Resonant Sticker Buttons: The Effect of Button Size and Feedback Latency on Perceived Button Weight and Vibrotactile Feedback Strength

Marco Kurzweg
Berlin University of Applied Sciences and Technology
Berlin, Germany
Marco.Kurzweg@bht-berlin.de

Jan Willms
Berlin University of Applied Sciences and Technology
Berlin, Germany
Jan.Willms@bht-berlin.de

Adrien Chaffangeon Caillet
Berlin University of Applied Sciences and Technology
Berlin, Germany
chaffangeon.adrien@gmail.com

Katrin Wolf
Berlin University of Applied Sciences and Technology
Berlin, Germany
katrin.wolf@bht-berlin.de

ABSTRACT

We use simple stickers on a wooden board to create ubiquitous touch interaction. Resonant frequencies make the board vibrate, creating tactile feedback when touching the stickers. In an experiment, we used stickers of three sizes to create Resonant Sticker Buttons and varied the delay of the feedback. Both size and feedback delay influenced the perceived weight of the buttons. While higher latencies result in a heavier perceived button, larger button sizes result in lighter perceived buttons and perceived feedback strength, and vice versa. Our findings suggest that touch interfaces with buttons of varying sizes, weights, and vibration strengths can be created on everyday surfaces, such as tables, by simply using stickers and speakers.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; **Empirical studies in HCI**; **Interaction design**.

ACM Reference Format:

Marco Kurzweg, Jan Willms, Adrien Chaffangeon Caillet, and Katrin Wolf. 2024. Resonant Sticker Buttons: The Effect of Button Size and Feedback Latency on Perceived Button Weight and Vibrotactile Feedback Strength. In *MUM '24: 23rd International Conference on Mobile and Ubiquitous Multimedia, December 1-4, 2024, Stockholm, Sweden.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3701571.3701591

1 INTRODUCTION

The availability of embedded sensors and actuators allows for integrating them into everyday surfaces and paves the way for the ubiquitous access of computers [42]. Touch is the dominant way we interact with interactive surfaces, and haptic feedback has been shown to improve the usability of touch interaction [5]. While augmented reality (AR) technologies, such as head-mounted displays, can support the interaction with everyday surfaces by visual and auditory feedback, the question of how to integrate haptics in ubiquitous computing and AR remains unclear.

As it is not practical to embed haptic actuators into everyday surfaces, it has been shown that using sound, in particular resonance frequencies, can bring a surface in vibration and, thus, provide it with vibrotactile touch feedback [22, 44, 48]. Using a simple speaker, commonly found in most households [24], can bring a surface vibrating in resonance frequencies. However, although design guidelines for vibrotactile feedback are well-known for devices, such as tablets, smartphones, and controllers, designing haptic feedback for vibrating everyday surfaces has not been explored in detail yet [30]. Simply adapting existing design guidelines is moreover often not possible [26, 39]. For example, specific frame rates or system latencies of resonance frequency-induced feedback have to be considered [21].

Thus, resonance frequency-induced haptic feedback, a practical solution for everyday surface interaction, needs further investigation. Making everyday surfaces vibrate at resonant frequencies when widgets (like our Resonant Sticker Buttons) placed on them are touched can create and enrich haptic feedback for interactions with these surfaces. This haptic feedback could be used when equipping everyday objects with additional actuators is impossible. Thus, Resonant Sticker Buttons could, in theory, be stuck on tables and surfaces, making them a ubiquitous user interface. Further, it could be combined with AR or VR applications to create a wide range of possible applications with an immersive user experience. Thus, Resonant Sticker Buttons can augment everyday surfaces with virtual keyboards and buttons using, e.g., AR glasses, which provide vibrotactile feedback for virtual input elements of different weights.

The buttons and keys of an interface usually differ in size. For example, the <enter> key is usually larger than the keys of digits and letters. Moreover, keys and buttons can differ in weight, which refers to the force needed to press a key or button. While this weight is a matter of material and hardware choice for physical buttons, digital buttons' weight can be modified through the delay of vibrotactile feedback induced by touch [16]. Furthermore, the size of an object influences its perceived weight [10, 27].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
MUM '24, December 1-4, 2024, Stockholm, Sweden
© 2024 Association for Computing Machinery.
ACM ISBN 979-8-4007-1283-8/24/12…$15.00
https://doi.org/10.1145/3701571.3701591

To better understand how Resonant Sticker Buttons can turn an everyday surface into a usable touch interface, we explored the impact of latency and size of Resonant Sticker Buttons. Thus, we run an experiment by placing stickers of three sizes that simulate buttons on a wooden board, which is brought into vibration using resonance frequency in seven latencies. Our study showed that latencies and button sizes influence the perceived button weight when being touched. The larger the button, the lower the perceived weight, and vice versa. In addition, we identified changes in weight perception induced by different latencies. 38 ms and 58 ms were perceived as significantly lighter, 118 ms, 138 ms, and 158 ms as significantly heavier than other latencies. Our results demonstrate how vibrotactile buttons of different weights and feedback strengths can be created using stickers and a speaker. This contributes to a richer and better usable ubiquitous calm computer interaction [42, 43] while keeping the setup simple.

2 RELATED WORK

This paper investigates whether a button's weight perception is influenced by a feedback delay and/or the button's size using buttons that are stuck onto everyday surfaces and brought into vibration using resonance frequencies. Therefore, we have reviewed (1) touch interactions on everywhere displays and (2) weight illusions in general, as there is little research on weight illusions for buttons.

2.1 Everywhere Touch Displays and Haptic Feedback

In the 1990s, Weiser envisioned the ideas of ubiquitous and calm computing imagining computers would be everywhere around us but interwoven into the fabric of everyday life objects and are indistinguishable from it [42, 43].

Concerning the vision of ubiquitous computing, Pinhanez proposed the concept of everywhere displays where surfaces can be transformed into a graphical interface [30]. Besides their advantages and profound use cases, designing haptic feedback for such displays was stated to be challenging and underexplored [30]. Examples of investigations where everyday life objects (furniture) were used as displays or interfaces are IllumiRoom and InForm [9, 14]. InForm is a table with actuated pins to physically display 3D information. While InForm's approach created promising feedback possibilities, it created much noise and did not follow the calm computing concept regarding Weiser. IllumiRoom uses a projector to augment a TV, displaying images on the wall and everyday objects around the TV. Thus, everyday objects can be visually animated to create the illusion of vibrations. However, IllumiRoom does not provide haptic feedback, as these objects are out of the user's reach.

Haptic feedback for hands-free interactions using AR or VR technology has been explored. Lopes et al. investigated if electrotactile stimuli can induce feedback in users' hands when interacting with AR interfaces placed on everyday objects [25]. Speicher et al. investigated pseudo-haptics to provide feedback for hands-free mid-air interactions on VR user interfaces [36]. While both approaches explored promising ways to provide feedback for everywhere displays, they did not research the possible influence of varying parameters like feedback latency or object size on object or widget perception, which is our research scope.

Ishii et al. explored surfaces that could change their shape and, by this, create different haptic sensations directly through the everyday object [13]. While they found such shape chain interfaces to be promising candidates for future interactions, they did not investigate the creation of haptic feedback on it despite the detection of perceiving different shapes. Besides, there are also works using resonant frequency to create haptic sensations on everyday objects [44, 48]. Piezoelectric actuators were used to create a working abstraction of a haptic feedback system inducing the stiffness of an object [48]. The piezoelectric actuator yielded better efficiency if it had a resonant frequency. Further, the resonant frequency was used for prototyping a haptic feedback assistive device for visually-impaired drivers [44]. A haptic device was built with pins of different diameters, lengths, and resonant frequencies ordered in a square layout. The pins were connected to the voice coil of a speaker. Slight differences in the pins' structure created a high-resolution haptic display.

A similar approach was used by Kurzweg et al. [21]. They explored how vibrotactile feedback can be provided on touch interactions with everyday surfaces when neither the environment nor the users can be equipped with additional hardware like actuators. As a solution, they created haptic illusions using only auditory and visual stimuli. Haptic illusions were also used to influence the perception of object properties like stiffness on rigid surfaces [46]. Therefore, different textures were projected onto a rigid surface and deformed visually. In addition, electrotactile stimuli were presented at the index finger's fingertip.

There is also research about the assignment of haptic feedback for touch interactions to objects stuck on everyday surfaces [22]. Different parameters that might influence the assignment were investigated, like the duration of the feedback, sound, and latency. It was shown that latency does not affect the feedback assignment. Nevertheless, it was not investigated if the latency impacted the perceived weight or resistance of the touched graphical object.

All these works investigated haptic feedback when interacting with everyday surfaces. However, no previous work explored the possibility of changing the perceived weight of touched widgets or objects relying on weight illusion.

While other important haptic sensations can influence the perception of a button, like its texture [2, 35], in this work, we investigate weight property.

2.2 Weight Illusion

Weight illusions have become a common way to create the perception of weight easily or quickly change weight perceptions when interacting with objects for different technologies, without the need to change the object or its weight. A physical object's weight is mostly determined by its density and volume and, thus, dependent on its material and/or size. When creating weight illusions, one of these parameters is mostly manipulated.

