The Case for Haptic Props: Shape, Weight and Vibro-tactile Feedback

Michael White
University of Cape Town
South Africa
mike.james.white@icloud.com

James Gain
University of Cape Town
South Africa
jgain@cs.uct.ac.za

Ulysse Vimont
University of Cape Town
South Africa
ulysse.vimont@hotmail.fr

Daniel Lochner
University of Cape Town
South Africa
daniel@daniellochner.com

ABSTRACT

The use of haptic props in a virtual environment setting is purported to improve both user immersion and task performance. While the efficacy of various forms of haptics has been tested through user experiments, this is not the case for hand-held tool props, an important class of input device with both gaming and non-gaming applications. From a cost and complexity of implementation perspective it is also worth investigating the relative benefits of the different types of passive and active haptics that can be incorporated into such props.

Accordingly, in this paper we present the results of a quantitative user experiment ($n = 42$) designed to assess a typical VR controller against passive, weighted, and active-haptic versions of a tracked prop, measured according to game experience, performance, and stance adopted by participants. The task involved playing a VR baseball game and the prop was a truncated baseball bat.

We found a statistically significant improvement (at $\alpha = 0.05$) with medium to large effect size ($r > 0.38$) for certain aspects of game experience (competence, immersion, flow, positive affect), performance (mean hit distance) and pose (two-handed grip) for the weighted prop over a generic controller, and in many cases over the unweighted passive prop as well. There was no significant difference between our weighted prop and the active-haptic version. This suggests that, for batting and striking tasks, tool props with passive haptics improve user experience and task performance but only if they match the weight of the original real-world tool, and that such weighting is more important than simple vibro-tactile style force-feedback.

CCS CONCEPTS

• **Computing methodologies → Virtual reality**.

MIG '19, October 28–30, 2019, Newcastle upon Tyne, United Kingdom

ACM ISBN 978-1-4503-6994-7/19/10…$15.00
https://doi.org/10.1145/3359566.3360058

KEYWORDS

Virtual Reality, Haptics, Game Experience

ACM Reference Format:
Michael White, James Gain, Ulysse Vimont, and Daniel Lochner. 2019. The Case for Haptic Props: Shape, Weight and Vibro-tactile Feedback. In *Motion, Interaction and Games (MIG '19), October 28–30, 2019, Newcastle upon Tyne, United Kingdom.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3359566.3360058

Since the very beginnings of Virtual Reality (VR) [Sutherland 1965] the sense of touch, or haptics, has been considered a crucial partner to auditory and visual elements in conveying a sense of realism, improving task performance, and enhancing user immersion. Unfortunately, effective haptic devices have proved far more complicated to develop than their audio-visual counterparts.

It is common to classify haptic devices according to whether they offer passive or active feedback [Azmandian et al. 2016; Cheng et al. 2018; Henderson and Feiner 2010], from an inert shape or an artificial force-response, respectively. This can be further subdivided into: (1) passive tactile feedback from cutaneous sensation of shape and texture; (2) passive kinesthetic feedback from the heft of the object dictated by its weight distribution; (3) active tactile feedback, such as a rumble response, that emulates an impact or pressure but without retarding movement, and (4) active force-reflecting feedback, which provides actual resistance to movement, usually requiring a servo-motor armature of some sort. This is, of course, far from a definitive classification of what is a complex and long-standing research endeavour. It does not, for instance, consider the difference between body-anchored and surface-anchored haptics, and neglects recent advances, such as mid-air haptics generated by air jets [Sodhi et al. 2013] and ultrasound fields [Inoue et al. 2015].

Our focus in this paper is on haptic props that act as proxies for hand-held real-world tools, such as bats, guns, knives, hammers, wrenches, hose-nozzles, clipboards, and the like, for gaming, visualization and training applications. These have a long history in Virtual Reality [Fujinawa et al. 2017; Hinckley et al. 1994; Jackson et al. 2013; Qi et al. 2005; Zenner and Krüger 2017] and generally fit into the category of passive tactile, passive kinesthetic and, to a lesser extent, active tactile haptic devices. Although they tend to be specialised to a particular context they have the advantage of being relatively inexpensive compared to typical active force-reflecting devices.

It must be recognised, however, that modern general-purpose VR controllers represent a challenge to the continued relevance of haptic props. These controllers enable position and orientation tracking, offer various button and trigger modalities, and support a measure of tactile feedback. Put bluntly, the question is: given the functionality of generic VR controllers is there any benefit to providing application-specific haptic props? And, if so, which properties are most important from a prop design perspective?

To answer these questions we undertook a within-subjects quantitative user experiment testing game experience and task performance in a VR baseball game variously using input from a generic VR controller and different versions of a haptic baseball bat prop.

Our choice to simulate a baseball bat was motivated by its familiarity (many potential users have either handled a baseball bat or similar sport's equipment, or seen it in use), dimensions (it is one of a class of tools that cannot be safely utilised at full size in VR), and lower dependence on force-reflection (ball-on-bat impact is transitory and thus suited to vibro-tactile feedback).

In the context of a baseball bat prop the most important haptic dimensions, as tested in our experiment, are: passive tactile feedback from shape and texture of the prop grip, passive kinesthetics from the heft of the prop and having to do with overall weight and its distribution, and active tactile feedback in the form of vibrational response to ball impact.

Based on this experiment, our findings are that a haptic baseball prop does indeed offer improved user experience and task performance over a generic VR controller, but only if appropriately weighted (i.e., with suitable passive kinaesthetic haptics). Adding vibro-tactile response to this weighted prop (i.e., active tactile haptics) does not afford statistically significant improvement. More generally, this points to the ongoing applicability of haptic props and offers some indication of weighting as a design priority for striking tasks requiring accuracy and where the analogous tool has non-negligible weight and transitory contact forces.

1 RELATED WORK

In this section we focus on a review of research into passive haptic props. For a more general coverage of haptics and pseudo-haptics the reader is referred to the taxonomy of Jeon and Choi [Jeon and Choi 2009] and overviews by Paterson [Paterson 2007] and Lécuyer [Lécuyer 2009].

The applicability of haptic props in the visualisation domain has been well demonstrated. For instance, Hinckley et al. [Hinckley et al. 1994] track a doll's head and planar prop to support orienting cutting planes through volumetric brain scan data; Qi et al. [Qi et al. 2005] provide a tangible cube, pen, and frame for interacting with more general volumetric data, and Jackson et al. [Jackson et al. 2013] use a patterned tube to orient and study thin fiber structures typical in bioimaging.

