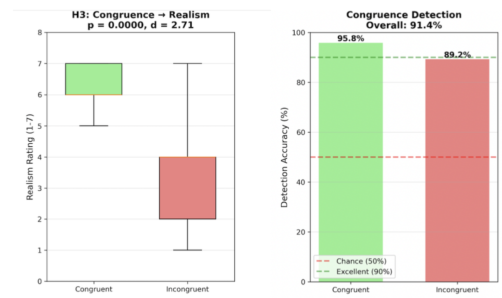

Figure 4: H3: (A) Realism ratings: congruent (M = 6.11) vs incongruent (M = 3.47), $p < .001$, $d = 2.71$. (B) Detection accuracy: 91.4% overall, above chance ($p < .001$).

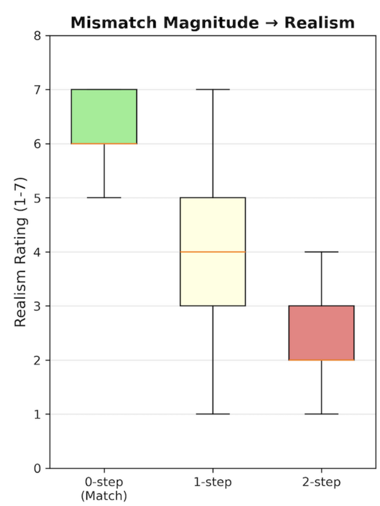

Figure 5: Realism ratings by mismatch magnitude: 0-step (M = 6.11) > 1-step (M = 4.05) > 2-step (M = 2.14). $F(2, 1077) = 1452.75$, $p < .001$, $\eta^2 = 0.37$.

6 Discussion and Conclusion

According to our study, we have found out that mismatch direction, not congruence, determines VR manipulation performance. Underestimation (visual > haptic) had 2.8x more corrections, 5.7x more overshoots, and 26% lower efficiency compared to overestimation. The results overall have shown strong support for H2 and limited support for H1. However, realism ratings (H3: $d = 2.71$) clearly dissociated between motor integration and perception.

H1's failure could be due to its design; incongruence contains two elements: underestimation and overestimation. As the hypothesis treated them as one uniform thing, they cancel each other out, making it no different than congruent conditions. Nevertheless, this highlights H2's results in the sense that the direction of the mismatch is consistent with visual dominance in motor planning. With all that being said, we were faced with several limitations, such as simple reaching task, unmeasured individual differences, or lack of haptic accuracy. The results leave us with main improvements to consider in the future. It is better to avoid simulating objects that are visually heavier than what the haptic feedback simulates. If mismatches are unavoidable, then it is better to have overestimation. Considering our failed hypothesis, the most valuable insight obtained is that visual dominance creates asymmetric costs. Underestimation heavily impaired performance, while overestimation did not.

Data and Code Availability

The experimental system (Unity project), analysis scripts (Python), and anonymized data are available at:
`https://github.com/YnsMt1/HCI.git`

References

[1] Benoît Geslain, Gilles Bailly, Sinan D Haliyo, and Corentin Duboc. Visuo-haptic illusions for motor skill acquisition in virtual reality. In *Proceedings of the 2021 ACM Symposium on Spatial User Interaction*, SUI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[2] Alexander Kalus, Johannes Klein, Tien-Julian Ho, and Niels Henze. Simulating object weight in virtual reality: The role of absolute mass and weight distributions. In *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology*, VRST '24, New York, NY, USA, 2024. Association for Computing Machinery.

[3] Yudai Tanaka, Arata Horie, and Xiang 'Anthony' Chen. Dualvib: Simulating haptic sensation of dynamic mass by combining pseudo-force and texture feedback. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*, VRST '20, New York, NY, USA, 2020. Association for Computing Machinery.

[4] Flavien Lebrun, Sinan D Haliyo, Malika Auvray, David Gueorguiev, and Gilles Bailly. Study of the impact of visual and haptic sensory sensitivities on the detection of visuo-haptic illusions in virtual reality. In *Proceedings of the 35th Conference on l'Interaction Humain-Machine*, IHM '24, New York, NY, USA, 2024. Association for Computing Machinery.

[5] Martin Feick, Niko Kleer, André Zenner, Anthony Tang, and Antonio Krüger. Visuo-haptic illusions for linear translation and stretching using physical proxies in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.