One of the most established methods to influence weight sensations is using the C/D ratio manipulation [23]. C/D ratio means that there is a change in the relationship between users' controls and what they receive as visual feedback on a display, i.e., a discrepancy between the real and virtual movement gains of an interactor. Especially in virtual reality (VR) applications, it can easily be used to make an object appear heavier or lighter than it is [32, 33]. If users lift a physical cube in real life while lifting a virtual representation of that cube in VR, the perceived weight of the physical cube can be influenced by the virtual cube's elevation. Users perceive the cube as heavier, the lower the virtual cube was lifted compared to the real one [33]. Such effects were also found for pseudo-haptics, and extended by findings stating that the position and the range in which the object was moved influenced the perception of the weight [18]. Such a method is normally used on VR systems, as this is the ideal technology to create a discrepancy between real hand and virtual hand movement. Furthermore, weight illusions have been applied to other technologies, like robotic teleoperation [31] and tablets [41].

The size-weight illusion is another common approach to manipulating weight perception. It is based on modifying the visual size of the handled object. Thus, the object's weight is perceived as heavier if the visual representation is smaller and vice versa. This illusion was investigated in AR [10], VR [27], and human-robot interaction [34]. Maehigashi et al. additionally explored the size-weight illusion with brightness-weight and material-weight illusions[27]. They found that, in contrast to the results obtained in real life, darker and heavier-looking materials were considered heavier in VR.

Another possible method to influence the perceived weight visually is to superimpose a movable object inside a physical prop, like a rolling ball or a liquid [19, 47]. Keller et al. showed that virtual objects manipulated on a tabletop interaction device could be augmented to provide an illusionary weight [19]. This was done by measuring the pressure applied with the fingers manipulating the superimposed object. Yamada et al. investigated the influence of the physical characteristics of virtual collision sound, such as the size and weight of the movable object. It was found that the weight perception changes according to the virtual collision sound [47].

The concept we rely on in this work is to alter the latency between a touch interaction, like pressing a button, and the resulting feedback that can be perceived at the fingertip. Kaaresoja et al. explored that a higher latency results in the perception that a button felt heavier when pressed on a tablet [16]. As most phenomena mentioned require special technologies, we can technically adapt this approach to everywhere surfaces stimulated with a resonant frequency. While Kaaresoja et al. always investigated the same button sizes, most interfaces or input technologies, have buttons of different sizes. Thus, we vary that parameter in the work presented here.

2.3 Summary

Designing haptic feedback for everywhere displays is challenging and yet underexplored. Previous research started to investigate some haptic sensations, but there are still research gaps that have not yet been explored, like the perceived weight of widgets in touch interactions. If weight was investigated as an object property, mostly only one possible modality to vary the weight was considered. In fact, while the combination of modalities can increase perceived feedback [7, 40], it is not always the case [46]. Further, the results for inducing weight differ depending on the technology used [26]. While research on weight and texture sensations using haptic technologies like gloves has been done [8, 29], our approach focuses on hands-free explorations without reducing the tactile sensation of the fingertips. Therefore, it remains unclear how existing findings can be transferred to everywhere displays and how the weight of a pressed widget is perceived when two phenomena (latency and size) are combined.

3 RESONANT STICKER BUTTON IMPLEMENTATION

In this work, we aim to investigate the impact of different feedback latencies and button sizes on the perceived weight of a button and, thus, its better suitability for everyday surfaces and calm computing.

Three thin wooden boards represented an everyday wooden surface, such as those found on tables, shelves, or cupboards. These wooden boards were identical in thickness (1 mm), size (26 cm x 53.5 cm), shape (rectangular), and material (chipboards). The wooden boards were placed over a loudspeaker to vibrate at a resonant frequency, producing perceptible vibrotactile feedback.

Resonant Frequency

Resonance frequencies are the frequencies at which an object, in this case, the wooden boards, vibrates at its maximum amplitude when stimulated by an external vibration, like sound. The boards' resonance frequency depends on their size, thickness, and material. The resonance frequency increases as the board thickness increases, requiring more energy to excite it. More energy means a higher frequency and a louder sound, which could be disruptive. Achieving high mobility (a stronger expression of vibration) without requiring high energy is desirable because it creates a pleasant haptic sensation and aligns with the concept of calm computing [43]. Therefore, the thinnest available board was chosen.

To determine the resonant frequency of the wooden boards, we struck them with a hammer and measured the impulse response using a piezoelectric sensor, as described in previous research [22]. The boards were stimulated at the same determined resonant frequency for all latencies in the experiment, resulting in consistent perceived haptic feedback.

Latencies

The study followed the methodology of Kaaresoja et al., using latencies in 20 *ms* intervals, beginning at 38*ms* up to a maximum latency of 158 *ms* [16]. We could not start at a latency of 18 *ms*, as in the work of Kaaresoja et al., due to the system's latency. In fact, we measured a latency of 35 *ms* for the round-trip latency of the analog-digital and digital-analog conversion of the computer, along with the additional external audio interface and connected Arduino. Without the external audio interface, we measured an even higher system's latency of (98 *ms*). Thus, we had to begin at the next possible latency value, which was 38 *ms* [16]. This limitation emphasizes that phenomena cannot be easily applied to other applications, technologies, or systems. To set up the desired latencies for the study, we used a synthesizer program to adjust the system latency by adding the differences between the latencies of prior work and our system latency.

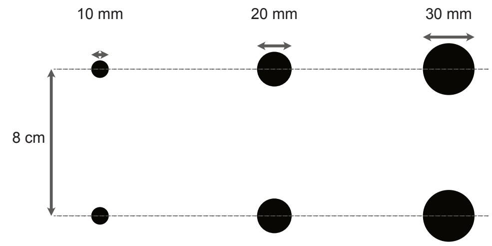

Figure 1: Three different button size representations were used in the study with 10 mm, 20 mm, and 30 mm diameters from left to right. The distance between the center points was 8 cm for all button sizes.

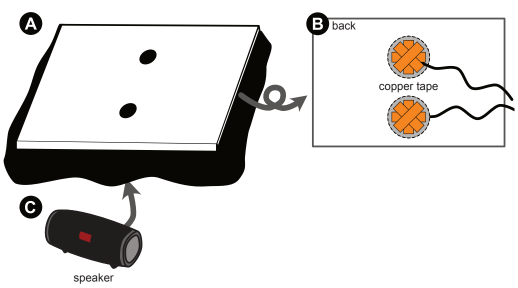

Figure 2: Overview of the prototype. A) wooden board with two stickers. B) The back of the board is covered with copper tape in the same position as the buttons. C) Speaker hidden underneath the table.

Buttons Sizes

The latencies were used to determine the perceivable haptic feedback after touching the stickers representing buttons on the wooden boards across all conditions. The buttons were kept as neutral as possible and, therefore, are round and colored solid black. The sizes of the buttons were based on the dimensions of computer keyboard keys, which have an average size of 1.8 to 1.9*cm*, which we rounded up to the next higher natural number of 2 *cm*. We then adjusted the size by decreasing and increasing it by 1 *cm*, resulting in button sizes of 1 *cm*, 2 *cm*, and 3 *cm* diameter, see Figure 1. This linear increment was chosen due to Steven's power law and is also perceived as linear in visual feedback [37]. This is important to consider when examining the possible size-weight illusion [23, 34]. The center point of all three button sizes was placed at the same spot but on a separate wooden board. This ensures consistent haptic feedback across all boards since they are identical and have the same measured resonant frequency The distance between the two center points was 8 *cm*.

Feedback Creation

The resonant frequency stimulates the entire board when played, allowing the vibration to be felt wherever the board is touched. To provide haptic feedback only when the button stickers are pressed, we utilized an Arduino and the additional 'CapacitiveSensor' library[1]. The capacitive sensors were placed on the back of the board at the same position as the buttons. The sensors were made of a copper cable affixed with copper tape. This approach ensures the natural surface appearance and haptic feedback [22, 30]. When the Arduino detects a touch, it sends a signal to the computer, which activates the speaker and plays the resonant frequency. Using a speaker makes the approach more scalable as it can stimulate the wooden board or other surfaces from any position. This reduces the number of additional actuators needed and keeps the environment "calm" [43]. While the main idea is to play the resonant frequency using infrasound so it can't be heard, this is impossible with commonly available speakers. Thus, we positioned the speaker 2*cm* directly beneath the wooden boards to minimize the volume needed when playing the audible resonant frequency. Higher volumes could have been disruptive during the study, and we wanted participants to fully concentrate on the perceived feedback without any distractions. To achieve this, we additionally utilized noise-canceling headphones and played white noise.

4 METHOD

We aimed to explore whether different latencies and different button sizes influence the perceived weight of a touched button placed on an everyday life surface. Therefore, we explored the research question: Does a feedback delay and/or the size of a button on an everyday surface influence the perception of the button feedback when being touched?

4.1 Experiment Design

As our goal was to explore if different variables can influence the perceived weight of a button, we tested different latencies, between the button is being touched and the perceived feedback, and different button sizes within our study.

Independent Variables. We designed a controlled experiment with a 7x3 within-subjects design. The independent variables are DELAY (38 *ms*, 58 *ms*, 78 *ms*, 98 *ms*, 118 *ms*, 138 *ms*, 158 *ms*) and BUTTONSIZE (1 *cm*, 2 *cm*, 3 *cm*). Using Gauss summations, the seven delays led to 28 conditions [3]. These conditions were counterbalanced using a Latin square design [4]. In addition, all 28 conditions were presented on three boards with different button sizes but in another order. These three boards were also counterbalanced using another Latin square, like it was done in other research [45].

Dependent Variables. In our experiment, we used the same dependent variables as in the work of Kaaresoja et al. [16]. The dependent variables are PERCEIVEDWEIGHT (to figure out if a button is perceived lighter or heavier compared to another one), and WEIGHTASSIGNMENT (exploring which button (latency) is perceived as the lightest and which button is perceived as the heaviest). Further, we investigated PLEASENTNESS (as a measurement for the pleasantness of the interaction and the perceived feedback), COMFORT (aiming to explore the comfort level of the interaction and the perceived feedback), REALISM (to figure out how realistic the feedback was perceived), and CONFIDENCE (exploring the confidence of the participants in their answers).

