Published in final edited form as:

Proc IEEE Symp 3D User Interfaces. 2013 March ; 2013: 79–86. doi:10.1109/3DUI.2013.6550201.

Redirected Touching: Training and Adaptation in Warped Virtual Spaces

Luv Kohli^{*}, Mary C. Whitton^{*}, and Frederick P. Brooks Jr.^{*}

The University of North Carolina at Chapel Hill

Abstract

Redirected Touching is a technique in which virtual space is warped to map many virtual objects onto one real object that serves as a passive haptic prop. Recent work suggests that this mapping can often be predictably unnoticeable and have little effect on task performance. We investigated training and adaptation on a rapid aiming task in a real environment, an unwarped virtual environment, and a warped virtual environment.

Participants who experienced a warped virtual space reported an initial strange sensation, but adapted to the warped space after short repeated exposure. Our data indicate that all the virtual training was less effective than real-world training, but after adaptation, participants trained as well in a warped virtual space as in an unwarped one.

Keywords

Virtual reality; haptics; perception; pointing; Fitts' law; ISO 9241–9; redirected touching

Index Terms

H.5.1 [Information Interfaces and Presentation]; Multimedia Information Systems-Artificial; augmented; virtual realities; H.5.2 [User Interfaces]; Haptic I/O; I.3.7 [Computer Graphics]; Three-Dimensional Graphics and Realism-Virtual reality

1 Introduction

Military aircraft pilots and maintenance crews must learn to perform cockpit procedures, such as the sequences of buttons and switches required for aircraft start-up, shut-down, and emergencies. Real aircraft and full flight simulators can be used to train these skills, but can cost hundreds of dollars or more per hour [29]. Full simulators are unavailable in deployed settings [1], and for many procedures they are not required; low-fidelity trainers with mock cockpits can effectively train cockpit procedures [20].

One of the challenges trainers face is training “muscle memory” and spatial knowledge [5, 34]. Pilots must use checklists for many procedures, but they must know the relative

locations of cockpit controls so that they can be accessed efficiently when needed. Mock cockpits enable trainees to reach for cockpit controls. However, they are designed for a specific aircraft. For a new aircraft model, another mockup must be built, or the existing one must be changed [33]. Reconfigurable trainers do exist, but they typically support reconfiguration by repositioning or swapping out components [18, 24, 31]. We instead approach reconfiguration using *Redirected Touching* [13, 15]: a virtual, perception-based technique that may enable a single quickly setup physical mockup to represent many virtual cockpits without the need to change it for new aircraft.

Mock cockpits are a form of passive haptic displays—physical props to which virtual objects are registered. Passive haptics are commonly used in virtual environments (VEs) to provide compelling touch feedback. Evidence indicates that passive haptics are generally preferred over absent touch cues and can improve both performance for precision tasks [17] and spatial knowledge [10]. However, they are inflexible: changes to virtual objects require corresponding changes to physical objects, and vice versa. Typically, passive haptic props are mapped one-to-one to their virtual counterparts. Recent work has warped virtual space to map many virtual objects onto a single physical object [2, 15], introducing a discrepancy between the seen and felt hands: the virtual hand's motion (in warped space) matches what is seen, and the real hand's motion matches what is felt (Figure 1, upper left and right). Vision usually dominates proprioception when the two conflict; people tend to believe their hand is where they see it, rather than where they feel it [4, 9, 21, 22].

We previously investigated angular discrepancy in the context of aircraft cockpit procedures training [15]. For a vertically oriented 10"×10" board, an angular discrepancy between seen and felt boards of 18° around the vertical axis yielded acceptable task performance relative to a one-to-one object mapping. We also saw users adapt to the discrepancy. Study participants experienced several different angular discrepancies for short periods of time, requiring frequent adaptation. However, in an actual training scenario, users would typically interact with only one discrepancy. For example, though a system using Redirected Touching might support many aircraft, a pilot learning a cockpit procedure would train on a specific cockpit. To confidently use Redirected Touching for training, we need to know how well users perform with a longer exposure to a single warped space.

Ideally, we would investigate task performance on a high-level task, such as remembering a long sequence of buttons and switches for a cockpit procedure. However, there is still much that we do not know about how Redirected Touching affects users. If users must adapt to a warped space, it is possible their attention will be divided between adaptation and learning the high-level skill, which could negatively impact learning. Also, real-world task performance may suffer until users readapt to the real world. We therefore conducted a study extending our earlier work, to more deeply investigate the effects of interacting with a warped space:

1. During a long exposure to a discrepant VE, how does task performance compare to an equivalent exposure to a one-to-one VE and a real environment?
2. Do users adapt to the discrepancy? How quickly?

