Multisensory Integration in Virtual Reality: Effect of Passive Haptic Stimulation

Master Thesis

submitted in fulfillment of the requirements for the degree
Master of Science (M.Sc.)
in the master's program "Mind and Brain"

Humboldt-Universität zu Berlin
Berlin School of Mind and Brain

Handed in by: Benjamin Dupré
Date of birth: 26.04.1986
 Address: Hoppestraße 16, 13409, Berlin

1. Supervisor: Dr. Michael Gaebler
2. Supervisor: Professor Dr. Arno Villringer

Berlin, den November 29, 2023

1. Introduction

1.1 Problem & Significance

In interoceptive research, a pivotal finding reveals the impact of visceral signals on our processing of external stimuli. Specific brain mechanisms responsible for predicting signals within the body, notably changes related to heart systole, play a crucial role in diminishing our perception of external signals like touch, vision, and auditory cues (Al, Iliopoulos, Forschack, et al. 2020; Al, Iliopoulos, Nikulin, et al. 2021; Grund et al. 2022; Motyka et al. 2019; Park et al. 2014). While this muting effect has been well-documented in psychophysical experimental setups and individual sense modalities, the influence of cardiac phases on cognitive processes like attention and memory has been under scrutiny from the early stages of research.

For instance, studies examining cognitive processes such as attention and memory have explored how the cardiac cycle impacts word retention. Words detected during systole were less effectively remembered compared to those detected during diastole (Garfinkel et al. 2013). Recent variations in experimental setups and stimuli have moved beyond phase-locked and passive stimuli, investigating active sampling of the world. Evidence suggests that more eye movement occurs during systolic phases while more fixation happens during diastolic phases (Galvez-Pol, McConnell, and Kilner 2018).

Furthermore, expanding into sensory integration, recent research has delved into how the cardiac phase influences multisensory integration (Saltafossi et al. 2023). Findings indicate that during diastole, there's a higher level of integration for Audio-Tactile and Visuo-Tactile stimuli compared to Audio-Visual stimuli when using the Race Model Inequality approach.

As the relevance of cardiac-cycle phases in sensory integration becomes more evident, research aims to develop explanations that encompass multiple observations. One such explanation involves understanding interoceptive processes through a predictive coding framework. A recent study employing a Markov decision process (MDP) using current cardiac cycle and visual stimuli suggests that this computational framework could explain observed phenomena (Allen et al. 2022), though further testing is required.

As research progresses into theoretical generalizations or practical applications, there's a growing need for ecological validity. To simultaneously test cognitive processes and multisensory integration, experiments need to align more closely with real-world scenarios. Improving ecological validity is crucial in understanding how body-brain phenomena translate into human psychology (Schmuckler 2001).

Immersive Virtual Reality (IVR) emerges as a tool facilitating a seamless transition from psychophysics findings to psychological phenomena. IVR, known for its effectiveness in studying cognitive processes within controlled yet complex scenarios, traditionally relies on visual displays and head-hand movement tracking. However, the integration of VR head-mounted displays with ECG and haptic devices presents new challenges, both practical and technical, as well as in terms of alignment with existing literature (Klotzsche et al. 2023).

This research aims to leverage Virtual Reality to bridge the gap between psychophysics findings and everyday experiences. It seeks to validate an experimental VR setup that extends two lines of research: multisensory modulation and the influence of cardiac cycles on environmental sampling. By doing so, it aims to bring these findings closer to everyday life and enhance ecological validity. However, assessing the true impact of this phenomenon on everyday experiences,

such as post-exercise sensations, poses a challenge. It necessitates a delicate balance between obtaining clear psychophysical results and integrating them with other cognitive functions.

1.2 Thesis Topic & Goal

The primary objective of this study is to investigate the feasibility of incorporating touch-cardiac-cycle modulation studies into Interactive Virtual Reality (IVR) setups. IVR, being a system that often involves visual, tactile, and proprioceptive senses, inherently engages multiple senses or is intentionally designed as a multisensory experience. To facilitate a comparative analysis of results, a relevant recent study by Martina Saltafossi on vision, touch, and hearing as multisensory pairs (Saltafossi et al. 2023) serves as a suitable reference. While there is limited research on two multisensory modalities and none to my knowledge using IVR, this study aims to bridge that gap. However, before delving into specific goals, it is necessary to define the concept of touch, as it encompasses various modes.