[1]https://www.arduino.cc/reference/en/libraries/capacitivesensor/

In addition, we had the dependent variable QUALITATIVEFEEDBACK (to possibly better understand our quantitative and qualitative results).

4.2 Measurements

We utilized a magnitude estimation method to measure the PERCEIVEDWEIGHT. Therefore, two buttons were provided to the participants in a vertical alignment on the wooden boards. The upper button served as a reference, while the lower button was to be judged in relation to the reference. In each round, the latency of the reference button was altered, and each latency was compared to the others [16]. For the weight judgment, the participants were instructed to use numbers between 0 and 100 for their answers.

For measuring the WEIGHTASSIGNMENT, the participants had to touch all seven different latency buttons again, and afterward, they had to judge which of the seven was perceived as lightest and as heaviest [16]. The participants first had to determine the lightest or the heaviest, and before determining the opposite one, they touched all seven buttons again. This was due to the fact the participants were able to focus on one characteristic at once, which would ease the process.

For further exploration of the perceived weight of the button, we asked the participants for the PLEASENTNESS, COMFORT, REALISM, and CONFIDENCE of the felt haptic feedback. Therefore, we used the same single-item questions, as done by the inspirational work [16]. Following the aforementioned work, this resulted in the following questions:

(1) How pleasant was the interaction with the button?
(2) How comfortable was the interaction with the button?
(3) How realistic was the interaction with the button?
(4) How confident are you in the interaction with the button?

All answers had to be answered on a 7-items Likert scale (1 - *not at all*, 7 - *totally*.

Afterward, to investigate the QUALITATIVEFEEDBACK, we asked following question:

(1) Do you have any feedback regarding the interaction with the different buttons and the perceived weight?

4.3 Participants

We recruited 28 participants (6 who identified themselves as female, 22 who identified themselves as male) aged 21 to 38 years with an average of 26.07 years (SD = 3.59). The participants were recruited via mailing lists from different scientific areas and professions.

4.4 Apparatus

We used a Laptop, i7 10750H/2,6 GHz, 16 GB RAM, RTX 3070, and an external Behringer UMC22 sound card for our apparatus. Via the audio software Waveform, we played the resonant frequency of the wooden board, created latencies, and generated the white noise that was used for masking the sound of the resonant frequency.

Noise-canceling headphones (Sony WH-1000XM4) covered frequencies not masked by the white noise. Both methods, either used alone or in combination, are commonly employed to mask sounds and frequencies [12, 17, 20, 28]. We applied both methods to ensure that all external sounds were masked. The wooden boards were 1 mm thick, providing adequate vibrations even with lower-volume sounds. All boards were identical and produced the same haptic feedback at the same locations. This was important to be able to compare the different BUTTONSIZES. To ensure that the vibrational behavior of each board remained unaffected, they were placed on four small wooden cubes with felt pieces on the bottom. The speaker was hidden underneath the two cardboard boxes on which the boards were placed. On top of each board, two black round-shaped stickers were placed in the middle of the wooden boards in a vertical line at the points where the participants had to touch the board to receive the haptic feedback, see Figure 3. These locations were consistent on all three wooden boards and related to the center of the different-sized buttons. The capacitive touch contact is recognized only when the buttons are touched and not at other parts of the wooden boards when adjusting the size of the conductive area. The speaker (JBL Charge 4) was directly placed under the wooden board with a distance of $3cm$ and could not be seen because, in front, it was covered by black fabric. While the white noise and the sound of the fly were played via the headphones, the speaker was used to excite the board with its resonant frequency.

Figure 3: Participant touching the bottom button during the study.

4.5 Task & Procedure

The experiment was conducted as a lab study. First, the participants were introduced to the study's purpose and then asked to agree to a consent form. They were informed that participation in the study was voluntary and that taking a break was possible. After filling in a demographic questionnaire, participants started with the study.

In each condition, the participants wore noise-canceling headphones and touched the graphical buttons with their right index finger, which let them perceive feedback according to the experimental conditions. In each round, the participants started with the weight assignment. For the WEIGHTASSIGNMENT, the participants could press the buttons as often as they wanted. No time limitation was given. After the weight assignment for all 28 comparisons, the participants judged which button was lightest and which was heaviest. The single-item questions were answered in the following, followed by the qualitative feedback. After exploring and answering all conditions and questionnaires for one button size, the next button size was presented to the participants using another board. We had prepared three boards, one per button size. The procedure was the same for all three different button sizes, resulting in 84 trials in total for the WEIGHTASSIGNMENT.

Overall, the participation in the study lasted 45 minutes. The participants were informed that participation was voluntary and that the study could be interrupted at any time.

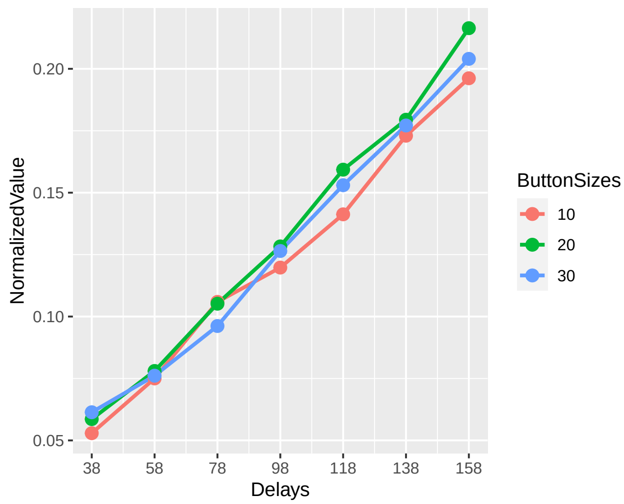

Figure 4: Overview of all three buttons and their estimated perceived weight for each latency. The perceived weight values had to be normalized as they were gained by magnitude estimation, and all participants judged the weight in relation to a reference value that could be chosen freely.

5 RESULTS

We first analyzed our quantitative data to learn how the different DELAY (38 ms, 58 ms, 78 ms, 98 ms, 118 ms, 138 ms, 158 ms) and BUTTONSIZE (1 cm, 2 cm, 3 cm) conditions affected the PERCEIVEDWEIGHT, WEIGHTASSIGNMENT, PLEASENTNESS, COMFORT, REALISM, and CONFIDENCE of the feedback. Subsequently, we analyzed the qualitative data to better understand our quantitative results.

5.1 Quantitative Results

For analyzing the quantitative results of the PERCEIVEDWEIGHT, we used an ANOVA to identify significant differences. The WEIGHTASSIGNMENT was analyzed with two different statistical tests. The analysis of the WEIGHTASSIGNMENT results for the LATENCY was made using a Cochran Q-test and a post-hoc McNemar test to compare which latency was chosen more often as lightest or heaviest and which was not across all buttons. This will end in a frequency distribution with a binary outcome (*Yes/No*). The different button sizes were analyzed using the Friedman test to identify significant differences, as we here investigate the distribution of the different latencies to the individual buttons. A post-hoc analysis was conducted using Wilcoxon signed-rank tests. The quantitative results of PLEASANTNESS, COMFORT, REALISM, and CONFIDENCE were analyzed using Friedman tests to identify significant differences. A post-hoc analysis with Wilcoxon Signed-Rank tests was conducted with a Bonferroni correction applied for the p-value, resulting in a significance level of 0.0024 for the latencies and 0.016 for the different button sizes.

5.1.1 *Perceived Weight.*

We first had to normalize the results of the PERCEIVEDWEIGHT to make them comparable and be able to perform a statistical analysis. For the normalized estimated weights for DELAY, the results indicated a significant positive correlation, suggesting that higher delays lead to heavier perceived buttons ($r(2350) = .468$, $p < .001$). For the BUTTONSIZE, the results showed no significant correlation ($r(2350) = .023$, $p = .270$). An ANOVA revealed statistical significant differences for the variable PERCEIVEDWEIGHT regarding DELAY ($F_{(6,\,2331)} = 110.011$, $p < .001$) and BUTTONSIZE ($F_{(2,\,2331)} = 3.575$, $p = .028$); see Figure 4. Further, no interaction effects could be found.

Post-hoc pairwise comparison (performed with Tukey-test) revealed significant differences between all compared delays, with the higher value always perceived with a significantly higher weight except between 38 ms and 58 ms; see Appendix Table 1. In addition, the post-hoc comparison indicated a significantly higher PERCEIVEDWEIGHT for the 20 mm button compared to the 10 mm button (p = .020); see Appendix Table 2.

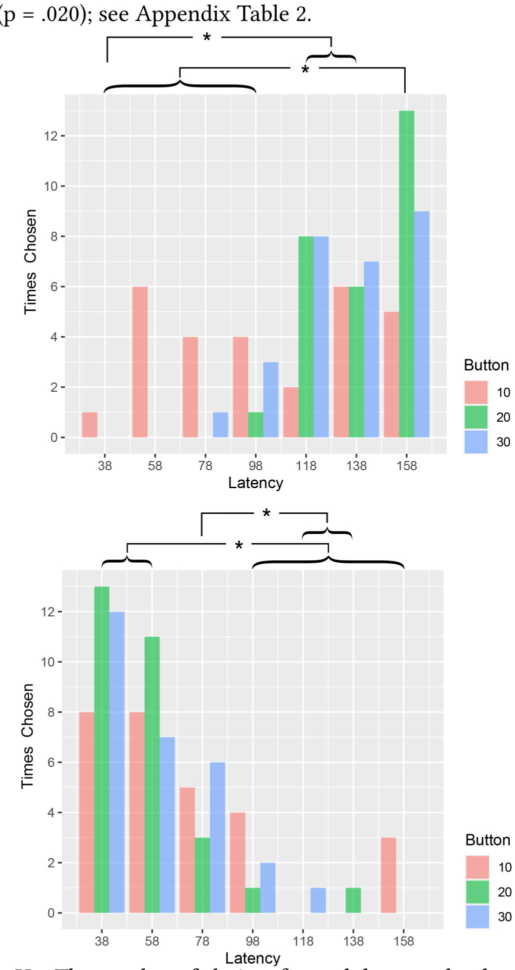

Fig. 5: Up: The number of choices for each latency that has been chosen to be perceived as the heaviest one. Down: The number of choices for each latency that has been chosen to be perceived as the lightest one. Both sides are also showing the choices split into each single button size.

