Short Papers

Gamifying Haptics User Studies: Comparison of Response Times From Smartphone Interfaces

Bushra Kudsi, Doris Xu, Umit Sen, Kyle T. Yoshida, *Member, IEEE*, Fabio Stroppa, *Member, IEEE*, Cara M. Nunez, *Member, IEEE*, and Mine Sarac, *Member, IEEE*

***Abstract*—Haptics user studies are often restricted to a set, physical location and use methods that do not captivate the user. Applying game design elements can create an entertaining environment and increase user engagement. Using ubiquitous tools, like smartphones, to conduct haptics user studies could allow researchers to access larger participant groups while a gamified approach could facilitate the data collection by making the experiment more enjoyable. To explore this concept, this work presents a gamified version of an existing psychophysical experiment that investigates response time to multisensory cues using a smartphone based on "Whac-A-Mole". We conducted a user study to compare our gamified interface with an existing psychophysical interface with thirteen participants exploring the response time from eighteen combinations of auditory, haptic, and visual stimuli at different levels of intensities and participant preferences for both interfaces. The results demonstrate that the gamified interface successfully captured similar trends in response times and significantly elevated participant enjoyment ($p < 0.003$), but did not result in equivalent response times to the original interface. This work shows the benefits and drawbacks of following a gamification approach when designing haptics user studies and discusses factors and trade-offs to consider when gamifying studies.**

***Index Terms*—Gamification, user studies, vibration, perception, smartphone, response time.**

I. Introduction

Haptics user studies are usually conducted in controlled laboratory environments with custom-made setups [1]. However, such setups restrict the participation pool to the physical location of the researchers making it difficult to conduct large-scale studies on diverse populations. Smartphones hold great promise as interfaces that can be used for broad vibrotactile studies in real-world environments given their ubiquity. They have been previously used to understand the role of vibrotactile feedback in at-home rehabilitation [2], to explore the effect of physical and cognitive activities on vibrotactile perception [3], and as a tool to measure vibration perception thresholds to monitor risks affiliated with peripheral neuropathy [4], [5], [6], [7].

Received 19 July 2024; revised 12 November 2024 and 24 January 2025; accepted 24 January 2025. Date of publication 30 January 2025; date of current version 20 June 2025. This work was supported by TUBITAK 2232-B International Fellowship for Early Stage Researchers, under Program 121C147. This article was recommended for publication by Associate Editor G. J. Gerling and Editor-in-Chief S. Choi upon evaluation of the reviewers' comments. *(Bushra Kudsi, Doris Xu, Cara M. Nunez, and Mine Sarac contributed equally to this work.) (Corresponding authors: Cara M. Nunez; Mine Sarac.)*

This work involved human subjects or animals in this research. Approval of all ethical and experimental procedures and protocols was granted by Kadir Has University Ethical Review Board under Application No 77611.

Bushra Kudsi, Fabio Stroppa, and Mine Sarac are with Kadir Has University, 34083 Istanbul, Türkiye (e-mail: mine.sarac@khas.edu.tr).

Doris Xu and Cara M. Nunez are with Cornell University, Ithaca, NY 14853 USA (e-mail: cmn97@cornell.edu).

Umit Sen is with Kadir Has University, 34083 Istanbul, Türkiye, and also with the University of Massachusetts Amherst, Amherst, MA 01003 USA.

Kyle T. Yoshida is with the Washington State University, Pullman, WA 99163 USA, and also with the University of California at Los Angeles, Los Angeles, CA 90095 USA.

Digital Object Identifier 10.1109/TOH.2025.3536885

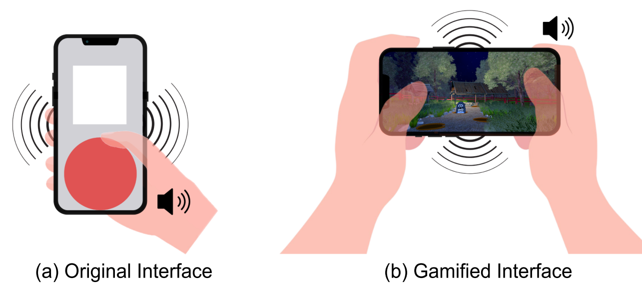

Fig. 1. Smartphone interfaces designed to conduct response time studies. (a) Original interface from [8]. (b) Gamified interface based on the arcade game "Whac-A-Mole".

One recent work used a smartphone to conduct research on the effect on response times to multimodal stimuli [8]. Understanding response times to multimodal stimuli is particularly important when designing human-machine interfaces, such as for collision avoidance [9] or surgical robotic systems [10]. Conventionally, researchers have analyzed response times through user studies with different tools such as joysticks [11] or wearable devices [12]. Yoshida et al. [8] designed a smartphone application that could relay multimodal stimuli (visual, auditory, and haptic) at different intensity levels (high, low, and none) and conducted a user study to understand the interaction between these stimuli. Their findings showed that they could use the smartphone platform to yield similar results to custom-made experimental tools, enabling future perceptual studies that extend to wider communities. However, their smartphone interface featured a basic design unlikely to entice participation from a large group. Their interface contained a large red button for the users to interact with to provide their responses and relayed visual stimuli in the form of a white square to the top-center of the screen (Fig. 1(a)). Participants needed to focus on the screen and the task for a 13-minute block before receiving a short break (and then continuing to complete 3 total blocks). The design of this original interface and study could result in users' boredom and loss of motivation, negatively impacting users' focus during the study and, thereby, the overall results.

