Redirected Touching: The Effect of Warping Space on Task Performance

Luv Kohli*  Mary C. Whitton*  Frederick P. Brooks, Jr.*

The University of North Carolina at Chapel Hill

Figure 1: A user touches a virtual board that is oriented differently than the real board providing passive haptic feedback

ABSTRACT
Passive haptic feedback in virtual environments is compelling, but changes to virtual objects require changes to associated real objects. Recent work suggests that by leveraging visual dominance, virtual space can be warped to map a variety of virtual objects onto a single real object. However, it is unknown whether users can interact with a warped virtual space as effectively as with an unwarped one. We present a study in which we measured task performance using the Fitts’-law-based ISO 9241-9 multidirectional tapping task. With a few caveats, results suggest that for certain tasks, warped virtual objects are no worse than unwarped virtual objects. We also present preliminary exploratory data on how well people can detect discrepancies due to space-warping.

KEYWORDS: Virtual reality, haptics, perception, pointing, Fitts’ law, ISO 9241-9.

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

1 INTRODUCTION
In virtual environments (VEs), a common way to provide users with touch feedback is to use passive haptics, or physical props to which virtual objects are mapped. This mapping is traditionally one-to-one. The result is compelling, because users touch a real object. However, passive haptic displays are inflexible; changing a virtual object requires changes to its associated real object.

Researchers have addressed this inflexibility mechanically, via Robotic Shape Displays [16]: when a user reaches for a virtual object, a robotic arm places a real object correctly in front of the user’s hand. One such robot has a Shape Approximation Device as its end-effector, which has several corners and curved and flat edges to approximate different shapes [22]. While impressive, these haptic displays are expensive and require sophisticated control mechanisms, and miscalculations and latency could be dangerous to users.

This paper instead further explores Redirected Touching, a perception-based technique for addressing the inflexibility of passive haptic displays [11]. The technique generates different mappings between real and virtual space such that a single real object can provide haptic feedback for many differently shaped virtual objects. These different mappings are created by warping virtual space. This warping introduces a discrepancy between a user’s real and virtual hand motion: the virtual hand moves in virtual directions different from its real-world motion, such that the real and virtual hands reach the real and virtual objects simultaneously. We call the case where real and virtual objects are the same one-to-one; otherwise, discrepant.

Discrepant stimuli are sometimes introduced in VEs to useful effect. For example, in Redirected Walking (from which our technique borrows its name), a discrepancy is injected between a user’s real-world head rotation and the virtual head rotation [17]. During head turns, the vestibular system is more sensitive than the visual system, and the injected discrepant visual rotation goes undetected. This discrepancy causes users to walk along real-world paths that are different from their virtual paths, enabling navigation in larger-than-tracked-space VEs.

*e-mail: {luv, whitton, brooks}@cs.unc.edu
We constructed a system (Figure 2) that tracks a user’s fingertip, a PhaseSpace IMPULSE optical motion capture system. To determine the user’s right index finger pose, we made a flexible ring out of sugru [33] and mounted three PhaseSpace LEDs on it (Figure 2, inset).

The system must calibrate for each user’s index finger. The finger is placed flat at a known location on a surface. The difference between the tracker data and the known location is later used to determine the user’s fingertip location.

The 10”x10” foam board also is tracked using three PhaseSpace LEDs, and is mounted on a motorized Directed Perception pan-tilt unit. The pan-tilt unit is rotated to present users with the real board at different angles. The VE is displayed on an NVIS nVisor SX head-mounted display (HMD) using Gamebase Co. Ltd’s Gamebyro game engine running on a dual quad-core 2.3GHz Intel Xeon machine with 8GB of RAM and an NVIDIA GeForce GTX 280 GPU. The rendered output is corrected for the HMD’s pin cushion distortion using a modified technique based on Kuhl’s work on HMD calibration [12]. A 3rdTech HiBall-3000 tracking system is used for head tracking. We use VRPN for tracker communication [23]. End-to-end system latency is ~50-60ms.

A contact microphone is mounted on the back of the foam board. A volume spike indicates that the board was touched, and tracker readings indicate where the board was touched.

![Image](319x341 to 557x519)

Figure 2: System constructed for space warping studies. Inset shows the finger tracking ring.