5.1.2 *Weight Assignment.*

A Cochran Q-test indicated significant differences between the seven latencies, regarding which latency was perceived as the heaviest ($\chi^2(6) = 44.3$, $p < .0001$) and which was perceived as the lightest ($\chi^2(6) = 83.7$, $p < .0001$). Post-hoc comparisons using a pairwise McNemar test revealed that the latency *158 ms* was perceived as statistically significantly heavier compared to the latencies *38 ms*, *58 ms*, *78 ms*, and *98 ms*. Further, the latencies *118 ms* and *138 ms* were perceived as statistically significantly heavier than the latency *38 ms*, see Appendix Table 3 and Figure 6, left. Regarding the assignment of which latency was perceived as the lightest, post-hoc comparisons using a pairwise McNemar test revealed that the latencies *38 ms* and *58 ms* were perceived as statistically significantly lighter compared to the latencies *98 ms*, *118 ms*, *138 ms*, and *158 ms*. Additionally, the latency *78 ms* was perceived as statistically significantly lighter than the latencies *118 ms* and *138 ms*, see Appendix Table 4 and Figure 6, right.

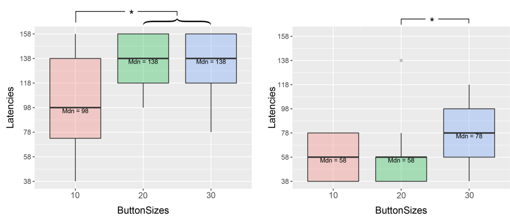

Fig. 6 Representation of which latencies were mainly perceived as the heaviest (left) and as the lightest (right) for the different button sizes.

Friedman tests revealed significant differences in the button sizes for the perception of the heaviest button ($\chi^2(2) = 13.31$, $p = 0.001$) as well as for the perception of the lightest button ($\chi^2(2) = 14.44$, $p < .001$).

Post-hoc comparisons using Wilcoxon Signed-Rank tests indicated that statistically significantly another range of latencies was perceived as heavier for the button with *10 mm* diameter compared to the buttons with *20 mm* ($p < .001$) and *30 mm* ($p = 0.0034$) diameter, see Figure 6, left and Appendix Table 5. Further, post-hoc comparisons using Wilcoxon Signed-Rank tests indicated that statistically significantly another range of latencies was perceived as lighter for the button with *30 mm* diameter compared to the button with *20 mm* ($p = 0.0028$) diameter, see Figure 6, right and Appendix Table 6.

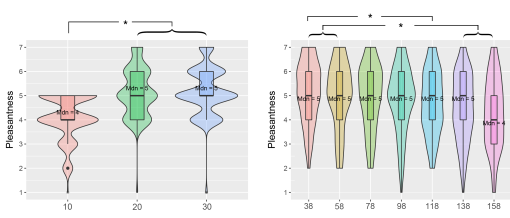

Fig. 7 Box plot representing the results for the question "How pleasant was the interaction with the button?". **Top:** The results are separated by the different button sizes. **Bottom:** The results are separated by the different latencies.

5.1.3 Pleasantness. For pleasantness, a Friedman test revealed a statistically significant difference for the different BUTTONSIZES ($\chi^2(2) = 132.67$, $p < .001$) and for the different LATENCIES ($\chi^2(6) = 40.103$, $p < .001$).

Post-hoc comparisons using Wilcoxon Signed-Rank tests indicated statistically significantly higher pleasantness for the buttons with *20 mm* and *30 mm* diameter compared to the button with *10 mm* diameter, see Figure 7 & **??** left and Appendix Table 7. Further, the post-hoc tests showed a statistically significant higher pleasantness for the latencies *38 mm* and *58 mm* compared to the latencies *138 mm* and *158 mm.* In addition, *38 mm* was perceived with a statistically significant higher pleasantness than *118 mm*, see Figure 7 right and Appendix Table 8.

5.1.4 Comfort. For the perceived comfort when interacting with the buttons, a Friedman test revealed a statistically significant difference for the different BUTTONSIZES ($\chi^2(2) = 52.631$, $p < .001$) and for the different LATENCIES ($\chi^2(6) = 48.636$, $p < .001$).

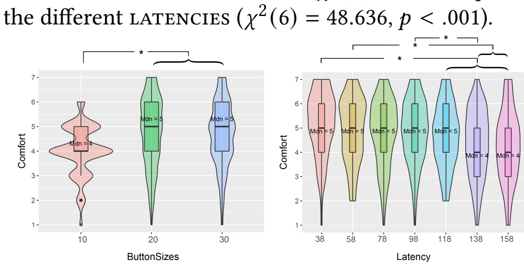

Fig. 8 Box plot representing the results for the question "How comfortable was the interaction with the button?". **Left:** The results are separated by the different button sizes. **Right:** The results are separated by the different latencies.

Post-hoc comparisons using Wilcoxon Signed-Rank tests indicated statistically significantly higher comfort for the buttons with *20 mm* and *30 mm* diameter compared to the button with *10 mm* diameter, see Figure 8 left and Appendix Table 9. Additionally, the post-hoc tests showed a statistically significant higher comfort for the latencies *38 mm*, *58 mm*, and *98 mm* compared to the latency *138 mm.* Further, *38 mm* and *58 mm* were perceived with a statistically significant higher comfort than *158 mm* and *38 mm* compared to *118 mm*, see 8 right and Appendix Table 10.

5.1.5 Realism.

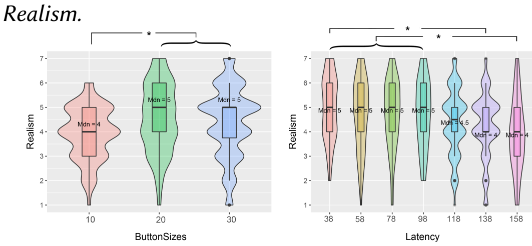

Fig. 9 Box plot representing the results for the question "How realistic was the interaction with the button?". **Left:** The results are separated by the different button sizes. **Right:** The results are separated by the different latencies.

A Friedman test revealed a statistically significant difference for the different BUTTONSIZES ($\chi^2(2) = 29.891$, $p < .001$) and for the different LATENCIES ($\chi^2(6) = 51.213$, $p < .001$), regarding the perceived realism when interacting with the buttons.

Post-hoc comparisons using Wilcoxon Signed-Rank tests showed a statistically significantly higher realism for the buttons *20 mm* and *30 mm* compared to the button *10 mm*, see Figure 9 left and Appendix Table 11. Further, the post-hoc tests showed a statistically significant higher realism for the latencies *38 mm*, *58 mm*, *78* *mm*, and *98 mm* compared to the latency *158 mm* and for *38 mm* compared to *138 mm*, see Figure 9 right and Appendix Table 12.

5.1.6 Confidence. For the confidence in the interaction with the buttons, a Friedman test revealed a statistically significant difference for the different BUTTONSIZES ($\chi^2(2) = 24.104$, $p < .001$) and for the different LATENCIES ($\chi^2(6) = 53.469$, $p < .001$).

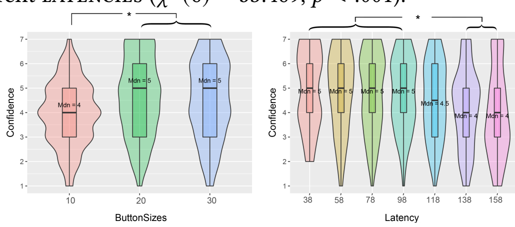

Fig. 10 Box plot representing the results for the question "How confident are you in the interaction with the button?". **Left:** The results are separated by the different button sizes. **Right:** The results are separated by the different latencies.

Post-hoc comparisons using Wilcoxon Signed-Rank tests showed significantly higher confidence in the buttons *20 mm* and *30 mm* compared to the button *10 mm*, see Figure 10 left and Appendix Table 13. For the different latencies, the post-hoc tests showed significant higher confidence in the buttons with *38 mm*, *58 mm*, *78 mm*, and *98 mm* latency compared to the buttons with the latencies *138 mm* and *158 mm*, see Figure 10 right and Appendix Table 14.

5.2 Qualitative Results

The qualitative data were coded using Grounded Theory [38]. Axial and selective coding was applied, building categories according to the questions that asked for positive or negative perceived aspects of a system [38]. Two researchers did the coding independently of each other and discussed their results afterward to develop common codes. That procedure aims to gain explanations for our quantitative findings through qualitative analysis. The qualitative results are separated into answers related to the different button sizes and the different latencies. The question was answered at the end of each button size. We gained a total of 110 answers that were split into 82 answers related to the button sizes and 28 answers related to the latencies.