Gamifying the user interface to collect perceptual data could potentially offer a more engaging and entertaining solution. Research has shown that gamification increases users' motivation and engagement, making the tasks more attractive, even without game mechanics [13]. It also makes tiresome and boring tasks more interesting, such as a data collection task for the creation of machine learning models [14]. Furthermore, making tasks more interesting and enjoyable improves users' level of engagement – especially for tasks with high cognitive load [15]. As a result of these findings, researchers from many different fields have started gamifying their user interfaces including creating

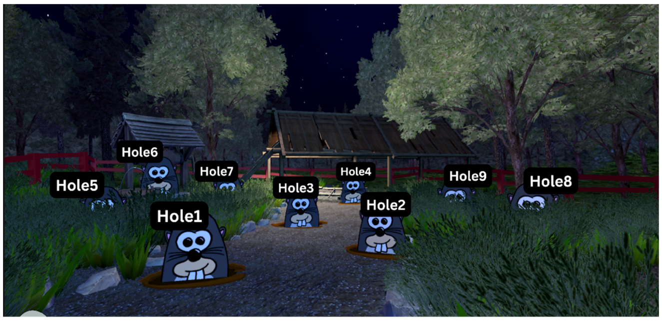

Fig. 2. Locations of the 9 possible target holes. High-level visual holes are located on the pathway (holes 1-4), and low-level visual holes are concealed within the grass (holes 5-9).

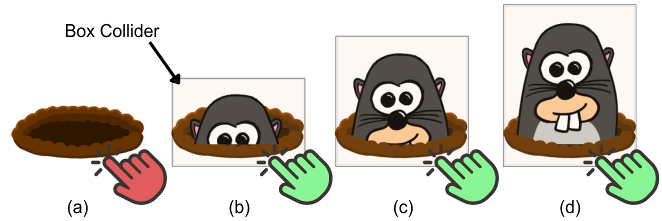

Fig. 3. Mole exposure levels. (a) When there is no visible mole, there is no box collider, and the user cannot interact with the hole. (b) - (d) When the mole ascends from the hole, a box collider starts ascending with it, dynamically extending its height from the lower edge based on the mole's position and allowing the user to interact with it at any time, leading to a successful hit.

	Stimulus	Off	Low	High
Original	Audio	0 dB	66 dB	86 dB
	Haptic	0g, 0Hz	0.036g, 230Hz	0.756g, 230Hz
	Visual			
Gamified	Audio	0 dB	70 dB	79 dB
	Haptic	0g, 0Hz	0.191g, 230Hz	0.820g, 230Hz
	Visual			

Fig. 4. Different levels of sensory stimuli rendered by the interfaces: visual (low or high), audio (off, low, or high), and haptic (off, low, or high).

sign language games for dataset generation [16], a game in which children have to "pop" (press) visual stimuli that appear and bounce around on a tablet screen to assist in diagnosing and monitoring amblyopia [17], and interactive, story-based games about dragons and police dogs to measure psychoacoustic thresholds [18]. Incorporating game design elements in haptics user studies, however, is still relatively unexplored as most studies involving gaming focus more on the role of haptic feedback in creating an immersive experience [19].

In this work, we aim to explore the impacts of gamifying user studies focused on vibrotactile perception. We propose a gamified version of the previous response time experiment interface from Yoshida et al. [8] based on the classic Japanese arcade game "Whac-A-Mole" (Fig. 1(b)). We hypothesize that our gamified interface will elicit similar performance in response times from multimodal stimuli as the original interface, but will have improved participant enjoyment, engagement, and overall user experience. This hypothesis is based on the design similarity of the multimodal cues across both interfaces, as well as the established appeal of mobile games, which is likely to increase user immersion without fundamentally altering the response trends. Section II describes the design of the gamified smartphone-based system and the conducted user study. Section III presents our analyses and results, and Section IV discusses the implications of our results and the comparison between the experimental and gamified interfaces, highlighting that our hypothesis was confirmed and the gamified interface was significantly more enjoyable for participants while obtaining similar response time trends. Lastly, Section V summarizes the work and discusses potential directions for future work.

II. Methods

A. Gamified Interface Design

We designed a version of the "Whac-A-Mole" game for smartphones as an iOS application (app) with Unity 2022.3.11f1. The scenery of the game was designed to imitate the ambiance of a peaceful summer night, with background sounds intended to immerse players in the scene.

During the game, the users are shown a scene with nine empty holes located either in the grass or on the path (Fig. 2). In each trial, one hole is determined as the target hole from which a mole appears. The location of the target hole is picked pseudo-randomly by the app every 2-5 s, such that moles pop out of each hole the same number of times. The appearance of the mole provides a visual cue, which is accompanied by a combination of audio and haptic stimuli at varying levels, resulting in eighteen unique combinations (Fig. 4):

- *Visual Stimuli:* There are two levels of visual stimuli (high or low) based on the location of the target hole. Holes that are located on the path (holes 1-4 in Fig. 2) are high-level visual stimuli as the mole is unobstructed and in clear view, and holes that are located in the grass (holes 5-9 in Fig. 2) are low-level visual stimuli as the mole is slightly obstructed by the grass. It is not possible to implement off-level visual stimuli within the gamified interface since a mole must appear for the player to *whack* it.
- *Audio Stimuli:* There are three levels of audio stimuli (high, low, and off) and two audio sounds depending on the target hole: *(1)* a digging sound effect [20] when the mole pops up from the pathway holes (holes 1-4) or *(2)* a rustling grass sound effect [21] when the mole appears from grassy holes (holes 5-9). Both sound effects are sourced from copyright-free SFX on YouTube and are played at one of three intensity levels adjusted through the "volume" variable (range of 0.0 to 1.0) of the AudioSource component in Unity: high-level audio was set to 1, low-level audio to 0.3, and off-level audio to 0.
- *Haptic Stimuli:* There are three levels of haptic stimuli (high, low, and off). The vibration waveforms were designed using the Interhaptics Haptic Composer tool and rendered using the Interhaptics Unity SDK [22]. The vibration stimuli are 150 Hz sine waves rendered for 0.7 s at three levels determined by the "amplitude" variable (range of 0.0 to 1.0): high-level haptic was set to 1, low-level haptic to 0.3, and off-level haptic was 0.