3. How quickly do users readapt to the real world?
4. Do users notice the discrepancy?

2 Measuring Task Performance

To measure task performance, we chose the standardized ISO 9241–9 multi-directional tapping task (Figure 2) [11, 15, 27, 28]. This task reasonably approximates pushing a button sequence and is arguably applicable to cockpit procedures training. The task presents the user with a ring of circular targets which light up individually in sequence. The user touches the lit target, the next target lights, the user then touches it, and so on. The sequence is always the same. A trial consists of a user's finger leaving one target and touching the next (specific) target at an *endpoint*. Speed and accuracy are measured.

The multi-directional tapping task is based on Fitts' law [7], a predictive model relating movement time to distance and accuracy in rapid aimed movements. Fitts' law is defined as:

$$MT = a + b \cdot ID \text{ where } ID = \log_2\left(\frac{A}{W} + 1\right) \quad (1)$$

MT is *movement time* (seconds), a and b are empirically determined constants, ID is the Shannon formulation of the *index of difficulty* (bits), A is *amplitude* (the distance between successive targets), and W is *target width*.

To compensate for users focusing on one of speed or accuracy, a common measure of Fitts' law task performance is the single combined metric *throughput*, or TP . TP is measured in bits per second (bps) and is defined as:

$$TP = \frac{ID_e}{MT} \text{ where } ID_e = \log_2\left(\frac{A_e}{W_e} + 1\right) \quad (2)$$

Because one user may tap very accurately and another very inaccurately, the difficulty of the task *performed* may be different from the task presented. Hence, ID is modified to the *effective index of difficulty* (ID_e) for each user, using *effective amplitude* (A_e) and *effective target width* (W_e). A_e is the mean movement distance between pairs of endpoints for a single target ring. W_e is defined as:

$$W_e = 4.133 \cdot SD_x \quad (3)$$

To compute W_e , we rotate all trials for a target ring to horizontal and project all endpoints onto the horizontal axis, yielding SD_x , the standard deviation of those values. This W_e calculation is the *adjustment for accuracy*, and accounts for the spread of touch endpoints for the task the user actually performed [19, 25].

3 Study

3.1 Apparatus

The experimental task is presented to the user on either a resistive touchscreen monitor when in the real world or in an NVIS nVisor SX head-mounted display (HMD) when in the VE. A 3rdTech HiBall-3000 tracks the user's head and a PhaseSpace IMPULSE tracking system tracks the right index finger and the touchscreen monitor. VPRN communicates with trackers [26]. We render using Gamebase Co. Ltd's Gamebryo game engine running on a hex-core 3.3GHz Intel i7-3960X machine with 16GB of RAM and an NVIDIA GeForce GTX 680 GPU. End-to-end system latency is ~50–60ms. We correct for pincushion distortion using the HMD calibration work by Kuhl et al. [16]. We overscan our texture input to the pincushion correction shader to avoid dark borders (unlit pixels).

The touchscreen monitor is mounted on a rotating base (Figure 1, lower left). The study task is presented in a 10"×10" area of the monitor to be consistent with our previous study [15].

3.2 Conditions

Typically in Fitts'-law studies, many target widths and distances are presented to test a wide index-of-difficulty range. Data are aggregated across those indices of difficulty to yield overall task performance for a particular experimental condition (e.g., a particular user interface being tested) [25]. In this study, we investigated angular discrepancy, so points near the target ring's center did not exhibit much discrepancy (Figure 1). We therefore chose to use multiple target widths with a single target distance. Our targets were 21cm apart and six different sizes: 10mm, 15mm, 20mm, 25mm, 30mm, and 35mm.

The study had three phases:

1. Pretest: 2 target rings, the largest and the smallest
2. Training: 72 target rings, presented in twelve blocks of six rings, one ring of each target width
3. Posttest: 18 target rings, in three blocks of six

The Pretest phase was very short to avoid too much training prior to the Training phase. In the Training and Posttest phases, the presentation order of different target widths was randomized without replacement. Target rings were presented in the same order to all participants.

The study used a between-subjects design. Participants were assigned to one of three Training-phase conditions:

1. Real (R): task done on real touchscreen monitor oriented at 18°, with no head-mounted display (HMD).
2. Virtual one-to-one (V_o): task done in HMD, with real and virtual monitors both oriented at 18°.
3. Virtual discrepant (V_d): task done in HMD, real monitor at 0° and virtual monitor at 18°. Virtual space was warped using thin-plate splines (Figure 3) [3, 6, 15, 23].