3 Warping space

We generate a mapping between real and virtual geometries by warping virtual space. The surface of the real geometry must be mapped to the surface of the virtual geometry, while smoothly and minimally warping the rest of the space.

Our system warps space using the well-known thin-plate spline technique commonly used in medical image analysis [1]. A thin-plate spline is a 2D interpolation method for passing a smooth and minimally bent surface through a set of points. The concept extends to higher dimensions; we use the 3D version [4, 19]. The method finds a function $f$ that passes through a set of points while minimizing bending energy, or the integral over $\mathbb{R}^3$ of the sum of squares of second-order partial derivatives of $f$. Displacements between corresponding real and virtual geometry points are computed. To generate the warp, the thin-plate spline method uses these displacements as points on which to operate.

Points on the surface of the real board are determined by interpolating the tracker data from the mounted PhaseSpace LEDs. Correspondences between real and virtual geometry points are predetermined. For the virtual board used in this work, the rendering PC completes the warp in about 150 milliseconds. Then, given an arbitrary finger position in real space, a warped
Early versions of our system did not warp the virtual hand’s orientation. The forces felt on the fingernail did not match what users expected based on the visual orientation of the virtual hand. Consequently, we now also warp hand orientation based on the warp field, creating an orientation discrepancy. At the board’s surface, the relative orientation between the virtual board and hand are made the same as the relative orientation between the real board and hand. The orientation discrepancy changes smoothly and dynamically as the user moves in the warp field.

**Figure 3:** Horizontal slices of unwarped space and real object (left) and warped space and virtual object (right), viewed from above.

## 4 Measuring Task Performance

In 1954, Fitts introduced a predictive model that relates movement time to distance and accuracy in rapid aimed movements, such as pointing [5]. It is now known as Fitts’ law, and is defined as:

\[ MT = a + b \cdot ID_e \]

where \( ID = \log_2 \left( \frac{A}{W} + 1 \right) \)

\( MT \) is movement time in seconds, \( a \) and \( b \) are empirically determined constants, and \( ID \) is known as the Shannon formulation of the index of difficulty, measured in bits [14]. \( A \) is amplitude, or the distance between targets, and \( W \) is target width. Increasing \( A \) or decreasing \( W \) makes an aiming task more difficult.

Fitts also introduced a composite measure of speed and accuracy called index of performance, now known as **throughput (TP)**. Human-computer interface researchers commonly use throughput to compare and evaluate different pointing devices. Throughput combines speed and accuracy and is unaffected by whether users decide to focus on one or the other [15].

The ISO 9241-9 document [9] describes a set of standardized tasks and measures for computing throughput. Recent work suggests using ISO 9241-9 tasks for evaluating VE user interfaces to improve consistency and enable comparisons across studies [24, 25]. We also have chosen to use the ISO 9241-9 multidirectional tapping task, with eleven targets (Figure 4). Users tap to improve consistency and enable comparisons across studies [24, 25].

![Figure 4: The multi-directional tapping task. Arrows are overlaid to show the sequence of targets touched.](image)

In Fitts'-law-based studies, it is common to look for statistically significant differences between different user interface techniques. However, even if there is a statistically significant difference between conditions, the magnitude of that difference could be very small in practical terms. What we really want to know is whether task performance while using a discrepant interface is good enough for a given application. We ideally want to show that task performance while using a discrepant interface is **no worse than** while using a one-to-one interface.

We can test whether one condition is no worse than another by statistically testing for noninferiority, a form of equivalence testing [28]. To do so, we need to define an indifference zone: the maximum difference between a discrepant condition and a one-to-one condition to be considered noninferior in the context of our application. Once we have an indifference zone, we analyze the mean difference of throughput between our conditions, and the one-tailed 95% confidence interval of that mean difference. If the mean difference and our entire confidence interval lie within our indifference zone, we consider the discrepant condition to be no worse than the one-to-one condition.

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

### 5.1 Indifference zone: throughput

We chose 1bps as the maximum allowable difference between discrepant and one-to-one conditions. This value was chosen because 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 [21]. Moreover, in Fitts’ 1954 paper, he labeled 10-12bps, a range of 2bps, as consistent [5].