The users should locate the mole appearing from one of the nine target holes and *whack* it as quickly as possible by tapping the screen at the mole's location. The interaction is implemented with an invisible bounding-box collider placed around the visible section of the mole, changing its height to fully contain the mole ascending from the hole – with a padding of 250 mm on the top and bottom side of the mole and 500 mm on the sides to facilitate interaction (Fig. 3). Users must tap within this box to correctly *whack* the mole. Each mole remains on the screen until it has been hit correctly.

B. Implementation of Interfaces

We implemented our gamified interface as well as the interface used to conduct user studies on response times from [8] on an iPhone SE (2nd generation). Below we describe the similarities and differences between the two interfaces:

1) Original Interface: This interface is from previous work [3] and is used as the baseline for comparison. It was designed to be displayed on the smartphone with a vertical orientation. Participants were instructed to hold the phone in their dominant hand as shown in Fig. 1 and to interact with the large red button at the screen using their dominant thumb after receiving stimuli. They received two levels of visual stimuli in the form of a white square (4.7 cm $\times$ 4.7 cm) that appears against a grey background for 0.5 s at either high transparency ($Opacity = 0.1$) or low transparency ($Opacity = 0.9$).

Participants received three levels of audio stimuli from a pre-recorded soundtrack of a 746 Hz tone [23] played using the AVF Audio framework (Apple Inc.) with an *AVAudioPlayer* volume of 0.1, 0.01, and 0. The output volume was measured using the Decibel X software on an iPhone 13 at the output of the speaker of the iPhone SE with maximum volume. The measurements were 86 dB, 66 dB, and 0 dB.

Participants also received three levels of haptic stimuli rendered for 0.1 seconds using the Core Haptics Framework (Apple Inc.) with *hapticSharpness* = 1.0 and *hapticIntensity* of either 1.0, 0.3, and 0.0 [24]. Output vibrations were measured using an accelerometer (Analog Devices, EVAL-ADXL354CZ) attached to the center of the screen on the iPhone SE that was placed on a benchtop screen-side up and a DAQ (National Instruments, NI6003) and processed and filtered with a bandpass frequency of 60-500 Hz in MATLAB (Mathworks). The vibrations have amplitudes of (0.756 g, 0.036 g, and 0 g) and a frequency of 230 Hz.

2) Gamified Interface: This interface is based on the "Whac-A-Mole" game as described in Section II-A. It was designed to be displayed on the smartphone with a horizontal orientation. Participants were instructed to hold the phone with both hands (Fig. 1), to interact with the moles using both thumbs after receiving stimuli, and to keep their thumbs above the screen to react quickly and complete the trial successfully.

Participants received two levels of visual stimuli: high-level visual stimuli from moles appearing in holes concealed within the grass and low-level stimuli appearing in holes on the pathways. They received three levels of audio stimuli with output volumes of 79 dB, 70 dB, and 0 dB for the digging sound and 78 dB, 70 dB, and 0 dB for the rustling grass sound as measured using the same method as described for taking the measurements with the original interface. They also received three levels of haptic stimuli with measured amplitudes of 0.820 g, 0.191 g, and 0 g using the method described above for the original interface. The output frequency of the vibrations was 230 Hz (matching the original interface) even though the vibration signal was 150 Hz likely because the actuator within the iPhone SE is a linear resonant actuator [25]. Therefore, the frequency of the haptic stimuli received by the participants was consistent across the two interfaces and had similar amplitudes for each haptic level.

C. User Study

We conducted a user study where participants completed response time experiments using the gamified and original interfaces on an iPhone SE (2nd generation) with iOS 16.5. The study protocol was approved by the Kadir Has University Review Board and is in accordance with the Declaration of Helsinki. All participants gave written informed consent.

1) Participants: We recruited 16 individuals, but 3 were removed from the analysis as outliers based on the Interquartile Range method [26], [27]. As such, the analysis included 13 participants (5 male and 8 female, aged 19-24). Upon arrival, participants were asked to fill out a pre-study survey containing questions about demographics. Regarding hand dominance, 2 participants reported being left-handed, and 11 reported being right-handed. Regarding their level of experience with human-machine interaction devices, 4 reported themselves as complete novices, 4 as beginners, 4 as intermediate users, and 1 as an expert.

2) Study Protocol: After providing informed consent, participants were asked to complete a pre-study survey regarding their demographic information and experience level with human-machine interaction devices (interfaces that resemble our study setup, including devices such as keyboards, mice, touchscreens, and other interactive technologies). They were informed about the study protocol, the type of stimuli they might experience during the study, and how to complete the trials using both interfaces. The fully charged smartphone was set to work with the maximum volume and maximum brightness.

The user study consisted of two phases (one for each interface) with a short three-minute break in between. The order of the phases was randomly determined and balanced with Latin Squares. In each phase, eighteen unique stimuli combinations were presented ten times. Each participant reacted to 180 stimuli at each phase and 360 stimuli overall. Each participant took approximately 35-40 minutes to complete the study, including the completion of two surveys.

3) Collected Data: We collected the following data for each provided stimulus by each interface:

- *Stimulus Timestamp:* The timestamp each stimulus is rendered to the participant from the start of the phase.
- *Response Timestamp:* The timestamp the participant reacts to the stimulus by tapping the screen (i.e., the mole or the target circle) from the start of the phase.
- *Tap Timestamp:* The timestamp the participant taps on the screen at any location to understand whether they were tapping on the screen randomly or intentionally.

Each timestamp is recorded with the resolution of 20 ms, which is Unity's default fixed timestep. This choice ensures precise capture of player interactions while avoiding frame rate drops and maintaining computational efficiency. The **response time** is calculated as the difference between the *Stimulus* and *Response Timestamp*s. In addition, we also recorded the following data *only* for the gamified interface:

- *Tap XY Coordinates:* The x and y pixel coordinates of the location where the participant tapped on the screen, recorded to understand whether participants were randomly tapping on the screen during the study or just missing the correct location of the mole.

