

---

Electronic Theses and Dissertations, 2004-2019

---

2007

## Accidental Inversion During 3d Rotation With 2-dof Input Devices

Derek Daniel Diaz  
*University of Central Florida*

 Part of the [Psychology Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact [STARS@ucf.edu](mailto:STARS@ucf.edu).

---

### STARS Citation

Diaz, Derek Daniel, "Accidental Inversion During 3d Rotation With 2-dof Input Devices" (2007). *Electronic Theses and Dissertations, 2004-2019*. 3142.

<https://stars.library.ucf.edu/etd/3142>

ACCIDENTAL INVERSION DURING 3D ROTATION  
WITH 2-DOF INPUT DEVICES

by

DEREK DANIEL DIAZ  
B.S. University of Florida, 1999

A dissertation submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy  
in the Department of Psychology  
in the College of Sciences  
at the University of Central Florida  
Orlando, Florida

Fall Term  
2007

Major Professor: Valerie K. Sims

© 2007 Derek Daniel Diaz

## ABSTRACT

This dissertation focuses on a human operator's ability to perform rotational control of a three-dimensional object using two-degrees of freedom (DOF) interface devices. Although input devices designed specifically for 3D interaction exist, devices traditionally used for two-dimensional user interaction, such as a mouse or joystick, have become ubiquitous to computer tasks. This research examines a particular human-computer interaction issue that arises from stimulus-response compatibility between three-dimensional stimuli spaces and 2-DOF response sets.

The focal point of this research is a phenomenon referred to here as accidental inversion. Accidental inversions occur when an operator erroneously moves a three-dimensional object in a direction opposite than was intended. Thus, the effect of accidental inversion results from a mismatch between the operator's intended and actual input. A key assumption in diagnosing the causal factors involved in the accidental inversion effect is contribution from both internal (i.e., having to do with the individual) and external (i.e., having to do with the environment) influences.

Three experiments were conducted to study accidental inversion. The first examined population stereotype, a measure of a target population's natural response tendencies to particular stimuli for a particular task. Results indicated a strong population stereotype for horizontal rotations (i.e., yaw) and weak stereotype for vertical rotations (i.e., pitch). This effect was mediated by whether the task was in the context of flight or ground-based movement. The second experiment analyzed the subjective preference for two opposite input-response (I-R) mappings (i.e., how the system responds to different input into the controlling device) for a task requiring control over

vertical rotation. Results indicated that subjective preferences for I-R mappings were not heavily polarized. The third experiment also focused on vertical rotational control and examined how subjective preference for a particular I-R mapping affected performance. Furthermore, this experiment also examined performance when interference was introduced in the form of a temporary interruption where the participant had to conduct the task using an opposite I-R mapping. Results indicated that, upon being interrupted with the opposite I-R mapping, the group who used the mapping they subjectively preferred did worse than the group who used the mapping they did not prefer.

This research has implications for the design of human-machine systems requiring human-in-the-loop three-dimensional rotational control. Some human-machine systems can have significant consequences from even a single mistake caused by a human-operator accidentally providing the wrong input. Findings from this research lead to two primary recommendations to the design of human-machine systems: a) an easily accessible and clearly indicated method to select input-response mapping which is provided before beginning the actual task, b) be informed of the current input-response mapping in use.

To Trish, my best friend,  
and to Sofia and Ryan, my inspiration.

## ACKNOWLEDGMENTS

First I would like to thank Valerie K. Sims for being my mentor and supporting me when things got tough. Thank you for the freedom and guidance you gave me in all my research endeavors.

For their patience and tolerance, I want to thank my committee members Mustapha Mouloua, Edward J. Rinalducci, and James Pharmer.

Next I would like to thank Richard D. Gilson for challenging me to always think out of the box and to seek solutions from different angles. I also want to thank Peter A. Hancock for driving me to aim high and strive to produce excellent work.

A special thanks goes to Peter and Aida Hamilton for always believing in me, to Tatiana Diaz for showing me how to succeed (it is a good thing that you are older sibling), and to my parents for their endless support.

Lastly I want to thank my wife, Trish, for motivating me throughout this project and for helping with the little things that made it, at times, especially challenging.

## TABLE OF CONTENTS

LIST OF FIGURES.....	x
LIST OF TABLES.....	xi
LIST OF ACRONYMS.....	xii
INTRODUCTION.....	1
Stimulus-Response Compatibility .....	4
Compatibility Effects.....	6
Design Implications from SRC Research.....	29
Rotational Control in Three Dimensional Space .....	33
Research Summary .....	39
Motivation for this Research.....	39
Implications of this Research .....	40
EXPERIMENT 1 – POPULATION STEREOTYPE.....	41
Introduction.....	41
Summary of Experimental Design.....	43
Method.....	47
Sampling Pool.....	47
Participant Assignment .....	48
Stimuli and Apparatus .....	48
Procedure .....	52
Results.....	53
Polarity .....	53
Reaction Time.....	55



Discussion .....	57
<b>EXPERIMENT 2 – SUBJECTIVE PREFERENCE FOR VERTICAL ROTATION</b>	
<b>CONTROL.....</b>	<b>60</b>
Introduction.....	60
Summary of Experimental Design.....	61
Method.....	63
Sampling Pool.....	63
Participant Assignment .....	63
Stimuli and Apparatus .....	63
Site Apparatus.....	68
Questionnaire.....	68
Procedure .....	69
Results.....	70
Discussion .....	76
<b>EXPERIMENT 3 – PREFERENCE AND INTERFERENCE .....</b>	
<b>79</b>	
Introduction.....	79
Summary of Experimental Design.....	80
Method.....	83
Sampling Pool.....	83
Participant Assignment .....	83
Stimuli and Apparatus .....	83
Procedure .....	84
Results.....	85

Training Phase .....	85
Switch and Return Phases .....	86
Discussion .....	89
CONCLUSIONS AND RECOMMENDATIONS .....	91
Conclusions .....	91
Theoretical Implications .....	93
Recommendations .....	94
Practical Implications .....	96
Recommendations for Future Research .....	97
APPENDIX A: EXPERIMENT MATERIALS .....	99
Experiment 1 .....	100
Informed Consent Form .....	100
Instructions .....	100
Experiments 2 & 3 .....	101
Informed Consent Form .....	101
Questionnaire .....	101
APPENDIX B: CORRELATION MATRIX FROM EXPERIMENT 2 .....	103
APPENDIX C: IRB HUMAN SUBJECTS PERMISSION LETTERS .....	107
REFERENCES .....	111

## LIST OF FIGURES

Figure 1: Pictorial Representation of Set and Element Level Compatibility.....	7
Figure 2: Stimuli and response set from Baur and Miller (1982) .....	13
Figure 3: S-R ensembles with results from Fitts and Seeger (1954). .....	19
Figure 4: 2D stimuli and response sets from Hoffmann (1997) .....	23
Figure 5: Examples of 3D stimuli and response sets from Hoffmann (1997).....	24
Figure 6: Representative results from Hoffmann (1997).....	26
Figure 7: Example of rotation along the Y and X axes. ....	41
Figure 8: Stimuli from Experiment 1. ....	49
Figure 9: Mean percent of matching responses for axis, context, and axis by context..	55
Figure 10: Mean response time for axis and axis by context.....	57
Figure 11: Experiment 2 screenshots of capturing a floor target. ....	66
Figure 12: Experiment 2 questionnaire responses. ....	71
Figure 13: Perspective versus control-map for question 1. ....	73
Figure 14: Control-mapping versus perspective for question 2. ....	74
Figure 15: Control-mapping versus perspective for question 3. ....	75
Figure 16: Accidental inversion errors training & baseline phases by group. ....	86
Figure 17: Overall means for accidental inversion errors for critical phases. ....	87
Figure 18: Accidental inversion errors for critical phases by I-R mapping group.....	88
Figure 19: Correspondence between I-R mapping groups.....	88

## LIST OF TABLES

Table 1: Major Stimulus-Response Compatibility Effects .....	27
Table 2: Spatial Compatibility Principles for Design of Controls and Displays .....	30
Table 3: Rotation/Vertical/Forward-Backward S-R Stereotypes (Hotta et al.,1981) .....	31
Table 4: Experiment 1 Stimuli and Trial Conditions.....	51
Table 5: Notional trial sequence for Experiment 2.....	67
Table 6: Results from question 4 of Experiment 2.....	76
Table 7: Phases in Experiment 2 .....	85
Table 8: Correlation Matrices for Experiment 2, Questionnaire Items 1 - 3 .....	104

## LIST OF ACRONYMS

2D	Two dimensional
3D	Three dimensional
CE	Compatibility effect
DOF	Degrees of freedom
DV	Dependant variable
I-R	Input-response
IV	Independent variable
S-R	Stimulus-response
SRC	Stimulus-response compatibility
UI	User interface

## INTRODUCTION

Three-dimensional (3D) graphics have become common in modern interfaces with computers and other machines. Despite the fact that off-the-shelf consumer systems are capable of 3D graphics, interface devices designed for generic 3D interaction have not become nearly as ubiquitous. As such, software utilizing 3D graphics will often use traditional 2D input devices such as a mouse, touchpad, trackball, or joystick (Scali, Wright, & Shillito, 2003). Applications of 2D interfaces for interaction with three dimensions range from being benign, such as video games, to those where precision and speed are critical factors and errors have significant consequences. Examples of such systems include remotely operated surgical tools, bomb-disposal robots, and other unmanned vehicles. As with most human-machine systems, usability is affected by the compatibility between information displayed to the operator and methods available by the operator to interact with the system. A usable interface is crucial in applications where significant consequences may arise as side effects of usability issues.

The relationship between stimulus and response (S-R) is a cornerstone of psychology and invariably plays a role in any system that includes a human as a component. Research in the area of stimulus-response compatibility (SRC) emphasizes the importance of the quality and nature of pairings between information presented to a human-operator and the type of responses that are available. Compatibility effects (CEs) are used to describe how these S-R pairings affect a human-in-the-loop system. Compatibility effects can have substantial impact on performance since performance is usually dependent on input by the human actor. Usability issues in human-computer

interaction may arise as a result of non-optimal S-R relationships which are often characteristic in user interfaces that are unintuitive, prone to user error, and difficult to learn (Vu & Proctor, 2001).

Certain stimulus-response relationships have received more research attention than others and, as a result, some types S-R effects can be predicted with reasonable accuracy in many situations. Additionally, sets of S-R pairs that would obviously result in negative effects can often be ruled out by common sense. For example, pressing an up-arrow key to move a text cursor down would not seem to be a logical mapping and would likely lead to poor performance due to operator error.

Given the limitless combinations of stimulus-response pairings, however, their effects are not always predictable. Factors such as prior experience and cultural norms (Norman, 1998) can complicate an S-R relationship. For example, domain knowledge can color an S-R pairing: one would expect nearly all people who have used a computer mouse to slide it directly away from their body in order to move the mouse cursor to the top of the computer display. However, knowledge about other domains can affect the strength of particular S-R pair. For example, given the aviation domain's strong tie between pushing and pulling a stick to affect vertical rotation (i.e., pitch), compatibility effects may be difficult to predict for users controlling an unmanned aircraft using a mouse and a computer display.

The compatibility effect of interest in this research has to do with a phenomenon that occurs when a user conducts a 3D spatial task involving rotation using an interface device with only 2-degrees of freedom. The phenomenon driving this research is a compatibility effect referred to here as accidental-inversion which, when it occurs, would

normally be considered an operator error (i.e., an error causally attributed to the human-operator rather than another component of the system). Accidental-inversion specifically describes an error an operator makes when, during a manual-spatial task involving rotation, an entity is rotated in a direction opposite than intended. Despite the fact that the operator has only two response options (i.e., for vertical rotations: rotate up or down, or for horizontal rotations: rotate left or right) the effect of accidental-inversion has the potential to impact performance. Moreover, hysteresis due to correction/over-correction looping and exacerbated by lag can be confusing and require a shift of attention away from the critical elements of the task in order to focus on manual control (e.g., Megaw, 1972). The accidental inversion phenomenon has been observed in fielded human-machine systems such as control of the infrared camera on the Navy P3 Orion Aircraft (Stephanie Hartin, personal communication, June 15, 2007) and across hundreds of human-subjects in research by the Games User Testing team at Microsoft Games Studios (Derek Diaz, personal observation, September 2003-2004).

The research conducted in this dissertation has three objectives. The first objective is to examine how people naturally (i.e., without instruction) respond to rotational stimuli and to examine factors that may influence these natural S-R mappings. This first objective provides insight into designing-out accidental-inversion to minimize its negative effects. Additionally, this objective investigates whether an optimal mapping between stimulus and response may exist and can thus be applied as a general rule to diminish the frequency of accidental inversion.

The second objective is to examine whether people subjectively prefer one type of S-R mapping over another, and, given a choice, which mapping is preferred. While



the first objective focuses on how users tend to respond, the second seeks to understand why users select one mapping over another when given the choice.

The third objective is to assess the degree to which compatibility effects from accidental inversion are mitigated by allowing an operator to select a particular S-R mapping for a task. Additionally, part of the third objective is to examine the effect of interference from unexpected use of an opposite S-R mapping on the frequency of accidental inversion.

The introduction section of this dissertation will provide an overview of stimulus-response compatibility theory, with special attention to research on spatial compatibility effects. Secondly, theory affecting interface design will be covered in terms of differences between 2D and 3D input devices and three factors that are theorized to influence accidental inversion: affordance, frames of reference, and spatial degrees of freedom.

### Stimulus-Response Compatibility

This research is focused on a compatibility effect that affects users working with a three-dimensional space while using an interface device with only two degrees of freedom (DOF). The issue, at heart, has to do with stimulus-response compatibility and thus a general overview is provided here. Note that this topic has been the subject of volumes of research and many overviews are available (e.g., Hommel & Prinz, 1997; Proctor, & Reeve, 1990).

Stimulus-response compatibility has to do with putting things together in ways that are intuitive, and, resultantly, improves performance. The concept is applicable to any human-in-the-loop system and is a basic tenet in human factors and cognitive

psychology. The general idea is that human-machine systems which provide intuitive pairings between the information they present and response options available to the user will improve performance. In contrast, performance is likely to decrease as the compatibility between S-R pairings decreases. Various definitions of compatibility can be found in the literature, a sample of which is provided here starting with one of the first:

"The ensemble of stimulus-response combinations comprising the task [which] results in a high rate of information transfer" (Fitts & Seeger, 1953, which can be traced back to a AM Small (1951)).