In all conditions, the Pretest and Posttest phases were done on the touchscreen monitor at 18° without the HMD.

3.3 Participants

42 paid participants (26 male, 16 female, aged 18–33, mean 21) took part; 14 were randomly assigned to each condition. All were right-handed and had normal or corrected-to-normal vision.

3.4 Analysis: Non-Inferiority

Our dependent variables measuring task performance were throughput, error rate, and movement time.

Although it is common to test for statistically significant differences between interface techniques, we want to know whether task performance while using a discrepant interface is *good enough* for a given application. We did not expect the V_d condition to produce better results than V_o or R. We would like to show that task performance in V_d is *no worse* than in V_o and R, which we test for by using a non-inferiority test, a form of equivalence testing [30].

Non-inferiority testing requires us first to define an indifference zone for our metric in the context of our application. We then calculate the mean difference of throughput between our conditions and its one-tailed 95% confidence interval. If the entire 95% confidence interval lies within our indifference zone, we consider the metric to be no worse than in the other condition.

3.5 Indifference zones

For our metrics, we chose the same indifference zones as in our previous study [15]. The maximum allowable difference between conditions for each metric was:

- Throughput: 1bps. A 2004 survey of ISO 9241–9 studies found that the range of throughputs for computer-mouse pointing in five studies was 3.7–4.9bps, a range of 1.2bps [25]. As before, we used the more conservative 1bps.
- Error rate: 1/11 missed targets (9%), the smallest unit of error we could measure.
- Movement time: 100ms. The human visual feedback loop is estimated to operate at 100–200ms [12] so we chose the low end of that range, 100ms.

3.6 Study procedure

Participants read and signed an informed consent form, entered the lab, and sat in front of the touchscreen. The touchscreen was oriented at 18° about the vertical axis (Figure 1, lower left). We explained the multi-directional tapping task. Participants were instructed to perform the task as quickly and as accurately as possible and to leave missed targets uncorrected. A white dot showed where users touched the touchscreen, and a red X was shown on missed targets.

Participants were fitted with the finger trackers, calibrated for finger length. In all three conditions, participants were presented with two Pretest target rings. R condition participants moved directly to the Training phase of the study. V_o and V_d participants were fitted with the HMD, which initially showed a white “focus” screen with a black rectangle. We faded participants from the focus screen into the VE, where they saw a virtual touchscreen monitor (at 18°) and their virtual hand (Figure 1, upper left).

After participants tried touching the virtual screen once or twice, they were faded out to the focus screen again so the experimenter could reiterate instructions. These instructions were used in both virtual conditions, but were designed particularly for the V_d condition. The V_d condition proceeded similarly except that the experimenter quietly rotated the real touchscreen monitor so that it was oriented at 0° (Figure 1, lower right). When V_o and V_d participants were faded back into the VE, they were shown a virtual touchscreen monitor at 18° . In the V_d case, space was warped to map the 18° virtual monitor onto the 0° real monitor.¹ Participants then did the multi-directional tapping task 72 times. Participants were required to take a one-minute break after 24 and 48 target rings in the Training phase. They could also take breaks if desired before beginning any target ring.

After the Training phase, R participants went on to the 18-ring Posttest. For the V_o and V_d participants, the VE again faded out to the focus screen. More instructions were given, and in the V_d condition, the touchscreen was quietly rotated back to 18° . The HMD was then removed, and participants did the Posttest.

Participants then filled out a short post-questionnaire. R condition participants were allowed to try out the HMD. Where R participants tried on the HMD, they did a shortened version of the V_d condition to yield more feedback on discrepancy. The real monitor was quietly rotated when they entered the VE, but the HMD was removed without the experimenter rotating the monitor back, to see each participant’s reaction to the rotated monitor.

All participants then had an open-ended interview, were told the study’s purpose, and were paid. The study lasted 45–60 minutes total per participant, the experimental conditions 15–25 minutes.