Participants also completed two surveys:

- *Short Task-Load Index Survey:* After each phase, they rated their physical, mental, and temporal demands, performance, frustration, engagement, and enjoyment of each interface on a 10-point Likert scale (inspired by NASA TLX).
- *Post-Study Survey:* After both phases, they compared the two interfaces and reported which interface they enjoyed more, felt more focused using, performed better with, were more motivated to complete, felt lasted longer, and would prefer to do again.

III. RESULTS

A. Response Time Data

Fig. 5 shows boxplots of the response times obtained by the original and gamified interfaces for all conditions – two levels of visual (high and low), three levels of audio (high, low, and off), and three levels of haptic stimuli (high, low, and off). For the original interface, response time was the fastest with high levels of visual, audio, and haptic stimuli (347 $\pm$ 49 ms). For the gamified interface, it was the fastest with high-level visual and audio stimuli and off-level haptic stimuli (763 $\pm$ 189 ms). The response times for both interfaces were the slowest with low-level visual, off-level audio, and off-level haptic

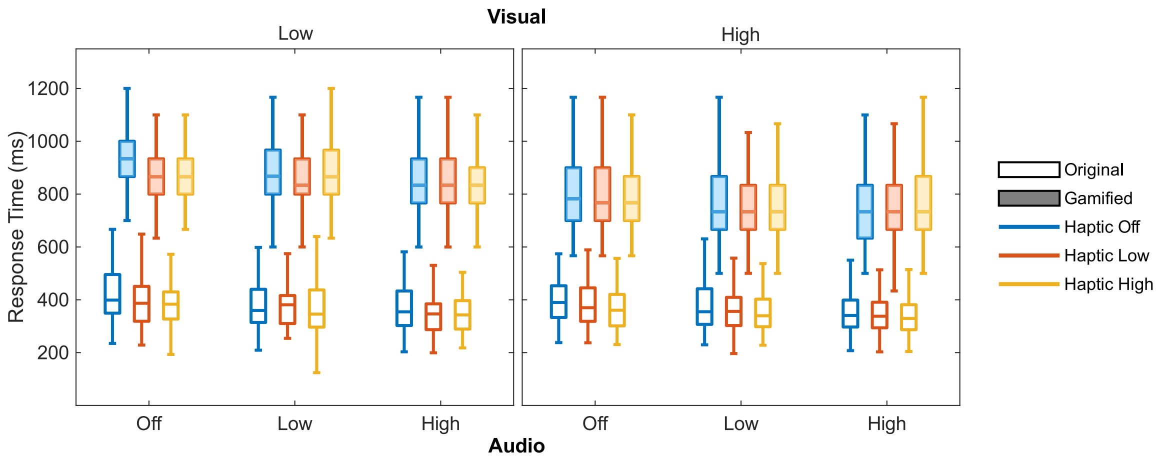

Fig. 5. Boxplots showing the response times for each combination of audio, haptic, and visual stimuli for each interface. Visual levels are denoted by subplots, audio levels are denoted by x-axis location, haptic levels are denoted by color, and interface type is denoted by the boxplot shading or lack thereof.

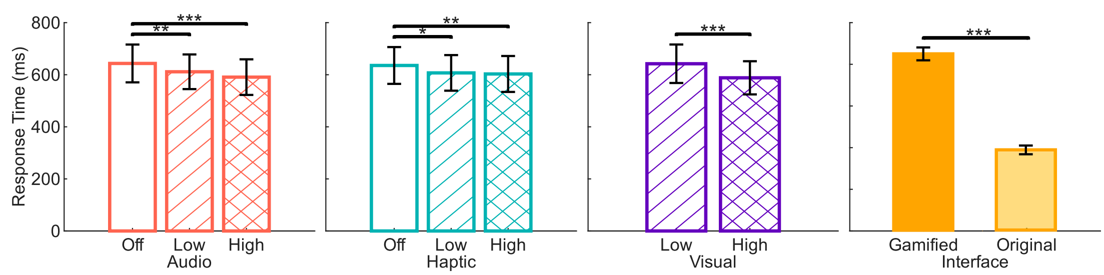

Fig. 6. Mean and standard error of the response times for each level of the main factors (audio, haptic, visual, and interface). Standard significance notation is used for p-values ($*** : p < 0.001$, $** : 0.001 < p < 0.01$, and $* : p < 0.05$).

stimuli (469 ± 128 ms for the original and 994 ± 275 ms for the gamified interfaces).

We conducted a four-way ANOVA with visual, audio, and haptic levels and interface as independent factors. Our results showed significant main effects for visual level ($F(1, 12) = 193.891$, $p < 0.001$, $\eta_p^2 = 0.942$), audio level ($F(2, 24) = 32.307, p < 0.001, \eta_p^2 = 0.729$), haptic level ($F(2, 24) = 19.004$, $p < 0.001$, $\eta_p^2 = 0.613$), and interface ($F(1, 12) = 1149.902$, $p < 0.001$, $\eta_p^2 = 0.99$). We performed post-hoc pairwise comparisons with Bonferroni corrections for each main effect to examine how the varying levels of each factor influenced response time. The statistical significance of these analyses are presented in Fig. 6.

These main effects were qualified by an interaction *only* between visual level and interface ($F(1, 12) = 113.292, p < 0.001, \eta_p^2 = 0.904$) and between audio and haptic levels ($F(4, 48) = 7.651$, $p < 0.001$, $\eta_p^2 = 0.389$). There was no significant interaction between audio and visual levels ($F(2, 24) = 0.522$, $p = 0.600$, $\eta_p^2 = 0.042$), haptic and visual levels ($F(2, 24) = 3.032$, $p = 0.067$, $\eta_p^2 = 0.202$), audio level and interface ($F(2, 24) = 3.007$, $p = 0.068$, $\eta_p^2 = 0.200$), and haptic level and interface ($F(2, 24) = 0.408$, $p = 0.669$, $\eta_p^2 = 0.033$). We then performed post-hoc pairwise comparisons with Bonferroni correction to explore the interaction effects (Fig. 7).