"The extent to which the ensemble of stimulus and response combinations comprising a task results in a high rate of information transfer." (Fitts & Seeger, 1953).

"Stimulus-response compatibility...refers to the fact that people respond more quickly and accurately with some mappings of stimuli to responses than with others." (Vu & Proctor, 2004).

"Stimulus-response compatibility refers to the fact that some tasks are easier or more difficult than others either because of the particular sets of stimuli and responses that are used or because of the way in which individual stimuli and responses are paired with each other". (Kornblum, Hasbroucq, & Osman, 1990)

The fact that the word ensemble appears in several definitions of SRC emphasizes two key points: a) SRC has everything to do with pairings between stimulus

and responses, and b) the fact that multiple responses are possible for a given stimulus, where some responses lead to better performance than others. Thus, ensembles refer to stimuli and responses that have been matched together into a particular S-R set with some expectation of their effect on the associated task. While several theoretical accounts of SR have been put forth since Fitts' (Fitts & Deininger, 1954; Fitts & Seeger, 1953) original accounts, the general state of the theory has not evolved far in the time since (Kornblum, Hasbroucq, & Osman, 1990). While the notion of compatibility is straight forward yet meaningful enough so as to provide guidelines for interface design, it is not a concept that lends itself well to being quantified. Researchers have observed that definitions of compatibility are often circular (e.g., Hoffmann, 1997; Sanders, 1980). For example, the following quote from Sanders (1980) reflects this observation,

“[SRC] refers to the degree of natural or overlearned relations between signal and responses... The weakness of the variable is that there is no clear underlying continuum of naturalness... Comparisons between studies on SRC are often difficult since the operational meaning of compatible and incompatible varies across experiments”.

### *Compatibility Effects*

How a S-R ensemble affects a system is called a S-R compatibility effect, or simply a compatibility effect (CE). Compatibility effects are well studied, especially in terms of their effects measuring performance of different S-R ensembles across the same task. The following section will discuss theoretical accounts of CEs and ways in which they may be produced to have a positive affect on performance.

*Element Level versus Set Level Compatibility*

Kornblum, Hasbroucq, & Osman (1990) distinguishes between element-level and set-level CEs, which was previously suggested by Fitts & Seeger (1953). Elements refer to individual stimuli and responses, and are thus composed of two groups: stimuli elements and response elements. Sets refer to collections of stimuli or collections of responses. A stimulus-response ensemble may be considered a super-set consisting of stimulus sets and response sets. Different mappings may be used between elements in the stimulus-set and the elements in the response set that are included in the S-R ensemble. Figure 1 depicts the relationships among elements, sets, and ensembles of stimuli and responses.

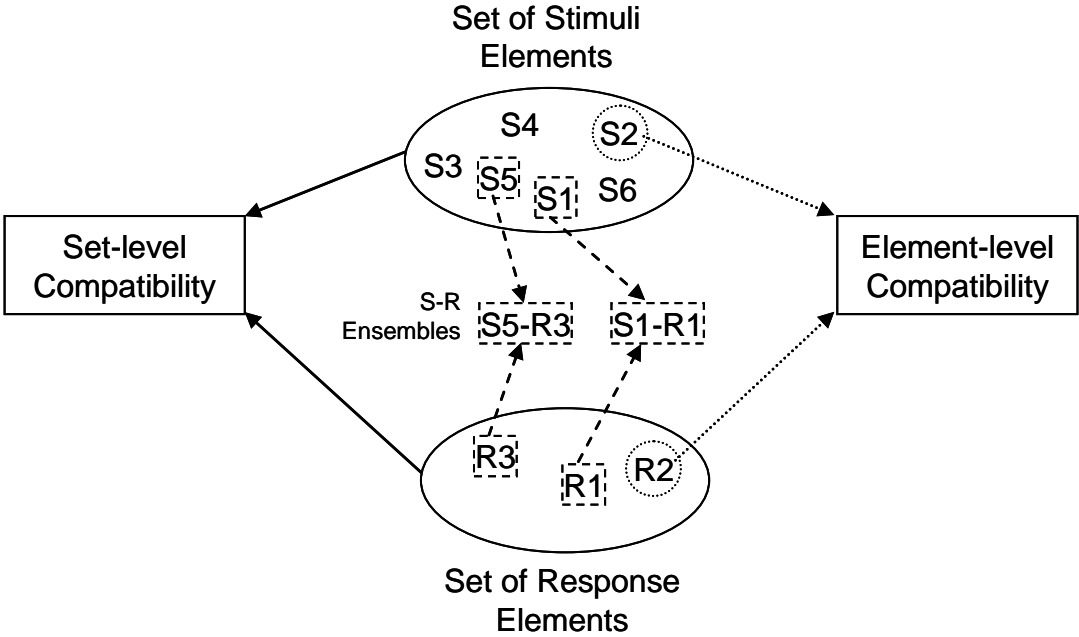


Figure 1: Pictorial Representation of Set and Element Level Compatibility.

Element level compatibility describes when performance varies as a function of the mapping of the individual stimuli and response elements within the same stimulus and response sets (Kornblum, Hasbroucq, & Osman, 1990). Similarly, feature-sharing pairs of individual stimuli and responses will allow for better performance. Element level compatibility has been the subject of numerous experiments where multiple task conditions are created by including multiple pairings of the stimuli and responses used in the study. A choice reaction task conducted by Vu and Proctor (2004) demonstrates element level CEs. This study involved two response alternatives, a left or right keypress depending on if the stimulus appeared on the left or right. Responses were faster and more accurate when the task was to respond to a stimulus on the left by pressing the key on the left rather than mapping the left stimuli to a response using the right key.

Set level compatibility refers to a difference in the compatibility of the overall stimulus set with the response set. Stimulus sets and response sets that share features will allow for better performance than those that do not. The concept of set-level compatibility is captured in Wicken's (1992) model of human performance which predicts that performance improves for tasks involving visuo-spatial stimuli with manual responses versus other response modalities, such as verbal. Similarly, a verbal stimulus is predicted to be better paired with a spoken response than with manual response.

### *Theoretical Accounts of SRC*

Theoretical accounts of SRC seek to explain why CEs occur and are useful to predict CEs for particular situations. Most accounts of SRC explain CEs based on the

ways that the stimuli and responses are similar. Different researchers have taken different approaches in organizing and analyzing these similarities. For example, Norman (1998) focused on spatial, biological, social, perceptual effect, Fitts & Seeger (1953) focused on spatial versus symbolic, and others have focused on the multi-dimensional nature of SRC (e.g., Alluisi, 1961; Kornblum 1990, 1995; and Allusi & Warm, 1990). Key to these accounts is the hypothesis that common attributes between a stimuli and a particular response lead to an advantage in reaction time and accuracy. The dimensional-overlap processing model (Kornblum, 1990, 1995; also see section on conceptual overlap) focuses on describing similarities between stimuli and responses in compatible and incompatible sets and also ties in response selection and execution. The approach provided by Proctor, Wang, and Pick (2004) is one that is consistent with most accounts of CEs and describes CEs in terms of three groups: (a) physical, (b) conceptual, and (c) structural. These three divisions will serve as a way to organize a discussion in the following few paragraphs on compatibility effects and theory.

*Physical compatibility.* Physical compatibility is used to describe when physical characteristics of the stimuli and response (e.g., color, shape, size, sound) share similarities. Physical correspondence refers to how similar the form of a stimulus is to the paired response and in most cases includes the concept of spatial compatibility. For example, if the stimulus is a red circular light, the following responses may be notionally listed in terms of decreasing physical similarity: pushing a red circular button, pushing a red square button, turning a switch to a red color patch, saying the word “stop”. An example of a spatial compatibility task was presented above in the description of the 2004 study by Vu and Proctor.

Some of the earliest work in SRC was focused on spatial CEs (e.g., Fitts & Seeger, 1953) and this topic has probably since received the most attention in SRC research. Fitts and Seeger's 1953 study (described in detail below) may have been the first to use the term compatibility. This study paired six combinations of three stimulus sets and three response sets. The stimuli consisted of lights in three patterns. The response sets consisted of manual-spatial controls with the same pattern as the lights. All six possible combinations of stimuli and responses were tested together. For each of these six S-R ensembles, reaction time and error rate scored significantly better when the stimulus set and the response control set shared the same pattern.

*Conceptual Correspondence.* Conceptual correspondence is broader than spatial correspondence and refers to various types of relationships between stimuli and responses. Alluisi and Warm (1990) point out that physical correspondence alone is limiting in explaining SR-C effects and suggests that conceptual correspondence offers a more precise explanation of the CE phenomena. The authors describe several examples of conceptual correspondence, including laterality, numerical codes, central processing, correspondence between alphabets, and dimensionality.

Proctor & Gilmour (1990) point out that CEs resulting from conceptual correspondence are typically smaller than when the dimensions match at a physical level as well. The reason, in part, has to do with set versus element level S-R compatibility: S-R ensembles that match physically, and in particular, spatially, are often of the same modality and thus have a high degree of element-level compatibility (e.g., a visual stimulus on the left side of a screen paired with moving a joystick to the left). Whereas, S-R ensembles that only share conceptual compatibility may not necessarily

be the same modality and thus may not be compatible at the set level (e.g., a visual stimulus paired with a verbal response).

Lateral correspondence is useful in accounting for spatial compatibility. Lateral correspondence refers to SR-C effects where performance is maximized when responses match the side where the stimulus was perceived. Wallace (1971) conducted a study where participants had to respond to a stimulus that appeared on their left or right side, or above or below a fixation point. Responses were made by pushing one of two buttons positioned on their left or right side. Participants responses with their hands either crossed or not. Results showed that key-presses to the left or right were faster when the stimuli matched the side of the button press, regardless if their hands were crossed or not.

A classic study by Morin and Grant (1955) had participants conduct a simple task where the goal was to press one button from among several arranged on a row when a light illuminated. Multiple lights were also arranged in a row and lining up directly above the buttons. Three response conditioned were tested: (a) direct, where the goal was to push the button directly below the light, (b) reverse, where the goal was to push the button that matched the horizontal position of the light, if the row of lights were reversed, and (c) random, where the mapping between the lights and button pushes were randomly created. Performance was best for the direct response condition, not much worse for the reverse response condition, and considerably degraded for the random response condition. These results demonstrate how coding affects performance, despite the fact the stimuli and responses are physically similar.



The Dimensional Overlap Processing model (Kornblum, Hasbroucq, & Osman, 1990; Kornblum & Lee 1995) provides one of the most complete accounts of SRC. The model explains relationships between stimulus and response sets and attempts to account for a wide range of compatibility effects. The term “dimensional overlap” refers to similarities at the set level (see Figure 1). For a given S-R ensemble, the stimulus elements and response elements will match (or not match) to a given a degree. For example, for a set of stimuli that consists of numbers and set of responses that consist of spoken numbers, S-R ensembles with matching elements will pair the visual representation of the number with the same number as the spoken response. The dimensional overlap model attributes set level compatibility effects to stronger automatic activation of the corresponding response when set-level compatibility is high than when it is low. When comparing stimulus and response sets with high versus low compatibility, the model predicts that sets with high compatibility will have: (a) faster responses are for highly compatible S-R ensembles (paired S-R elements) and, (b) slower responses for lesser compatible S-R ensembles.

*Structural Correspondence.* Structural compatibility is described in the literature as another contributor to compatibility effects in addition to physical and conceptual similarity (Cho, & Proctor, 2003; Kornblum & Lee 1995; Reeve & Proctor, 1984; Proctor & Vu, 2005). One may view performance being affected by the correspondence in the structure of the stimulus and response sets, even in the absence of physical or conceptual similarity. For example, Proctor and Gilmour (1990) found that sequential mappings of digits to each of the ten fingers on both hands had better performance than randomly mapping digits to fingers. Additionally, Baur and Miller (1982) conducted a

study where they varied stimuli that appeared in an upper or lower position with left or right responses as well as stimuli appearing in an left or right positions with upper or lower responses (see Figure 2). Results showed an overall advantage for up-right/down-left S-R ensembles. Moreover, the effect persisted across the various modalities that were tested (manual vs. vocal, unimanual, vs. bimanual, spatial vs. symbolic). Their finding was that assigning up stimuli to right responses and bottom stimuli to left responses was easier for participants than the reverse mappings (Up-Left, Down-Right). Of particular interest was that this SRC effect occurred when there was no spatial correspondence between the stimuli and responses, and thus no obvious basis to code the S-R elements.

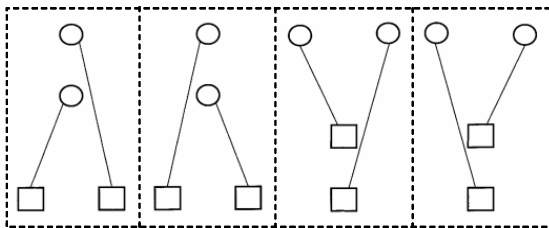


Figure 2: Stimuli and response set from Baur and Miller (1982). Circles represent stimulus positions and squares represent response positions. The solid lines between the stimuli and responses indicate the four S-R ensembles used in the study.

### *Population Stereotype*

Thus far, SRC effects have been primarily attributed to similarities intrinsic between the stimulus and response. Another factor that contributes to CEs is the degree to which a S-R ensemble agrees with characteristics prevalent among a given population, i.e., a population stereotype. Population stereotypes measure the consistency of response across a target population (Hoffmann, 1997). Population

stereotypes are used to describe the tendency that members of a representative target population exhibit a particular response to a particular stimulus. A fundamental rule applicable to the design of any human-machine interface is to take advantage of population stereotypes.

Fitts (1959) discusses population stereotypes when he describes patterns of reaction time and error results as involving the "transformation, translation, and receding of information, [all of which] are assumed to vary in ... the time required, and the likelihood of errors, as a function of unlearned and/or highly overlearned behavior patterns" (p. 17). Later, he goes on: "We shall forego use of the concept of habit strength and shall attempt to predict compatibility effects on the basis of the concept of population stereotype" (Fitts, 1959, p. 19). Fitts continued his description with, "The degree of population stereotype [is defined as] a function of the uniformity of the responses made by a representative sample of people when they are placed in a standard test situation without any special instruction or training that would bias them in favor of any one of the several responses possible in that situation. Population stereotype is denned such that the larger the proportion of individuals who make identical responses to identical stimuli in such a situation, the stronger is the population stereotype".