5.2.1 Button Sizes. Of the 82 answers given concerning the different button sizes, 31 were given for the button with a 10 mm diameter. In these 31 answers, the participants mentioned that the 10 mm button was the most uncomfortable of all three buttons (15 out of 31 answers), was perceived as heavier in comparison to the other two buttons (9 out of 31 answers), and the perceived feedback felt stronger in comparison to the other two buttons (7 out of 31 answers):

- The buttons felt very similar. I found that the buttons were harder to operate and felt heavier (Participant 7)
- Felt heavier than the other buttons — haptic feedback felt stronger than with the other buttons (Participant 27)

Further, we gained 23 answers regarding the 30 mm button size. In contrast to the 10 mm button, the 30 mm button was perceived as pleasant (11 out of 23 answers). The 30 mm was perceived as the lightest of all buttons (6 out of 23 answers), and the feedback of the button was rated as the weakest of all three buttons (6 out of 23 answers):

- The surface area of the buttons was larger, but the feedback felt weaker and the buttons lighter (Participant 11)
- I generally felt the buttons had weaker feedback (Participant 3)

Finally, 28 of the 82 responses on button sizes referred to the button size of 20 mm. Also, here, the participants mentioned three different topics regarding the feedback on the 20 mm button. 14 out of the 28 answered that the perceived feedback was very comfortable, and additionally, the feedback was well perceivable and not too strong or weak (9 out of 28 answers). Further, the participants mentioned that the 20 mm button was the best out of all three tested buttons (5 out of 28 answers).

- With a few exceptions, the feedback was good and felt pleasant and overall the best (Participant 4)
- I found it much easier to identify clear differences between the two buttons presented at the same time, and I would say that, on average, I have much more confidence in the buttons (Participant 14)

5.2.2 Latencies. Regarding the different LATENCIES, we gained a total of 28 answers. In 20 out of the 28 answers, the participants mentioned, across all button sizes, most of the time, it was hard to differentiate between the latencies:

- The delays were often very difficult to distinguish from one another (Participant 1, 10 mm)
- It was difficult to recognize different vibrations between the individual sequences (Participant 10, 30 mm)

Related to the fact that most latencies are hard to differentiate, in 8 out of 28 answers, the participants mentioned that the different latencies felt more like three different weight areas that could be differentiated. This was only the case for the buttons with 20 mm and 30 mm diameter:

- Relaxed operation difficult to recognize which should be stronger. Rather, three ranges and, therefore, rare differences (Participant 16, 30 mm)
- Felt like a real button Only big difference in latency was really noticeable; it felt like about three different levels (participant 17, 20 mm)

6 DISCUSSION

Aiming to explore if we can influence the perceived weight of a graphical button placed on an everyday-life object, we investigated different LATENCIES and BUTTONSIZES. In addition, we aimed to figure out possible differences or similarities in the influence of latencies compared to previous research on a tablet [16]. Our results indicate that both the delay of the perceptible feedback and the different button sizes can influence the perceived weight of the touched button. The results also suggest as an additional insight that the perceived weight and the perceived strength of the feedback seem to depend on the same factors. We lastly regard the limitations of our design and implementation and how these might be addressed by future research.

6.1 Latencies

Just like on tablets [16], we found different latencies being able to induce different weight percepts for graphical buttons on an everyday surface when being touched. Even if we could not investigate latencies under 38 ms because of the setup where we created the feedback through resonant frequency, we still identified three areas of weight perception. While latencies of 38 ms and 58 ms are perceived as significantly lighter compared to others, latencies of 118 ms, 138 ms, and 158 ms are perceived as significantly heavier compared to others; see Section 5. These findings are also consistent with the knowledge from other work and models about when latencies are perceived at all in different interactions, where 50 ms and 100 ms are essential boundaries where the perception of latencies changes [6, 15].

Conversely, the three areas of perceived weight are different for the light perceived latencies compared to previous work [16]. This is not only because we could not provide the same latencies but also because 58 ms is perceived as "light" in our results, which was not the case on the tablet. Further, we found more significant differences for the perceivedWeight compared to the results on the tablet [16]. We found more often significant differences for latencies right next to each other compared to previous work; see Table 1. This was the case for all latencies except for the two lowest latencies that are under or directly at the boundary where latencies can be noticed.

The differences in the results could be due to the interaction with different devices on the one hand and the different latencies presented on the other. Furthermore, we did not examine "conventional" buttons but kept them neutral in appearance. We also extended the design by examining three different sizes, all of which have an influence and can lead to a different overall result than just one button size. However, we were able to reconstruct and extend existing findings and gain insights for interaction designs on everyday surfaces in terms of providing weights for touch interactions.

6.2 Button Sizes

As an extension to the inspirational work [16], we also investigated the potential influence of different button sizes on the perceived weight of the button for touch interactions. Even if it is not much, research exists on the size-weight illusion. When it has been studied, it was mainly for objects being held in the hand [23]. In addition, technologies requiring head-mounted displays (HMDs) mainly were used, such as AR [10] or VR [27]. In our work, we used the size-weight illusion for interactions on everywhere displays. Our results indicated that the size of the touched button can influence the button's perceived weight. The results suggest that larger button sizes are perceived as lighter and vice versa. This can also be seen in the participants' feedback, mentioning the largest button size (30 mm) as the lightest of all touched button sizes and the smallest (10 mm) as the heaviest. Further, our results show that the button size should be large enough not to be covered by the user's finger when touched. Users cannot judge the perceived weight or differentiate between the latencies if this is the case. This was the case for the smallest button (10 mm). This is in line with research on touchscreens investigating, for example, a non-detectable offset of the mouse cursor so that it is still visible on touch interactions. Covering the mouse cursor or widget that the user wants to interact with lowers the user experience [1, 11].

As an additional insight, the qualitative results suggest that the size impacts the button's perceived weight and the vibrational feedback's perceived strength, which, in our work, is provided by resonance frequency. The perception is similar in the way that a larger button size leads to the perception of weaker feedback and vice versa.

6.3 Interplay of Size and Latency

Even with different sizes, the latencies are still perceived differently. However, there is already a shift in that the gradient is lower with larger buttons. However, this is only a first impression and cannot yet be fully verified by the results. This is because "only" 3 different sizes have been tested so far, of which the smallest did not work well. However, a tendency can be recognized by the quantitative results, which are confirmed and reinforced by the qualitative feedback. This means that a large variety of perceived weight can be conveyed via the size in combination with the latency. In addition, the results suggest that the size also influences the perceived strength of the feedback. In combination, this also contributes to further design possibilities for interfaces on everyday objects.

6.4 Design Recommendations

Our results indicated that buttons/stickers used on everywhere displays should have at least a size that they will not be covered by users' fingers during the touch interactions. Consequently, the smallest button size in our study (10 mm) was judged significantly worse than the other two button sizes, and it was also difficult to perceive the different latencies. Therefore, we recommend using a button size with a minimum of 20 mm diameter. Further, our results show that different button sizes lead to different weight perceptions of the buttons. Larger-sized buttons will be perceived as lighter and vice versa. As a result, we recommend using different button sizes for interfaces on everywhere displays, just if different weights should be perceived. The same can be seen for the different latencies. Higher latencies between touching a button/sticker on an everywhere display will create the illusion of a higher button weight and vice versa. Our results suggest that there are three different areas of weight perception for the latencies. Latencies until a value of 58 ms are likely judged as light, and latencies from 118 ms and above are considered heavy. In between, it can be understood as "neutral". With both latencies and button sizes, we can create the perception of different button weights. As an additional insight, the qualitative results indicate that the button sizes influence the perceived weight of a touched button/sticker and its perceived feedback strength. In this case, the strength of the feedback is similar to the perceived weight, as larger button sizes lead to weaker perceived feedback and vice versa.

6.5 Limitations & Future Work

While this work investigated vibrotactile feedback created through vibrating wooden boards, interacting with graphical objects on everyday surfaces might sometimes be challenging using this technology. Only if surfaces are thin, a conventional speaker can vibrate

	38 ms	**58 ms**	**78 ms**	**98 ms**	**118 ms**	**138 ms**	**158 ms**
38 ms	–	0.6238	**< .001**	< .001	< .001	< .001	< .001
58 ms	–	–	**0.0280**	**< .001**	< .001	< .001	< .001
78 ms	–	–	–	**0.0210**	< .001	< .001	< .001
98 ms	–	–	–	–	**< .001**	**< .001**	< .001
118 ms	–	–	–	–	–	**< .001**	**< .001**
138 ms	–	–	–	–	–	–	**< .001**

Table 1: Comparison to previous findings for the perceived weight of touched buttons on a tablet. The boldly marked entries are significant differences that were not found on the tablet but on an everyday surface.

it using sound. Otherwise, it could be combined with other technologies like AR. With such an approach, the vibration of objects could be created using vision and sound, which can lead to an illusion of vibration that can be realistically felt [21].

7 CONCLUSION

This work explores whether latency between touching a graphical button of different sizes influences its perceived weight. If this is the case, it can help to understand how to design interfaces on everyday life surfaces and for everywhere displays. This can help to design and build calm interfaces that contribute to ubiquitous computing, especially by stimulating the surfaces in a scalable way with common technology like a speaker. This would also be important in cases where haptic feedback technology is not feasible. For our investigation, we placed three different graphical buttons of different sizes on different wooden boars, representing a wooden everyday surface. These boards were completely identical and were placed above a common speaker that let the wooden boards vibrate using resonant frequencies. Also, we used different latencies that were used in previous research on other technologies. Our results lead to the following design recommendations: (1) The size of a graphical button should be larger than 10 mm and reach a size that the button is not covered by the finger when being touched. (2) A larger button size leads to a lighter perceived button and vice versa. (3) Frequencies until 58 ms are perceived as light. (4) Frequencies equal to or larger than 118 ms are perceived as heavy. We were able to extend previous findings on touch-based interface design and transfer them to the use for everywhere displays. We also proved that just a minimum of additional hardware is needed to create feedback sensations, like inducing different button weights.