The extended classification of tactile sensation provides a useful framework for understanding touch, categorizing it into five different modes based on the presence or absence of voluntary movement: (1) tactile (cutaneous) perception, (2) passive kinesthetic perception, (3) passive haptic perception, (4) active kinesthetic perception, and (5) active haptic perception (Healy and Proctor 2003). For this thesis, touch is defined as passive haptic perception generated by a vibrating Data-Glove. Based on this definition, three main goals are derived:

(i) Assess the impact of passive haptic stimuli on reported immersion. This investigation quantifies the influence of passive haptic stimuli on questionnaire scores, shedding light on touch's role in creating a sense of presence. Saltafossi's study refers to this as "body illusions induced by multisensory conflicts between exteroceptive sensory modalities, such as vision and touch."

(ii) Confirm findings supporting the redundancy signal effect (RSE) and some aspects of the race model inequality (RMI). Checking for example the order of the cumulative distibution functions (CDF) for different RT. Additionally comparing single modal condition visual (F_V) and bimodal conditions were touch matches ($F_{V=T}$) or doesn't match ($F_{V \neq T}$) the visual space.

By examining how passive haptic stimuli influence the response time in the motor-memory task, this study seeks to replicate the overall findings of the reference paper (2023) and validate behaviorally the effects of passive haptic stimuli in a IVR motor-memory task. In a wider frame, this research aims to enhance our understanding of IVR as a research tool and further validate findings on multisensory integration and perception.

2. Methods

2.1 Participants

From an original pool of 23 participants that presneted to do the task we finish we 20 participats. Two of them did not fill all questionnaires and a third one was above the age limit so we had

to exclude them from the analysis, leaving us with 20 participants. The call for participants was answered considerably more by women (15) than men (5). The age across the sample was consistently around 25 years old, with the exception of one participant ($\mu = 25.1, \sigma = 6.3$).
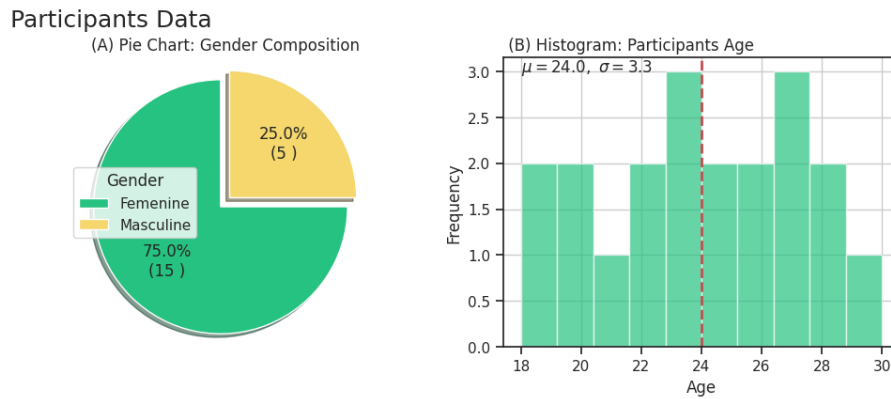

Figure 1: Participants Composition

2.2 Materials

2.2.1 Electrocardiogram (ECG):

Heart rate data is collected using an Arduino Uno and a SparkFun Single Lead Heart Rate Monitor - AD8232. The collected data is transferred through a USB 2.0 connection and integrated into the Unity log file at a frequency of 133 Hz. Based on live ECG device (Möller et al. 2022).

2.2.2 Head Mounted Display & Lighthouses:

The VR setup includes an HTC Vive head-mounted display (HMD) with two lighthouses. The headset specifications include a Dual AMOLED 3.6" diagonal display, with 1080 x 1200 pixels per eye (2160 x 1200 pixels combined), a 90 Hz refresh rate, and a 110-degree field of view. The lighthouses are equipped with SteamVR Tracking, G-sensors, gyroscopes, and proximity sensors. Both the HMD and lighthouses are connected using USB 2.0. For this study, the VR controllers were not used, and instead, hand tracking was performed using the Leap Motion sensor.