Many population stereotypes exist (for examples see Woodson & Conover, 1970, and Wickens, 1987). Stimulus-response ensembles could have become a stereotype for any number of reasons, including culture and practice (Alluisi & Warm, 1990). For example, the standard position of a light switch to signal the off and on positions are opposite between the United States and the United Kingdom. Also, Brebner (1976),

described examples of S-R ensembles for different situations (e.g., clockwise turns means to increase) result from cultural conventions. An important characteristic of population stereotypes is that they are subject to change not only across cultural divisions but across time as well. Thus, experience (i.e., practice, familiarity, etc) can foster a stereotype (Alluisi & Warm, 1990).

A method used to quantify a population stereotype is by eliciting responses from members of the target population to sets of stimuli without suggesting what types of responses are correct or preferable in any way (Alluisi & Warm, 1990). Data from this type of free-response paradigm might take the form of frequencies of responses to the stimuli that were tested. It should be noted that, when measuring a stereotype for a particular stimuli, the stimuli and responses are usually constrained to a relatively small set of what is practical or reasonable to be assessed in the context of an experiment. One risk in conducting such studies is that the stereotype discovered is only directly associated with the stimulus and response sets demonstrated in the experiment. Measurement of population stereotype is always not conducted in studies on SRC effects. Rather, assumptions are made on what S-R ensembles best and least represent the population (Alluisi & Warm, 1990).

#### *Processing and Action Selection in SRC*

It is important to understand the cognitive processes involved in how SRC effects influence performance. According to some researchers (c.f., Wickens, 1987, 1992; Atkinson & Shiffrin, 1968) cognition generally follows the flow of sensation–recognition–response selection–response execution. Sensation is the process where a distal stimulus is received by a sensory organ. During recognition, the proximal stimulus gains

meaning through processing that require some level of cognitive resources (e.g., attention) and short-term (i.e., working) memory. Response selection and execution are the serial processes where a particular response is chosen and then performed, both which require some level of cognitive resources.

Research in SRC typically explains CEs in terms of two paths that occur when processing stimuli that have high and low compatibility with responses that are available. The first process represents the path for high compatibility stimuli. This process is automatic and results in direct activation the most compatible response. This expedited path has minimal memory and cognitive resource requirements.

Electrophysiological evidence supports this theory that response activation is automatic rather than voluntary. Eimer (1995) conducted a study where arrows were presented that pointed to the left or right and recorded event-related brain potentials (ERPs). Eimer detected ERPs in cortical areas associated with the most compatible response (i.e., spatially congruent: left-left, right-right).

The second process occurs for stimuli without highly compatible responses and is neither voluntary nor automatic, and thus requires cognitive resources and memory to select and execute the response. Cognitive processing is required in this path since some rule must be applied for selecting the response appropriate for the stimuli was perceived. Performance (e.g., reaction time) is negatively affected for incompatible S-R ensembles due to automatic activation of the compatible response. Similarly, performance improves for compatible S-R ensembles due to the automatic activation of the compatible response.

Alluisi & Warm (1990) reasoned that population stereotypes are maintained in long-term memory and are mediated by a central-processing mechanism different than the two processes described above. For SRC, long-term memory may be conceptualized as frequencies or probabilities of S-R ensembles that have been through the central-processing system.

Reaction time to make a response to a given stimulus is thus the time it takes between perception and response execution, where variations may be attributed to the level or amount of cognitive processing required for the S-R ensemble. Response time elicited as a result of the expedited path will be faster than responses that require more processing. To that end, reaction time is also indicative of the strength of the stereotype associated with the S-R ensemble, where responses time is expected to decrease as the strength of the stereotypes increases. Accuracy, a measure that is based from rules for a particular task, will be higher when the S-R ensemble matches a strong stereotype than when the stereotype is violated.

#### *Population Stereotype SRC Studies*

As mentioned above, different accounts of SRC have been put forth and numerous studies have been conducted examining a variety of SRC effects. Three particular research projects on SRC will be discussed in this section: Fitts and Seeger (1954), Hoffmann (1997), and Vu and Proctor (2003).

*Stereotype Effects in Spatial Compatibility.* Deininger and Fitts conducted some of the first published works on SR-C and produced results that demonstrate the characteristics that are accounted for by the theories discussed above. One study by

Fitts and Seeger (1954) will be described here since its results span across several of the compatibility effects and principles that have been discussed thus far.

Fitts and Seeger (1954) assessed SR-C effects by measuring reaction time and response accuracy for a spatial task where pairings of stimuli and responses were varied. The paradigm the authors used has become common practice in S-RC research since. Four sets of stimuli were used, where two were spatial and two were non-spatial (see Figure 3). The spatial stimuli sets consisted of 8 lights, arranged in either a circle (2-dimensional) or a straight line (1-dimensional). One of the non-spatial stimuli sets consisted of the four-digit numbers meant to represent 8 hours on a clock face at 12:00, 1:30, 3:00, 4:30, 6:00, etc. The other non-spatial set of stimuli consisted of three letter first names (e.g., VIC, BEN, ROY). The participant's objective was to respond to each stimuli using the correct response option as quickly as possible. Participants responded by sliding a stylus in one of 8 possible directions radiating around a center point at 45 degree angles. The study paired three different response sets to each of the stimuli sets except the three-letter non-spatial which only had one response set. Thus, ten S-R ensembles were created based on pairings of stimulus type and response method:

- |                        |                                   |
|------------------------|-----------------------------------|
| 1. Spatial 2D-optimum  | 4. Symbolic 2D-optimum            |
| 2. Spatial 2D-mirrored | 5. Symbolic 2D-mirrored           |
| 3. Spatial 2D-random   | 6. Symbolic 2D-random             |
| 7. Spatial 1D-optimum  |                                   |
| 8. Spatial 1D-mirrored |                                   |
| 9. Spatial 1D-random   | 10. Symbolic (non-spatial)-random |

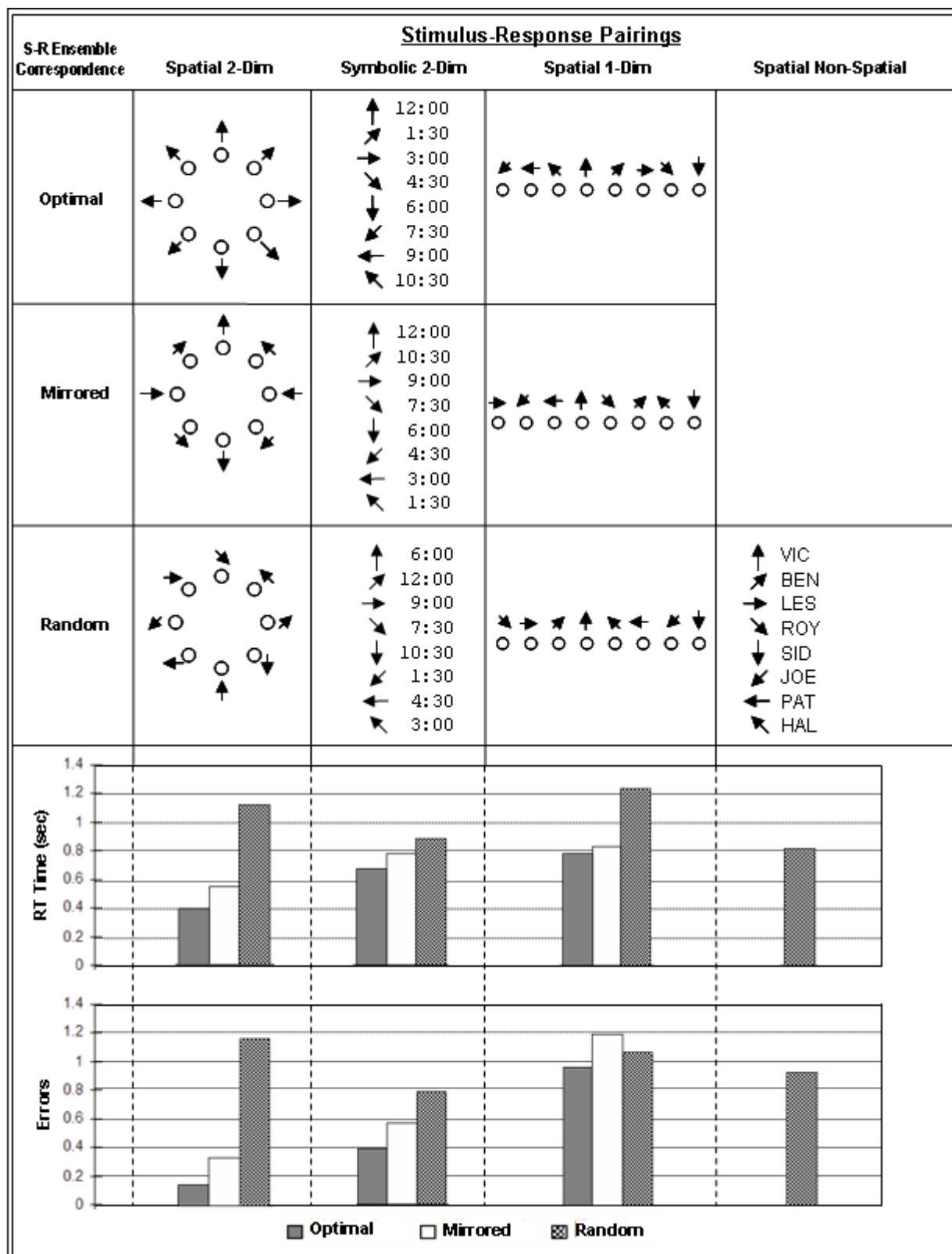


Figure 3: S-R ensembles with results from Fitts and Seeger (1954).



The task was to simply slide the stylus to make the correct response when a stimulus appeared. Reaction time and accuracy were measured. Additionally, all participants took part in the study across two days and only experienced one of the ten response conditions.

As expected, the quickest reaction times and fewest errors were observed for the spatial 2D stimuli with optimal corresponding responses. This condition resulted in an error rate of 1.6%, which came to only 2 mistakes over 128 trials. Additionally, response time was consistent under this condition, having a standard deviation of only .03 seconds. Results indicated that S-R ensemble used in the spatial 2D-optimal condition appeared to have been a highly compatible pairing.

The condition with the most errors and second slowest response-time was with the same spatial 2D stimuli but with random response pairings. The performance difference between the two conditions was substantial, where responses took three times as long and were eight times less accurate for the random response condition. Moreover, this result demonstrates an important SR-C effect: S-R ensembles with a strong population stereotype tend to be more negatively affected by nonsensical deviations from the stereotype than are S-R ensembles based on weak stereotypes. It can be reasoned that participants had difficulty with the spatial-random condition because correct responses went against their natural S-R pairings, ostensibly formed as a result of past experience.

The spatial 2D-mirrored condition saw performance that was similar to the maximum response condition for the stimuli and had the second best performance overall. Participants were quite capable of applying the simple S-R mapping rule of

"inversion" and made fast and accurate responses. Thus, the spatial 2D condition was a usable ensemble for the task. Reasoning based on SR-C literature explains this result in terms of S-R coding. Both the maximum and mirrored response conditions created usable S-R ensembles when paired with the spatial stimuli because of the strong population stereotype and, in the case of the mirrored condition, the logical and easy to remember code of mapping stimuli to their natural response opposites.

Interference with the population stereotype plays a role in random response condition for the spatial 2D stimuli. The use of the complicated S-R mapping to violate the population stereotype in the random response condition caused the poor performance.

Thus, results were consistent with common sense: the maximum response condition has the strongest stereotype response of the three response conditions to the spatial 2D stimuli. Participants in the random response condition had to not only properly code (and remember) the nonsensical S-R mapping, but also suffered interference from the strong population stereotype. The mirrored response condition fared well because the S-R code was simple and logically mapped to the stereotype response.

For the symbolic-2D condition, the fewest errors occurred for the maximum response condition, followed by the mirrored, with random being last at twice that than the maximum. Reaction times followed the same pattern but were lesser in magnitude.

Why was it expected that the spatial 2D response condition with maximum correspondence would have the best performance? The likely answer is because the purpose of the study was to test a range of different S-R ensembles and, of the

response conditions used in the experiment, the 2-dimensional response with maximum correspondence condition seemed to the authors offer the most stimulus-response compatibility. While the a-priori hypothesis was entirely qualitative, it appears to have considerable face-validity that the response condition is stereotypical to the population to whom the research seeks to generalize.

When a population stereotype is strong, high physical correspondence between stimulus and response, especially spatial correspondence is critical. In the case of a strong stereotype, monotonic S-R mappings are often effective (e.g., if the stimulus is an up arrow and the response is to press a button with a picture of an up arrow). Simple codes such as inversion (e.g., stimulus is an up arrow and the response is to press a button with a picture of a down arrow) will also usually be effective. Difficult, nonsensical codes will result usually result in poor performance (e.g., stimulus is an up arrow and the response is to say the number 5), and will likely result in worse performance than would be obtained for any S-R ensemble that does not have a strong population stereotype.

*Stereotype Effects in Rotation Responses.* Hoffman (1997) conducted a series of studies on SRC which were unique in that one goal was calculate equations for the strengths of population stereotypes based on the empirically measured strengths of various display/control arrangement principles from the literature. His first study had participants from two groups, students of engineering or psychology, respond to 64 different arrangements of displays and controls, where each instance consisted of a knob and horizontal meter (see Figure 4). Along with the 8 knob placements, the meter was shown with a strength line (neutral) indicator or an arrow (directional) indicator, and

values decreasing or increasing from left to right to represent the direction of the scale. Using a formula, Hoffman calculated the strength of the stereotype for each of the 64 arrangements based on SRC principles.