In a classic series of experiments, Insko [Insko 2001] showed that large-scale haptic props placed to reflect environmental surroundings can improve presence and wayfinding. For the former, a simple wooden ledge was positioned to simulate a cliff-edge, and for the latter, foam and similar barriers were configured to match a virtual maze.

For interaction tasks where a surface, such as a table, provides positional constraints on a haptic prop, such as a wooden cube [Wang and MacKenzie 2000] or buttons for a memory game [Viciana-Abad et al. 2010], it is clear from user experiments that haptic props offer presence and performance benefits over unconstrained free-space or stylus-based alternatives, and can even enhance narrative impact [Harley et al. 2017].

Fujinawa et al. [Fujinawa et al. 2017] investigate the general case where haptic props cannot for safety or other reasons match the size and shape of their real-world equivalent (of which our experiment is a specific instance). They provide a data-driven computational framework for designing hand-held haptic props that, while smaller, have perceptually similar wielding characteristics. This is done by strategically placing weights to emulate the mass properties of the original.

1.1 Haptic Retargeting

More recently, research in passive haptics has sought to overcome the limitations inherent in a single static shape. One approach is to allow multiple virtual attachments on a single handle or trigger grip with force feedback [Ortega and Coquillart 2005]. Another tack is to build on the notion of substitutional reality [Simeone et al. 2015] where objects in the physical environment, such as chairs, desks, and sofas, are paired with similar but not identical virtual representations. This is extended by haptic retargeting [Azmandian et al. 2016], which allows a single passive haptic prop, such as a wooden block, to substitute for multiple objects in a virtual environment. As a user reaches for a new block they are subtly redirected through visual distortions in VR to touch the original. This was shown to improve presence over controller-based interaction.

Cheng et al. [Cheng et al. 2017] extend on this principle by employing a faceted parabolic dish. The system analyzes a user's gaze and hand motions to infer their target and then distorts the VR scene in a variety of ways so that they reach to touch a section of the dish with a planar surface whose orientation matches the virtual target, all without the user being aware of the visual illusion in play. This is best suited to tactile exploration where the user pokes at objects with a single fingertip.

These represent some examples in what is already a rich vein of research in haptic retargeting and redirected touch [Kohli et al. 2013; Sait et al. 2018; Spillmann et al. 2013; Zhao and Follmer 2018].

1.2 Haptic Reconfiguration

Rather than redirecting the user's movement, the environment can also be reconfigured. For example, a sophisticated, but expensive, alternative proposed by Vonach et al. [Vonach et al. 2017] is to employ a robot arm to reposition passive haptic surfaces to match the virtual environment.

Another form of reconfiguration is offered by Shifty [Zenner and Krüger 2017], TorqueBar [Swindells et al. 2003], ElastiArm [Achibet et al. 2015], and HaptoBend [McClelland et al. 2017] which are proofs of concept for a new category of dynamic passive props. Shifty, for instance, is a rod-like device with an internal pulley capable of shifting a weight up and down its length, thereby dynamically changing the moment of inertia. It can simulate objects that change continuously in length or thickness as well as, to a

(a) (b) (c) (d)

Figure 1: Input devices used in the experiment: (a) HTC Vive Controller; (b) Simple unweighted prop; (c) Prop with housing for weights and force-feedback electronics; and, (d) Close up of vibro-tactile electronics.

lesser extent, instantaneous changes such as picking up additional virtual objects with the tool. It is thus capable of simulating to some extent the weight distribution of a baseball bat and the impact of a hit, but not the passive haptics of the handle shape. ElastiArm is an elastic armature connecting the user's shoulder and hand, providing measurable and progressive resistance as the arm is extended. Its applications are primarily in selection and navigation tasks. HaptoBend provides a planar surface with rigid panels that can be folded by the user into a variety of configurations.

Henderson and Feiner [Henderson and Feiner 2010] avoid creating haptic props altogether by opportunistically co-opting aspects of the physical environment. User interface elements such as buttons and dials are overlaid in Augmented Reality onto existing objects and surfaces. Similarly, iTurk [Cheng et al. 2018] enables the user to interact with configurable or dynamic objects, such as a ball on a pendulum wire, which can take on different roles in the environment as the user's actions alter their physical state, such as direction of swing.

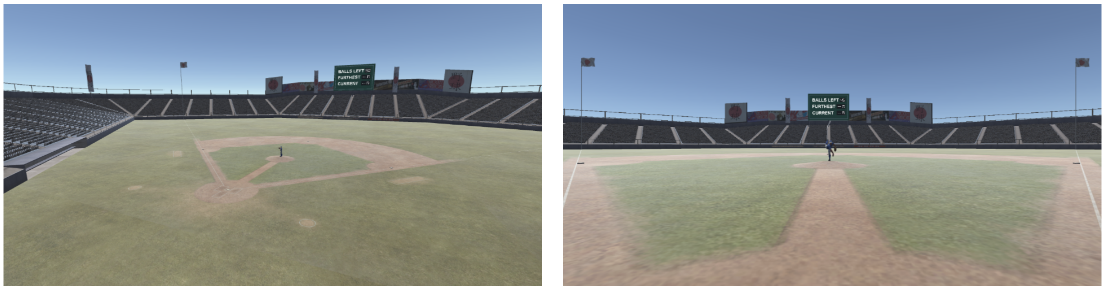
In terms of research that is most similar to our own, Knoerlein et al. [Knoerlein et al. 2007] present a system for collaborative active haptic feedback in an entertainment context. Their example application is an augmented reality ping-pong game. However, their work does not provide any experimental evaluation. Ryge et al. [Ryge et al. 2017] do undertake quantitative user experiments for a simple baseball task. However, their focus is on a generic controller retrofitted with a high fidelity vibro-tactile device rather than haptic props. In agreement with our own findings, of seven sub-items in their Likert-scale questionnaire only one (responsiveness) showed statistically significant improvement for vibro-tactile feedback.

2 INTERACTION DEVICES

We developed a system with a choice of four input devices (our treatments) to support a virtual reality baseball game (our task). The first of these was a standard VR controller with vibro-tactile feedback disabled, while the remaining three represented haptic props of increasing sophistication: unweighted, weighted, and weighted with vibro-tactile feedback.