We conservatively chose a tighter indifference zone bound of 1bps.¹

### 5.2 Indifference zone: error rate

To be conservative, we chose 9% as the maximum allowable difference between discrepant and one-to-one conditions. 9% corresponds to our smallest unit of measurement for error rate: one out of eleven targets missed.

¹ The throughput formula used in Fitts’ paper is different from that specified in the ISO 9241-9 standard. However, a reanalysis of Fitts’ data [14] using the Shannon formulation of ID and adjustment for accuracy suggests that the throughput range would not differ much from Fitts’ range, even though the absolute values change.
5.3 Movement time
It was not clear what indifference zone to choose for movement time. We therefore analyzed movement time to look for significant differences between conditions.

6 STUDY

6.1 Participants
22 paid participants (11 male, 11 female, aged 18-28, mean 21) took part. There was a mix of majors of both undergraduate and graduate students, and a mix of video game experience. All participants were right-handed and had normal or corrected-to-normal vision.

6.2 Notation
To determine values for which using discrepant objects is no worse than using one-to-one objects, we tested a range of discrepancies. A particular condition is represented as a tuple of the form (real angle, virtual angle), in degrees. Counter-clockwise rotations about the vertical axis are positive. For each discrepant condition, we had an analogous one-to-one condition. Discrepant conditions always had a real board angle of 0°. For example, a one-to-one condition with both the real and virtual angles at -18° is (-18, -18), and its discrepant counterpart is (0, -18). As in other recent research [24], a trial is a single target touch. A target circle is the set of all eleven targets in a particular instance of the multi-directional tapping task.

6.3 Angle range
In a pilot study, participants performed the multi-directional tapping task at various angles in the range -24° to 24°, and we determined that 24° and 24° were too discrepant to be used effectively (in terms of throughput) by pilot participants. We discretized the range into six angles: -18°, -12°, 0°, 12°, 18°, and 24° (24° was kept to verify pilot results). These angles yielded eleven conditions: (0, -18), (-18, -18), (0, -12), (-12, -12), (0, 0), (0, 12), (12, 12), (0, 18), (18, 18), (0, 24), and (24, 24). Figure 5 shows examples of one-to-one and discrepant conditions.

Figure 5: Left: front view of real board at 18°; Middle: virtual board at 18°; Right: real board at 0°. Left and middle panels make up a one-to-one condition, and right and middle panels a discrepant one.

6.4 Adaptation
During the pilot study, we found that after each change of conditions, throughput values went down for one or two target circles, but then went higher and became more consistent. In our full experiment, we excluded two adaptation target circles at the beginning of each condition block from analysis.

6.5 Target distance and width
In Fitts’ law tasks, it is typical to include many different target distances and widths to explore a wide index of difficulty range. Recent work suggests that the same results can be obtained by using a single target distance and varying only target width [7, 31]. In our angle discrepancy scenario, targets that are near the center of the board do not present the user with much discrepancy. We therefore decided to use a single target distance to minimize the number of conditions, and to maximize the effect of the discrepancy. The distance between targets was 21cm.

Target widths were chosen based on the US Department of Defense design criteria standard for human engineering MIL-STD-1472F [3], which specifies criteria for design and development of military systems. The standard states that push buttons (e.g., in cockpits) should have a diameter between 10mm and 25mm for bare fingertips, and at least 19mm for gloved fingertips. We chose six target widths: 10mm, 15mm, 20mm, 25mm, 30mm, and 35mm. The last two widths, despite being larger than described in the standard, were included to make the set of conditions more varied.

The chosen target distance and widths resulted in six distinct IDs ranging from ~2.8 to 4.5 bits.

6.6 Trials and counterbalancing
All participants did all conditions. The study used a within-subjects design, and conditions were counterbalanced using a Williams design [27, 29]. Each target circle had eleven targets. Because participants do not start with their hands in a standard position, timing information is not available for the first target. Thus, there were ten trials per target circle. Conditions were presented in eleven blocks of eight target circles. The first two target circles in each block were used for adaptation. The target widths on the last six target circles were randomized from the set of six target widths, without replacement. Including all participants, a total of 14,520 trials were recorded.