4 Results: Task Performance

Our results represent data after removing outliers (undetected or double touchscreen touches, forgotten targets, and overly long motion paths). In the graphs below, the horizontal axis is divided by study phase and target ring block: *Pre 01–02* is the Pretest, *Trn* are the twelve Training blocks, and *Post* are the three Posttest blocks. Each data point is the mean for a particular block, across all participants in a particular condition. The error bars show 95% confidence intervals. The vertical dashed lines are the beginning of the Training phase,

¹During target rings, the warped space was computed using only the real and virtual monitor geometries and a distant boundary beyond which space was not warped; the table was not used. Consequently, when touching the real table straight on, the virtual finger would be shown at an angle. So that participants would not notice this discrepancy during breaks, we changed the warping behavior between target rings. The hand pose was interpolated between the warped and one-to-one spaces: the closer the finger to the monitor, the more warped the space, and the closer to the table edge, the more one-to-one the space.

the target rings after the first break, the target rings after the second break, and the Posttest phase.

Figures 4–6 show mean throughput, error rate, and movement time respectively across all participants.

4.1 Posttest vs. Pretest

Effective training for the real-world task should show as a Posttest-Pretest difference. Figure 4 shows a general upward trend in all three conditions, although there is a drop at the beginning of the Training phase for both virtual conditions, probably related to users adjusting to the VE. We suspected readaptation to the real world affected performance during the first Posttest target ring block. Our focus here is on performance after (re)adaptation, so that block was removed from analysis.

We tested for significant differences between the Posttest and Pretest blocks using an ANOVA (condition: between-subjects; target ring block: within-subjects). There were significant differences in throughput (Table 1) and movement time (Table 3) for all conditions. For error rate, there was a significant difference for only the R condition (Table 2). In all cases where there was a significant difference, the bound of the associated one-tailed 95% confidence lies outside our predetermined indifference zone, so non-inferiority was not shown.

In Figure 5, the R Pretest error rate appears different from V_o and V_d (though within 2/11 missed targets). Compared to Training and Posttest, Pretest was more susceptible to between-subject variability: there were fewer Pretest rings. Pretest included only the largest and smallest target widths, and error rates were consistently highest on the smallest. Because a target is either hit or missed, small variations in movement speed can lead to very different error rates. Pretest movement speeds were similar, but R participants varied more. We are not confident that the R Posttest-Pretest error rate represents a true difference. However, Pretest throughput values across all participants were similar (Figure 4), so we are more confident in the other results.

4.2 Non-inferior task performance

We have seen that there was some training between the Pretest and Posttest phases of each condition. Ideally, we would also find V_d no worse than R and V_o . We computed the mean differences between each pair of conditions and plotted their one-tailed 95% confidence intervals, as well as the indifference zones. With that data, we can test for non-inferiority, and also see when during the sequence of target rings non-inferiority is achieved, if at all.

Figure 7 shows mean throughput differences and confidence intervals. The shaded region is the 1bps indifference zone. For the Pretest, the confidence intervals all lie within the indifference zone, meaning that V_d is no worse than R and V_o , and V_o is no worse than R. However, as soon as the Training phase begins, we can no longer say this. V_d and V_o get better over time, but never quite get to be no worse than R. After about 18 target rings, V_d becomes no worse than V_o and remains that way for most of the Training phase. Upon

starting the Posttest, V_d participants' performance degrades and then improves as they readapt to the real world, until the end when V_d is again no worse than V_o .

Figure 8 shows mean error rate differences and confidence intervals, with our 9% indifference zone shaded. We generally see non-inferiority between V_d and V_o , and V_o and R. However, the V_d -R confidence intervals lie outside the indifference zone for the period before the first Training break, and then hover near the boundary of the indifference zone. By the end of the Posttest, V_d is no worse than V_o and R, and V_o is no worse than R.

Figure 9 shows mean movement time differences and confidence intervals, with our 100ms indifference zone shaded. In the Training and Posttest phases, V_d and V_o are never no worse than R, and V_d is never no worse than V_o . These results are consistent with our observations that participants moved faster in the R condition, and we suspect that adaptation to discrepancy led to slower movement times in the V_d condition.

5 Measuring Adaptation by Path Variability

Most participants in the V_d condition reported a strange sensation when touching the monitor: something felt odd about their depth perception when first exposed to the discrepant environment, and again when returning to the real world, but the feeling mostly went away after doing the task repeatedly. These reports are consistent with Gibson's findings [9]: participants wearing vision-distorting lenses adapted to a distorted visual field, and after removing the lenses, participants experienced a negative aftereffect (distortion in the opposite direction) until they readapted to the undistorted real world. We suspect that our participants experienced something similar.

Adaptation has potential implications for training: if users adapt to a discrepant environment, they should readapt quickly to the real environment so that they are not mistrained. We used path variability to measure whether participants adapted to the discrepant VE and how long it took them to adapt and readapt.