2.2.3 Leap Motion Controller:

The Leap Motion Controller has a field of view of 150x120 degrees, with a variable range of roughly 80 cm (arm's length). It weighs 32 grams and is mounted on the HMD. The device features two 640x240 infrared cameras with a frame rate of 120 fps.

2.2.4 Data Gloves:

The data gloves used in the study are equipped with magnetic sensors and connected to Unity using a microUSB connection. These gloves provide haptic feedback through 10 vibrotactile

actuators, offering a wide range of tactile sensations with 1,024 levels of intensity. The gloves also incorporate complete finger tracking using six 9-axis Inertial Measurement Units (IMUs). These IMUs enable precise tracking of finger movements, allowing for accurate gesture recognition and enhanced interaction in virtual environments.

2.3 Task

After receiving information about the COVID rules and the experiment, participants engaged in the following tasks:

(i) **Questionnaires:** Before the IVR experience, participants completed the Edinburgh Handedness Questionnaire and the PRE-Cybersickness Questionnaire. Following the IVR experience, participants filled out the Virtual Reality Subjective Evaluation Questionnaire and POST-Cybersickness Questionnaire.

(ii) **Heartbeat Count Task (HCT):** Participants performed a one-minute heartbeat count task before and after the IVR task. Note that this task is not considered within this thesis.

Following the completion of the initial questionnaires and HCT, participants moved to another room where the IVR (Immersive Virtual Reality) equipment was set up. This included a head-mounted display (HMD), data gloves, and an ECG (Electrocardiogram) device. Participants received a brief training session before proceeding with the heartbeat count task and the IVR memory-motor task.

The IVR Memory-Motor Task comprised the following steps:

a. Participants stood in front of a virtual table, allowing ample time to acclimate to the virtual environment, as depicted in Figure 2.

css

b. Prior to each trial, a new calibration process started. Participants placed their palms facing up within the shadowed hands, ensuring the standard positioning. Refer to Figure 2 (b) for the calibration setup.

c. Participants viewed a two-dimensional sketch (Figure 2 (c)) displayed in the virtual environment, prompting them to memorize the red ball's position. During this phase, participants kept their virtual hands open with palms facing up.

d. The sketch disappeared, and a red ball appeared from the top, landing in either of the participant's hands. Participants then observed a template on the table, resembling the initial sketch they memorized. Their task was to place the ball in the correct location on the template as swiftly as possible.

In 108 trials, three conditions—congruent ($V = T$), incongruent ($V \neq T$), and visual-only (V)—were randomly presented 36 times each, appearing rapidly in succession. The initiation of each trial was marked by placing the ball in a hand, and its conclusion was noted when the ball reached its designated location on the template (see Figure 2).Additionally, as mentioned before after thte IVR test final round of questionnaires intended for repeated measures on sickness and initial assessment of the immersion experienced with the haptic glove and hands.
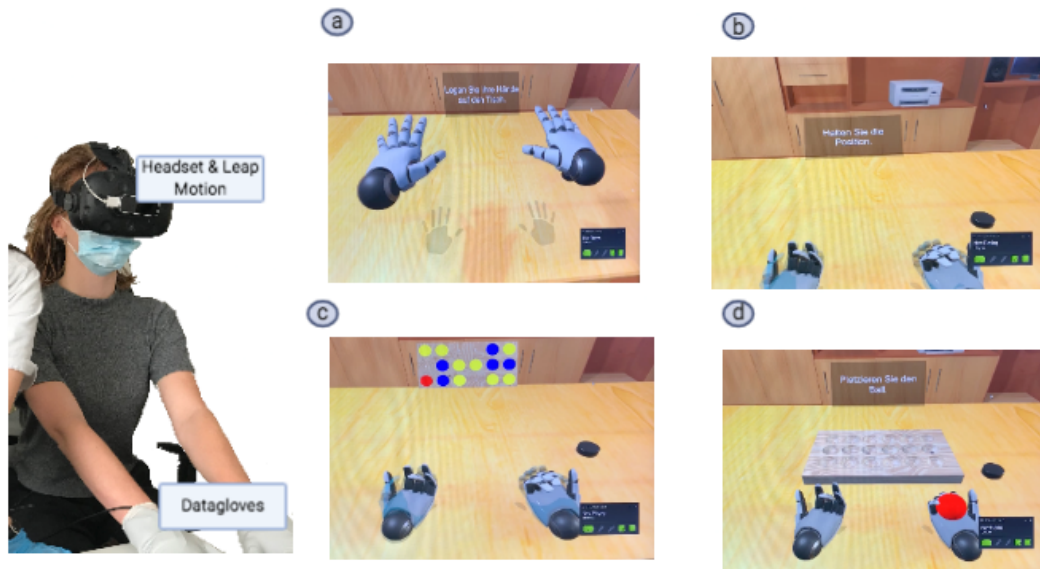
4