6.7 Procedure
Participants sat on a chair inside the tracking cage in front of the real board (Figure 1). To minimize arm fatigue during the experimental task, the chair was adjusted such that each participant’s shoulders were higher than the top of the board. We explained the multi-directional tapping task by showing a printout of a target circle. Typical of Fitts’ law studies, participants were instructed to touch each highlighted target as quickly and as accurately as possible, and to move on to the next target even if targets were missed. A small white dot was shown where the user touched the board, and a red X was shown on missed targets. Participants were told that things might occasionally feel strange (referring indirectly to the angular discrepancy), and that they should do the best they could in those situations.

Participants were fitted with the finger tracking equipment and HMD and completed seven training target circles to get used to the equipment and task. The training target circles consisted of some one-to-one conditions and some discrepant conditions, and were the same for all participants. After asking any questions they had, participants then proceeded with the eleven condition blocks of target circles. They were given a break of at least one minute after the fourth and eighth blocks. After each target circle, the real and virtual boards rotated to the proper angles for the next target circle. In the VE, the virtual board immediately switched to the new angle while the real board rotated. The real board rotated regardless of whether the real angle of the next target circle was different. Because participants might determine how far the board rotated based on the sound of the pan-tilt unit, all rotations were performed in fourteen steps, each either +3° or -3°. For example, rotating from 12° to 18° involved two +3° rotations, followed by twelve alternating +3° and -3° rotations.

After the VE session, participants completed a short questionnaire, and the experimenter conducted an open-ended interview to elicit comments and feedback. Participants were then debriefed and paid.
7 RESULTS

7.1 Spatial outliers

In Fitts’ law studies, it is customary to remove spatial outliers from the target touch data. In the literature, spatial outliers have been defined as errors in which movement was less than half the nominal target distance $A$, or in which the touch endpoint was more than twice the target width $W$ from the target center [31]. There were no outliers of the first form in our data.

We chose to not remove spatial outliers of the second form. In cases where participants had not yet adapted to a condition, or where the angular discrepancy was large, touches would often land more than twice the target width from the target center. However, from direct observation of participants, these misses were an indication that participants were unable to reach a target when desired. Removing these data points would artificially inflate throughput values in cases where a miss was due to a true difficulty with using the interface.

7.2 Learning effect and analysis

Upon visual inspection of our aggregate throughput data, we discovered that there was a substantial learning effect that lasted considerably longer than we had anticipated based on pilot data (Figure 6). When viewed in isolation, this learning effect is more pronounced in discrepant conditions than in one-to-one conditions, but the effect is present in both. Consequently, an effect for group membership (first three vs. last eight target circle blocks) was included in our statistical model.

Various outcomes were modeled using the generalized linear model with the generalized estimating equations to adjust standard errors and hypothesis tests for multiple observations within subjects. We compared discrepant vs. one-to-one conditions.

7.3 Throughput

Mean throughputs are shown in Figure 7. The differences between the mean throughputs of discrepant/one-to-one condition pairs and associated one-tailed 95% confidence intervals are shown in Table 1. Given our indifference zone limit of 1bps, the mean difference and the lower bound of the confidence intervals should be no less than -1bps for a discrepant condition to be considered no worse than a one-to-one condition.

All discrepant conditions were found to be no worse than their associated one-to-one conditions to within 1bps of throughput.

Table 1: Differences between discrepant and one-to-one throughputs (bps)

<table>
<thead>
<tr>
<th>Condition pair</th>
<th>Mean difference</th>
<th>One-tailed 95% conf. interval</th>
<th>Std. error</th>
<th>Noninferiority comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, -18) - (-18, -18)</td>
<td>-0.77</td>
<td>&gt; -0.998</td>
<td>0.141</td>
<td>-0.998 &gt; -1.0</td>
</tr>
<tr>
<td>(0, -12) - (-12, -12)</td>
<td>-0.28</td>
<td>&gt; -0.49</td>
<td>0.132</td>
<td>-0.49 &gt; -1.0</td>
</tr>
<tr>
<td>(0, 12) - (12, 12)</td>
<td>-0.56</td>
<td>&gt; -0.79</td>
<td>0.139</td>
<td>-0.79 &gt; -1.0</td>
</tr>
<tr>
<td>(0, 18) - (18, 18)</td>
<td>-0.62</td>
<td>&gt; -0.93</td>
<td>0.185</td>
<td>-0.93 &gt; -1.0</td>
</tr>
<tr>
<td>(0, 24) - (24, 24)</td>
<td>-0.71</td>
<td>&gt; -0.94</td>
<td>0.142</td>
<td>-0.94 &gt; -1.0</td>
</tr>
</tbody>
</table>