ACKNOWLEDGMENTS

This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 425869442 and is part of Priority Program SPP2199 Scalable Interaction Paradigms for Pervasive Computing Environments.

REFERENCES

[1] Lynn Y Arnaut and Joel S Greenstein. 1986. Optimizing the touch tablet: The effects of control-display gain and method of cursor control. *Human Factors* 28, 6 (1986), 717–726.

[2] Olivier Bau, Ivan Poupyrev, Ali Israr, and Chris Harrison. 2010. TeslaTouch: electrovibration for touch surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology.* 283–292.

[3] Bruce C Berndt and Ronald J Evans. 1981. The determination of Gauss sums. *Bull. Amer. Math. Soc.* 5, 2 (1981), 107–129.

[4] James V Bradley. 1958. Complete counterbalancing of immediate sequential effects in a Latin square design. *J. Amer. Statist. Assoc.* 53, 282 (1958), 525–528.

[5] Stephen Brewster, Faraz Chohan, and Lorna Brown. 2007. Tactile feedback for mobile interactions. In *Proceedings of the SIGCHI conference on Human factors in computing systems.* 159–162.

[6] Stuart Card, Thomas Moran, and Allen Newell. 1986. The model human processor-An engineering model of human performance. *Handbook of perception and human performance.* 2, 45–1 (1986), 1 – 35.

[7] Youngjun Cho, Andrea Bianchi, Nicolai Marquardt, and Nadia Bianchi-Berthouze. 2016. RealPen: Providing Realism in Handwriting Tasks on Touch Surfaces Using Auditory-Tactile Feedback. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 195–205. https://doi.org/10.1145/2984511.2984550

[8] Inrak Choi, Heather Culbertson, Mark R. Miller, Alex Olwal, and Sean Follmer. 2017. Grabity: A Wearable Haptic Interface for Simulating Weight and Grasping in Virtual Reality. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 119–130. https://doi.org/10.1145/3126594.3126599

[9] Sean Follmer, Daniel Leithinger, Alex Olwal, Akimitsu Hogge, and Hiroshi Ishii. 2013. InFORM: Dynamic Physical Affordances and Constraints through Shape and Object Actuation. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 417–426. https://doi.org/10.1145/2501988.2502032

[10] Satoshi Hashiguchi, Shohei Mori, Miho Tanaka, Fumihisa Shibata, and Asako Kimura. 2018. Perceived Weight of a Rod under Augmented and Diminished Reality Visual Effects. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 12, 6 pages. https://doi.org/10.1145/3281505.3281545

[11] Sean T. Hayes and Julie A. Adams. 2017. Control-display ratio enhancements for mobile interaction. *International Journal of Human-Computer Studies* 99 (2017), 57–71. https://doi.org/10.1016/j.ijhcs.2016.10.006

[12] Seongkook Heo, Jaeyeon Lee, and Daniel Wigdor. 2019. PseudoBend: Producing Haptic Illusions of Stretching, Bending, and Twisting Using Grain Vibrations. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 803–813. https://doi.org/10.1145/3332165.3347941

[13] Hiroshi Ishii, Dávid Lakatos, Leonardo Bonanni, and Jean-Baptiste Labrune. 2012. Radical Atoms: Beyond Tangible Bits, toward Transformable Materials. *Interactions* 19, 1 (Jan. 2012), 38–51. https://doi.org/10.1145/2065327.2065337

[14] Brett R. Jones, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. 2013. IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 869–878. https://doi.org/10.1145/2470654.2466112

[15] Topi Kaaresoja, Stephen Brewster, and Vuokko Lantz. 2014. Towards the Temporally Perfect Virtual Button: Touch-Feedback Simultaneity and Perceived Quality in Mobile Touchscreen Press Interactions. *ACM Trans. Appl. Percept.* 11, 2, Article 9 (June 2014), 25 pages. https://doi.org/10.1145/2611387

[16] Topi Kaaresoja, Eve Hoggan, and Emilia Anttila. 2011. Playing with Tactile Feedback Latency in Touchscreen Interaction: Two Approaches. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part II* (Lisbon, Portugal) *(INTERACT'11)*. Springer-Verlag, Berlin, Heidelberg, 554–571.

[17] Nicholas Katzakis, Jonathan Tong, Oscar Ariza, Lihan Chen, Gudrun Klinker, Brigitte Röder, and Frank Steinicke. 2017. Stylo and Handifact: Modulating Haptic Perception through Visualizations for Posture Training in Augmented Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (Brighton, United Kingdom) *(SUI '17)*. Association for Computing Machinery, New York, NY, USA,

58–67. https://doi.org/10.1145/3131277.3132181

[18] Takahiro Kawabe and Yusuke Ujitoko. 2023. Pseudo-Haptic Heaviness Influenced by the Range of the C/D Ratio and the Position of the C/D Ratio Within a Given Range. *IEEE Transactions on Haptics* (2023), 1–6. https://doi.org/10.1109/TOH.2023.3266494

[19] Chantal Keller, Jérémy Bluteau, Renaud Blanch, and Sabine Coquillart. 2012. PseudoWeight: Making Tabletop Interaction with Virtual Objects More Tangible. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces* (Cambridge, Massachusetts, USA) *(ITS '12)*. Association for Computing Machinery, New York, NY, USA, 201–204. https://doi.org/10.1145/2396636.2451335

[20] Sunjun Kim and Geehyuk Lee. 2013. Haptic Feedback Design for a Virtual Button along Force-Displacement Curves. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 91–96. https://doi.org/10.1145/2501988.2502041

[21] Marco Kurzweg, Maximilian Letter, and Katrin Wolf. 2023. Vibrollusion: Creating a Vibrotactile Illusion Induced by Audiovisual Touch Feedback. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (<conf-loc>, <city>Vienna</city>, <country>Austria</country>, </conf-loc>) *(MUM '23)*. Association for Computing Machinery, New York, NY, USA, 185–197. https://doi.org/10.1145/3626705.3627790

[22] Marco Kurzweg, Simon Linke, Yannick Weiss, Maximilian Letter, Albrecht Schmidt, and Katrin Wolf. 2023. Assignment Of a Vibration To a Graphical Object Induced By Resonant Frequency. In *Human-Computer Interaction – INTERACT 2023: 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part I* (York, United Kingdom). Springer-Verlag, Berlin, Heidelberg, 523–545. https://doi.org/10.1007/978-3-031-42280-5_33

[23] Marco Kurzweg, Yannick Weiss, Marc O. Ernst, Albrecht Schmidt, and Katrin Wolf. 2024. A Survey on Haptic Feedback through Sensory Illusions in Interactive Systems. *ACM Comput. Surv.* (feb 2024). https://doi.org/10.1145/3648353 Just Accepted.

[24] Roberto Longo, Thomas Delaunay, Didier Laux, Mohammed El Mouridi, Olivier Arnould, and Emmanuel Le Clezio. 2012. Wood elastic characterization from a single sample by resonant ultrasound spectroscopy. *Ultrasonics* 52, 8 (2012), 971–974.

[25] Pedro Lopes, Sijing You, Alexandra Ion, and Patrick Baudisch. 2018. Adding Force Feedback to Mixed Reality Experiences and Games Using Electrical Muscle Stimulation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174020

[26] Akihiro Maehigashi, Akira Sasada, Miki Matsumuro, Fumihisa Shibata, Asako Kimura, and Sumaru Niida. 2021. Virtual Weight Illusion: Weight Perception of Virtual Objects Using Weight Illusions. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Yokohama</city>, <country>Japan</country>, </conf-loc>) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 344, 6 pages. https://doi.org/10.1145/3411763.3451842

[27] Akihiro Maehigashi, Akira Sasada, Miki Matsumuro, Fumihisa Shibata, Asako Kimura, and Sumaru Niida. 2021. Virtual Weight Illusion: Weight Perception of Virtual Objects Using Weight Illusions. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 344, 6 pages. https://doi.org/10.1145/3411763.3451842

[28] Cara M. Nunez, Sophia R. Williams, Allison M. Okamura, and Heather Culbertson. 2019. Understanding Continuous and Pleasant Linear Sensations on the Forearm From a Sequential Discrete Lateral Skin-Slip Haptic Device. *IEEE Transactions on Haptics* 12, 4 (2019), 414–427. https://doi.org/10.1109/TOH.2019.2941190

[29] Jérôme Perret and Emmanuel Vander Poorten. 2018. Touching virtual reality: a review of haptic gloves. In *ACTUATOR 2018; 16th International Conference on New Actuators*. VDE, 1–5.

[30] Claudio S. Pinhanez. 2001. The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces. In *Proceedings of the 3rd International Conference on Ubiquitous Computing* (Atlanta, Georgia, USA) *(UbiComp '01)*. Springer-Verlag, Berlin, Heidelberg, 315–331.