To track position and orientation an HTC Vive Tracker was attached to each haptic prop. This tracking technology was chosen because it integrated well with our prop design and demonstrated a combination of low latency and high accuracy. This, in turn, dictated use of the HTC Vive head-mounted display and controller.

(1) *Vive Controller (CTL):* Most commercial VR systems now come bundled with an input device with 6DOF positional tracking, and multiple buttons and triggers. The Vive Controller (see Figure 1(a)) is representative of this class of devices in terms of functionality and was therefore included as a control condition. In our environment, the virtual bat was aligned vertically with respect to the pistol grip of the controller since this represents the closest correspondence to real-world use. Although supported by the controller, we explicitly excluded vibro-tactile feedback in this case because it is not universally available and we wanted to establish a baseline.
(2) *Unweighted Prop (PRP):* The first, most-basic prop simply mimics the shape and texture of a baseball-bat handle, without compensating for other factors, such as weight. To produce this prop, we cut down a wooden baseball bat to a length of 28.5cm — long enough to ensure a comfortable grip and induce passive tactile haptics, but not so long as to cause safety concerns when swung vigourously in an enclosed space where the user is effectively blind to their real-world surroundings. As pointed out by Fujinawa et al. [Fujinawa et al. 2017], such truncation is often necessary for longer haptic props.
To mount the Vive Tracker, a hole was drilled into the prop and a 1/4" metal thread inserted and secured using epoxy. This allowed rapid attachment and detachment of the tracker during experiments (see Figure 1(b)). The final combined weight of the prop was relatively light at 150g compared against major league baseball bats, which are at least 907g by regulation and typically in the range 935 – 964g
(3) *Weighted Prop (W-PRP):* To enable a more realistic swing sensation and overall heft, weight was added to the end of the simple unweighted prop (PRP) by threading metal washers onto the mounting point. This was enclosed by a PVC housing and secured with a wingnut (see Figure 1(c)). A final weight of 850g was chosen for the prop by polling participants as part of a pre-experiment. It should be noted, however, that matching weight alone does not provide our prop with the same swing characteristics as a real bat. Since the weight distribution in the prop is closer to the handle additional weight is needed to match a bat's moment of inertia. We did not adopt this approach due to concerns



(a) (b)

Figure 2: Virtual environment representing a baseball stadium as used for the task: (a) Aerial view of the stadium, and (b) First-person participant perspective

about fatigue and because it was not favoured by the pre-experiment participants.

(4) *Weighted Prop with Active Haptic Feedback (H-PRP):* Our final version incorporated a vibro-tactile device (see Figure 1(d)) within the housing of the weighted prop, in order to support testing of the effect of active tactile feedback. The vibration motors of this system were activated when a ball was struck in the virtual environment, with their amplitude dictated by the force of the hit.

In terms of device design there are effectively three feature axes in play, namely the presence or absence of approximately-correct weighting, vibro-tactile feedback, and passive tactile shape. The full space of treatments is thus $2^3 = 8$, of which we have selected 4 particular feature combinations and omitted, among others, a VR controller and unweighted prop with vibro-tactile feedback. This was necessary to cap the duration of the planned within-subjects experiment to less than an hour, thereby reducing fatigue effects. Given this constraint we chose a commonly available controller option (CTL), two low-cost alternatives (PRP and W-PRP), and the highest fidelity choice in the design space (H-PRP).

2.1 Implementing Vibro-tactile Feedback

Our force-feedback system was designed around the Wemos D1 Mini, a WiFi-enabled circuit board based on the ESP8266EX microchip. This was chosen because it was cost-effective, capable of wireless transmission and readily available. Bluetooth would have been more power-efficient, but we were unable to source a similarly cost-effective Bluetooth-compatible board at the time of production. Once a hit was registered in the virtual environment, this was signaled over WiFi to the board, which activated the four attached vibration motors. The system was supplied by two 9V batteries connected in parallel, whose power output was controlled by four MOSFETs connected to the circuit board, which in turn fed the motors (which were the same 7.6V motors used in PlayStation DualShock 2). Initially, standard transistors were used but these resulted in sluggish motor performance. A motor was secured to each of the four faces of a wooden block and this was in turn attached to the top of the bat handle (i.e., at 28.5cm from the base). Although the HTC Vive tracker introduces a 22ms latency, no appreciable delay was discernible in pre-testing or reported by users in the experiments.

The control software was written in C++ using the Arduino platform and flashed onto the circuit board. UDP was used as the transmission protocol because devices communicated on a private network, guaranteeing minimal congestion.

The circuit board was capable of pulse width modulation, which allowed the strength of motor vibration to be adjusted based on a WiFi-transmitted magnitude. This enabled force feedback to be adjusted to low, medium or high depending on the force of impact between bat and ball in the virtual environment.

3 METHOD

3.1 Experimental Design

We performed a randomised single-factor repeated-measures experiment with 46 university students. Thus, each participant experienced all four treatments, in a random order. These participants were recruited through convenience sampling using advertising with posters on campus and through social media. To encourage participation a small monetary incentive was offered. Our target population was computer literate adults, because past experiments have shown us that those lacking familiarity with computers are overwhelmed by the VR experience and unable to focus on a task effectively. At the same time, we chose to exclude, through pre-screening, subjects who reported extensive experience with dedicated VR controllers (such as an Oculus Touch or HTC Vive Controller), since we felt that their participation would bias the outcome.

To ensure computer literacy, and record potential confounding factors, participants completed a pre-experiment questionnaire, which captured age, gender, computer literacy, frequency playing video games, and proficiency with ball sports, in general, and baseball, in particular. Participants were warned about the potential for simulator sickness and allowed to recuse themselves during the experiment, although none did so.

Experiments were performed in a dedicated experiment room with a single participant at a time. The experiment was run on a desktop computer with Windows 10 and Unity3D [Technologies [n. d.]] equipped with an Intel® Core *i*5 with 6 cores, clocked at 2.8 GHz with 16GB of RAM, and an NVidia GTX 1070 graphics card. The VR equipment consisted of an HTC Vive 2018 headset, a single HTC Vive controller for the base treatment, and an HTC Vive tracker for the treatments involving haptic props.