7.4 Error rate

Mean error rates are shown in Figure 8. The differences between the mean error rates of discrepant/one-to-one condition pairs and associated one-tailed 95% confidence intervals are shown in Table 2. Given our indifference zone limit of 9%, indifferent discrepant conditions must have an error rate no worse than 9% (1/11 targets) higher than associated one-to-one conditions.

Figure 6: A learning effect is seen in the throughput data, most evident between the first three and last eight blocks of target circles.

Figure 7: Throughput by virtual angle. Error bars are 95% confidence intervals.

Figure 8: Error rate by virtual angle. Error bars are 95% confidence intervals.
Table 2: Differences between discrepant and one-to-one error rates

<table>
<thead>
<tr>
<th>Condition pair</th>
<th>Mean difference</th>
<th>One-tailed 95% conf. interval</th>
<th>Std. error</th>
<th>Noninferiority comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 18) - (-18, 18)</td>
<td>2.8%</td>
<td>&lt; 4.9%</td>
<td>0.0128</td>
<td>4.9% &lt; 9%</td>
</tr>
<tr>
<td>(0, 12) - (-12, 12)</td>
<td>-0.14%</td>
<td>&lt; 2.0%</td>
<td>0.0131</td>
<td>2.0% &lt; 9%</td>
</tr>
<tr>
<td>(0, 12) - (12, 12)</td>
<td>3.5%</td>
<td>&lt; 5.7%</td>
<td>0.0134</td>
<td>5.7% &lt; 9%</td>
</tr>
<tr>
<td>(0, 18) - (18, 18)</td>
<td>2.8%</td>
<td>&lt; 4.8%</td>
<td>0.0118</td>
<td>4.8% &lt; 9%</td>
</tr>
<tr>
<td>(0, 24) - (24, 24)</td>
<td>5.7%</td>
<td>&lt; 8.3%</td>
<td>0.0162</td>
<td>8.3% &lt; 9%</td>
</tr>
</tbody>
</table>

All discrepant conditions were found to be no worse than their associated one-to-one conditions to within 9% of error rate.

7.5 Movement time

Mean movement times between targets are shown in Figure 9. Because we did not know of a reasonable value to use, we did not choose an indifference zone for movement time. We instead analyzed the data for significant differences using the Wald test. Within the context of the full model, several planned contrasts were done. A significant difference between movement times was found for all pairs of discrepant/one-to-one conditions (Table 3). Movement times are represented in milliseconds in the figure and table, but in seconds for throughput computation.

![Figure 9: Movement time by virtual angle. Error bars are 95% confidence intervals.](image)

Table 3: Differences between discrepant and one-to-one movement times (ms)

<table>
<thead>
<tr>
<th>Condition pair</th>
<th>Mean difference</th>
<th>95% conf. interval</th>
<th>Std. error</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 18) - (-18, 18)</td>
<td>74.7</td>
<td>(43.9, 105.5)</td>
<td>0.0157</td>
<td>$t = 22.63, p &lt; 0.0001$</td>
</tr>
<tr>
<td>(0, 12) - (-12, 12)</td>
<td>35.6</td>
<td>(4.4, 66.7)</td>
<td>0.0159</td>
<td>$t = 5.02, p = 0.0251$</td>
</tr>
<tr>
<td>(0, 12) - (12, 12)</td>
<td>41.2</td>
<td>(9.2, 73.3)</td>
<td>0.0163</td>
<td>$t = 6.36, p = 0.0117$</td>
</tr>
<tr>
<td>(0, 18) - (18, 18)</td>
<td>61.9</td>
<td>(27.8, 95.9)</td>
<td>0.0174</td>
<td>$t = 12.66, p = 0.0004$</td>
</tr>
<tr>
<td>(0, 24) - (24, 24)</td>
<td>109.4</td>
<td>(70.4, 148.4)</td>
<td>0.0199</td>
<td>$t = 30.17, p &lt; 0.0001$</td>
</tr>
</tbody>
</table>

8 DISCUSSION

The results above indicate that with noninferiority indifference zone limits of 1bps and 9% for throughput and error rate respectively, the discrepant conditions are no worse than their one-to-one counterparts, but significant differences were found between discrepant/one-to-one pairs in movement time.