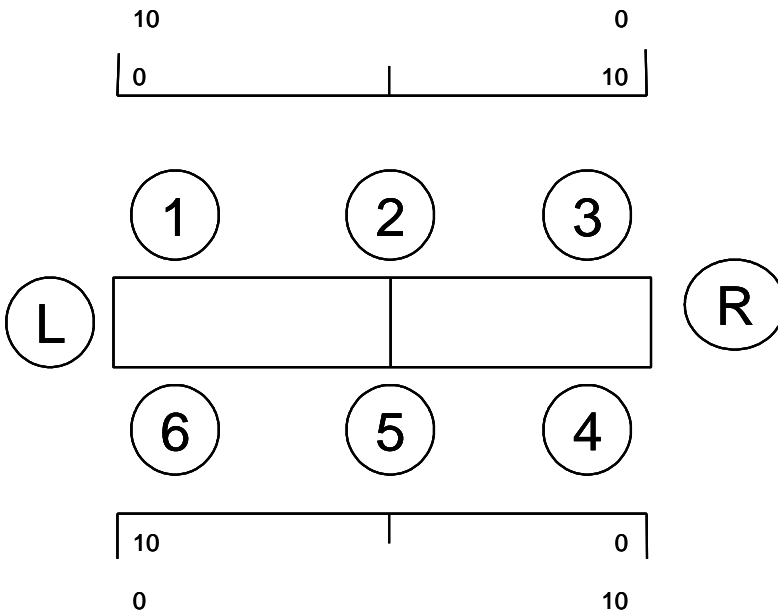


Figure 4: 2D stimuli and response sets from Hoffmann (1997). Composite picture of the arrangements of 2D stimuli and response sets used in the experiment. Circles represent the 8 knob positions used (1, 2, 3, 4, 5, 6, L, and R). The vertical line in the horizontal meter in the center was affected by the turning a knob. Note that a second condition used an arrow (not shown) as a marker rather than the vertical line. Four different ways to represent the direction of the scale were tested (anchors 10 and 0 positioned on either the upper or lower side of the meter).

The task was for the participant was to indicate how he or she would turn the knob in order to increase the value represented on each display arrangement. Participants were instructed to respond as quickly as possible (about 5 seconds were allowed). Reaction time was faster for arrangements with strong population stereotypes than for arrangements with weak population stereotypes. One of the more interesting

findings in this study was that differences between the psychology and engineering groups were substantial and statistically significant. Hoffmann suggests that these differences could possibly be attributed to mechanical knowledge and biases, thus suggesting that these two groups have different population stereotypes.

Hoffmann's second study utilized three-dimensional arrangements of displays and controls. Two conditions were tested: (a) pictures of different arrangements of the knob and meter drawn on paper, and, (b) actual physical mockups (see Figure 5). The drawing condition also varied the viewing perspective (i.e., angle) from which each arrangement was presented while the hardware condition did not. Participants indicated the direction he or she would turn the knob to increase the value on the meter. Responses were made in the hardware condition by actually turning the knob. The primary dependant measure was the proportion of clockwise movements. Participants were all drawn randomly from university students. Viewing angle had a strong effect when the axis of the control knob was parallel to the display.

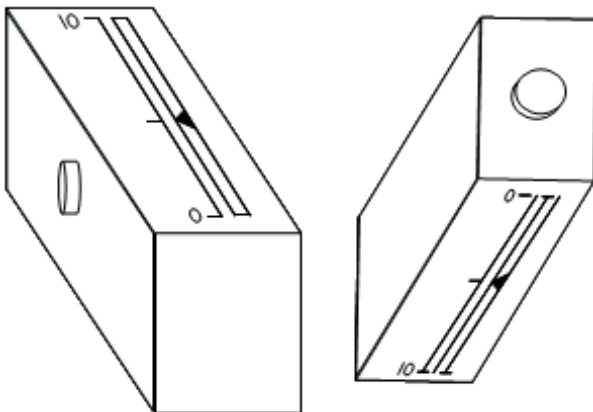


Figure 5: Examples of 3D stimuli and response sets from Hoffmann (1997).

Hoffman's third study was again similar to the first but also included vertical meter arrangements (see Figure 5) and only included engineering students. Participants were shown the different two-dimensional pictures of display/control arrangements and asked to respond by rotating a physical knob (located flat on the desk in front of them) in order to increase the value on the meter. Using his method to calculate stereotype strength, Hoffman compared stereotype strength against response time (Figure 6). The most interesting finding was a difference for horizontal and vertical displays. For horizontal displays, Hoffman found that reaction time decreased as the strength of the stereotype increased. However, for vertical displays, there was no relationship between reaction time and the strength of the stereotype. Hoffmann reasoned that reaction time may not necessarily be the best measure of stereotype. While this explanation has merit, given that this study (like most others) did not separate reaction time based on stages of information processing and response execution. However, another explanation may be that the equations applied to calculate stereotype did not capture difference between vertical and horizontal arrangements used in the study. Thus, stereotypes for horizontal arrangements may actually be far stronger than stereotypes for vertical arrangements, despite Hoffmann's calculation that equated the stereotype strength between the two.

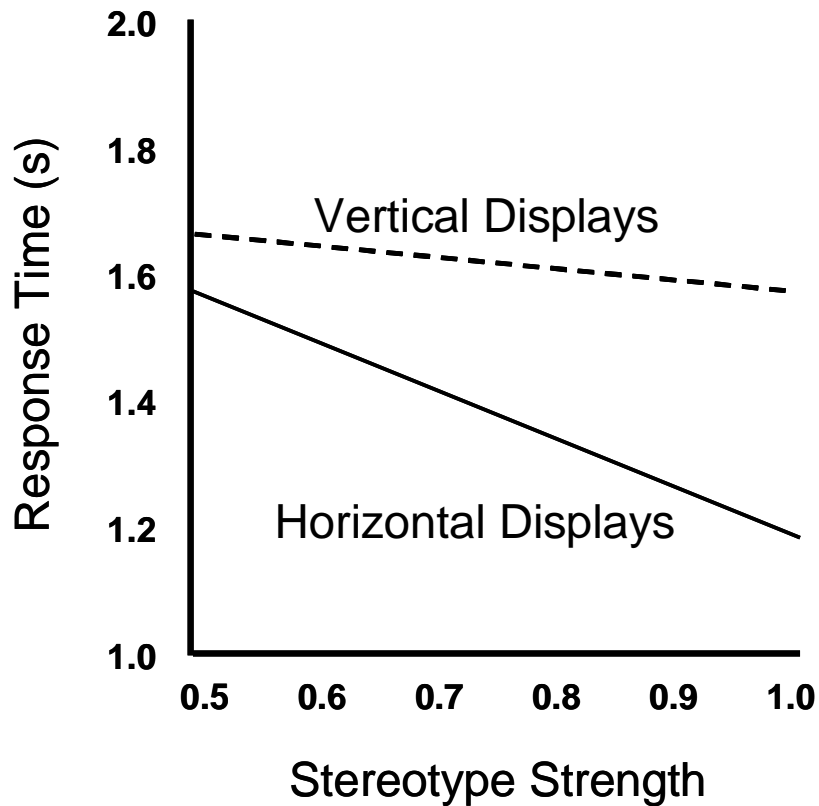


Figure 6: Representative results from Hoffmann (1997). Mean response time versus stereotype strength for vertical (dashed) and horizontal (solid) display arrangements. As stereotype strength increased, response time decreased for horizontal S-R ensembles. The nearly horizontal line indicates a weak relationship between stereotype strength and response time for “vertical” S-R control-display arrangements.

### *Common Compatibility Effects*

Given the amount of research on the topic of SRC, the nature of several specific CEs are well established. Below is a list of several major CE, represented here with citations:

Table 1: Major Stimulus-Response Compatibility Effects

Reaction time increases for optimal S-R mapping between ensembles when the pair has dimensional overlap.	From Kornblum et al. (1990) who cited Brainard, Irby, Fitts, & Alluisi, 1962; Brebner, Shepard, & Cairney, 1972; Fitts & Seeger, 1953; Hawkins & Underhill, 1971; Shulman & McConkie, 1973.
The overall mean reaction time is faster with congruent than with incongruent mapping.	From Kornblum et al. (1990) who cited Blackman, 1975; Duncan, 1977a, 1977b; Morrin & Grant, 1955; Sanders, 1970; Schwartz, Pomerantz, & Egeth, 1977; Simon, 1967, 1969; Simon & Craft, 1970; Sternberg, 1969.
The greater the dimensional overlap, the greater the reaction time difference between congruent and incongruent mapping.	From Kornblum et al. (1990) who cited Simon & Small, 1969; Wallace, 1971.
The difference between congruent and incongruent mapping is greater for nonrepetitions than for repetitions.	From Kornblum et al. (1990) who cited Bertelson, 1963.
The increase in mean reaction time when the number of alternatives is increased is greater the less the S-R compatibility, whether it is varied by changing the degree of dimensional overlap or the mapping.	From Kornblum et al. (1990) who cited Bainard et al., 1962; Davis, Moray, & Treisman, 1961; Hawkins & Underhill, 1971; Leonard, 1959; Morrin; Konick, Troxell, & McPherson, 1965; Theios, 1975.
The effects of varying dimensional overlap or mapping with irrelevant dimensions are similar to those with relevant dimensions.	From Kornblum et al. (1990) who cited Broadbent & Gregory, 1965; Costa, Horwitz, & Vaughan, 1966; Kornblum, Hasbroucq, & Osman, 1984; Smith, 1977; Sternberg, 1969; Whitaker, 1979.
When S-R sets can be coded with respect to more than one frame or reference, which pairings of stimuli and responses are most compatible is dependent on upon the frames on which the coding was based.	Proctor, Wang, & Pick, 2004.
When stimuli and responses vary along orthogonal spatial dimensions, the mapping of an upper stimulus location to a right response and lower stimulus location to a left response often produces better performance than the alternative mapping because it maintains correspondence between the positive and negative alternatives of the two dimensions.	Cho & Proctor, 2003.



### *Characteristics of Compatibility Effects*

Research (e.g., Loveless, 1962) describes general characteristics common to a variety of CEs, which will be elaborated upon below: (a) CE are large in magnitude, (b) CE are stable and reliable, (c) CE are exacerbated by stress, (d) CE have weak relationships between speed and accuracy, and (e) spatial CEs tend to be reversible (Chan & Chan, 2003; Chan & Chan, 2007). These characteristics are more often found for CEs that are based on a strong population stereotype and/or possess large degrees of dimensional overlap.

Compatibility effects are generally large in magnitude. Studies (Alluisi & Warm, 1990) have demonstrated CEs to have effects greater than from those obtained from practice. Substantial differences in reaction time and error rate have been reported in many studies. For example, results presented from a SRC study by Fitts and Seeger (1954) in Figure 3 presents reaction times that differ by a factor of 5 between different S-R ensembles.

Compatibility effects tend to be relatively stable and reliable in their magnitude (Fitts & Seeger, 1953). Although CEs may decrease in magnitude with practice, substantial effects have been demonstrated to remain even after extended practice (Vu & Proctor, 2003; Dutta & Proctor, 1992).

Research in SR-C (e.g., Garvey & Knowles, 1954) has shown that CEs are exacerbated by stress such that performance is negatively affected for S-R ensembles with low compatibility when under stress than for S-R ensembles that are have higher compatibility.

A negative relationship between response speed and accuracy is often observed in studies on human performance. The general reason for this is that accurate responses tend to take more time and rapid responses tend to be more prone to error. This negative relationship between speed and accuracy does not appear to be prevalent in many S-RC studies. For example, in Deiniger and Fitts' 1955 study, S-R ensembles that cause an operator to respond more slowly also tended to be less accurate.

Compatibility effects, especially when they are spatial in nature, are often reversible (Alluisi & Warm, 1990). For example, Hoffmann's 1997 study found both clockwise-right and clockwise-left CE.

#### *Design Implications from SRC Research*

The compatibility of S-R ensembles affect performance because operators tend to make fewer errors using highly compatible ensembles which, in turn, lead to improved system performance since system performance often depends on human performance. Stimulus-response ensembles are thus a critical part of an operator-machine interface. An understanding of the mechanisms underlying SRC effects is crucial to the design of a human-computer system for optimal performance (Alluisi & Warm, 1990). Thus, one goal of the designer of a human-computer interface is to pair stimuli and responses in such a way as to optimize performance for the user.

#### *Principles for Design*

Human-machine interface design principles have been derived from SRC research, especially in terms of spatial arrangement and manual controls. The following list (Table

2) is a sub-set of that provided by Proctor & Vu (2005) that relate to spatial arrangement and manual control design characteristics of a human-computer interface:

Table 2: Spatial Compatibility Principles for Design of Controls and Displays

---

Spatial Compatibility

---

- Compatible mappings of stimuli assigned to their spatially corresponding responses typically yield better performance.
- Better performance occurs when the mapping of stimuli to responses can be characterized by a rule or relation than when it is random.

Movement Compatibility

---

- The motion of the display should move in the same direction as the motion of the control.
- Clockwise movement is used to indicate upward movement or an increase in magnitude of the display.

Proximity Compatibility

---

- Controls should be placed closest to the display they are controlling.
- Controls and displays should be arranged in functionally corresponding groups.
- Control and displays should be sequentially arranged.

Other Spatial

---

- The up-right/down-left mapping is often better than the up-left/down-right mapping.
- Pure tasks of a single stimulus-response mapping produce better performance than mixed tasks with multiple mappings.

Hotta et al. (1981) presented data on common direction of motion stereotypes for a variety of tasks using different types of controls (see Table 3). Of particular interest in this research are Hotta's categorization of how rotation-based interface devices are typically used to accomplish different tasks.

Table 3: Rotation/Vertical/Forward-Backward S-R Stereotypes (Hotta et al.,1981)

<b>Purpose</b>	<b>Plane</b>	<b>Rotary Knob</b>	<b>Rotary Lever</b>	<b>Button</b>	<b>Slide Lever</b>		<b>Two Buttons</b>
Door	Front Side		CC	Pull			
Water/Gas	Front Side	CW	CC		Downward		
	Top Side		CC		Backward		
	Bottom Side	CW	CW	Pull	Backward	Right	
	Right Side	CW	CW	Pull	Downward	Backward	
	Left Side	CC	CC				
Electricity	Front Side	CW	CC	Push	Downward		
	Top Side	CW	CC	Push	Backward		Backward
	Bottom Side	CW	CW		Backward	Right	Backward
	Right Side		CW	Push	Downward	Backward	
	Left Side		CC	Push	Downward		Upward
Increase	Front Side	CW	CC		Upward		
	Top Side	CW	CC		Forward	Right	
	Bottom Side	CW	CW	Pull	Forward		
	Right Side	CW	CW	Pull	Forward		
	Left Side		CC		Backward		

Note: CC: Counterclockwise, CW: Clockwise

*Optimizing SRC.* One of the most salient characteristics of the stimuli and response are their physical form. The operator must be able to clearly identify the stimulus in order to select the desired response. Similarly, the operator must be able to distinguish the desired response from other possibilities (Fitts & Deininger, 1954). The number of distinguishable responses available to select from has been shown to have a direct effect on response time. This SRC effect has been demonstrated across a number of experiments and shows an increase in response time as the number of response possibilities also increases (e.g., Brainard et al., 1962, Davis, Moray, & Treisman, 1961; Hawkins & Underhill, 1971).