Figures 10–11 show one V_d participant's motion paths for the first Training target ring and a target ring just after the first break. Note that in the first target ring, the paths are not particularly smooth. Figure 12 shows the paths viewed from above. The first target ring's left-to-right (e.g., target 4 to target 5) and right-to-left (e.g., target 5 to target 6) paths exhibit a skew. Because the real monitor was oriented at 0° and the virtual monitor at 18° , participants expected to need to pull their real finger back toward them when moving left, when in fact the finger needed to move directly left to stay near the monitor. As a result, we see sharp turns back towards the monitor on the left side when participants realize their finger is not moving where they expect. On the right side, participants often hit the monitor early. In the target ring just after the first break, the paths are much smoother, and there is no apparent skew despite the continued 18° discrepancy.

The paths in Figure 10 are quite different from one another, whereas the paths in Figure 11 are alike. These differences between motion paths were seen for all V_d participants, indicating that participants were (consciously or otherwise) adapting to the warped space. We hypothesized that as participants adapt to a discrepant environment, their motion paths

become more similar. We tested this hypothesis using a path variability metric we devised earlier [14], similar to that of Georgopoulos et al. [8].

All of the paths in a given target ring move in different directions and have different lengths and movement times. To compare them, we normalized them:

- Each path was rotated onto a common task axis (Figure 13).
- Each path was then resampled at 50 equally spaced time intervals between its start and end touch points. Position data were interpolated for these sample points.
- For each of the 50 sample points, we collected the samples from all rotated paths and computed their 3D standard deviation as a measure of path variability at that time sample (Equation 4).
- We summed these path variability samples to get overall path variability for a single target ring.

The 3D standard deviation equation was as presented by Wobbrock, et al. [32]:

$$SD_{x,y,z} = \sqrt{\frac{\sum_{i=1}^N (\sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2 + (z_i - \bar{z})^2})^2}{N - 1}} \quad (4)$$

We hypothesized that path variability would decrease with more exposure to the discrepant condition, and that after adaptation, V_d path variability would be no higher than V_o and R path variability. We used the R condition data to define our indifference zones.

We computed path-variability standard deviation for each R-condition target ring block, and used one standard deviation as that block's indifference zone. Consequently, our indifference zones change slightly from block to block.

6 Results: Adaptation

Figure 14 shows mean path variability across all participants, by block. Note that the path variability starts at about the same level in the Pretest, but the path variability is much higher in V_d than in V_o and R when the Training phase starts. With repeated exposure to the discrepancy, V_d path variability decreases and appears to converge with V_o and R. When the Posttest starts, there is another spike in V_d path variability (and a smaller spike in V_o as well), but the three conditions appear to converge again after the first Posttest target ring block.

Figure 15 shows mean path variability differences and confidence intervals, with our indifference zones shaded. We see that in the Pretest, V_d and V_o are no worse than R, and V_d is no worse than V_o . When the Training phase begins, V_o is no worse than R, but V_d-V_o and V_d-R are substantially outside the indifference zone. However, with continued exposure to the task, path variability across all three conditions begins to look more alike. After the first break time, V_d is mostly no worse than V_o , and V_d is no worse than R. We see small spikes in path variability at the first and second break times. We suspect these spikes are due

to having to readapt slightly to the discrepant environment after having rested. Another spike is evident at the beginning of the Posttest, which indicates that participants needed to readapt to the real world. That is consistent with our observations and participant comments. After the first Posttest target ring block, V_d is once again no worse than V_o and R, and V_o is no worse than R.

It is interesting to note that in general, V_o path variability was lower than R path variability. We suspect this is because R participants tended to move faster than V_o participants.

7 Discussion

We cannot definitively say that Posttest error rates were better than Pretest error rates for any condition; but for all conditions, Posttest throughput and movement time were better than in the Pretest. These results suggest that training for the real task occurred in all conditions and that training in the virtual world transferred to the real world. However, training in the real condition seemed more effective than in either virtual condition.

For the virtual conditions, Posttest performance was possibly due to training during Posttest rather than during Training. However, there was a performance drop when Posttest started, particularly in the V_d condition. Path variability metrics and participant comments indicate this drop was due to readaptation to the real world. The magnitude of the subsequent performance increase throughout Posttest is unlikely to be explained by continued real-world training. We argue that the initial decrease in task performance in the V_d Posttest was due to readaptation, and the overall increase relative to Pretest was due to repeated VE task exposure. However, we cannot be certain without a different study design. Ideally, we would test with a higher-level cognitive task that could not be learned in the span of the Posttest.