Controller type is our single factor, with four treatments (as outlined in section 2): a conventional HTC Vive controller as the base case, a simple tactile prop consisting of a sawed-off baseball bat with HTC Vive tracker attached, a weighted version of this prop with a more accurate moment of inertia, and, finally, the weighted version with a simple active force-feedback device incorporated to signal a successful hit. To prevent learning effects the order of presentation of the controllers was randomised between participants.

Participants were provided verbal instruction at the start of the experiment indicating that their aim should be to achieve as high a score as possible in the virtual baseball game that they would be playing.

For each controller treatment, participants performed a baseball task with the specified controller (which was not shown to them beforehand) and their hit performance was captured. Furthermore, during the course of completing the task their characteristic pose in terms of grip and body orientation was recorded photographically (for which prior consent was obtained). After each treatment task participants assessed the controller by completing the core module of the game experience questionnaire [IJsselsteijn et al. 2013]. The rest time between treatments varied between participants, but was generally at least $3 - -4$ minutes in duration since it required removing and donning the VR equipment, and completing the GEQ Core questionnaire inbetween.

The primary measures (see Section 3.3) for each task were thus: game score, adopted pose and game experience.

During several experiment sessions the active force feedback device failed due to a depleted battery. We chose to exclude the four participants affected by this leaving 42 for analysis of game score and game experience. One subject did not give consent to be photographed, so there were 41 sets of poses available.

3.2 Task Design

A virtual reality baseball game was constructed in the Unity game engine to serve as the experiment task. Participants were positioned at the home plate in a virtual baseball stadium (see Figure 2) with an animated pitcher figure 18m away on the pitcher's mound. A baseball bat model was keyed to follow the position and orientation of either the Vive controller or tracker, depending on the input device. Although the bat was visible, no hand or body avatar was displayed. Since we were unable to automatically detect the configuration of a user's grip on the haptic prop we chose to disable this feature for all treatments to avoid introducing an additional independent variable.

The stated goal for participants was to hit as many balls as far as possible within the bounds of play. Participants were given some time to acclimatise while an introductory theme played before an audio cue of 'play ball' signalled the start of pitching. Ball pitches travelling from the animated pitcher to the strike zone varied from normal straight balls, to curved balls with variable turning points in their parabolic trajectory. All pitches were legal and participants did not have to concern themselves with no balls, merely the task of hitting. In early versions of the game the pitcher could deliver a ball at up to 144 km/h (the speed of an average major-league fastball), but this speed was adjusted downwards to a maximum of 90 km/h (56 mph) because it made the game too difficult.

In the distance, behind the pitcher and near the edge of the stadium, a large digital scoreboard was used to display the number of balls remaining, the travel distance of the previous hit and furthest hit overall. On the backend the outcome of all pitches was recorded in terms of both state (legal, out of bounds, or missed) and travel distance to provide performance measure (see Section 3.3).

Participants were provided with various forms of audio and visual feedback. Hits, both legal and illegal, and misses were signalled by audio from an imaginary crowd murmuring in disappointment or affirmation. Audio cues from the crowd and commentator were also based on how hard the player hit the ball and the distance it travelled, with home runs (hits further than 100m) being a special case.

The sound of a ball-on-bat impact was set proportionate to the displacement between the point of impact and the sweet-spot of the bat. This also correlated with the vibration amplitude signaled to the force-feedback chip in the case of the active haptic prop (H-PRP).

On a successful hit, a trail effect and text banner were used to help participants track the ball's trajectory and travel distance, respectively. The banner was attached to the ball, scaled with distance from the viewer, and display set to the current traversed distance. Three seconds after a ball came to rest, the pitcher threw another.

3.3 Measures

(a) (b)

Figure 3: Examples of pose coding: (a) one-handed grip with front-facing stance (1F), and (b) two-handed grip with side-facing stance (2S).

Per controller, we captured: the overall experience using IJsselsteijn *et al.'s* Game Experience Questionnaire [IJsselsteijn et al. 2013], performance using score attained in the baseball game task, which was based on the travel distance of successful hits, and the pose adopted by participants during gameplay. Below we define these measures more explicitly.

Game Experience: We used the core module of the Game Experience Questionnaire (GEQ) [IJsselsteijn et al. 2013] in order to measure factors such as engagement, immersion and enjoyment. In this instance, we chose game experience over presence or realism measures because it encompasses a broad range of experiential aspects.

There is a surfeit of choice in game experience and immersion instruments, including: the Player Experience of Need Satisfaction (PENS) [Ryan et al. 2006], the Immersive Experience Questionnaire (IEQ) [Jennett et al. 2008], and the Game Engagement Questionnaire (GEngQ) [Brockmyer et al. 2009], and, unsurprisingly, there is significant correlation between them [Denisova et al. 2016].

We chose the GEQ because it has been widely used, and in a similar context [Nacke 2010]. Furthermore, it is relatively quick to administer, which is important given the number of treatments in our experiment and our desire to avoid subject fatique, and does not focus too heavily on interaction between players or story elements, both of which were absent from our sports-oriented task.

A common criticism of the GEQ is that it lacks validation [Norman 2013] but, in fact, it has been validated both by the authors [Poels et al. 2007] and independently [Johnson et al. 2018].

The GEQ Core is a 33-item scale for measuring player experience across 7 dimensions, namely: competence (performance against game goals), sensory and imaginative immersion (connection to the game), flow (lack of awareness of time and effort), tension (annoyance experienced during play), challenge (degree of difficulty), negative affect (a bad emotional experience), positive affect (a good emotional experience). Respondents are asked to rate short statements about their experience (e.g., 'I was good at it.' and 'I felt pressured') while playing the game on a Likert scale from 0 ('not at all') to 4 ('extremely'). Two items were removed from the version of the questionnaire that we administered, namely: (3) 'I was interested in the game's story'; (19) 'I felt that I could explore things', because they were not relevant to the task in question and would have confused participants. These deleted items contribute to the immersion dimension, which in our study is a mean of 4 items rather than 6, and consequently less weight should be given to this aspect. Also germane to the analysis is that Johnson *et al.* [Johnson et al. 2018] report strong overlap between dimensions, particularly those focusing on negative experience (negative affect, tension, challenge).

Performance: For each pitch i during the game task we recorded whether the player achieved a fair ball (a legal hit inside the field of play), foul ball (a hit landing behind the foul lines), or strike (miss). For fair ball hits we also captured the horizontal distance (d_i) from the batting plate to the first bounce of the ball. There were 50 pitches in total during each task. From this two performance measures are derived that reflect the player's achievement: hit ratio (number of hits, both foul and fair / 50) and mean hit distance ($\sum_i d_i/50$).