The first post-hoc analysis focused on the visual stimuli level for each interface. Our results show that participants had faster response times with high-level visual stimuli than low-level stimuli with the gamified interface ($p < 0.001$), but not with the original interface ($p = 0.256$). Additionally, the original interface resulted in significantly faster response times than the gamified interface with both low-level visual stimuli ($p < 0.001$) and with high-level visual stimuli ($p < 0.001$). The second post-hoc analysis focused on the audio stimuli level for each level of haptic stimuli. With off-level haptic stimuli, off-level audio stimuli resulted in statistically significantly slower response times compared to low-level ($p < 0.001$) and high-level audio stimuli ($p < 0.001$). There was no significant difference between low-level and high-level audio stimuli with off-level haptic stimuli ($p = 0.306$). With low-level and high-level haptic stimuli, there were no significant differences between the three levels of audio stimuli (all $p > 0.05$). In addition, with off-level audio stimuli, participants were statistically significantly slower with off-level haptic stimuli compared to low-level ($p = 0.002$) and high-level haptic stimuli ($p < 0.001$), but there was no statistically significant difference between low-level and high-level haptic stimuli ($p > 0.05$). All other combinations of audio-haptic interactions were found to be not significantly different from each other (all $p > 0.05$).

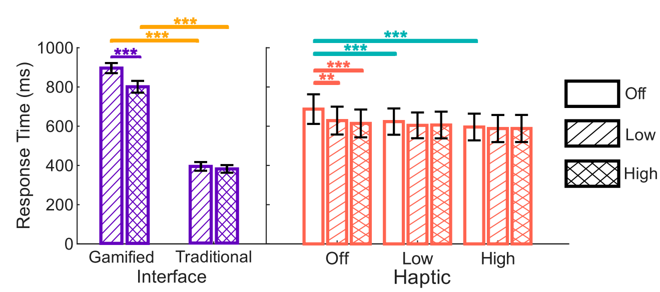

Fig. 7. Mean and standard error of the response times for the visual level for each interface (left) and audio level for each haptic level (right) corresponding with the interaction effects. Statistical significance of post-hoc pairwise comparisons of the interaction effects are shown in purple for differences between visual levels for each interface and in orange for differences between interfaces for each visual level (left), and in coral for differences between audio levels for each haptic level and cyan for differences between haptic levels for each audio level (right). Standard significance notation is used for p-values ($*** : p < 0.001$ and $** : 0.001 < p < 0.01$).

TABLE I
RESULTS FROM POLYNOMIAL CONTRAST ANALYSIS AND PAIRWISE INTERACTION CONTRASTS

Figure	Interface	Visual	Audio	Slope	Difference in slope	SE	Df	T-ratio	P-value
(a)	Gamified	Low	Off	-0.109		0.022	420	-5.040	**<0.001**
	Original	Low	Off	-0.085		0.022	420	-3.903	**<0.001**
	Pairwise Interaction Contrast				-0.025	0.031	420	-0.803	0.422
(b)	Gamified	Low	Low	0.002		0.022	420	0.105	0.916
	Original	Low	Low	-0.019		0.022	420	-0.884	0.377
	Pairwise Interaction Contrast				0.021	0.031	420	0.699	0.485
(c)	Gamified	Low	High	-0.022		0.022	420	-1.000	0.317
	Original	Low	High	-0.016		0.022	420	-0.729	0.466
	Pairwise Interaction Contrast				-0.006	0.031	420	-0.193	0.847
(d)	Gamified	High	Off	-0.037		0.022	420	-1.726	0.085
	Original	High	Off	-0.060		0.022	420	-2.768	**0.006**
	Pairwise Interaction Contrast				0.023	0.031	420	0.737	0.462
(e)	Gamified	High	Low	-0.032		0.022	420	-1.452	0.147
	Original	High	Low	-0.021		0.022	420	-0.970	0.333
	Pairwise Interaction Contrast				-0.010	0.031	420	-0.341	0.734
(f)	Gamified	High	High	0.030		0.022	420	1.379	0.169
	Original	High	High	-0.023		0.022	420	-1.037	0.300
	Pairwise Interaction Contrast				0.052	0.031	420	1.708	0.088

Given the significant main effects of interface and haptic level were not qualified by an interaction ($p = 0.669$, $\eta_p^2 = 0.033$), we conclude that the trends of the participants' response times across different levels of haptic stimuli does not suggest a significant difference between the gamified interface and the original interface. Similarly, the significant main effects of audio stimuli and interface were not qualified by a significant interaction ($p = 0.068$, $\eta_p^2 = 0.200$), which does not suggest a significant difference between different levels of audio stimuli across both interfaces. To further explore the similarity in the trends, we conducted a polynomial contrast analysis to assess the linear trends of the response time between the haptic levels within each combination of visual and audio levels across both interfaces. Subsequently, we conducted a pairwise interaction contrast to determine how the trends in response time between the haptic levels for each level of visual and audio stimuli compare between the gamified and traditional interfaces. We present the results of our analyses in Table I and Fig. 8. Our analyses revealed no significant differences between the interfaces ($p > 0.05$, for all combinations of audio and visual stimulus levels).

B. Survey Data

Fig. 9 shows the mean and standard error of the participant responses to the task-load index surveys and survey questions regarding engagement and enjoyment. Wilcoxon signed-rank tests were conducted for each survey question, and only the responses about engagement were found to be statistically significant ($p = 0.003$, $r = -0.827$).

Fig. 10 shows the responses to the post-study survey comparing the two interfaces. 92.3% of participants (12 out of 13) found the gamified interface more enjoyable than the original one. Similarly, 69.2% of participants (9 out of 13) stated that they were more focused, 100% of participants (13 out of 13) perceived better performance, and 92.3% of participants (12 out of 13) reported they were more motivated while using the gamified interface than the original interface. Additionally, 69.2% of participants (9 out of 13) stated they thought the trials lasted longer while using the original interface than the gamified interface. Overall, 92.3% of participants (12 out of 13) preferred the gamified interface compared to the original interface.