As previously described in the context of population stereotype, domain knowledge and experience can influence the CE. For example, a compatibility effect often seen in the automobile accident avoidance literature documents a human tendency to respond in the direction away from a negative stimulus, such as an obstacle on a collision course. In a skid, turning the wheel away from the skid is the most compatible response in a population of “normal” drivers. However, the opposite would be expected for an off-road race car driver with training to turn the wheel into the direction of the skid.

General guidelines for the design of human-machine systems garnered from the SRC literature suggest that that SR ensembles should go together in a meaningful way, be easy to learn, and easy to remember. Coding refers to how responses are mapped to stimuli and its complexity, which can vary, is an important factor to an effective S-R ensemble. The simplest S-R code is the one where the stimulus and response are identical or nearly identical. An example of a SR code that is identical is where the stimulus is a word spoken out loud and the response is to also speak the same word out loud. An example of a SR code that is slightly less congruent is when the responder must point to a particular word in a list that matches the word that was spoken.

Previous experiments have often used spatial pairings between arrays of lights that served as the stimuli and buttons in specific spatial locations as responses. In these studies, the most compatible response was usually defined button located directly adjacent to the stimulus light. So long as the button is clearly identifiable as being the closest to the stimulus light, this paradigm would be expected to have a simple coding requirement. In contrast, random mapping between the placement of the response

buttons and their associated light stimulus would be expected to have a complex coding requirement. Simple codes free the operator from devoting cognitive resources (e.g., attention and working memory) to select a response. S-R codes that follow simple, logical rules, such as pressing left upon seeing a right arrow and vice-versa would be expected to require less cognitive resources than ensembles with random, nonsensical mapping between stimuli and responses.

#### *Main Points from SRC Research*

Two key findings in S-R compatibility research is that the degree of compatibility is not so much dependant upon the particular stimuli or response, but rather upon (a) the interactive effect of the two as a stimulus-response pair, and (b) amount of correspondence between a S-R ensemble and population stereotypes. It is important to note that the notion of population stereotype is related to the S-E ensemble, rather than specifically to the stimulus or response. Factors relating to only the stimulus or response may contribute to the SRC effect, but population stereotype has to do with how well the S-R pair matches the S-R pairing that is most prevalent in the population for one reason or another. In other words, the influence of population stereotype on SRC deals with the relationship between SR ensembles rather than individual stimulus and response elements.

#### *Rotational Control in Three Dimensional Space*

Three factors are theorized to affect accidental-inversion, the CE of primary interest in this dissertation: (a) degrees of freedom: the number of dimensions that a human operator can control in a 3D space, (b) affordance: how the design of an object

influences how a human operator uses it, and (c) frames of reference: how the human operator relates himself or herself to the system or object under control.

### *Degrees of Freedom in Human-Computer Interface Devices*

Fully specified spatial control capability in a three-dimensional space involves a total of six dimensions, three for translation, and three for rotation (see Poupyrev & Ichikawa, 1999). Translation refers to the position of the object in space in reference to three axes: horizontal (X), vertical (Y), and height (Z). Rotation refers to an object's orientation in the place where it is in terms of pitch, yaw, and roll. A problem in the usability of a device with two DOF to control an object in a 3D space is that there are not enough control axes to map to each dimension. As a result, software applications often provides the user with the capability to toggle a mode between translation and rotation, thus, at any one time, control input is only used to manipulate the former or the later set of dimensions. Moding, however, is notorious for introducing usability problems related to mode awareness (Reitinger et al, 2006). Thus, the two available DOFs on a mouse, joystick, or trackball pose a challenge to the interface designer in terms of how to map controls, and also to the user in terms of using them effectively. While devices specifically tailored for 3D control offer advantages, the sheer ubiquity of standard 2D input devices is reason to seek an optimal design using for these more limited devices.

*Two versus Three DOF Input Devices.* Compared to 2D user interfaces (UI), 3D-UIs are generally more complicated and require greater efforts to achieve a high level of usability (Bowman, Kruijff, LaViola, & Poupyrev, 2004). While computer users may reasonably be expected to be familiar with standard interaction methods for traditional

2D-UIs, in part due to the strong population stereotypes for 2D human-computer interaction, 3D-UIs do not benefit from the same standards, metaphors, and stereotypes (Bowman, 2001).

*Degrees of Freedom in Input.* Computer software has traditionally been tailored for two dimensional spaces and interacted with using devices with only 2 DOFs. Two-DOF control devices such as the mouse, keyboard, and joystick have evolved their own standard set of interaction methods (e.g., WIMP). As 3D applications proliferate, these UI components developed for 2D interaction are commonly being used to interact with 3D software. Some computer applications allow the user to interact with both 2D and 3D content simultaneously or interchangeably. For example, a user may browse a museum's web site (using a 2D interface) in order to gain access to a virtual walkthrough of a pyramid thus switching to a 3D-UI.

In terms of translation control (i.e., control of movement along the X, Y, and Z axes), research suggests that the added depth dimension during 3D over 2D manual control is a frequent source of error. A common observation in 3D tracking studies is that accuracy along the Z dimension is often worse than for the X and Y dimensions (Erp & Oving, 2002).

In terms of rotation control, studies have been conducted that demonstrate the benefit of control devices with more than 2 DOFs for tasks requiring the rotation of 3D objects. Hinckley et al. (1997) conducted a study specifically examining the usability of various methods to rotate a 3D object. Their study compared three interaction techniques. The first was the virtual trackball and required the user to use a mouse to manipulate a 2D interface that simulated a physical trackball. As the user click and



dragged on the virtual trackball, the 3D object rotated as if it was encapsulated in the trackball. Another technique, called the Arcball, was similar to the Virtual Trackball, but provided more realistic transformation the 2D input into the 3D space. Both techniques only allow two dimensions to be manipulated at once (since the mouse only allows for two DOFs). The third technique, the 3D ball, simply allowed the user to rotate the 3D virtual object by rotating a physical ball in their hand. Unlike a trackball, the ball was completely free-floating. A fourth method, the tracker, was identical to the 3D ball but did not have the sphere housing around but rather had the participant hold the rotation sensor directly. The task required participants to match the orientation of a stimulus using the four different control devices. Thus, essentially this study compared two 2-DOF devices with 2 3-DOF devices. Results showed that participants were able to perform the task faster without sacrificing accuracy using the 3-DOF input techniques than they could using the 2-DOF input techniques.

### *Affordance*

The term affordance is common in a variety of domains in psychology and other disciplines including ecological psychology, learning, visual perception, cognitive psychology, artificial intelligence, and robotics. The origin of the concept of affordance may be traced back to the field of motion perception (Gibson, 1966) and is often described in such a way as to emphasize “direct awareness”, that is, that the organism inherently behaves a certain way in regard to and because of characteristics of its environment. In respect to the research in this dissertation, the term affordance is used to describe how characteristics inherent in an object (e.g., size and shape) interact with characteristic of the human user (e.g., intentions, goals, and physical capabilities) to

result in specific behaviors by the user. The following definition captures this interpretation of an affordance: “The affordance of anything is a specific combination of properties of its substance and its surfaces taken with respect to an animal” (Sahin, Girgin, & Ugur, 2006).

Research indicates that the physical form of the device affects how the human attempts to use it to accomplish their goal (Hickley et al., 1997; Zhai et al., 1996; Jeannerod, 1981; Ellis & Tucker, 2000). At least two factors may influence the affordance of a hand-held device: (a) how the device is intended to be used, and (b) the nature of the grip on the device.

The physical form of the input device used in a human-computer interface can suggest both function, what it can be used for, and behavior, how it can be used. For example, Hickley et al. (1997) had success in using the head of a doll as a free-floating orientation control for a task where participants attempted to rotate a 3D object on computer display to match a specific orientation. The authors reasoned that the users tend to use it properly without training because the doll head naturally provided a clear sense of orientation. A sphere with only minimal cues for orientation was also tested and proved to be less usable. While the device was intended to be held in the hand and rotated around, some participants attempted to roll the sphere on a desk.

Another way that the physical form of an input device affects user’s expectation of function and behavior is by how the user would naturally grasp the device. Different types of grasps on hand-held devices are naturally associated with different ways to use the device (Hotta et al., 1981; Jeannerod, 1981; Ellis & Tucker, 2000). For example, gripping with the thumb and pointer finger tends to suggest precision control, while

gripping with the whole hand by making a fist around the device suggests grosser control. The power grasp (Mackenzie, 1994) emphasizes strength and security of the grip and the precision grasp emphasizes dexterity and tumbling of the device. Zhai et al. (1996) analyzed types of grasps using a six degree of freedom docking task and found faster performance for precision grasps emphasizing use of fine muscle groups. Moreover, the muscle groups involved in a grasp suggest how movement patterns to apply to the device.

### *Frames of Reference*

A reference frame is a means of representing the position of objects in the environment. Multiple frames of reference can be assumed when described a visual scene. For example, a scene can be described as seen by an actor that is also within the scene, or as seen from any other direction within the space. One primary difference between the two above examples is that in the former, the actor is cannot be seen (because the actor is the one observing), and in the latter, the actor may be appear in the scene. The term egocentric refers to a frame where entities are represented in respect to the observer, whereas the term allocentric refers to a representation where entities are referred to external to the observer and independent of his or her position (Klatzky, 1998).

It is critical to know the frame of reference in order to understand a description of object placements in a space (e.g., Chua, Weeks, Ricker, & Poon, 2001). Reference frames are theorized to affect 3D rotational control because a human operator's response depends on an understanding of the directional layout of the space. Egocentric versus allocentric frames of reference have been shown to affect how a

human operator manipulates objects and controls movement in 3D space (Klatzky, 1998). The egocentric frame is akin seeing the world through the eyes of the actor in the space and interacting with the world as if one was only a single entity within. In the egocentric frame of reference, one is controlling the actor in the world. The allocentric frame of reference allows the human operator to assume a global perspective of the world. In the allocentric frame of reference, the human operator may assume they are controlling the entire space in relation to their actor.

### *Research Summary*

This research examines an important issue that arises from use common 2D UIs for interacting with 3D objects. The focus of this research, a phenomenon referred to as the inversion effect, describes a specific type of error a user is prone to make while manually controlling the rotation of a 3D object. The inversion effect occurs when a user rotates a 3D object in the direction opposite than was intended. Based on a broad range of research on stimulus-responses compatibility, cognitive psychology, and human-computer interface design, this research postulates four factors that contribute to the inversion effect: affordance, context, visual reference frame, and axis.

### *Motivation for this Research*

One of the primary motivations for this research is to seek empirical data on the strength of population stereotypes for rotational control for objects in a 3D space using interfaces devices with 2 degrees of freedom. While theories of stimulus-response compatibility would suggest that the aforementioned population stereotype would be weak, this particular paradigm and associated effects of subjective preference for control methods and implications on training have not been captured empirically. The

primary hypotheses in this research are: (a) that the population stereotype for vertical rotational control is weaker than for horizontal (lateral) rotational control, (b) that the stereotype for subjective preference over opposite S-R mappings is weak, and (c) that usage of subjectively-preferred S-R mappings does not completely mitigate the occurrence of accidental inversion errors.

#### *Implications of this Research*

This research is primarily applicable to the design of systems when 2-DOF interface devices are used. A variety of tasks utilizing such systems occur across a multitude of domains. Examples of such tasks include remote operation of a camera system, computer-aided drawing, laparoscopic surgery, and controlling an avatar in a video game. It is expected that accidental inversion has the greatest consequences in human-machine systems where a single mistake or only a handful of mistakes can lead to substantial consequences. For example, recent medical advances are leading toward the development of systems that allow surgeons to operate on a patient via robotic apparatus controlled via an interface device that provides a three-dimensional representation of the procedure (see Huber, Taffinder, & Darz, 2003; Reitinger 2005; and Reitinger, Schmalstieg, Bornik, & Beichel (2006)

## EXPERIMENT 1 – POPULATION STEREOTYPE

### Introduction

Population stereotypes are useful to measure the level of consistency that a target population exhibits for a particular stimulus. One way that has been used to measure population stereotypes in previous research has been to present members for a target population with stimuli in the context of a task but without any indication of the correctness of responses (e.g., Hoffmann, 1997). The objective of this first study was to assess the population stereotype associated with rotating a three-dimensional object on two orthogonal axes (see Figure 7) mapped to rotations on the vertical axis (i.e., yaw) and horizontal axis (i.e., pitch).

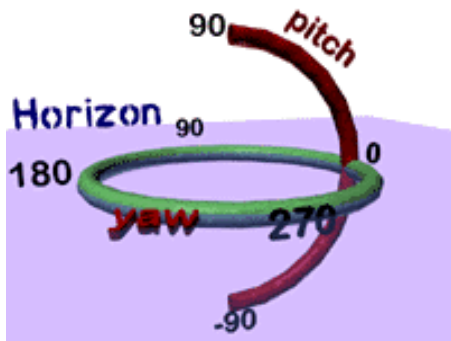


Figure 7: Example of rotation along the Y (horizontal/yaw) and X (vertical/pitch) axes.

This study examined the aforementioned S-R stereotype using a task that presented participants with a three-dimensional stimuli and required response using a two degree of freedom input device. The stimuli used were videos of a human figure (avatar) moving down a corridor and eventually rotating along the X or Y axis. The primary hypothesis was that a strong population stereotype would not be observed for

vertical rotations but will be for horizontal rotations. This is reasoned because previous research suggests that response stereotypes for manual spatial tasks tend to be weaker when referring to spatial relationships that are vertically oriented in a three dimensional space compared to horizontally oriented spatial relationships (e.g., Hoffmann, 1997). For two-dimensional spatial relationships, stereotypes tend to be very strong for both axes (e.g., results by Fitts & Seeger, 1953).

This research also examines four factors that may affect a human operator's response to three-dimensional stimuli: visual frame of reference (egocentric or allocentric), situational context (walking on the ground or flying through the air), and control device (joystick or mouse input device).