[31] Pragathi Praveena, Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2020. *Supporting Perception of Weight through Motion-Induced Sensory Conflicts in Robot Teleoperation*. Association for Computing Machinery, New York, NY, USA, 509–517. https://doi.org/10.1145/3319502.3374841

[32] Michael Rietzler, Gabriel Haas, Thomas Dreja, Florian Geiselhart, and Enrico Rukzio. 2019. Virtual Muscle Force: Communicating Kinesthetic Forces Through Pseudo-Haptic Feedback and Muscle Input. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 913–922. https://doi.org/10.1145/3332165.3347871

[33] Majed Samad, Elia Gatti, Anne Hermes, Hrvoje Benko, and Cesare Parise. 2019. *Pseudo-Haptic Weight: Changing the Perceived Weight of Virtual Objects By Manipulating Control-Display Ratio*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300550

[34] Jonas Schmidtler and Klaus Bengler. 2018. Influence of Size-Weight Illusion on Usability in Haptic Human-Robot Collaboration. *IEEE Transactions on Haptics* 11, 1 (2018), 85–96. https://doi.org/10.1109/TOH.2017.2757925

[35] Youngbo Aram Shim and Geehyuk Lee. 2018. Demonstrating Gamepad with Programmable Haptic Texture Analog Buttons. In *Adjunct Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 194–196.

[36] Marco Speicher, Jan Ehrlich, Vito Gentile, Donald Degraen, Salvatore Sorce, and Antonio Krüger. 2019. Pseudo-Haptic Controls for Mid-Air Finger-Based Menu Interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3312927

[37] Stanley S Stevens. 1960. The psychophysics of sensory function. *American scientist* 48, 2 (1960), 226–253.

[38] Anselm Strauss and Juliet M Corbin. 1997. *Grounded theory in practice*. Sage.

[39] R A Syaputra and A Widiyanto. 2020. Implementation of the level of detail method on augmented reality android-based applications. *Journal of Physics: Conference Series* 1517, 1 (apr 2020), 012090. https://doi.org/10.1088/1742-6596/1517/1/012090

[40] Luca Turchet, Maud Marchal, Anatole Lécuyer, Rolf Nordahl, and Stefania Serafin. 2010. Influence of Auditory and Visual Feedback for Perceiving Walking over Bumps and Holes in Desktop VR. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology* (Hong Kong) *(VRST '10)*. Association for Computing Machinery, New York, NY, USA, 139–142. https://doi.org/10.1145/1889863.1889893

[41] Yusuke Ujitoko, Yuki Ban, Takuji Narumi, Tomohiro Tanikawa, Koichi Hirota, and Michitaka Hirose. 2015. Yubi-Toko: Finger Walking in Snowy Scene Using Pseudo-Haptic Technique on Touchpad. In *SIGGRAPH Asia 2015 Emerging Technologies* (Kobe, Japan) *(SA '15)*. Association for Computing Machinery, New York, NY, USA, Article 29, 3 pages. https://doi.org/10.1145/2818466.2818491

[42] Mark Weiser. 1991. The Computer for the 21 st Century. *Scientific american* 265, 3 (1991), 94–105.

[43] Mark Weiser and John Seely Brown. 1996. Designing calm technology. *PowerGrid Journal* 1, 1 (1996), 75–85.

[44] Daehan Wi, Angela Sodemann, and Riley Chicci. 2017. Vibratory haptic feedback assistive device for visually-impaired drivers. In *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*. IEEE, 1–5.

[45] Jan Willms, Maximilian Letter, Emile Marchandise, and Katrin Wolf. 2023. Pull Outperforms Push as Vibrotactile Wristband Feedback for Mid-Air Gesture Guidance. In *Proceedings of Mensch Und Computer 2023* (<conf-loc>, <city>Rapperswil</city>, <country>Switzerland</country>, </conf-loc>) *(MuC '23)*. Association for Computing Machinery, New York, NY, USA, 138–148. https://doi.org/10.1145/3603555.3603579

[46] Katrin Wolf and Timm Bäder. 2015. Illusion of Surface Changes Induced by Tactile and Visual Touch Feedback. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI EA '15)*. Association for Computing Machinery, New York, NY, USA, 1355–1360. https://doi.org/10.1145/2702613.2732703

[47] Taiki Yamada, Fumihisa Shibata, and Asako Kimura. 2018. Analysis of the R-V dynamics illusion behavior in terms of auditory stimulation. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 70, 2 pages. https://doi.org/10.1145/3281505.3281595

[48] Chien-Hsien Yeh, Fong-Chin Su, Yan-Shen Shan, Marat Dosaev, Yury Selyutskiy, Irina Goryacheva, and Ming-Shaung Ju. 2020. Application of piezoelectric actuator to simplified haptic feedback system. *Sensors and Actuators A: Physical* 303 (2020), 111820.

A APPENDIX

Sample 1 (S1)	Sample 2 (S2)	Mean (S1)	Mean (S2)	Cohen's d (effect-size)	p-value
38ms	58ms	0.057	0.076	-0.03	0.6238
38ms	**78ms**	**0.057**	**0.102**	**-0.09**	**< .001**
38ms	**98ms**	**0.057**	**0.124**	**-0.14**	**< .001**
38ms	**118ms**	**0.057**	**0.151**	**-0.19**	**< .001**
38ms	**138ms**	**0.057**	**0.176**	**-0.24**	**< .001**
38ms	**158ms**	**0.057**	**0.205**	**-0.30**	**< .001**
58ms	**78ms**	**0.076**	**0.102**	**-0.06**	**0.0280**
58ms	**98ms**	**0.076**	**0.124**	**-0.13**	**< .001**
58ms	**118ms**	**0.076**	**0.151**	**-0.20**	**< .001**
58ms	**138ms**	**0.076**	**0.176**	**-0.27**	**< .001**
58ms	**158ms**	**0.076**	**0.205**	**-0.35**	**< .001**
78ms	**98ms**	**0.102**	**0.124**	**-0.07**	**0.0210**
78ms	**118ms**	**0.102**	**0.151**	**-0.15**	**< .001**
78ms	**138ms**	**0.102**	**0.176**	**-0.23**	**< .001**
78ms	**158ms**	**0.102**	**0.205**	**-0.32**	**< .001**
98ms	**118ms**	**0.124**	**0.151**	**-0.09**	**< .001**
98ms	**138ms**	**0.124**	**0.176**	**-0.18**	**< .001**
98ms	**158ms**	**0.124**	**0.205**	**-0.28**	**< .001**
118ms	**138ms**	**0.151**	**0.176**	**-0.10**	**< .001**
118ms	**158ms**	**0.151**	**0.205**	**-0.21**	**< .001**
138ms	**158ms**	**0.176**	**0.205**	**-0.12**	**< .001**

Table 1: Results of the WEIGHTASSIGNMENT for the different LATENCIES analyzed with an ANOVA and a post-hoc Tukey test. Significant results are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Mean (S1)	Mean (S2)	Cohen's d (effect-size)	p-value
10mm	**20mm**	**mean**	**mean**	**-0.04**	**0.0207**
10mm	30mm	mean	mean	-0.02	0.4258
20mm	30mm	mean	mean	0.02	0.3279

Table 2: Results of the WEIGHTASSIGNMENT for the different BUTTONSIZES analyzed with an ANOVA and a post-hoc Tukey test. Significant results are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Cohen's g (effect-size)	p-value
38ms	58ms	0.346	1
38ms	78ms	0.318	1
38ms	98ms	0.365	0.956
38ms	**118ms**	**0.334**	**0.003**
38ms	**138ms**	**0.387**	**0.005**
38ms	**158ms**	**0.312**	**< .001**
58ms	78ms	0.377	1
58ms	98ms	0.398	1
58ms	118ms	0.318	0.344
58ms	138ms	0.334	0.519
58ms	**158ms**	**0.387**	**0.010**
78ms	98ms	0.388	1
78ms	118ms	0.376	0.167
78ms	138ms	0.365	0.258
78ms	**158ms**	**0.315**	**0.004**
98ms	118ms	0.354	1
98ms	138ms	0.371	1
98ms	**158ms**	**0.396**	**0.044**
118ms	138ms	0.369	1
118ms	158ms	0.392	1
138ms	158ms	0.394	1

Table 3: Results of the perceivedWeight for the different latencies analyzed with Cochran-Q test an post-hoc McNemar tests. The results represent which latency was perceived as the heaviest more often than the others. Significant results are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Cohen's g (effect-size)	p-value
38ms	58ms	0.365	1
38ms	78ms	0.372	0.182
38ms	**98ms**	**0.361**	**< .001**
38ms	**118ms**	**0.354**	**< .001**
38ms	**138ms**	**0.351**	**< .001**
38ms	**158ms**	**0.331**	**< .001**
58ms	78ms	0.397	1
58ms	**98ms**	**0.367**	**0.016**
58ms	**118ms**	**0.369**	**< .001**
58ms	**138ms**	**0.368**	**< .001**
58ms	**158ms**	**0.344**	**< .001**
78ms	98ms	0.390	1
78ms	**118ms**	**0.391**	**0.041**
78ms	**138ms**	**0.397**	**0.041**
78ms	158ms	0.342	0.321
98ms	118ms	0.359	1
98ms	138ms	0.367	1
98ms	158ms	0.365	1
118ms	138ms	0.364	1
118ms	158ms	0.354	1
138ms	158ms	0.358	1

Table 4: Results of the perceivedWeight for the different latencies analyzed with Cochran-Q test an post-hoc McNemar tests. The results represent which latency was perceived as the lightest more often than the others. Significant results are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
10mm	**20mm**	**98 / 59.30**	**138 / 29.65**	**-1.978**	**< .001**
10mm	**30mm**	**98 / 59.30**	**138 / 29.65**	**-1.690**	**0.003**
20mm	30mm	138 / 29.65	138 / 29.65	-0.956	0.097