IV. DISCUSSION

A. Response Time Data

Our results showed that the presence of multisensory stimuli (i.e., audio or haptic stimuli provided in addition to visual stimuli) results in

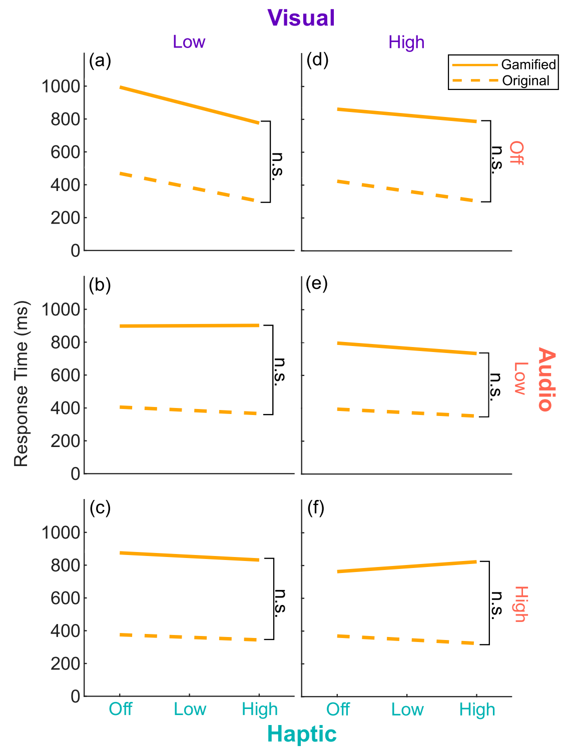

Fig. 8. Plots from the results of the polynomial contrast analysis and pairwise interaction contrasts (Table I) illustrating the trends in response times across haptic stimuli levels for the gamified and original interfaces under different combinations of visual and audio stimuli levels. There are no significant differences between the trends of the interfaces for any combination of visual and audio stimuli levels.

faster response times than visual stimuli only (off-level audio and haptic stimuli conditions) and the combination of these multisensory stimuli also influences their response times for both the gamified and original interfaces. Yet, our analysis showed that participants were significantly slower with the gamified interface than the original. This indicates that while a gamified approach to response time research *can determine relationships and trends* between multisensory variables, it *may not result in equivalent values*.

Even though the gamified interface was designed to mimic several aspects of the original one, such as the type of stimuli and intensity levels, participants were significantly slower with the gamified interface compared to the original. There might be a few possible reasons for such a difference. Our first speculation concerns how participants hold and interact with the phone. As shown in Fig. 1, they held the phone horizontally and used both their thumbs for the gamified interface, while they held the phone vertically in their dominant hand and used their dominant thumb for the original interface. There is likely more neural processing involved when the participant needs to decide which thumb to move to respond, as opposed to only moving their dominant thumb, which would explain the longer response times from the gamified interface.

Our second speculation concerns the idle thumb position as participants wait for the stimuli. With the original interface, they positioned their thumb over the area while waiting for the stimuli and only move downwards after the stimuli. With the gamified interface, they kept their thumbs off to the sides of the phone so as not to occlude the screen while waiting. When a mole appears, they must move one of their thumbs to the correct spot on the screen and downwards to make contact with the mole. It is possible that this additional thumb movement played a role in the longer response times with the gamified interface. We define reaction time as the time elapsed from the delivery of an imperative stimulus until the initiation of the physical action, and movement time

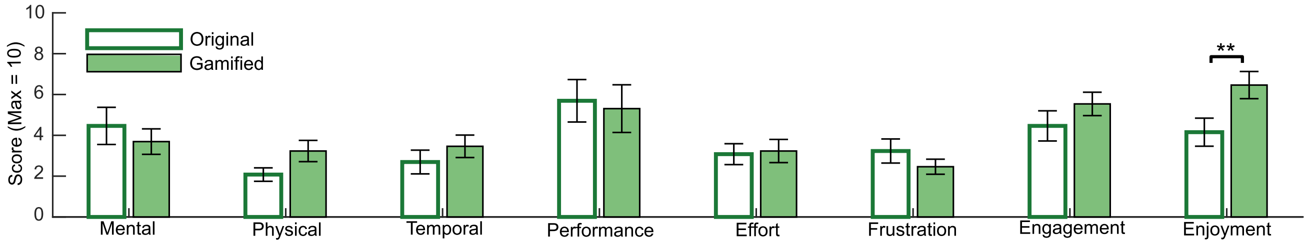

Fig. 9. Mean and standard error of participant responses to the task load index surveys pertaining to mental, physical, and temporal demands, performance, effort, frustration, engagement, and enjoyment. ($** : 0.001 < p < 0.01$).

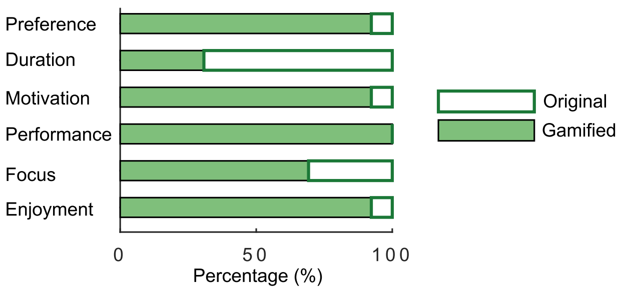

Fig. 10. Percentage of participant selections during the post-study surveys comparing the two interfaces pertaining to their enjoyment, focus, perceived performance, motivation, and perceived duration and preferred interface.

as the time elapsed from the initiation of a prescribed action until the action is completed [28]. In our study, our measured response time is the combination of the participants' reaction time and movement time. Thus, the longer response time in the gamified interface does not necessarily mean longer reaction time, but could be a result of the longer movement time, caused by the thumb traveling across the screen.