We generally observed that participants performed much better once adapted to the warped space: after adaptation, V_d throughput was generally no worse than in V_o , and V_d error rates were generally no worse than in V_o and R. However, V_d and V_o participants were not able to perform as quickly as in R. Aside from the discrepancy, factors such as HMD weight and limited FOV may have caused participants to move more slowly.

A decrease in path variability coupled with an increase in task performance clearly indicates that participants adapted to the discrepancy. Though it usually took until the first break time (~5–6 min.) for participants to fully adapt, most of the adaptation occurred during the first six Training target rings (~1–2 min.). We suggest that training with Redirected Touching may be effective if a user is first given time to adapt to a discrepant environment.

7.1 Participant comments: discrepancy detected?

We asked participants to describe their study experience. We wanted to find out whether people noticed the 18° discrepancy between the real and virtual monitors, particularly for the V_d participants and those R participants who experienced an abbreviated version of the V_d condition.

All 14 V_d participants felt a strange sensation when first starting the virtual task and when first starting the Posttest. A typical comment was that depth perception felt a little off, but

that by the first break time or sooner, it felt more normal. Several participants mentioned that until they grew accustomed to the VE, it felt like the virtual hand lagged behind their real hand. This perceived lag was likely due to the warped space causing discrepant real and virtual hand motion velocities. 10 of the 14 V_d participants could not identify that an orientation discrepancy caused the strange sensation; the other 4 did.

Seven R participants tried the abbreviated V_d condition. Five did not notice that the real monitor was oriented differently than the virtual monitor, and when the HMD was removed, were surprised to see the monitor's orientation. Two correctly identified the different orientation.

All but 2 of the 21 participants said that their perception in the VE began to feel more normal after doing the task a few times. One V_d participant noted that about half-way into the Training phase, he no longer had to think about his hand motion—it became more instinctive. All V_d participants found the task to become easier in the Posttest after an initial strange feeling, indicating readaptation.

7.2 Other participant comments

Two V_o participants felt they moved more slowly in the virtual world. Our observations and data confirmed that. About half responded that they did not experience any fatigue. Others experienced only moderate fatigue; only one indicated that it may have slightly affected their performance. Otherwise, participants found rest opportunities sufficient.

The most common complaints in the virtual conditions were that participants had to look down to see the entire virtual monitor (likely due to the HMD's limited field of view), and the HMD was too heavy, leading to some neck strain. Some would have preferred the chair lowered.

2 of the 28 virtual condition participants felt dizzy near the end of the experience, but felt fine shortly. One V_o participant felt dizzy around the first Training break time and asked to have the HMD removed. He then drank some water, and he did not subsequently mention feeling dizzy.

8 Conclusions and Future Work

We cannot claim that V_o and V_d are generally no worse than R, but evidence indicates that V_d is no worse than V_o for throughput, error rate, and path variability. Whether training in a discrepant environment is useful depends largely on whether training in a VE at all is useful. Also, it is unknown what effect discrepant environments have on spatial knowledge acquisition. We suspect that Redirected Touching would work better for a higher-level task that does not require as much rapid motor control. For example, a button sequence may be easier to learn in a discrepant environment than tasks that require rapid motion from point to point on a virtual object.

If participants were given time to adapt to the discrepant environment before Training and to readapt to the real environment before Posttest, V_d would likely end up even closer to V_o in all performance metrics. These adaptation phases could be as simple as tapping the warped

virtual object or the real object for about a minute. Also, foreknowledge of the discrepancy may reduce adaptation time.

We believe that Redirected Touching continues to be worth exploring. Despite a few indications of degraded performance in discrepant environments, evidence suggests that after adaptation, users can perform tasks in a discrepant VE generally no worse than in a one-to-one VE.

Acknowledgments

We thank the Link Foundation and the NIH National Institute for Biomedical Imaging and Bioengineering for supporting this work. We thank John Thomas for building our study's physical setup; Rohit Saboo for help with warping code; Chris Wiesen for statistical consulting; Kan Anant at PhaseSpace for technical support; and Rob Teather, Eric Burns, Sharif Razzaque, Sachin Patil, and Lt. Col. Paul Holst, USAF for valuable discussion. We thank the anonymous reviewers for their helpful suggestions. The hand model is © 2013 CyberGlove Systems.