Visual reference frames have been shown to affect performance in three dimensional (e.g., Bowman et al., 2001) as well as two-dimensional (e.g., Pennel, et al., 2002) motor-spatial tasks. Klatzky (1998) describes that allocentric and egocentric perspectives differ in terms of how the human perceiver spatially relates points within the three dimensional environment. Allocentric representations promote spatial relationships in terms of an internal Cartesian plane, where the perceiver is contained within, egocentric representations lead to spatial representations relative to the perceiver who is at the center of the environment. Thus, directional references (and spatial responses toward) a common point in a three-dimensional environment may differ based on the reference frame adapted by the perceiver.

The control device used to respond to the stimuli is examined in this study because research indicates that the physical form of the device used to respond to a stimulus can affect what type of response the human operator exhibits (Hickley, Paush,

and Proffitt, 1997; Zhai et al., 1996). The two control devices used in this study were selected because their design suggests response mechanisms. The mouse, held gripped with the palm and controlled with the fore-arm is held using a power-grip, which emphasizes gross motor responses (Mackenzie, 1994). The joystick used in this study, in contrast, was controlled with the thumb which emphasizes precise motor responses.

The reason context is explored in this research is to analyze for the effect of existing stereotypes related to spatial responses to three-dimensional stimuli. In particular, it is reason that some people may expect certain stimulus-response relationships for orientational (i.e., rotational) control when they are flying versus walking on the ground. The reason for this expectation is due to the strong association between moving an input device toward and away from the body to affect the pitch of an aircraft. The standard yoke control in aviation causes the aircraft to nose down when the yoke is pushed away from the body and, conversely, to nose up when the yoke is pulled toward the body.

### *Summary of Experimental Design*

#### *Independent Variables*

Two between-subjects independent variables were manipulated in this mixed design experiment. The first included two types of visual perspective used to present the VE, as follows:

Perspective (two levels, between-subjects)

1. Egocentric – This perspective utilized a first person point of view. This view was presented as seen through the eyes of the participant's avatar. Thus, the body of the avatar was never visible.



2. Allocentric – This perspective utilized a third person point of view. This view was slightly behind participant’s avatar. Thus, the body of the avatar was always visible.

The second between-subjects independent variable included two types of control devices used to respond in the experiment, as follows:

Control device (two levels, between-subjects)

1. Mouse – most common device used with graphic user interfaces on computers
2. Joystick – commonly used in video games and for remote operation of model vehicles (planes, etc) and robots

Each participant experienced only one level from both of these two IVs (perspective and control device) throughout the entire experiment. In addition, two within-subjects (repeated measures) IVs were included in the design of this experiment. The first, axis-type, included two types of stimuli, as follows:

Axis (two levels, within-subjects)

1. Vertical-rotation – the participant’s avatar rotates from 0 degrees along the X axis to face +70 or – 70 degrees.
2. Horizontal-rotation – the participant’s avatar turns from 0 degrees along the Y axis to face +70 or – 70 degrees.

The second within-subjects IV, context, included two levels, as follows:

Context (two levels, within-subjects)

1. Walk – the participant’s avatar jogged along the ground.
2. Fly – the participant’s avatar flew through the air.

Two other within-subjects IVs, inter-trial interval and block, occurred in the design of the study. The first was inter-trial interval (ITI) and had three levels: (a) 2 seconds, (b) 4 seconds, and (c) 6 seconds. This IV determined how much time passed between trials, starting from the point that a response was made. The 3 levels of ITI were included in order to reduce the predictability of the onset time of the critical stimuli. The second was block-order, which had two levels that were defined by the order in which participants were presented with blocks of trials using the flying or walking contexts. Neither ITI nor block-order were anticipated to have an effect on the dependant variables.

### *Dependant Variables*

The dependant variables that were measured in this experiment were polarity and reaction time. Polarity is a measure calculated based on how the participant responded to the different rotations presented in each trial. A participant could respond to a trial in one of five ways: a) push the control device forward, b) push the control device backward, c) push the control device right, d) push the control device left, and e) no response at all. Polarity was calculated as the percentage of responses where: a) for trials where the stimulus rotated +90 degrees vertically, the participant responded by moving the control device forward, b) for trials where the stimulus rotated -90 degrees vertically, the participant responded by moving the control device backward, c) for trials where the stimulus rotated +90 degrees horizontally, the participant responded by moving the control device left, d) for trials where the stimulus rotated -90 degrees horizontally, the participant responded by moving the control device right. The four stimulus-response pairs described above will be referred to as matching. Another group

of stimulus-response pairs, inverted, is used to describe when the control device is moved in the opposite direction for the same stimuli described above. Thus, polarity refers to the percentage of matching S-R pairs and was calculated separately for rotations on the horizontal and vertical axes. For example, a participant who provided matched S-R pairs for every trial where the stimuli consisted of vertical rotations would receive a vertical polarity rating of 100%. A participant who provided half matched and half inverted S-R pairs for trials where the stimuli consisted of vertical rotations would receive a vertical polarity rating of 50%.

The second dependant measure, reaction time was calculated as the length of time it took the participant to respond from the moment that their avatar began rotating away from neutral position (0 degrees of rotation on both the horizontal and vertical axes). Reaction time was calculated the same for rotations on both axes.

### *Hypotheses*

The following are the hypotheses for this study:

1. The primary hypothesis is that, for rotations along the horizontal axis, there will not be a strong stereotype.
2. Rotations along the vertical axis will exhibit a strong stereotype.
3. Response time to vertical rotations will be slower than response time to horizontal rotations.
4. Perspective (allocentric and egocentric) will segment the stereotype response for vertical polarity.
5. Affordance (joystick versus mouse) will segment the stereotype response for vertical polarity.

6. Context (flying versus walking) will segment the stereotype response for vertical polarity

## Method

### *Sampling Pool*

A total of 96 participants from undergraduate psychology, sociology, and digital media classes at the University of Central Florida took part in the study. Participation was voluntary and students were offered extra credit as an incentive. The age range of participants was 18 to 46 (mean = 20.79 years, median = 20 years, standard deviation = 3.93 years). Females accounted for 59 of the participants. Although all students who volunteered were allowed to participate, usable data was limited to 80 participants who produced an error rate no greater than 20% of total responses. The rationale for excluding these participants was that they did not properly understand the task or instrumentation error occurred, which accounted for their high rate of error. A response was considered an error if met at one or more of the following four criteria:

1. The participant made no response during a trial.
2. The participant responded before the critical stimuli occurred. The critical stimulus was defined as the moment that the avatar turned rotated away from the 0 degree “neutral” orientation.
3. The participant responded 5 seconds or later after the critical stimuli occurred.
4. The participant’s responded to a vertical rotation presented by moving their control device horizontally or responded to a horizontal rotation presented by moving their control device vertically.

### *Participant Assignment*

Participants were randomly assigned to one of four groups, based on two types of controllers (mouse or joystick) and two visual perspectives (egocentric or allocentric): (a) mouse control with an egocentric perspective, (b) mouse control with an allocentric perspective, (c) joystick control with an egocentric perspective, or (d) joystick control with an allocentric perspective. The sampling pool was divided, then, so that each of these four groups received 40 participants.

### *Stimuli and Apparatus*

#### *Virtual Environment and Avatar*

A simple virtual environment (VE) was constructed using the freely available Unreal 2 Runtime Engine (Epic Games, 2004). The VE consisted of a long rectangular corridor designed to appear like a tunnel of infinite length (see Figure 8). An animated human avatar was created using the freely available models from Demiurge Studios (2004). Software was written using UnrealScript to automate the movement of the avatar in the VE. An application called UnrealEd (Epic Games, 2004) was used to construct the VE. A set of precise animations (see Table 4) were created using the VE and avatar for playback using the Unreal 2 Runtime Engine. A digital video recorder was used to sample the scenes at a rate of 30 frames per second and at a resolution of 720 x 480 (4:3 aspect ratio) into digital video files. Scripts were produced for all possible combinations of perspective (egocentric and allocentric), context (flying and walking), and axis-type (pitch-up, pitch-down, yaw-left, yaw-right), including three different ITIs. In total, 48 video files were created from the scripts (see Table 4).

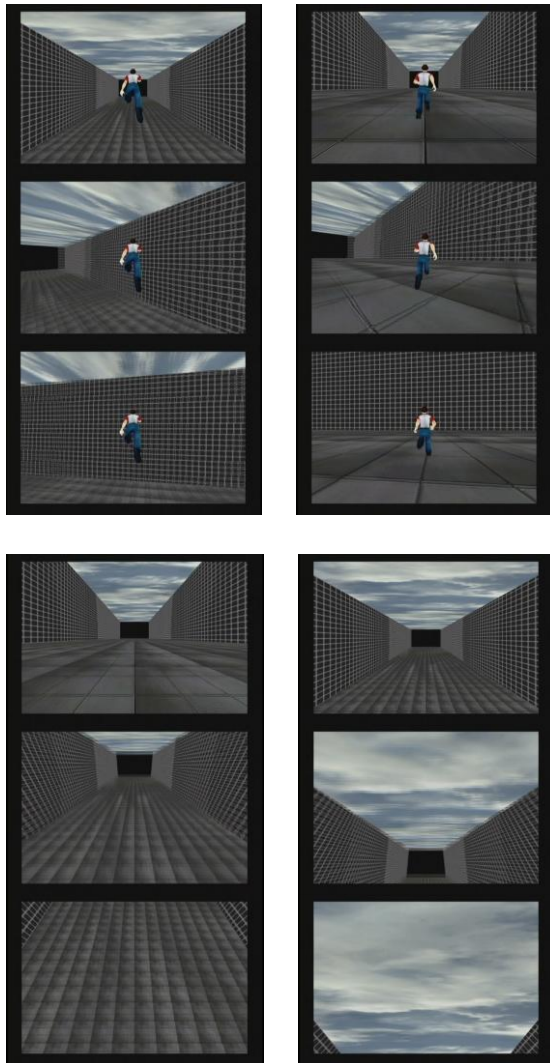


Figure 8: Stimuli from Experiment 1. Clockwise from top-left: allocentric-fly with right turn, allocentric-walk with right turn, egocentric-fly with turn toward the sky, and egocentric-walk with turn toward floor.

### *Stimulus Presentation and Data Collection*

Inquisit (Millisecond Software, Seattle, WA), software designed to present digital stimuli and record responses, was used to control the presentation of the video files (stimuli) and record responses from participants. Custom scripts for Inquisit were written to implement the experiment. The stimuli were presented on a 17 inch diagonal CRT

monitor with a 4:3 aspect ratio. The accuracy of Inquisit for displaying images at the prescribed time has been measured to have a standard error 0.010 msec, and error rang between -1.27 msec and 1.4 msec (De Clercq, Crombez, Buysse, & Roeyers, 2003). The error range for recording reaction time via Inquisit has been previously measured to vary between 1.20 to 3.77 msec, with a mean of 2.79 msec

Two pieces of software were used to overcome technical limitations of Inquisit. First, custom software was made to convert mouse movements into keyboard input. Consideration was given to error tolerance (e.g., processing of the timing, direction, and speed of mouse movements) in the design of this software. Secondly, another custom software application was made to convert joystick movements into keyboard input. The impact of these programs on the performance of the computer system was minimized because both operated in a separate processing thread while Inquisit ran the experiment.

### *Site Apparatus*

The virtual environment (VE) was presented on a computer with a 17 inch diagonal color monitor (32 bit color, 100 Hz vertical refresh rate). A Microsoft X-box controller (thumb actuated joystick) was used in the joystick conditions and a Microsoft optical mouse was used in the mouse conditions. Participants either held the joystick on their lap or used the mouse situated to the right of the monitor.

Table 4: Experiment 1 Stimuli and Trial Conditions

Video File	Perspective	Context	Axis & Direction of Turn	Inter-trial Interval
1	Egocentric	Fly	Horizontal: top	2
2	Egocentric	Fly	Horizontal: top	4
3	Egocentric	Fly	Horizontal: top	6
4	Egocentric	Fly	Horizontal: bottom	2
5	Egocentric	Fly	Horizontal: bottom	4
6	Egocentric	Fly	Horizontal: bottom	6
7	Egocentric	Fly	Vertical: left	2
8	Egocentric	Fly	Vertical: left	4
9	Egocentric	Fly	Vertical: left	6
10	Egocentric	Fly	Vertical: right	2
11	Egocentric	Fly	Vertical: right	4
12	Egocentric	Fly	Vertical: right	6
13	Egocentric	Walk	Horizontal: top	2
14	Egocentric	Walk	Horizontal: top	4
15	Egocentric	Walk	Horizontal: top	6
16	Egocentric	Walk	Horizontal: bottom	2
17	Egocentric	Walk	Horizontal: bottom	4
18	Egocentric	Walk	Horizontal: bottom	6
19	Egocentric	Walk	Vertical: left	2
20	Egocentric	Walk	Vertical: left	4
21	Egocentric	Walk	Vertical: left	6
22	Egocentric	Walk	Vertical: right	2
23	Egocentric	Walk	Vertical: right	4
24	Egocentric	Walk	Vertical: right	6
25	Allocentric	Fly	Horizontal: top	2
26	Allocentric	Fly	Horizontal: top	4
27	Allocentric	Fly	Horizontal: top	6
28	Allocentric	Fly	Horizontal: bottom	2
29	Allocentric	Fly	Horizontal: bottom	4
30	Allocentric	Fly	Horizontal: bottom	6
31	Allocentric	Fly	Vertical: left	2
32	Allocentric	Fly	Vertical: left	4
33	Allocentric	Fly	Vertical: left	6
34	Allocentric	Fly	Vertical: right	2
35	Allocentric	Fly	Vertical: right	4
36	Allocentric	Fly	Vertical: right	6
37	Allocentric	Walk	Horizontal: top	2
38	Allocentric	Walk	Horizontal: top	4
39	Allocentric	Walk	Horizontal: top	6
40	Allocentric	Walk	Horizontal: bottom	2
41	Allocentric	Walk	Horizontal: bottom	4
42	Allocentric	Walk	Horizontal: bottom	6
43	Allocentric	Walk	Vertical: left	2
44	Allocentric	Walk	Vertical: left	4
45	Allocentric	Walk	Vertical: left	6
46	Allocentric	Walk	Vertical: right	2
47	Allocentric	Walk	Vertical: right	4
48	Allocentric	Walk	Vertical: right	6



## *Procedure*

### *Informed Consent*

Participants were told that the study examined reaction time to moving pictures. All participants voluntarily signed an informed consent form.

### *Instructions*

Participants were presented with instructions for their task on the computer screen (see Appendix A). Participants were instructed to respond using their control device as if they caused the rotational changes that they observe the avatar to make. Since pilot testing found that people thought it was odd not to have a mouse pointer during the study, participants in the mouse condition practiced responding by moving the mouse five times. No stimuli were presented during the practice responses. Participants began the study immediately after completing the instructions.