Finally, our third speculation concerns a difference between the two interfaces regarding the cognitive load. Prior work has shown that reaction time depends on the complexity of the task, and that reaction times are slower as complexity increases [29]. In the original interface, participants receive all visual stimuli in the same, fixed location. In the gamified interface, the visual stimuli are provided in locations distributed throughout the screen. This requires participants to constantly scan the environment for a mole, instead of focusing on a single location on the screen, increasing the task complexity. Additionally, while the original interface has a plain, static, gray background, the gamified interface has more interactive scenery that could distract the participants.

Additionally, our results showed that changing the visual stimuli from low-level to high-level led to faster response time while using the gamified interface, but not the original one. While the difference in visual levels for the original interface was from changes in the opacity of the displayed visual stimuli, the difference between visual levels with the gamified interface was a bit more complex. Low and high-level visual stimuli for the gamified interface were determined based on the location of the hole (in plain sight in the path or occluded by the grass). This implementation of visual stimuli levels for the gamified interface incorporates many of the elements previously discussed in this section (thumb selection, additional thumb movement, and additional cognitive load) which could explain the differences between the interfaces.

B. Survey Data

The results demonstrate that the gamified approach offered participants a statistically significantly more enjoyable experience as the enjoyment Likert scale survey scores were significantly higher for the gamified interface and 92.3% of the participants reported the gamified interface was more enjoyable than the original. While the results do not indicate a statistically significant increase in engagement, there is a higher score for engagement of the gamified interface than for the original. Additionally, 69.2% of participants reported feeling more focused on the task, and 92.3% reported they were more motivated to complete the task using the gamified interface. Similarly, despite the tasks for each interface taking the same time, 69.2% of participants thought that the session with the original interface lasted longer than the gamified interface. Lastly, given that 92.3% of participants preferred the gamified interface to the original interface, it is clear that using a gamified approach improves the user experience and would likely increase user participation in future work. We believe that ensuring participants enjoy the tasks would lead to a better user experience. A better user experience could lead to increased motivation of participants, leading to a larger participant pool or to complete the experiments with more repetitions. A larger sample size is key to drawing meaningful insights, making this gamified approach a promising tool for long-term psychophysical research.

C. Factors to Consider When Gamifying Studies

The main goal of this work was to explore the potential of gamifying haptics user studies. Our gamified interface obtained similar trends in the recorded response time results, even though the response times for the gamified interface were significantly higher than those from the original interface. As previously discussed, potential reasons for these differences include how the phone was held, the distributed visual stimuli from the mole locations, and interactive scenery. However, these factors likely also resulted in the increased enjoyment and engagement experienced with the gamified interface. For example, the gamified interface could have been designed with a vertical orientation with one target mole location in the center of the screen to more directly mimic the original interface. On the other hand, this does not fully capture the fun aspect of the "Whac-A-Mole" game and would likely be less engaging. Another example of a difference between the interfaces was the simple tone sound played by the original interface compared to the digging and rustling grass sound effects played by the gamified interface; these game-specific tones likely contributed to the increased enjoyment of the gamified interface but could have negatively impacted the response times. This demonstrates that designers may need to consider trade-offs between certain factors when designing their interfaces and posing their research questions.

V. Conclusion

In this work, we designed a gamified version of an existing psychophysical experiment investigating the impact of multisensory cues on response time using a smartphone inspired by "Whac-A-Mole" to explore if gamification can be used to improve participant experience and increase their engagement. We conducted a human subjects user study to compare response times to multimodal stimuli from our gamified design to the original one and to collect survey data regarding the user experience. Our findings indicated that both interfaces exhibit no significant differences in trends and behaviors regarding the relationships between multisensory variables, even though participants took significantly higher times to respond with the gamified interface compared to the original one. Despite eliciting slower response times, the gamified interface showed increased participant enjoyment and engagement compared to the original interface. Higher levels of enjoyment and engagement could ultimately result in recruiting larger pools of participants and minimizing participant fatigue during studies, enabling them to maintain focus for a longer period of time resulting in the collection of more reliable data.

In the future, we will conduct intensive user studies to investigate our speculations regarding the differences between the two interfaces. We will simplify the gamified interface such that the phone is held vertically with one hand and there is a single hole from which the mole can appear, ensuring a more consistent user experience across both interfaces. We will then introduce one element at a time (horizontal phone with two thumbs, multiple holes, etc.) to understand how each component affects the participants' response time. We will also explore alternative gamified designs that could result in response times similar to the original interface. Once we have a stronger understanding of the impacts of different factors pertaining to gamification, we plan to conduct large scale, in-the-wild studies to explore the impact of gamification on participant recruitment and the fatigue experienced during the study. Finally, we will expand this gamification approach to other topics within haptic perception beyond response times, exploring the effects on discrimination performance metrics, such as just noticeable differences. This expanded research will help us better understand the broader applicability of gamification in haptics studies and its impact on various psychophysical measurements beyond response times.

Acknowledgment

The authors would like to thank Elif Batci for her assistance in conducting the user study.

References

[1] J. R. Blum et al., "Getting your hands dirty outside the lab: A practical primer for conducting wearable vibrotactile haptics research," *IEEE Trans. Haptics*, vol. 12, no. 3, pp. 232–246, Jul.–Sep. 2019.

[2] Q. Mourcou, A. Fleury, B. Diot, and N. Vuillerme, "iProprio: A smartphone-based system to measure and improve proprioceptive function," in *Proc. 38th IEEE Int. Conf. Eng. Med. Biol. Soc.*, 2016, pp. 2622–2625.

[3] K. T. Yoshida et al., "Cognitive and physical activities impair perception of smartphone vibrations," *IEEE Trans. Haptics*, vol. 16, no. 4, pp. 672–679, Oct.–Dec. 2023.

[4] R. A. G. Adenekan, A. J. Lowber, B. N. Huerta, A. M. Okamura, K. T. Yoshida, and C. M. Nunez, "Feasibility of smartphone vibrations as a sensory diagnostic tool," in *Proc. Haptics: Sci., Technol., Appl.*, 2022, pp. 337–339.