References

1. Andre TS, Bennett WR, Castillo RA, McClain MD, Purtee DP. Generalized operations simulation environment for aircraft maintenance training. RTO-MP-HFM-101: Advanced Technologies for Military Training. 2004
2. Ban Y, Kajinami T, Narumi T, Tanikawa T, Hirose M. Modifying an identified angle of edged shapes using pseudo-haptic effects. *EuroHaptics*. 2012:25–36.
3. Bookstein FL. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans Pattern Anal Mach Intell*. 1989 Jun.11:567–585.
4. Burns E, Razzaque S, Panter AT, Whitton MC, McCallus MR, Brooks FP Jr. The hand is more easily fooled than the eye: users are more sensitive to visual interpenetration than to visual-proprioceptive discrepancy. *Presence: Teleoper. Virtual Environ*. 2006 Feb.15(1):1–15.
5. Degani A, Wiener EL. Cockpit checklists: Concepts, design, and use. *Human Factors*. 1993 Jun. 35(2):345–359.
6. Geometric Tools Interpolation. 2012 [Online]. Available:<http://www.geometrictools.com/LibMathematics/Interpolation/Interpolation.html>.
7. Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol*. 1954 Jun.47(6):381–391. [PubMed: 13174710]
8. Georgopoulos AP, Kalaska JF, Massey JT. Spatial trajectories and reaction times of aimed movements: effects of practice, uncertainty, and change in target location. *J Neurophysiol*. 1981 Oct.46(4):725–743. [PubMed: 7288461]
9. Gibson JJ. Adaptation, after-effect and contrast in the perception of curved lines. *J Exp Psychol*. 1933 Feb.16(1):1–31.
10. Insko, BE. Ph.D. dissertation. UNC Chapel Hill: Dept of Comp Sci; 2001. Passive Haptics Significantly Enhances Virtual Environments.
11. International Organization for Standardization. ISO 9241–9:2000, Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices. Multiple. Distributed through American National Standards Institute (ANSI); 2007.
12. Jones, LA.; Lederman, SJ. *Human Hand Function*. 1st ed.. USA: Oxford University Press; 2006.
13. Kohli, L. *IEEE 3DUI*. Washington, DC, USA: 2010. Redirected touching: Warping space to remap passive haptics; p. 129–130.
14. Kohli, L.; Waegel, K.; Whitton, MC.; Brooks, FP, Jr. Redirected touching: Space warping, task performance, and adaptation. Dept of Comp Sci; UNC Chapel Hill: 2013. Tech. Rep. TR13-001
15. Kohli L, Whitton MC, Brooks FP Jr. Redirected touching: The effect of warping space on task performance. *IEEE 3DUI*. 2012:105–112.
16. Kuhl SA, Thompson WB, Creem-Regehr SH. HMD calibration and its effects on distance judgments. *ACM APGV*. 2008:15–22.

17. Lindeman RW, Sibert JL, Hahn JK. Towards usable VR: An empirical study of user interfaces for immersive virtual environments. *ACM CHI*. 1999:64–71.
18. Lockheed Martin. Multi-Function Training Aid. [Online]. Available: <http://www.lockheedmartin.com/us/products/MFTA.html>.
19. MacKenzie IS. Fitts' law as a research and design tool in human-computer interaction. *Hum.-Comput. Interact.* 1992 Mar.7(1):91–139.
20. Prophet WW, Boyd HA. Device-task fidelity and transfer of training: Aircraft cockpit procedures training, HumRRO. Technical Report 70–10. 1970 Jul.
21. Pusch A, Martin O, Coquillart S. HEMP-hand-displacement-based pseudo-haptics: A study of a force field application and a behavioural analysis. *Int J Hum-Comput St.* 2009 Mar.67(3):256–268.
22. Rock I, Victor J. Vision and touch: An experimentally created conflict between the two senses. *Science*. 1964 Feb.143(3606):594–596. [PubMed: 14080333]
23. Saboo, R. Ph.D. dissertation. UNC Chapel Hill: Dept of Comp Sci; 2011. Atlas Diffeomorphisms via Object Models.
24. Schiefele J, Albert O, van Lier V, Huschka C. Simple force feedback for small virtual environments. *Proceedings of SPIE*. 1998 Aug.3367(1):100–110.
25. Soukoreff RW, MacKenzie IS. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *Int J Hum-Comput St.* 2004 Dec.61:751–789.
26. Taylor RM II, Hudson TC, Seeger A, Weber H, Juliano J, Helser AT. VRPN: A device-independent, network-transparent VR peripheral system. *ACM VRST*. 2001:55–61.
27. Teather RJ, Natapov D, Jenkin M. Evaluating haptic feedback in virtual environments using ISO 9241–9. *IEEE VR*. 2010:307–308.
28. Teather RJ, Stuerzlinger W. Pointing at 3D targets in a stereo head-tracked virtual environment. *IEEE 3DUI*. 2011:87–94.
29. Vincenzi, DA.; Wise, JA.; Mouloua, M.; Hancock, PA., editors. *Human Factors in Simulation and Training*. 1st ed.. CRC Press; 2008.
30. Wellek, S. *Testing Statistical Hypotheses of Equivalence and Noninferiority*. 2nd ed.. Chapman and Hall/CRC; 2010.
31. Williams GW, Lawrence K, Weeks R. Modeling and simulation technologies: Reconfigurable flight simulators in modeling and simulation, in. *AIAA Modeling and Simulation Technologies*. 2004
32. Wobbrock JO, Shinohara K, Jansen A. The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models. *ACM CHI*. 2011:1639–1648.
33. Wurpts M. Poster: Updating an obsolete trainer using passive haptics and pressure sensors. *IEEE 3DUI*. 2009:155–156.
34. JSUPT Phase I - Academics and Preflight Training. [Online]. Available: http://www.baseops.net/militarypilot/phase1_upt.html.