Table 5: Results of the PERCEIVEDWEIGHT for the different BUTTONSIZES showing which latencies were perceived as the heaviest. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.016. Significant results smaller than the corrected α are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
10mm	20mm	58 / 29.65	58 / 29.65	-0.813	0.158
10mm	30mm	58 / 29.65	78 / 29.65	-1.389	0.017
20mm	**30mm**	**58 / 29.65**	**78 / 29.65**	**-1.725**	**0.002**

Table 6: Results of the PERCEIVEDWEIGHT for the different BUTTONSIZES showing which latencies were perceived as the lightest. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.016. Significant results smaller than the corrected α are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
10mm	**20mm**	**4 / 1.48**	**5 / 1.48**	**-5.332**	**< .001**
10mm	**30mm**	**4 / 1.48**	**5 / 1.48**	**-5.399**	**< .001**
20mm	30mm	5 / 1.48	5 / 1.48	-0.318	0.580

Table 7: Results of the PLEASANTNESS for the different BUTTONSIZES. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.016. Significant results smaller than the corrected α are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
38ms	58ms	5 / 1.48	5 / 1.48	-0.204	0.3489
38ms	78ms	5 / 1.48	5 / 1.48	-0.282	0.1963
38ms	98ms	5 / 1.48	5 / 1.48	-0.466	0.0324
38ms	**118ms**	**5 / 1.48**	**5 / 1.48**	**-0.681**	**0.0018**
38ms	**138ms**	**5 / 1.48**	**5 / 1.48**	**-0.789**	**< .001**
38ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.902**	**< .001**
58ms	78ms	5 / 1.48	5 / 1.48	-0.051	0.8142
58ms	98ms	5 / 1.48	5 / 1.48	-0.176	0.4189
58ms	118ms	5 / 1.48	5 / 1.48	-0.487	0.0254
58ms	**138ms**	**5 / 1.48**	**5 / 1.48**	**-0.684**	**0.0016**
58ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.852**	**< .001**
78ms	98ms	5 / 1.48	5 / 1.48	-0.153	0.4821
78ms	118ms	5 / 1.48	5 / 1.48	-0.372	0.0878
78ms	138ms	5 / 1.48	5 / 1.48	-0.550	0.0117
78ms	158ms	5 / 1.48	4 / 1.48	-0.656	0.0026
98ms	118ms	5 / 1.48	5 / 1.48	-0.270	0.2153
98ms	138ms	5 / 1.48	5 / 1.48	-0.509	0.0196
98ms	158ms	5 / 1.48	4 / 1.48	-0.623	0.0043
118ms	138ms	5 / 1.48	5 / 1.48	-0.260	0.2328
118ms	158ms	5 / 1.48	4 / 1.48	-0.557	0.0106
138ms	158ms	5 / 1.48	4 / 1.48	-0.301	0.1668

Table 8: Results of the PLEASANTNESS for the different LATENCIES. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.0024. Significant results smaller than the corrected α are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
10mm	**20mm**	**4 / 1.48**	**5 / 1.48**	**-3.631**	**< .001**
10mm	**30mm**	**4 / 1.48**	**5 / 1.48**	**-3.536**	**< .001**
20mm	30mm	5 / 1.48	5 / 1.48	-0.501	0.384

Table 9: Results of the Comfort for the different buttonSizes. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.016. Significant results smaller than the corrected α are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
38ms	58ms	5 / 1.48	5 / 1.48	-0.471	0.0308
38ms	78ms	5 / 1.48	5 / 1.48	-0.474	0.0297
38ms	98ms	5 / 1.48	5 / 1.48	-0.579	0.008
38ms	**118ms**	**5 / 1.48**	**5 / 1.48**	**-0.803**	**< .001**
38ms	**138ms**	**5 / 1.48**	**4 / 1.48**	**-1.004**	**< .001**
38ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.898**	**< .001**
58ms	78ms	5 / 1.48	5 / 1.48	-0.069	0.7592
58ms	98ms	5 / 1.48	5 / 1.48	-0.085	0.6963
58ms	118ms	5 / 1.48	5 / 1.48	-0.432	0.0477
58ms	**138ms**	**5 / 1.48**	**4 / 1.48**	**-0.737**	**< .001**
58ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.730**	**< .001**
78ms	98ms	5 / 1.48	5 / 1.48	-0.047	0.8285
78ms	118ms	5 / 1.48	5 / 1.48	-0.293	0.179
78ms	138ms	5 / 1.48	4 / 1.48	-0.591	0.0067
78ms	158ms	5 / 1.48	4 / 1.48	-0.571	0.0088
98ms	118ms	5 / 1.48	5 / 1.48	-0.323	0.1387
98ms	**138ms**	**5 / 1.48**	**4 / 1.48**	**-0.734**	**< .001**
98ms	158ms	5 / 1.48	4 / 1.48	-0.637	0.0035
118ms	138ms	5 / 1.48	4 / 1.48	-0.447	0.0405
118ms	158ms	5 / 1.48	4 / 1.48	-0.449	0.0393
138ms	158ms	5 / 1.48	4 / 1.48	-0.129	0.5535

Table 10: Results of the Comfort for the different latencies. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.0024. Significant results smaller than the corrected α are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
10mm	**20mm**	**4 / 1.48**	**5 / 1.48**	**-2.798**	**< .001**
10mm	**30mm**	**4 / 1.48**	**5 / 1.48**	**-2.416**	**< .001**
20mm	30mm	5 / 1.48	5 / 1.48	-0.772	0.180

Table 11: Results of the Realism for the different buttonSizes. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.016. Significant results smaller than the corrected α are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
38ms	58ms	5 / 1.48	5 / 1.48	-0.208	0.3387
38ms	78ms	5 / 1.48	5 / 1.48	-0.271	0.2131
38ms	98ms	5 / 1.48	5 / 1.48	-0.376	0.0846
38ms	118ms	5 / 1.48	4.5 / 0.74	-0.649	0.0029
38ms	**138ms**	**5 / 1.48**	**4 / 1.48**	**-0.875**	**< .001**
38ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.963**	**< .001**
58ms	78ms	5 / 1.48	5 / 1.48	-0.063	0.7702
58ms	98ms	5 / 1.48	5 / 1.48	-0.018	0.934
58ms	118ms	5 / 1.48	4.5 / 0.74	-0.440	0.0436
58ms	138ms	5 / 1.48	4 / 1.48	-0.657	0.0026
58ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.856**	**< .001**
78ms	98ms	5 / 1.48	5 / 1.48	-0.007	0.9717
78ms	118ms	5 / 1.48	4.5 / 0.74	-0.392	0.0721
78ms	138ms	5 / 1.48	4 / 1.48	-0.589	0.0069
78ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.814**	**< .001**
98ms	118ms	5 / 1.48	4.5 / 0.74	-0.455	0.0371
98ms	138ms	5 / 1.48	4 / 1.48	-0.634	0.0037
98ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.806**	**< .001**
118ms	138ms	4.5 / 0.74	4 / 1.48	-0.328	0.1319
118ms	158ms	4.5 / 0.74	4 / 1.48	-0.534	0.0142
138ms	158ms	4 / 1.48	4 / 1.48	-0.375	0.0853

Table 12: Results of the Realism for the different latencies. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.0024. Significant results smaller than the corrected α are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
10mm	**20mm**	**4 / 1.48**	**5 / 1.48**	**-2.815**	**< .001**
10mm	**30mm**	**4 / 1.48**	**5 / 1.48**	**-2.533**	**< .001**
20mm	30mm	5 / 1.48	5 / 1.48	-0.566	0.326

Table 13: Results of the Confidence for the different buttonSizes. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.016. Significant results smaller than the corrected α are printed in bold.

Sample 1 (S1)	Sample 2 (S2)	Median (S1) / Mad	Median (S2) / Mad	r (effect-size)	p-value
38ms	58ms	5 / 1.48	5 / 1.48	-0.113	0.6016
38ms	78ms	5 / 1.48	5 / 1.48	-0.093	0.6699
38ms	98ms	5 / 1.48	5 / 1.48	-0.164	0.4498
38ms	118ms	5 / 1.48	4.5 / 2.22	-0.485	0.0260
38ms	**138ms**	**5 / 1.48**	**4 / 1.48**	**-0.722**	**< .001**
38ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.794**	**< .001**
58ms	78ms	5 / 1.48	5 / 1.48	-0.024	0.9092
58ms	98ms	5 / 1.48	5 / 1.48	-0.026	0.9032
58ms	118ms	5 / 1.48	4.5 / 2.22	-0.4446	0.0410
58ms	**138ms**	**5 / 1.48**	**4 / 1.48**	**-0.744**	**< .001**
58ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.935**	**< .001**
78ms	98ms	5 / 1.48	5 / 1.48	-0.071	0.7421
78ms	118ms	5 / 1.48	4.5 / 2.22	-0.389	0.0744
78ms	**138ms**	**5 / 1.48**	**4 / 1.48**	**-0.700**	**0.0013**
78ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.742**	**< .001**
98ms	118ms	5 / 1.48	4.5 / 2.22	-0.372	0.0882
98ms	**138ms**	**5 / 1.48**	**4 / 1.48**	**-0.742**	**< .001**
98ms	**158ms**	**5 / 1.48**	**4 / 1.48**	**-0.839**	**< .001**
118ms	138ms	4.5 / 2.22	4 / 1.48	-0.417	0.0556
118ms	158ms	4.5 / 2.22	4 / 1.48	-0.602	0.0058
138ms	158ms	4 / 1.48	4 / 1.48	-0.293	0.1781

Table 14: Results of the Confidence for the different latencies. The results were analyzed with a Friedman test and post-hoc Wilcoxon-Signed Rank tests. A Bonferroni correction for the significance level led to an (α) of 0.0024. Significant results smaller than the corrected α are printed in bold.