[5] W. O. Torres, M. E. Abbott, Y. Wang, and H. S. Stuart, "Skin sensitivity assessment using smartphone haptic feedback," *IEEE Open J. Eng. Med. Biol.*, vol. 4, pp. 216–221, 2023.

[6] R. A. G. Adenekan, A. G. Reyes, K. T. Yoshida, S. Kodali, A. M. Okamura, and C. M. Nunez, "A comparative analysis of smartphone and standard tools for touch perception assessment across multiple body sites," *IEEE Trans. Haptics*, vol. 17, no. 4, pp. 970–977, Oct.–Dec. 2024.

[7] R. A. G. Adenekan et al., "Reliability of smartphone-based vibration threshold measurements," in *Proc. IEEE Haptics Symp.*, 2024, pp. 25–32.

[8] K. T. Yoshida, J. X. Kiernan, A. M. Okamura, and C. M. Nunez, "Exploring human response times to combinations of audio, haptic, and visual stimuli from a mobile device," in *Proc. IEEE World Haptics Conf.*, 2023, pp. 121–127.

[9] J. Scott and R. Gray, "A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving," *Hum. Factors*, vol. 50, pp. 264–275, 2008.

[10] A. R. Peon and D. Prattichizzo, "Reaction times to constraint violation in haptics: Comparing vibration, visual and audio stimuli," in *Proc. IEEE World Haptics Conf.*, 2013, pp. 657–661.

[11] L. Scalera, S. Seriani, P. Gallina, M. Di Luca, and A. Gasparetto, "An experimental setup to test dual-joystick directional responses to vibrotactile stimuli," *IEEE Trans. Haptics*, vol. 11, no. 3, pp. 378–387, Jul.–Sep. 2018.

[12] I. Karuei, K. E. MacLean, Z. Foley-Fisher, R. MacKenzie, S. Koch, and M. El-Zohairy, "Detecting vibrations across the body in mobile contexts," in *Proc. ACM SIGCHI Conf. Hum. Factors Comput. Syst.*, 2011, pp. 3267–3276.

[13] A. Lieberoth, "Shallow gamification: Testing psychological effects of framing an activity as a game," *Games Culture*, vol. 10, no. 3, pp. 229–248, 2015.

[14] S. Alaghbari, A. Mitschick, G. Blichmann, M. Voigt, and R. Dachselt, "Achiever or explorer? Gamifying the creation process of training data for machine learning," in *Proc. ACM Mensch und Comput.*, 2020, pp. 173–181.

[15] M. A. Friehs, M. Dechant, S. Vedress, C. Frings, and R. L. Mandryk, "Effective gamification of the stop-signal task: Two controlled laboratory experiments," *JMIR Serious Games*, vol. 8, no. 3, 2020, Art. no. e17810.

[16] D. Bragg, N. Caselli, J. W. Gallagher, M. Goldberg, C. J. Oka, and W. Thies, "ASL sea battle: Gamifying sign language data collection," in *Proc. 2021 CHI Conf. Hum. Factors Comput. Syst.*, Yokohama Japan, May 2021, pp. 1–13.

[17] D. Elfadaly, S. T. Abdelrazik, P. B. M. Thomas, T. M. Dekker, A. Dahlmann-Noor, and P. R. Jones, "Can psychophysics be fun? Exploring the feasibility of a gamified contrast sensitivity function measure in amblyopic children aged 4–9 years," *Front. Med.*, vol. 7, 2020, Art. no. 469.

[18] V. Vanden Abeele, J. Wouters, P. Ghesquière, A. Goeleven, and L. Geurts, "Game-based assessment of psycho-acoustic thresholds: Not all games are equal!," in *Proc. ACM Symp. Comput.- Hum. Interact. Play*, 2015, pp. 331–341.

[19] U. Söderström, W. Larsson, M. Lundqvist, O. Norberg, M. Andersson, and T. Mejtoft, "Haptic feedback in first person shooter video games," in *Proc. ACM Eur. Conf. Cogn. Ergonom.*, 2022, pp. 1–6.

[20] Sound Effect Warehouse, "Shovel digging - sound effect," 2023. Accessed: Jul. 19, 2024. [Online]. Available: https://www.youtube.com/watch?v=nImg4gwtxqM

[21] DJ Sound FX, "Walking on grass-sound effect," 2021. Accessed: Jul. 19, 2024. [Online]. Available: https://www.youtube.com/watch?v=4eVmKIfZFpY

[22] Interhaptics A Razer Company, "Version v1.2.1," 2024. Accessed: Jul. 19, 2024. [Online]. Available: https://www.interhaptics.com/download/

[23] Ethraiel, "Blip," 2016. Accessed: Jul. 19, 2024. [Online]. Available: https://freesound.org/people/Ethraiel/sounds/351569/

[24] Inc Apple., "Core Haptics Developer Documentation," Accessed: Jul. 19, 2024. [Online]. Available: developer.apple.com/documentation/corehaptics

[25] A. O'Camb, C. Miesch, and A. Diaz-Kokaisl, "iPhone SE 2020 taptic engine replacement," 2024. Accessed: Jul. 19, 2024. [Online]. Available: https://www.ifixit.com/Guide/iPhoneSE2020TapticEngineReplacement/133892

[26] A. Berger and M. Kiefer, "Comparison of different response time outlier exclusion methods: A simulation study," *Front. Psychol.*, vol. 12, 2021, Art. no. 675558.

[27] C. C. Aggarwal, *Outlier Analysis*, 2nd ed. Cham, Switzerland: Springer, 2017.

[28] R. A. Bjørklund, "Reaction time and movement time measured in a key-press and a key-release condition," *Perceptual Motor Skills*, vol. 72, no. 2, pp. 663–673, Apr. 1991.

[29] T. Darbutas, V. Juodžbalienė, A. Skurvydas, and A. Kriščiūnas, "Dependence of reaction time and movement speed on task complexity and age," *Medicina (Kaunas)*, vol. 49, no. 1, pp. 18–22, 2013.