### *Test Session*

A total of 144 trials were presented. The context IV was used to separate the study into 2 blocks, thus half (72 trials) of the trials were in the fly context and half were in the walk context. The order of the 2 blocks was counter balanced across conditions. In each block, half of the trials were pitch and half were yaw. Furthermore, half of the horizontal rotation trials were left turns and half were right turns. Similarly, half of the vertical rotation trials were upward turns and half were downward turns. Eighteen trials of each turn were presented in each block. Thus, each participant experienced 18 up, down, left, and right turns in each of the two blocks. Of the 18 trials per type of turn, there were 6 that occurred after an ITI of 2 seconds, 6 that occurred after an ITI of 4 seconds, and 6 that occurred after an ITI of 8 seconds. All trials were randomly ordered

for each participant in each block. The study took between 18 and 22 minutes. There were no breaks.

Participants responded with their control device according to the instructions. The screen remained black for 2 seconds once a response was made.

## Results

Analyses were done on data for polarity and reaction time. An alpha level of .05 was used for all analyses.

### *Polarity*

Polarity data were first analyzed using a 2 (controller type) x 2 (perspective) x 2 (context) x 2 (axis) ANOVA. Controller type (mouse versus joystick) and perspective (allocentric vs. egocentric) were between-subjects variables, and context (fly vs. walk) and axis (vertical versus horizontal rotation) were within-subjects variables.

Significant main effects were observed for axis,  $F(1, 76) = 17.03, p < .05, \eta_p^2 = .18$ , and context,  $F(1, 76) = 55.65, p < .05, \eta_p^2 = .42$ . The comparison of the pair of means for axis indicate that, on average, participants were significantly more polarized toward matched responses for horizontal rotation trials ( $M = .91, SD = .22$ ) than vertical rotation trials ( $M = .61, SD = .39$ ). The difference between the means was .30 with a standard deviation of .42. A paired-sample t-test further indicated that these means were significantly,  $t(159) = -9.10, p < .01$ , different from each other.

Similar results occurred for the main effect of context. Responses to fly trials ( $M = .81, SD = .33$ ) were slightly, although statistically significant, more polarized toward matched responses than were responses to walk trials ( $M = .71, SD = .38$ ). The difference between the means was .10 with a standard deviation of .29. A paired-

sample t-test further indicated that this difference in the means was significant  $t(159) = -4.00, p < .01$ ).

In addition, the interaction between context and axis was also significant,  $F(1, 76) = 19.92, p < .01, \eta_p^2 = .21$ . To further analyze this finding, separate paired-sample t-tests were conducted on context-axis data pairings (i.e., walk-vertical vs. fly-vertical, and walk-horizontal vs. fly-horizontal). A significant effect was found for walk-vertical vs. fly-vertical mean comparison,  $t(79) = -4.40, p < .01$ , representative of the effect of context type on trials that required a vertical-rotation response. Responses for walk-vertical ( $M = .51, SD = .39$ ) received an almost even proportion of matched and inverted responses. Matched responses were slightly more prevalent for the fly-vertical conditions ( $M = .71, SD = .39$ ). The difference between the means for the walk-vertical vs. fly-vertical comparison was .19 with a standard deviation of .39. There was not a significant effect of walk-horizontal vs. fly-horizontal (i.e., the effect of context on trials requiring a horizontal response).

A 2 (block) x 2 (controller type) x 2 (perspective) x 2 (context) x 2 (axis) ANOVA determined that order of blocks (whether the first block was flying or walking context) did not have a significant effect on polarity.

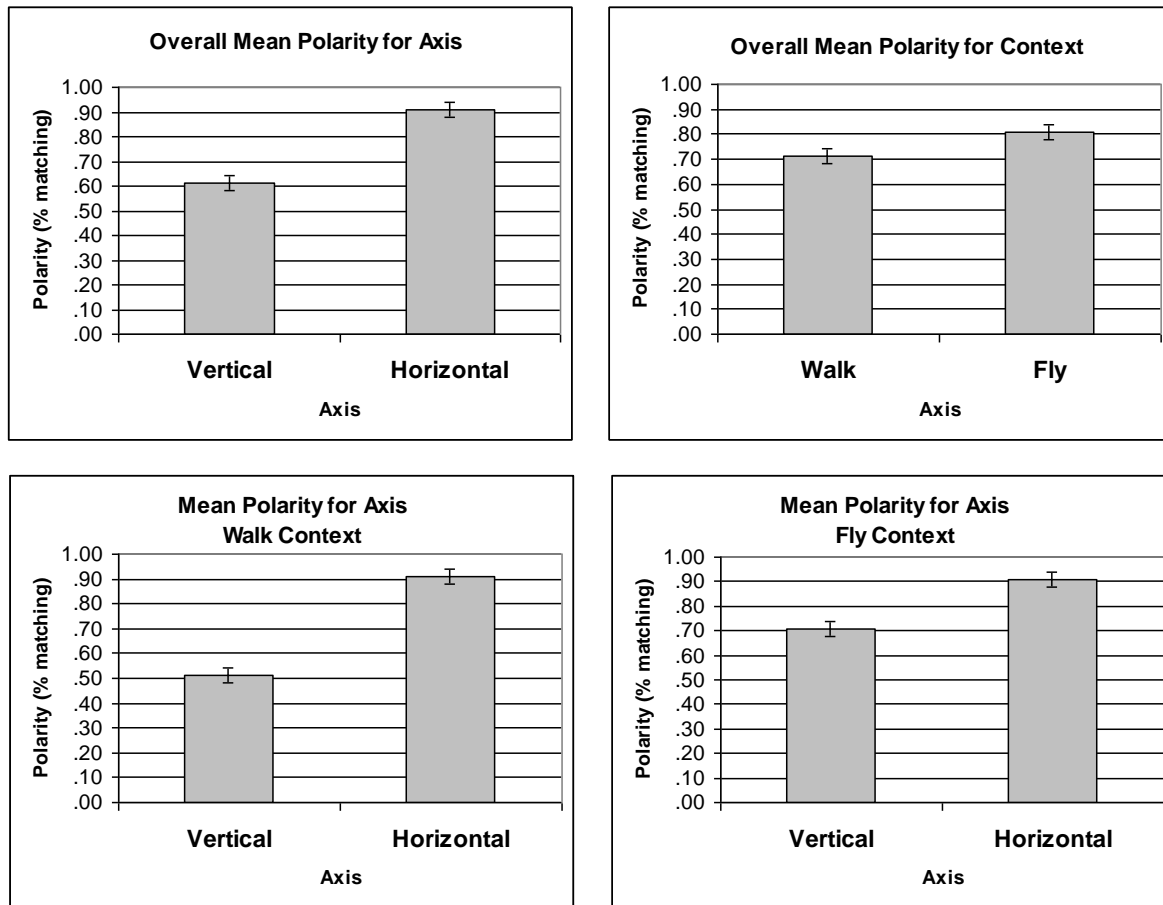


Figure 9: Mean percent of matching responses for axis, context, and axis by context.

### *Reaction Time*

A 2 (controller type) x 2 (perspective) x 2 (context) x 2 (axis) mixed ANOVA was conducted on the response time dependant variable. Controller type (mouse or joystick) and perspective (allocentric or egocentric perspectives) were between-subjects variables and context (fly or walk) and axis (vertical or horizontal).

Significant main effects observed were for axis,  $F(1, 76) = 48.30, p < .01, \eta_p^2 = .39$ , and context  $F(1, 76) = 17.21, p < .01, \eta_p^2 = .19$ . The means from the axis comparison indicate that, on average, participants responded more slowly to the

vertical-rotation stimuli ( $M = 776.02$ ,  $SD = 323.72$ ), than the horizontal-rotation stimuli ( $M = 617.00$ ,  $SD = 136.4$ ). In addition, the means from the context comparison indicate that, participants responded, on average, more slowly to the fly stimuli ( $M = 657.32$ ,  $SD = 172.04$ ), than the walk stimuli ( $M = 735.70$ ,  $SD = 321.50$ ).

A significant interaction was found between context and axis,  $F(1, 76) = 6.61$ ,  $p < .01$ ,  $\eta_p^2 = .01$ . Separate paired-sample t-tests indicated that the pairs of means for both axis,  $t(159) = 7.65$ ,  $p < .01$ , and context  $t(159) = -4.19$ ,  $p < .01$  were significantly different from each other. The difference between the means for the axis comparison was 159.01 with a standard deviation of 263.02. The difference between the means for the context comparison was 78.34 with a standard deviation of 263.37.

Separate paired-sample t-tests were also conducted for walk-vertical ( $M = 838.50$ ,  $SD = 410.99$ ) and fly-vertical ( $M = 713.54$ ,  $SD = 184.63$ ) to further analyze these findings. This refers to the effect of context on trials requiring a vertical-rotation response. A significant effect was found,  $t(79) = 3.57$ ,  $p < .01$ , with a paired difference 124.96 and a standard deviation of 312.74. There was not a significant effect found in a paired sample t-test of walk-horizontal and fly-horizontal (i.e., the effect of context on trials requiring a horizontal response).

A 2 (block) x 2 (controller type) x 2 (perspective) x 2 (context) x 2 (axis) ANOVA determined that order of blocks (whether the first block was flying or walking context) did not have a significant effect on response time.

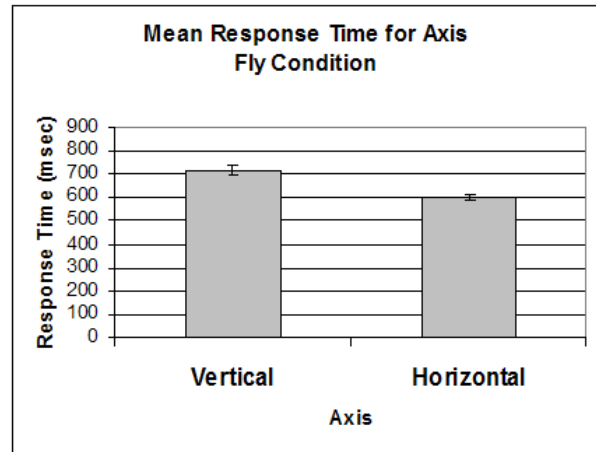
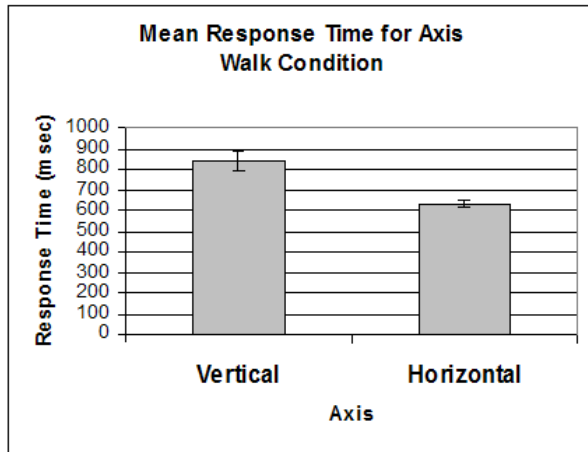
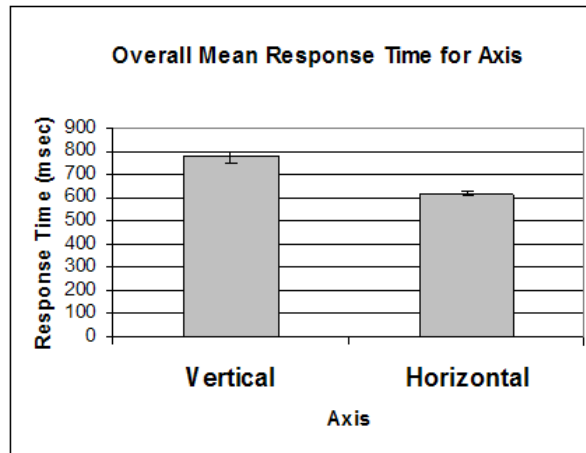


Figure 10: Mean response time for axis and axis by context.

### Discussion

The clearest finding is that the results indicate a strong stereotype for horizontal rotation (as predicted in hypothesis B) and a weak stereotype for vertical rotation (as predicted in hypothesis A). These findings are evidenced by both the polarity and the response time measures. For the polarity measure, results show that over 90% of responses made to horizontal-rotations were matched and less than 10% were inverted. This pattern for horizontal-rotations was consistent across both levels of flying and walking contexts. Results for vertical-rotations indicated that, overall, about 60% of

responses were matched and 40% were inverted. However, unlike for horizontal rotations, the proportion of matched to inverted responses was affected by whether or not the context was flying or walking. Under the walking context, about half of the responses were matched and half were inverted. This result shows that not even the slightest stereotype was found for vertical-rotations under the walking condition. In contrast, a slight, although weak, stereotype, was found for the vertical rotations in the flying condition, where about 70% of responses were matched and 30% were inverted.

The response time data also confirm hypotheses A and B. Overall, responses to horizontal-rotation trials were significantly faster than responses to vertical-rotation trials. According to stimulus-response compatibility theory, this finding can be attributed to information processing and response selection. Responses to horizontal-rotations were faster because of a strong stereotype response to (matched S-R) thus allowing response selection to benefit from automatic activation. In contrast, it may be reasoned that responses to the vertical-rotation stimuli took longer due to extra time spent between stimulus recognition and action selection. Responses time data was consistent with Hoffmann (1997) finding that responses to vertical spatial stimuli, in a three dimensional environment took longer than responses to horizontal spatial stimuli.

A surprising finding was that input device type and perspective did not have significant effects on polarity nor response time. It was hypothesized that perspective and controller-type would have a segmenting effect on the proportion of matched to inverted responses. One reason for this negative finding may be that the control devices used, a mouse and joystick, did not strongly afford a particular behavioral pattern that

affected the response. A stronger affordance might be obtained with, for example, a flight yoke versus a hand lever.

The finding that perspective did not have a significant effect on polarity may have been because the variable did not cause the participants to acquire two different spatial frames of reference. Further research may benefit from more immersive conditions such as a virtual-reality display that covers the entire field of view.