Figure 1.

A user sees an 18° monitor in an HMD (upper left) while touching a 0° touchscreen monitor (upper right); Lower left: touchscreen oriented at 18° ; Lower right: touchscreen at 0° .

Figure 2.
The multidirectional tapping task. Overlaid arrows show the target sequence.

Figure 3.
Viewed from above, horizontal slices of: unwarped space and real monitor (left); warped space and virtual monitor (right)

Figure 4.

Mean throughput across all participants for R, V_o, and V_d conditions. Error bars are 95% confidence intervals.

Figure 5.

Mean error rate across all participants for R, V_o, and V_d conditions. Error bars are 95% confidence intervals.

Figure 6.

Mean movement time across all participants for R, V_o, and V_d conditions. Error bars are 95% confidence intervals.

Figure 7.

Mean throughput differences between pairs of conditions. Error bars are one-tailed 95% confidence intervals. Shaded area represents non-inferiority indifference zone.

Figure 8.

Mean error rate differences between pairs of conditions. Error bars are one-tailed 95% confidence intervals. Shaded area represents non-inferiority indifference zone.

Figure 9.

Mean movement time differences between pairs of conditions. Error bars are one-tailed 95% confidence intervals. Shaded area represents non-inferiority indifference zone.

Figure 10.

A V_d participant's motion paths, first Training target ring.

Figure 11.
A V_d Training target ring just after the first break.

Figure 12.

A V_d participant's motion paths (viewed from above) for the first Training target ring (left) and just after the first break (right).

Figure 13.

A V_d participant's rotated motion paths for the first Training target ring (left) and just after the first break (right).

Figure 14.

Mean path variability across all participants for R, V_o , and V_d conditions. Error bars are 95% confidence intervals.

Figure 15.

Mean path variability differences between pairs of conditions. Error bars are one-tailed 95% confidence intervals. Shaded area represents non-inferiority indifference zone.

Table 1

ANOVA results for mean throughput differences

Condition	Mean throughput difference (Post-Pre)	One-tailed 95% conf. interval	Result
R	1.82 bps	< 2.16 bps	$F(1,711)=76.91$ $p < 0.0001$ significant
V_o	1.15 bps	< 1.50 bps	$F(1,711)=31.02$ $p < 0.0001$ significant
V_d	0.70 bps	< 1.05 bps	$F(1,711)=10.96$ $p = 0.001$ significant

Table 2

ANOVA results for mean error rate differences

Condition	Mean error rate difference (Post-Pre)	One-tailed 95% conf. interval	Result
R	-7%	> -10.8%	$F(1,711)=9.80$ $p = 0.0018$ significant
V _o	-3%	> -6.8%	$F(1,711)=1.88$ $p = 0.17$ not significant
V _d	-0.1%	> -3.9%	$F(1,711)=0.0025$ $p = 0.96$ not significant

Table 3

ANOVA results for mean movement time differences

Condition	Mean movement time difference (Post-Pre)	One-tailed 95% conf. interval	Result
R	-162ms	> -195ms	$F(1,711)=63.52$ $p < 0.0001$ significant
V _o	-107ms	> -141ms	$F(1,711)=27.98$ $p < 0.0001$ significant
V _d	-78ms	> -112ms	$F(1,711)=13.91$ $p = 0.0002$ significant