8.1 Caveats

From directly observing participants, we would hesitate to say (0, 24) is no worse than (24, 24). While in the (0, 24) condition many participants said they had difficulty using the interface. Can we be sure that 1bps is a valid indifference zone limit?

Truthfully, we cannot be sure. Even though ranges of about 1-2bps have been labeled consistent in the past, there are many studies in the literature that show significant differences in throughput between conditions with mean throughput differences less than 1bps. Despite showing noninferiority under our chosen indifference zone, the graphs of mean throughput show a clear trend of performance degradation as discrepancy increases. Throughput, while being useful as a metric that combines speed and accuracy, is not nearly as concretely understandable as either metric alone. It is important to evaluate speed and accuracy along with throughput in the context of the application.

Let us assume for the moment that 1bps is too generous a value for the throughput indifference zone. Now consider error rate and movement time. In all condition pairs other than (0, 24)-(24, 24), the highest confidence interval bound for error rate difference is ~6%. The highest confidence interval bound for movement time difference is ~100ms (again excluding (0, 24)-(24, 24)).

Certainly, only the clients of a VR system that employs space warping can decide what differences are acceptable. However, these small differences in error rate and movement time suggest that participants were able to perform the given task almost as well in discrepant conditions as in one-to-one conditions. In other words, users may not be able to touch virtual buttons as precisely in discrepant conditions as in one-to-one conditions, but we believe they can touch them precisely enough to activate them when desired, and that may be sufficient for many tasks.

8.2 Alternative throughput computation

The effective width computation defined in section 4 projects touch endpoints onto the vector between task targets. The adjustment for accuracy thus takes into account only one dimension of endpoint deviation. Because the ISO 9241-9 multi-directional tapping task is a 2D task, it has been suggested that it may be more appropriate to compute endpoint deviation in two dimensions to account for deviations orthogonal to the task axis [31]. Wobbrock et al. present a formula for this 2D deviation:

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

\((x_i, y_i)\) are the endpoints, and \((\bar{x}, \bar{y})\) is the centroid of the endpoints. We reanalyzed our data using \(SD_{x,y}\). As with Wobbrock et al.’s results, throughput values were systematically lower than with the 1D computation (Figure 10). This is expected because we are now incorporating additional deviation that was not present in the 1D computation.

Mean differences and confidence intervals are shown in Table 4. Using the same throughput indifference zone as before, 1bps, all discrepant conditions except for (0, 24) were no worse than their associated one-to-one conditions. These results are more consistent with our observations that (0, 24) is unlikely to be no
worse than (24, 24). 1bps might mean something different because throughput is computed differently, but throughput computed using $SD_{x,y}$ may better fit the data.

![Figure 10: Throughput computed using 2D endpoint deviation. Error bars are 95% confidence intervals.](image)

Table 4: Differences between discrepant and one-to-one throughputs (bps) computed using 2D endpoint deviation

<table>
<thead>
<tr>
<th>Condition pair</th>
<th>Mean difference</th>
<th>One-tailed 95% conf. interval</th>
<th>Std. error</th>
<th>Noninferiority comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, -18) - (-18, 18)</td>
<td>-0.63</td>
<td>&gt; -0.83</td>
<td>0.12</td>
<td>-0.83 &gt; -1.0</td>
</tr>
<tr>
<td>(0, -12) - (-12, -12)</td>
<td>-0.28</td>
<td>&gt; -0.47</td>
<td>0.116</td>
<td>-0.47 &gt; -1.0</td>
</tr>
<tr>
<td>(0, 12) - (12, 12)</td>
<td>-0.55</td>
<td>&gt; -0.75</td>
<td>0.121</td>
<td>-0.75 &gt; -1.0</td>
</tr>
<tr>
<td>(0, 18) - (18, 18)</td>
<td>-0.57</td>
<td>&gt; -0.83</td>
<td>0.159</td>
<td>-0.83 &gt; -1.0</td>
</tr>
<tr>
<td>(0, 24) - (24, 24)</td>
<td>-0.85</td>
<td>&gt; -1.07</td>
<td>0.137</td>
<td>-1.07 &lt; -1.0</td>
</tr>
</tbody>
</table>

8.3 Participant comments and other observations

After the training part of each experiment session, no participants asked any questions. One participant did show signs that he discovered the nature of the discrepancy during the first experiment block. He asked whether the real board always rotated or if it sometimes did not. Indeed, after the experiment, that participant stated that the real and virtual boards did not always seem to be at the same orientation.