Figure 2: Illustration of the IVR Memory-Motor Task: (a) represents the acclimation period during initial setup; (b) showcases the first-person perspective of the calibration phase; (c) displays the 2D sketch for memorizing ball position; (d) presents the appearance of the ball and the 3D template for ball placement, marking the beginning of each trial.

2.4 Mesurements

2.4.1 Immersive Virtual Reality (VR):

The VR experience was presented and tracked using an HTC Vive head-mounted display (HMD), two lighthouses, a Leap Motion sensor, and Haptic Data Gloves. Movement data from the data gloves, Leap Motion device, and the HMD was collected. For movement analysis, only the wrist movements tracked by the Leap Motion device were considered, excluding the fingertips' magnetic tracking sensor data. All movements were recorded in a Euclidean coordinate system (X, Y, Z) with the original calibrating point set at (0, 0, 0). This provided a total of nine streaming sources of data (e.g., Headset X, Headset Y, Headset Z, and so on). Notably, rotational data was not included in the analysis.

Additionally, in the game output data, there are flags that signal if a button was pressed, if the ball is placed in the holder, and when the trial started.

2.4.4 Electrocardiogram (ECG):

Heart rate data was collected using a USB cable and integrated into Unity Engine. This information was coupled with all other Unity data at a frequency of 133 Hz.

2.4.3 Questionnaires:

(i) Virtual Reality Subjective Evaluation Questionnaire: This self-made questionnaire consists of 26 items oriented towards capturing whether the VR experience felt real for the participant or not. It explores the level of engagement, hand movement, task difficulty, and other controlling factors. The questionnaire uses a Likert scale ranging from one to seven.

(ii) PRE/POST-Cybersickness Questionnaire: The version applied in this study is a shorter adaptation of the simulator sickness questionnaire (SSQ) (1993). It employs a Likert scale ranging from one to four, with labels going from "not present," "somewhat," "clearly," and "very strongly." The questionnaire includes 16 items based on symptoms, such as "fatigue" and "general discomfort," among others.

2.5 Data Analysis

All response times were measured from the moment the ball entered the scene in the trial until the moment it "exploded". Trials where the error button was pressed were excluded from the analysis. Prior to analysis, outlier correction was performed by removing data points that deviated more than 3 × 1.4826 times the median absolute deviation (MAD) from the median, equivalent to 3 standard deviations assuming a normal distribution (Innes and Otto 2019). No responses were removed due to being too fast. However, 13% of trials were eliminated due to excessive slowness. The final sample consisted of 1826 response times (approximately 24 per condition per participant). After the removal of outliers, the response times for each condition were converted into rates (1/RT). Prioritizing normality in the error distribution over other factors, such as the scale of properties or interacting effects, was considered crucial for this proof of concept, emphasizing the essential validity of the experimental conditions.

To enable comparisons with Saltafossi's reference paper, our aim was to confirm the well-established redundant signal effect — indicating faster response times to bimodal stimulation compared to single stimulation — and partially construct the Race model Inequality to validate its applicability across all time periods t.