Pose: In the interests of investigating whether the type of controller influenced players to assume a classic baseball batter's stance, we photographed the characteristic pose adopted by the participant during each treatment. Our process per treatment was to photograph all distinct poses and retain the one used for the most hits. In most cases, players ended up adopting a single stance for a given treatment. A single image does not, of course, capture the full dynamic motion of a baseball swing, but it does allow a simple coding (see Figure 3, as follows: grip – number of hands used to hold the device (1 or 2), and facing – orientation relative to the pitcher (F = front facing, S = side on). These codings are orthogonal. For instance, while a front-facing two-handed stance ($2F$) might be regarded as awkward, it was employed in some instances.

4 DATA ANALYSIS AND RESULTS

We analyzed data from 42 participants in the case of performance and game experience, and 41 participants for pose. In demographic terms, 12 participants identified as female and 30 as male, with ages ranging from 17 to 40, although the majority (83%) were between 18 and 25. For prior experience most participants considered themselves proficient with computers (7% slightly, 12% moderately, 36% fairly, 45% extremely) and had little or no prior experience playing baseball (57% none, 43% a little, 0% a lot). There was a spread of proficiency at ball sports in general (12% not at all, 36% slightly, 28% moderately, 17% fairly, 7% extremely) and in typical hours spent playing computer games in a week (17% none, 21% < 1 hour, 26% 1–3 hours, 14% 3–6 hours, 10% 6–10 hours, 12% > 10 hours).

Each participant performed a task with 50 observations (baseball pitches) and completed a game experience questionnaire for each of the four controllers. For all but one participant, a photograph of their pose was also taken during each controller treatment. This provided 8400 observations, 168 game experience questionnaires, and 164 pose photos for analysis.

As a first step in our analysis of continuous measures (Game Experience and Performance), we applied the Shapiro-Wilks test of normality on the residuals. In cases where we found with statistical significance that an outcome followed a non-parametric distribution, we subsequently applied the Friedman test to determine if a significant difference between controllers existed. Otherwise, an assumption of normality held and we used a repeated-measures ANOVA. In cases where a significant difference was found, we followed up with a post-hoc Tukey's HSD test (for normal data) or Wilcoxon Signed Rank test (for non-parametric data), both with Bonferroni correction to determine which particular treatments differed.

For binary measures (Pose) Cochran's Q Test was applied followed, where indicated, by post-hoc analysis using pairwise McNemar's test with Bonferonni correction. Note that in all tests for significance we use a value of $\alpha = 0.05$ as discriminant.

4.1 Game Experience

Table 2 and Figure 4 show GEQ results as tabulated means and standard deviations, and box-and-whisker plots, respectively.

In analyzing the GEQ data, the Competence, Immersion and Challenge dimensions passed the test of normality and so we proceeded to within-subjects ANOVA. On the other hand, Flow, Tension, Negative Affect and Positive Affect were found not to obey a normal distribution ($p < 0.05$) and so the Friedman test was applied.

Competence ($F = 12.448$, $p < 0.001$), Immersion ($F = 8.849$, $p < 0.001$), Flow ($\chi^2 = 18.71$, $p < 0.001$), Tension ($\chi^2 = 13.62$, $p = 0.003$) and Positive Affect ($\chi^2 = 22.23$, $p < 0.001$) all showed a significant effect for controller, but not Challenge and Negative Affect.

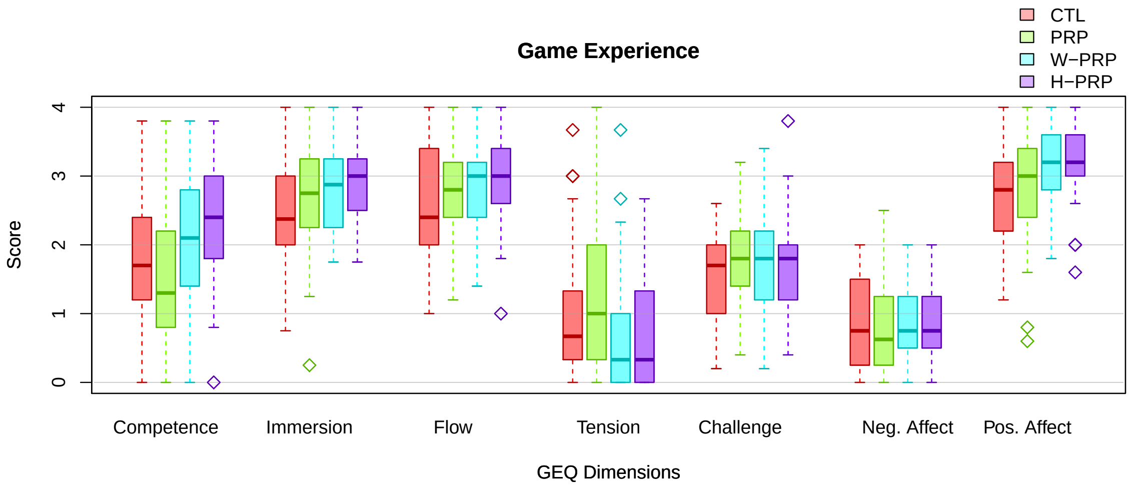


Figure 4: Box-and-whisker plot of Game Experience Questionnaire results for 7 GEQ dimensions, based on items ranked from 0 to 4.

Table 1: Significance (p) and effect size (Pearson's r) results for post-hoc tests comparing treatments for dimensions of the Game Experience Questionnaire (Values below 5% significance threshold in red). Note that Challenge and Negative Affect are omitted because ANOVA and Friedman tests, respectively, did not point to significant differences between treatments.

Treatment Comparisons	Competence		Immersion		Flow		Tension		Pos. Affect	
	p	r	p	r	p	r	p	r	p	r
CTL:PRP	0.372		0.304		0.960		1.000		1.000	
CTL:W-PRP	0.086		**<0.001**	0.807	**0.008**	0.656	0.089		**0.013**	0.507
CTL:H-PRP	**<0.001**	0.389	**<0.001**	0.586	**0.022**	0.470	0.390		**0.004**	0.383
PRP:W-PRP	**<0.001**	0.573	0.105		0.376		**0.021**	0.522	**0.012**	0.654
PRP:H-PRP	**<0.001**	0.401	**0.026**	0.543	0.163		0.054		**0.019**	0.411
H-PRP:W-PRP	0.338		0.951		1.000		1.000		1.000	

Table 2: Mean scores for dimensions of the Game Experience Questionnaire (with standard deviation in brackets) for each treatment.