During the post-experiment interview, most participants had similar comments. All participants said that they experienced a strange feeling at times (the discrepant conditions). Five participants identified that the real and virtual boards were not always at the same orientation. The most commonly mentioned cue that something was strange was that participants had to reach farther than (or not as far as) they expected given the virtual board they saw. Another common cue was that when participants would move their finger from one side of the board to the other, the finger would often hit the board prematurely due to the discrepancy. Participants reacted by pulling their hand back a little farther before touching the next target. Most participants remained directly in front of the board during the task. One participant often reoriented his head to face the virtual board head-on, and another occasionally rotated his chair. We did not notice any difference in performance due to these strategies.

Participants rated arm fatigue from 1 (not tired at all) to 5 (very tired). Three participants rated arm fatigue as 3, and the rest rated it 1 or 2. The most common complaint, from about half of the participants, was that the HMD was too heavy or uncomfortable. Some participants said their neck was tired from looking down at the board. One participant felt disoriented after the first block of target circles. The HMD was removed for a short time during the first break, and there were no difficulties after that. No other participants mentioned feeling nauseated or dizzy.

9 Conclusions and Future Work

We have presented the results of a study investigating whether task performance changes when users are presented with discrepant real and virtual objects. Discrepant conditions yielded throughputs and error rates that are no worse than analogous one-to-one conditions using our 1bps throughput and 9% error rate noninferiority indifference zones. Significant differences were found in movement time. Our choice of 1bps for throughput indifference might be too large, but the small mean differences in error rate and movement time suggest that discrepant objects can be used almost as well as one-to-one objects.

It may be useful to choose one or two discrepancy levels and conduct another study. Participants would experience each condition longer and we could determine how well they learn to use a discrepant interface without confounding conditions.

The actual task, in a virtual cockpit for example, is unlikely to be a rapid multi-directional tapping task like the one used here. A user’s task may involve only occasionally pressing virtual buttons. If button presses are occasional, will task performance suffer due to fading adaptation? How quickly?

We also want to investigate at what discrepancy levels task-engaged users detect discrepancies, as well as task performance and detectability for discrepancies other than orientation.

9.1 Exploratory detectability data

As a preliminary exploration of what levels of discrepancy users can detect, six members of the research group experienced several conditions in blocks similar to our main study. Participants were presented with nine blocks of target circles (four discrepant and five one-to-one generated from the angles -24, -12, 0, 12, and 24). After every target circle, they judged whether what they saw and what they felt were the same or different, and gave a confidence value for their judgment (3 for most confident, 1 for least). The same/different responses were weighted by the confidence values and normalized giving a percentage—100%, all participants responded different with confidence 3 for all target circles in a given condition; 0%, all same with confidence 3; and 50%, same and different were equally frequent (Figure 11).
There are many potential sources of bias in these data, including aggregating these data across participants who have different sensitivities and different definitions of confidence. However, the results inform future studies. These data suggest that there is indeed a non-zero amount of discrepancy that is undetectable. Formal studies to investigate detectability are planned.

10 ACKNOWLEDGMENTS

We thank the Link Foundation and the NIH National Institute for Biomedical Imaging and Bioengineering for supporting this work. We thank John Thomas, Rohit Saboo, Sachin Patil, Peter Lincoln, and Anish Chandak for their help with hardware, warping code, video, pan-tilt unit code, and microphone code respectively; Chris Wiesen for statistical consulting; Kan Anant at PhaseSpace for technical support; and Rob Teather, Eric Burns, Sharif Razzaque, and Lt. Col. Paul Holst, USAF for valuable discussion. The hand model is © 2011 CyberGlove Systems.

REFERENCES