To achieve this, I followed the methodologies outlined in Ulrich, Miller, and Schröter 2007; Innes and Otto 2019. Given one individual stimulus and two bimodal stimuli, a simplified version without the "bounding sum" . The process involved three steps. First, empirical cumulative density functions (CDFs) were constructed for all three conditions (Bimodal pairs: $(F_{V=T})$, $(F_{V\neq T})$, and Single Signal: (F_V)). Second, percentiles were determined. Third, aggregation across participants involved computing separate paired t-tests at each percentile under examination. For a detailed breakdown of the calculations, please refer to Ulrich, Miller, and Schröter 2007.

3. Results

3.1 Assessing the Impact of the Haptic Globe on Reported Immersion

Nineteen respondents answered 27 questions, rating them on a scale from 1 (Does Not Apply) to 7 (Totally Applies). The key findings from the questionnaire are as follows:

1. In question number 24, the perceived increase in immersion due to the haptic globes was high (Mdn $= 6$, $\mu = 6$, $\sigma = 0.76$). Furthermore, the task was considered enjoyable at least some of the time (Mdn $= 6$, $\mu = 5.4$, $\sigma = 1.04$).

2. Question 26 revealed that haptic feedback was perceived as either not significantly improving performance or having a neutral effect (Mdn $= 3$, $\mu = 3.6$, $\sigma = 1.49$). Similarly, in question 12, the perception of haptic feedback improving response time was generally rated as neutral to not applicalbe (Mdn $= 3$, $\mu = 3.7$, $\sigma = 1.74$). In question 7, when asked about the impact on results, haptic feedback was perceived as not applicable (Mdn $= 6$, $\mu = 6$, $\sigma = 0.76$).

3. Notably, the assertion that it was very challenging to remember the position of the red ball (Mdn $= 2$, $\mu = 2.9$, $\sigma = 1.45$) was generally disagreed upon. Conversely, in question 8, which asked about the ease of remembering the location of the red ball, responses were as expected (Mdn $= 5$, $\mu = 5.1$, $\sigma = 1.37$).

4. The highest variance was observed in question 13 (Mdn $= 5$, $\mu = 4.2$, $\sigma = 1.79$), indicating that haptic feedback made it easier to place the ball. Similarly, question 18, which inquired about the ease of counting heartbeats at the beginning of the experiment, also showed significant variance (Mdn $= 4$, $\mu = 4$, $\sigma = 1.79$).

For more details please go to the supplement section

3.2 Touch stimuli influence the response time

3.2.1 General Performance

First, I examined the mistake rates in ball placement per condition to ensure they did not significantly affect the experiment. In the visual-only condition (V), the mistake rate was 6.35% ($\pm 0.99\%$, SEM), for the visual incongruent touch condition ($V \neq T$), the mistake rate was 5.79% ($\pm 0.59\%$), and for the visual congruent touch condition ($V = T$), it was 6.91% ($\pm 1.93\%$). However, a one-way ANOVA (feedback type) revealed no significant effects ($F \leq 0.08$, $p \geq 0.91$). Although the task is notably more complex than mere stimulus recognition, these mistake rates remain notably low and were not subjected to further analysis.

To assess whether the experimental manipulations mimicked the diversity of Redundancy Signal Effect (RSE) studies, we next tested A one-way repeated-measures ANOVA. Using the previously mentioned transformations, the test demonstrated a significant main effect for stimulus ($F(2, 38) = 3.4$, $p \leq 0.043$, $\eta p^2 = 0.15$). The median response time for the 'Congruent' condition ($V = T$) was the fastest (3236 ms ± 456 ms), followed by the 'Incongruent' ($V \neq T$) condition (3268 ms ± 470). The slowest condition was observed in the absence of haptic stimuli 'None' (V) (3284 ms ± 463). Hence, the experimental conditions effectively influenced response times as shown in figure 3.