Treatment	Competence	Immersion	Flow	Tension	Challenge	Neg. Affect	Pos. Affect
(CTL) Vive controller	1.73 (0.87)	2.49 (0.75)	2.56 (0.87)	0.99 (0.98)	1.58 (0.60)	0.86 (059)	2.79 (0.73)
(PRP) Simple Tracked Prop	1.49 (0.95)	2.64 (0.78)	2.71 (0.77)	1.11 (1.02)	1.79 (0.67)	0.75 (0.62)	2.87 (0.79)
(W-PRP) Prop with Weights	2.10 (0.91)	2.85 (0.58)	2.87 (0.72)	0.71 (0.88)	1.78 (0.68)	0.80 (0.52)	3.15 (0.62)
(H-PRP) Prop with Weights & Haptics	2.35 (0.91)	2.89 (0.62)	2.96 (0.69)	0.71 (0.85)	1.72 (0.66)	0.82 (0.55)	3.23 (0.57)

Table 1 shows the results of post-hoc tests of differences between treatments. The effect size (Pearson's r-value) is also included in cases where a significant difference was found. Generally, effect sizes are in the medium to large range.

4.2 Performance

Performance results for the batting task appear in Figure 5 (box-whisker plots) and Table 4 (means and standard deviations). Both sub-measures followed a normal distribution. In a within-subjects ANOVA test for controller effect, Mean Distance exhibited a significant effect ($\chi^2 = 29.2, p < 0.001$) but Hit Ratio did not ($\chi^2 = 4.96$, $p = 0.175$).

Post-hoc tests of differences in mean hit distance between treatments (Table 3) show statistically significant improvement with the weighted prop (W-PRP) over both the control (CTL) and unweighted prop (PRP), and with the Haptic Prop (H-PRP) over the unweighted prop (PRP). In all cases the effect sizes were medium to large ($r > 0.3$).

4.3 Pose

The proportion of participants favouring a two-handed bat grip and the traditional side-facing batters stance for each treatment is tabulated in Table 4. Both Grip ($Q = 43.8, p < 0.001$) and Facing ($Q = 8.4, p = 0.038$) registered controller effect with Cochran's Q

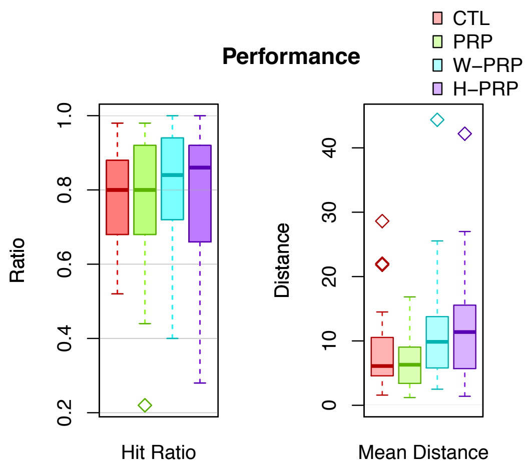


Figure 5: Box-and-whisker plot of Game Score for Hit Ratio and Mean Distance dimensions.

Table 3: Significance (p) and effect size (Pearson's r) results for post-hoc tests comparing treatments for Mean Distance, Grip and Facing (Values below 5% significance threshold in red). The performance measure Hit Ratio is omitted because a significant difference between treatments was not established.

Treat. Cmp.	Mean Dist.		Grip		Facing
	p	r	p	ϕ	p
CTL:PRP	0.591		**<0.001**	0.188	0.152
CTL:W-PRP	**0.024**	0.522	**<0.001**	0.133	1.000
CTL:H-PRP	0.106		**0.002**	0.171	1.000
PRP:W-PRP	**0.001**	0.385	1.000		0.500
PRP:H-PRP	**0.001**	0.448	1.000		0.273
H-PRP:W-PRP	1.000		1.000		1.000

Table 4: Mean scores and standard deviations for Performance and proportion adopting a particular Pose for each treatment.

Treatment	Score		Pose	
	Hit Ratio	Mean Dist.	2-hand Grip	Side Facing
CTL	0.77 (0.13)	8.31 (5.73)	0.58	0.61
PRP	0.77 (0.18)	6.27 (3.56)	0.98	0.73
W-PRP	0.80 (0.17)	11.00 (7.71)	0.98	0.66
H-PRP	0.79 (0.19)	11.78 (7.87)	1.00	0.63

Test. However, post-hoc tests (Table 3) reveal that the only significant difference was a shift towards greater use of a two-handed grip for the prop treatments (PRP, W-PRP, H-PRP) with small effect size ($\phi < 0.3$).

5 DISCUSSION

Our experimental results indicate that a haptic baseball prop generally improves the positive dimensions of Game Experience, namely Competence, Immersion, Flow and Positive Affect, and aspects of Performance, specifically mean hit distance, but only if the prop is appropriately weighted. In the case of the positive GEQ measures, the improvement is not entirely universal (CTL vs. W-PRP for Competence; PRP vs. W-PRP for Immersion; PRP vs. W-PRP and PRP vs. H-PRP for Flow show no significant difference), but it does represent an identifiable trend. We posit that participants are better able to control their swing, with concomitant improved game experience and performance, when the weighted prop accords with their real-world expectations and experience. Based on our demographics, participants had relatively little direct exposure to baseball play, but their experience with other bat-based sports would likely have transferred.

Across the board, there is no significant difference in any of our Game Experience or Performance measures for the Controller (CTL) as against the Unweighted Prop (PRP), or, for that matter, between the Weighted (W-PRP) and Active-haptic Prop (H-PRP). This further indicates that the passive shape-based haptics of the unweighted prop are not sufficient to overcome the discrepancy in expected weight. While the vibro-tactile feedback of H-PRP does not significantly improve in our measures, this represents a relatively simple form of active haptics. More sophisticated (and expensive) approaches might very well still be beneficial.

The decrease in tension (with a large effect size) when moving from an unweighted to weighted prop is surprising. Clearly this does not represent the tension associated with physical effort because then we would expect the trend to be reversed. Rather, this is likely due to the cognitive dissonance of swinging a baseball bat with the correct shape and texture, but without the correct weight.

The pose measures, as recorded in photographs, indicate that there was a shift from a slight majority (58%) with the controller towards almost universal adoption ($\geq$ 98%) with the various haptic props of a two-handed grip. We originally intended this as a proxy measure of presence (similar to behavioural reactions, such as sway, used in other studies [Freeman et al. 2000]). However, it could equally be explained by affordance: a two-handed grip is more comfortable for the haptic prop than the controller. Participants may also have seen a controller being wielded one-handed and thus be primed to adopt this grip. Any argument in favour of presence being the explanation is further weakened by the lack of significant change in stance (side-facing) among participants between treatments.

We also provided an opportunity for comments in the GEQ questionnaire. This qualitative feedback bore out the importance of appropriately weighting the prop. For instance, one participant commented: 'It's much easier to swing without weight but I found it more difficult to time the hitting and with no weight the perception of the amount of power I put into a swing is lost.' Among the 103 sets of comments, 35 specifically praised the weighting, while only 8 mentioned the vibration favourably. Although regarded as beneficial there was concern in some instances over the fatigue induced by the extra weight (e.g., 'Adding the extra weight helped make it feel more realistic, however, it also added to the physical endurance required.') This suggests that it might be beneficial to design haptic props with adjustable weights.

In terms of generalisation, we believe that these results may well apply to batting, striking or hitting tasks where the physical analog has a certain heft, contact forces are transitory, and accuracy is a factor. This category includes objects such as bats, swords, axes, hammers, and the like. In such cases it seems likely that appropriately-weighted haptic props will improve elements of task performance and user experience in both a gaming and non-gaming context.

5.1 Limitations

This study has two significant limitations. First, we investigate a relatively narrow class of haptic props. There are many other real-world tools exhibiting various combinations of tool heft, surface contact, and required precision for which our guidelines on prop design may well not be applicable. For instance, the design tradeoffs for lightweight precision props that mimic pens and paintbrushes are likely different. Second, our study tested only four of eight possible device configurations in order to ameliorate fatigue issues that could have impacted the experiment. A weighted controller, vibro-tactile controller, weighted vibrating controller and unweighted vibrating prop remain untested, and it is therefore not possible to make definitive statements about their effectiveness. In both cases, further investigation is warranted.

6 CONCLUSIONS

We undertook the quantitative evaluation of a generic VR controller as compared to progressively more sophisticated forms of a haptic baseball prop (unweighted, weighted, and weighted with vibro-tactile feedback) for a baseball game task. The weighted prop was found to significantly improve, with medium to large effect, aspects of both game experience and task performance as compared to the VR controller and unweighted prop. There was no significant difference between the weighted prop and its vibro-tactile enhancement.

Our results indicate that developing haptic props corresponding to the particular case of striking, batting or hitting tools with non-negligible weight, transitory contact, and a requirement for accuracy is worthwhile, but that correct shape and texture are insufficient, while vibro-tactile feedback provides only marginal benefit. The most important aspect is to approximate the weight of the original striking tool. This has implications for the design of certain haptic proxy tools for VR-based computer games, training and visualisation.

There are many further aspects of this design space worth exploring in follow-up work, such as the impact of more sophisticated active haptics, the case where real-world tools are unfamiliar to participants, or the transfer of skills obtained with haptic props to a real-world context. For instance, long-term improvement in skill through VR training with professional baseball players has been demonstrated [Gray 2017], but this was with a full-sized bat prop. Another avenue of future work is to explore similar considerations for props with different weight, accuracy, and contact characteristics.

ACKNOWLEDGMENTS

This project was funded by the THRIP program under grant number TP14080788335.

REFERENCES

Merwan Achibet, Adrien Girard, Anthony Talvas, Maud Marchal, and Anatole Lécuyer. 2015. Elastic-Arm: Human-scale passive haptic feedback for augmenting interaction and perception in virtual environments. In *2015 IEEE Virtual Reality (VR)*. 63–68.

Mahdi Azmandian, Mark Hancock, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. 2016. Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. 1968–1979.

Jeanne H. Brockmyer, Christine M. Fox, Kathleen A. Curtiss, Evan McBroom, Kimberly M. Burkhart, and Jacquelyn N. Pidruzny. 2009. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology* 45, 4 (2009), 624 – 634.

Lung-Pan Cheng, Li Chang, Sebastian Marwecki, and Patrick Baudisch. 2018. iTurk: Turning Passive Haptics into Active Haptics by Making Users Reconfigure Props in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Article 89, 10 pages.

Lung-Pan Cheng, Eyal Ofek, Christian Holz, Hrvoje Benko, and Andrew D. Wilson. 2017. Sparse Haptic Proxy: Touch Feedback in Virtual Environments Using a General Passive Prop. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. 3718–3728.

Alena Denisova, A. Imran Nordin, and Paul Cairns. 2016. The Convergence of Player Experience Questionnaires. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '16)*. 33–37.

Jonathan Freeman, S. E. Avons, Ray Meddis, Don E. Pearson, and Wijnand IJsselsteijn. 2000. Using Behavioral Realism to Estimate Presence: A Study of the Utility of Postural Responses to Motion Stimuli. *Presence: Teleoperators and Virtual Environments* 9, 2 (2000), 149–164.

Eisuke Fujinawa, Shigeo Yoshida, Yuki Koyama, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. 2017. Computational Design of Hand-held VR Controllers Using Haptic Shape Illusion. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (VRST '17)*. Article 28, 10 pages.

Rob Gray. 2017. Transfer of Training from Virtual to Real Baseball Batting. *Frontiers in Psychology* 8, Article 2183 (2017), 11 pages.

Daniel Harley, Aneesh P. Tarun, Daniel Germinario, and Ali Mazalek. 2017. Tangible VR: Diegetic Tangible Objects for Virtual Reality Narratives. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. 1253–1263.

Steven Henderson and Steven Feiner. 2010. Opportunistic Tangible User Interfaces for Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 16, 1 (Jan 2010), 4–16.

Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. Passive Real-world Interface Props for Neurosurgical Visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*. 452–458.

Wijnand IJsselsteijn, Yvonne de Kort, and Karolien Poels. 2013. *The Game Experience Questionnaire*. Technische Universiteit Eindhoven.