However, post hoc Tukey HSD tests did not show any specific pairwise differences. While there is an overall difference between conditions, the specific pairwise differences were not identified

Given the absence of identified specific pairwise differences in the post hoc analysis, I decided to conduct a more comprehensive investigation using a Generalized Linear Mixed Model (GLMM).
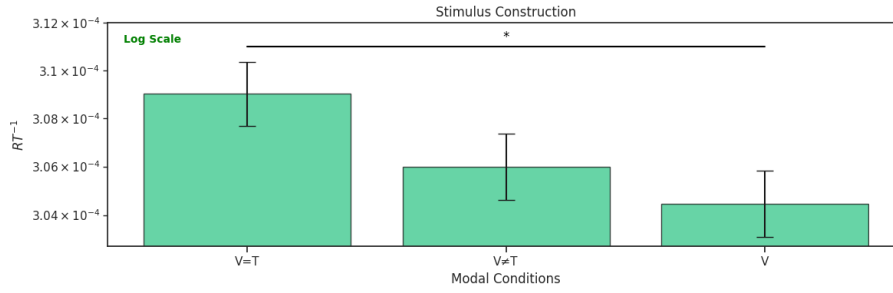
7

Figure 3: One-way repeated-measures ANOVA using RT^{-1}. The test revealed a significant main effect for the conditions ($F(2, 38) = 3.4, p \leq 0.043, \eta p^2 = 0.15$)

Although this method was not used in the referenced paper, it offers numerous advantages in multilevel research designs. Particularly, as it addresses the following issue:

"A linear relationship between the standard deviation of RTs and mean RT demonstrated in many previous studies of RT in binary choice tasks. This linear relationship is also evident in plots of the residuals, indicating heteroscedasticity in LMM analyses, characterized by an increasing spread in residuals for longer predicted RT" (Lo and Andrews 2015).

By using GLMM, rather than imposing normality and eliminating error deviation, we allow to use a distributions that matchs the properties of the mesuared RT (Lo and Andrews 2015). To do this, I utilized the *statsmodels* statistical package in Python, the *mixedlm* function. Other similar studies use the same method (Swinkels, Veling, and Schie 2021).

A mixed linear model analysis was conducted to assess the impact of feedback types ($(V = T), (V \neq T), (V)$) on response times. The model included feedback type as a fixed effect and participant as a random effect. The model coefficient for the 'none' (V) feedback type ($\beta = 45.726, SE = 22.594, p = 0.043$) reached statistical significance at the conventional level ($\alpha = 0.05$) when compared to the 'congruent' ($V = T$) feedback type. This suggests a significant difference in response times between the 'none' and 'congruent' feedback types. The coefficient for 'incongruent' feedback type ($\beta = 32.975, SE = 22.686, p = 0.146$) did not reach conventional levels of significance.

3.2.2 Redundancy Signal Effect and Multisensory Integration

4. Discussion
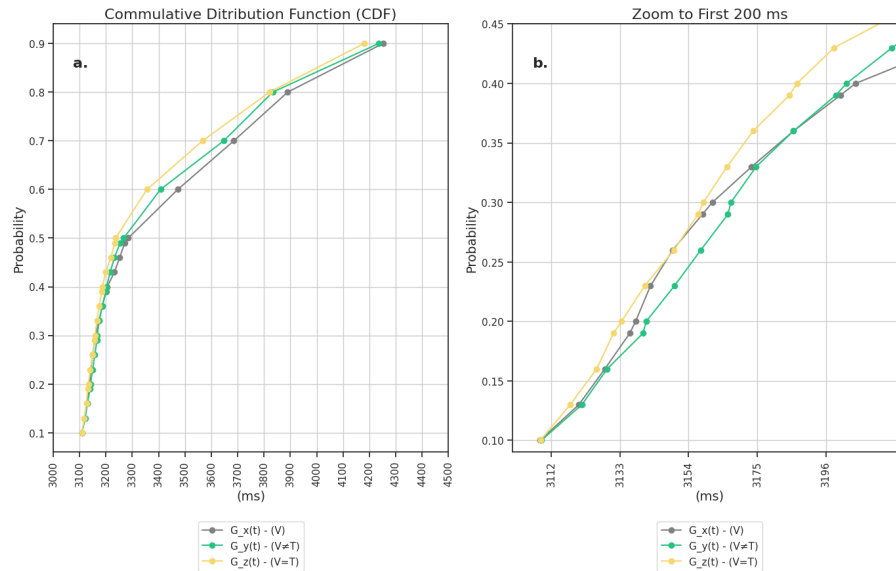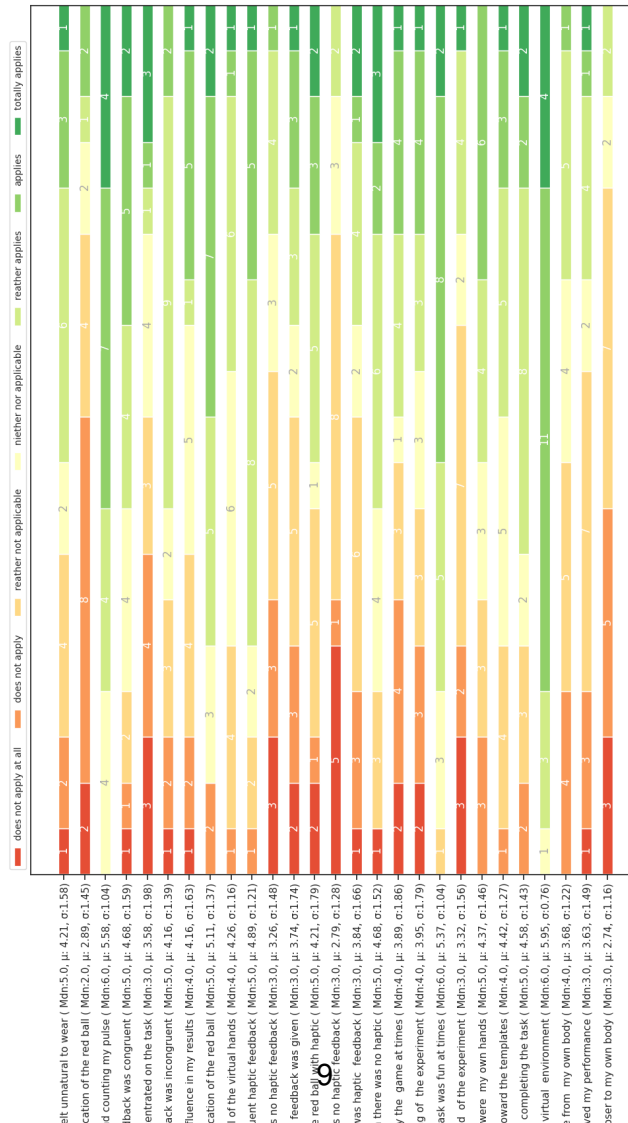
Figure 4: (a.) Displays the 10 estimated percentile points for each of the three functions of interest: Gx, Gy, Gz, considering all participants. (b.) Is a zoom in the first 200 ms. According to the RSE, the visual V type should be under Gy, Gz at all moments. Our data shows that is not the case for the RT in the first 200 ms.

Supplement

References

Al, Esra, Fivos Iliopoulos, Norman Forschack, et al. (2020). "Heart–brain interactions shape somatosensory perception and evoked potentials". In: *Proceedings of the National Academy of Sciences* 117.19, pp. 10575–10584. DOI: 10.1073/pnas.1915629117. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.1915629117. URL: https://www.pnas.org/doi/abs/10.1073/pnas.1915629117.

Al, Esra, Fivos Iliopoulos, Vadim V. Nikulin, et al. (2021). "Heartbeat and somatosensory perception". In: *NeuroImage* 238, p. 118247. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2021.118247. URL: https://www.sciencedirect.com/science/article/pii/S1053811921005243.

Allen, Micah et al. (Sept. 2022). "In the Body's Eye: The computational anatomy of interoceptive inference". In: *PLOS Computational Biology* 18.9, pp. 1–27. DOI: 10.1371/journal.pcbi.1010490. URL: https://doi.org/10.1371/journal.pcbi.1010490.