Seki Inoue, Yasutoshi Makino, and Hiroyuki Shinoda. 2015. Active touch perception produced by airborne ultrasonic haptic hologram. In *2015 IEEE World Haptics Conference (WHC)*. 362–367.

Brent Edward Insko. 2001. *Passive Haptics Significantly Enhances Virtual Environments*. Ph.D. Dissertation. Advisor(s) Brooks,Jr., Frederick P.

Bret Jackson, Tung Yuen Lau, David Schroeder, Kimani Toussaint, and Daniel Keefe. 2013. A Lightweight Tangible 3D Interface for Interactive Visualization of Thin Fiber Structures. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec 2013), 2802–2809.

Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies* 66, 9 (2008), 641 – 661.

Seokhee Jeon and Seungmoon Choi. 2009. Haptic Augmented Reality: Taxonomy and an Example of Stiffness Modulation. *Presence: Teleoperators and Virtual Environments* 18, 5 (2009), 387–408.

Daniel Johnson, M. John Gardner, and Ryan Perry. 2018. Validation of two game experience scales: The Player Experience of Need Satisfaction (PENS) and Game Experience Questionnaire (GEQ). *International Journal of Human-Computer Studies* 118 (2018), 38 – 46.

Benjamin Knoerlein, Gábor Székely, and Matthias Harders. 2007. Visuo-haptic Collaborative Augmented Reality Ping-pong. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE '07)*. 91–94.

Luv Kohli, Mary C. Whitton, and Frederick P. Brooks. 2013. Redirected Touching: Training and adaptation in warped virtual spaces. In *2013 IEEE Symposium on 3D User Interfaces (3DUI)*. 79–86.

Anatole Lécuyer. 2009. Simulating Haptic Feedback Using Vision: A Survey of Research and Applications of Pseudo-Haptic Feedback. *Presence: Teleoperators and Virtual*

Environments 18, 1 (2009), 39–53.

John C. McClelland, Robert J. Teather, and Audrey Girouard. 2017. Haptobend: Shape-changing Passive Haptic Feedback in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17)*. 82–90.

Lennart E. Nacke. 2010. Wiimote vs. Controller: Electroencephalographic Measurement of Affective Gameplay Interaction. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology (Futureplay '10)*. 159–166.

Kent Norman. 2013. GEQ (Game Engagement/Experience Questionnaire): A Review of Two Papers. *Interacting with Computers* 25, 4 (July 2013), 278–283.

Michaël Ortega and Sabine Coquillart. 2005. Prop-based haptic interaction with co-location and immersion: an automotive application. In *IEEE International Workshop on Haptic Audio Visual Environments and their Applications*. 6 pp.–. https://doi.org/10.1109/HAVE.2005.1545646

Mark Paterson. 2007. *The Senses of Touch: Haptics, Affects and Technologies*. London, UK: Bloomsbury.

Karolien Poels, Yvonne de Kort, and Wijnand IJsselsteijn. 2007. *Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games*. Technische Universiteit Eindhoven.

Wen Qi, Jean-Bernard Martens, Robert van Liere, and Arjan Kok. 2005. Reach the virtual environment: 3D tangible interaction with scientific data. In *Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future*. Computer-Human Interaction Special Interest Group (CHISIG) of Australia, 1–10.

Richard M. Ryan, C. Scott Rigby, and Andrew Przybylski. 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30, 4 (01 Dec 2006), 344–360.

Andreas Ryge, Lui Thomsen, Theis Berthelsen, Jonatan Hvass, Lars Koreska, Casper Vollmers, Niels Nilsson, Rolf Nordahl, and Stefania Serafin. 2017. Effect on high versus low fidelity haptic feedback in a virtual reality baseball simulation. In *2017 IEEE Virtual Reality (VR)*. 365–366.

Mohamed Suhail Mohamed Yousuf Sait, Shyam Prathish Sargunam, Dustin T. Han, and Eric D. Ragan. 2018. Physical Hand Interaction for Controlling Multiple Virtual Objects in Virtual Reality. In *Proceedings of the 3rd International Workshop on Interactive and Spatial Computing (IWISC '18)*. 64–74.

Adalberto L. Simeone, Eduardo Velloso, and Hans Gellersen. 2015. Substitutional Reality: Using the Physical Environment to Design Virtual Reality Experiences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. 3307–3316.

Rajinder Sodhi, Ivan Poupyrev, Matthew Glisson, and Ali Israr. 2013. AIREAL: Interactive Tactile Experiences in Free Air. *ACM Transactions on Graphics* 32, 4, Article 134 (July 2013), 10 pages.

Jonas Spillmann, Stefan Tuchschmid, and Matthias Harders. 2013. Adaptive Space Warping to Enhance Passive Haptics in an Arthroscopy Surgical Simulator. *IEEE Transactions on Visualization and Computer Graphics* 19, 4 (April 2013), 626–633.

Ivan Sutherland. 1965. The Ultimate Display. In *Proceedings of the Congress of the International Federation of Information Processing (IFIP)*. 506–508.

Colin Swindells, Alex Unden, and Tao Sang. 2003. TorqueBAR: An Ungrounded Haptic Feedback Device. In *Proceedings of the 5th International Conference on Multimodal Interfaces (ICMI '03)*. 52–59.

Unity Technologies. [n. d.]. Unity 3D. https://unity3d.com

Raquel Viciana-Abad, Arcadio Reyes Lecuona, and Matthieu Poyade. 2010. The Influence of Passive Haptic Feedback and Difference Interaction Metaphors on Presence and Task Performance. *Presence: Teleoperators and Virtual Environments* 19, 3 (2010), 197–212.

Emanuel Vonach, Clemens Gatterer, and Hannes Kaufmann. 2017. VRRobot: Robot actuated props in an infinite virtual environment. In *2017 IEEE Virtual Reality (VR)*. 74–83.

Yanqing Wang and Christine L. MacKenzie. 2000. The Role of Contextual Haptic and Visual Constraints on Object Manipulation in Virtual Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*. 532–539.

André Zenner and Antonio Krüger. 2017. Shifty: A Weight-Shifting Dynamic Passive Haptic Proxy to Enhance Object Perception in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 23, 4 (April 2017), 1285–1294.

Yiwei Zhao and Sean Follmer. 2018. A Functional Optimization Based Approach for Continuous 3D Retargeted Touch of Arbitrary, Complex Boundaries in Haptic Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Article 544, 12 pages.