Galvez-Pol, Alejandro, Ruth McConnell, and James M Kilner (2018). "Active sampling in visual search is coupled to the cardiac cycle". In: *Cognition* 196. URL: https://api.semanticscholar.org/CorpusID:91565050.

Garfinkel, Sarah N et al. (Mar. 2013). "What the heart forgets: Cardiac timing influences memory for words and is modulated by metacognition and interoceptive sensitivity". en. In: *Psychophysiology* 50.6, pp. 505–512.

Grund, Martin et al. (2022). "Respiration, Heartbeat, and Conscious Tactile Perception". In: *Journal of Neuroscience* 42.4, pp. 643–656. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.0592-21.2021. eprint: https://www.jneurosci.org/content/42/4/643.full.pdf. URL: https://www.jneurosci.org/content/42/4/643.

Healy, Alice F. and Robert W. Proctor (2003). "Handbook of Psychology, volume 4, Experimental Psychology". In: 4, pp. 147–176.

Innes, Bobby R. and Thomas U. Otto (2019). "A comparative analysis of response times shows that multisensory benefits and interactions are not equivalent". In: *Scientific Reports* 9. URL: https://api.semanticscholar.org/CorpusID:67862549.

Kennedy, Robert S et al. (1993). "Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness". In: *The International Journal of Aviation Psychology* 3.3, pp. 203–220. DOI: 10.1207/s15327108ijap0303_3. URL: https://doi.org/10.1207/s15327108ijap0303_3.

Klotzsche, Felix et al. (2023). "Visual short-term memory related EEG components in a virtual reality setup". In: *bioRxiv*. DOI: 10.1101/2023.01.23.525140. eprint: https://www.biorxiv.org/content/early/2023/01/23/2023.01.23.525140.full.pdf. URL: https://www.biorxiv.org/content/early/2023/01/23/2023.01.23.525140.

Lo, Steson and Sally Andrews (Aug. 2015). "To transform or not to transform: using generalized linear mixed models to analyse reaction time data". en. In: *Front Psychol* 6, p. 1171.

Möller, T. et al. (2022). "An Arduino Based Heartbeat Detection Device (ArdMob-ECG) for Real-Time ECG Analysis". In: *2022 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1–3. DOI: 10.1109/SPMB55497.2022.10014819.

Motyka, Paweł et al. (2019). "Interactions between cardiac activity and conscious somatosensory perception". In: *Psychophysiology* 56.10, e13424. DOI: https://doi.org/10.1111/psyp.13424. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/psyp.13424. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/psyp.13424.

Park, Hyeong-Dong et al. (Apr. 2014). "Spontaneous fluctuations in neural responses to heartbeats predict visual detection". In: *Nature Neuroscience* 17.4, pp. 612–618. ISSN: 1546-1726. DOI: 10.1038/nn.3671. URL: https://doi.org/10.1038/nn.3671.

Saltafossi, Martina et al. (2023). "The impact of cardiac phases on multisensory integration". In: *Biological Psychology*, p. 108642. ISSN: 0301-0511. DOI: https://doi.org/10.1016/j.biopsycho.2023.108642. URL: https://www.sciencedirect.com/science/article/pii/S0301051123001606.

Schmuckler, Mark A (2001). "What is ecological validity? A dimensional analysis". In: *Infancy* 2.4, pp. 419–436.

Swinkels, Lieke M. J., Harm Veling, and Hein T. van Schie (2021). "The Redundant Signals Effect and the Full Body Illusion: not Multisensory, but Unisensory Tactile Stimuli Are Affected by the Illusion". In: *Multisensory Research* 34.6, pp. 553–585. DOI: https://doi.org/10.1163/22134808-bja10046. URL: https://brill.com/view/journals/msr/34/6/article-p553_1.xml.

Ulrich, Rolf, Jeff Miller, and Hannes Schröter (May 2007). "Testing the race model inequality: An algorithm and computer programs". In: *Behavior Research Methods* 39.2, pp. 291–302. ISSN: 1554-3528. DOI: 10.3758/BF03193160. URL: https://doi.org/10.3758/BF03193160.