## EXPERIMENT 2 – SUBJECTIVE PREFERENCE FOR VERTICAL ROTATION CONTROL

### Introduction

Experiment 1 provided evidence that rotational control in a three-dimensional space, when controlled via a two degree of freedom interface device, has only a strong population stereotype for vertical axis rotations and that, for horizontal axis rotations, the stereotype is not only weaker, but can also be mediated by situational factors such as the context of the task being performed. Two stimulus-response mappings were examined for both vertical rotations and horizontal rotations in the first experiment. These two mappings were opposites, such that one mapping associated rotations to the left with control responses and the other associated rotations to the left with control responses to the right. The same pattern of mapping was also tested for vertical rotations with forward and backward control responses. Results from the first experiment suggested that, for the population that was studied, nearly half of the members have one stimulus-response expectation while the other half has the opposite stimulus-response expectation.

The purpose of this second experiment is to assess whether the same population examined in the first experiment subjectively prefers one stimulus-response mapping over the other after having experience with both. The rationale for this study is that many human-computer interfaces allow operators to select the S-R mappings for vertical rotation. For example, the infrared camera on the United States Navy P3 Orion aircraft allows the operator to select between two opposite mappings for vertical rotations. This study examines whether one S-R mapping tends to be selected over

another when the human-operator has experience using more than one S-R mapping. A secondary goal in this study is to investigate whether perspective and context play a role in the subjective preferences. Since, as predicted, the most important results from the first study focused on vertical rotations, this second experiment does not examine the horizontal rotation axis.

### *Summary of Experimental Design*

#### *Independent Variables*

Three within-subjects independent variables were manipulated in this experiment: perspective, context, and input-response (I-R) mapping. The first two IVs, perspective and context, were implemented in the same manner as in Experiment 1.

Perspective included two types of views for displaying the VE:

#### Perspective (two levels, within-subjects)

1. Egocentric – This perspective utilized a first person point of view. This view was presented as seen through the eyes of the participant's avatar. Thus, the body of the avatar was never visible.
2. Allocentric – This perspective utilized a third person point of view. This view was slightly behind participant's avatar. Thus, the body of the avatar was always visible.

The second independent variable context, included two levels, as follows:

#### Context (two levels, within-subjects)

1. Walk – the participant's avatar jogged along the ground.
2. Fly – the participant's avatar flew through the air.

The third independent variable, I-R mapping, included two levels, as follows:

Input-response mapping (two levels, within-subjects)

1. Matching– moving the mouse forward caused the view to pitch up to at most 90 degrees (looking straight at the ceiling). A downward movement caused the view to pitch down to a maximum of 90 degrees (looking straight at the floor).
2. Inverted – moving the mouse backward caused the view to pitch down to a maximum of 90 degrees (looking straight at the floor). A downward movement caused the view to pitch up to a maximum of 90 degrees (looking straight at the ceiling)

Each participant experienced eight conditions of these three IV's (perspective, context, and I-R mapping) twice (see Table 5) throughout the experiment.

#### *Dependant Variables*

All dependant variables in this experiment were self reported and consisted of perceived task difficulty, task performance, and preference for I-R mapping. A questionnaire, presented in Appendix A, was used to gather data on the DVs. Trials were organized in sequential pairs referred to as trial-pairs. Both trials of a trial-pair consisted of the same conditions for context and perspective, but one trial of the pair used the matching I-R mapping and the other used the inverted I-R mapping. After both trials of a trial-pair, participants were asked to answer three questions which asked about perceived task difficulty and performance as it related to the I-R mapping used for that trial. Participants answered two additional questionnaire items after having experienced both trials of a trial-pair, which meant he or she had experienced both IR mappings for the otherwise same sets of conditions. These two questions both asked

about preference between the two IR mappings. The first question asked the participant to rate their preference on a 7-point scale and the second required explicit indication (i.e., two choice) of which I-R mapping was preferred.

### *Hypotheses*

The main hypothesis for this study is that preference for I-R mappings will be neutral. In addition, context will have an effect on preference for I-R Mapping. Perspective is not predicted to have an effect on preference for I-R Mapping do to the negative results from Experiment 1.

## Method

### *Sampling Pool*

A total of 50 participants from undergraduate psychology classes at the University of Central Florida took part in the study. Participation was strictly voluntary, and students were offered extra credit as an incentive. The age range of participants extended from 18 to 28 ( $M = 20.24$  years, median = 20 years,  $SD = 2.14$  years). A total of 16 males and 34 females took part.

### *Participant Assignment*

This study used only within-subjects IV's so participants were not placed into separate groups.

### *Stimuli and Apparatus*

#### *Virtual Environment*

A virtual environment (VE) was created that was similar to the VE used in Experiment 1. The VE used in this experiment was interactive and allowed the participant to control the vertical gaze angle of their avatar. The VE was presented as if

through the eye's of the participant's avatar in the egocentric condition and from approximately 3 real world feet (1 foot = 16 VE units) behind the avatar in the allocentric condition. The avatar was only visible in the allocentric condition. The VE presented a corridor made up of side walls and a floor. The ceiling was open and showed a blue sky with clouds (see Figure 11). The same avatar used in the first experiment was also used in this experiment. The participant could not affect the movement speed nor movement direction (e.g., heading) of the avatar.

#### *Stimulus Presentation and Data Collection*

Each trial in the study consisted of a corridor with a random arrangement of targets that were initially invisible. There were 8 combinations of trial conditions, made up of the three IVs used in this study (see Table 5). Once the trial began, the avatar moved down the corridor at a constant rate. Targets appeared at random times and ceiling or floor locations in front of the avatar as the avatar moved down the corridor. The task was simply to use a standard computer mouse to move the avatar's vertical line of sight so as to gaze directly at targets as they appeared on either the floor or ceiling. Moving the mouse forward or backward caused the line of sight to rotate vertically, pivoting on the avatar's head. Each corridor had six targets, three of which appeared on the ceiling and three which appeared on the floor. The targets were blue rectangles. The panels were not visible until the participant's avatar came within a short distance of them, at which time they appeared immediately. The avatar, which never stopped moving down the corridor during a trial, completely moved passed a target 3 seconds after it appeared. A randomization algorithm was used to create thirty-six corridors with different placements of the targets.

Custom software was written to record data in real-time. This program recorded data at a frequency of one sample every 10 millisecond and achieved a high rate of reliability. Another piece of custom software was used by the experimenter to trigger the start of each trial and coordinate the correct conditions for each trial. A third piece of custom software was written to process raw output data into a usable format after the study had completed.

The experiment consisted of 16 trials composed of the following IVs: context (fly vs. walk), perspective (allocentric versus egocentric), and I-R mapping (normal versus inverted) controls. Eight conditions (see Table 5) were created by combining these three IVs. Trials were presented as pairs referred to here as trial-pairs. Both trials of a trial-pair consisted of the same conditions for context and perspective, but one trial of the pair used the matching I-R mapping and the other used the inverted I-R mapping. Thus, participants experienced each of the eight conditions twice. For example, if egocentric-flying-matched were the conditions of the first trial in a pair, the next trial would be egocentric-flying-inverted. The order of trial-pairs was randomized for each participant.

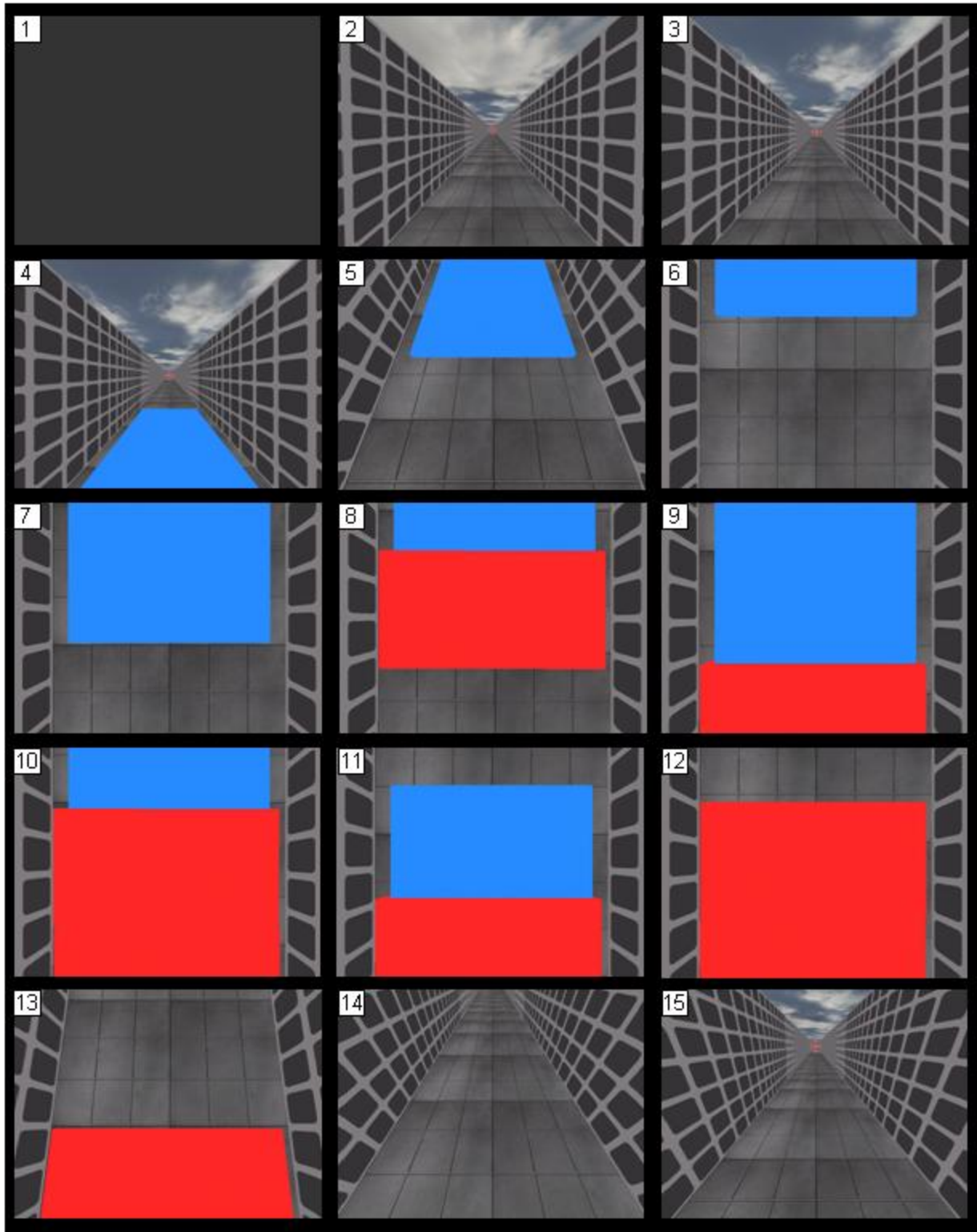


Figure 11: Experiment 2 screenshots of capturing a floor target.

Table 5: Notional trial sequence for Experiment 2.

Actual Trial	Trial Pair	Order of Trials in Pair	Perspective	Context	First Trial Control-mapping	Second Trial Control-mapping
1	1	1 <sup>st</sup>	Egocentric	Flying	Matching	Inverted
2		2 <sup>nd</sup>				
3	2	1 <sup>st</sup>	Egocentric	Flying	Inverted	Matching
4		2 <sup>nd</sup>				
5	3	1 <sup>st</sup>	Egocentric	Walking	Matching	Inverted
6		2 <sup>nd</sup>				
7	4	1 <sup>st</sup>	Egocentric	Walking	Inverted	Matching
8		2 <sup>nd</sup>				
9	5	1 <sup>st</sup>	Allocentric	Flying	Matching	Inverted
10		2 <sup>nd</sup>				
11	6	1 <sup>st</sup>	Allocentric	Flying	Inverted	Matching
12		2 <sup>nd</sup>				
13	7	1 <sup>st</sup>	Allocentric	Walking	Matching	Inverted
14		2 <sup>nd</sup>				
15	8	1 <sup>st</sup>	Allocentric	Walking	Inverted	Matching
16		2 <sup>nd</sup>				

The goal of the task was to look up or down at targets as they appeared as quickly as possible while automatically moving through each a corridor. Each trial



consisted of a single corridor. Each corridor had 6 targets, 3 of which appeared on the floor and 3 of which appeared on the ceiling. The location of targets along each corridor as well as the order of ceiling and floor targets was randomized. Once a target appeared participants were instructed to keep looking at it until they completely moved past the target upon which time an audio cue was heard. Once a trial began, the participant's avatar moved down the corridor automatically at a constant speed. The avatar flew through the corridor for the flying context and walked through for the walking context. The participant had control over the vertical rotation of the line of sight by moving the mouse forward and backward. The I-R mapping condition defined whether an upward or downward rotation occurred during a trial when the participant moved the mouse either forward or backward.

#### *Site Apparatus*

The virtual environment was implemented on the same computer and displayed used for Experiment 1. Participants used a mouse for responding. Speakers presented audio cues at about 60 decibels. The experimenter sat at workstation that was adjacent to the participant and hidden behind a wall. This workstation consisted of a display that cloned the participant's monitor in real-time and a wireless keyboard and mouse linked to the computer that ran the study. The experimenter used the keyboard to trigger scripts that automated the sequence trials and the triggering of data recording.

#### *Questionnaire*

The following three questions were asked after every trial:

1. How difficult was the last tunnel?
2. How difficult was it to control where you were looking in the last tunnel?

3. When responding, how many times did you accidentally turn the wrong way, even for a brief moment?

The following two questions were asked after every pair of trials:

4. Which method of control did you prefer?
5. If you were forced to choose a method to use for now on, which would you prefer?

A 7-point Likert scale was used for questions 1 and 2. Question 5 had only two response options. Appendix A provides details on the questionnaire items.

### *Procedure*

#### *Informed Consent*

Participants were told that the study would measure their opinion about different types of controls and displays. All participants voluntarily completed an informed consent form.

#### *Instructions*

Participants were presented with instructions for their task first verbally (See Appendix A) and were then shown a video clip of the task being performed. Participants were instructed to respond by using the mouse to look at each target when they appeared.

#### *Test Session*

After the first trial of each trail-pair, participants answered three questions and verbally indicated when they were done. Responding to these three questions generally took less than one minute. The experimenter then began the second trial of the pair, which used a I-R mapping opposite to that which was used in the first trial of the pair.

The participant then answered the same 3 questions as they pertained to the second trial of the pair and then answered two more questions that compared the two trials of the pair together.

## Results

Analyses were done on data for questionnaire items with an alpha level of .05 used for all analyses.

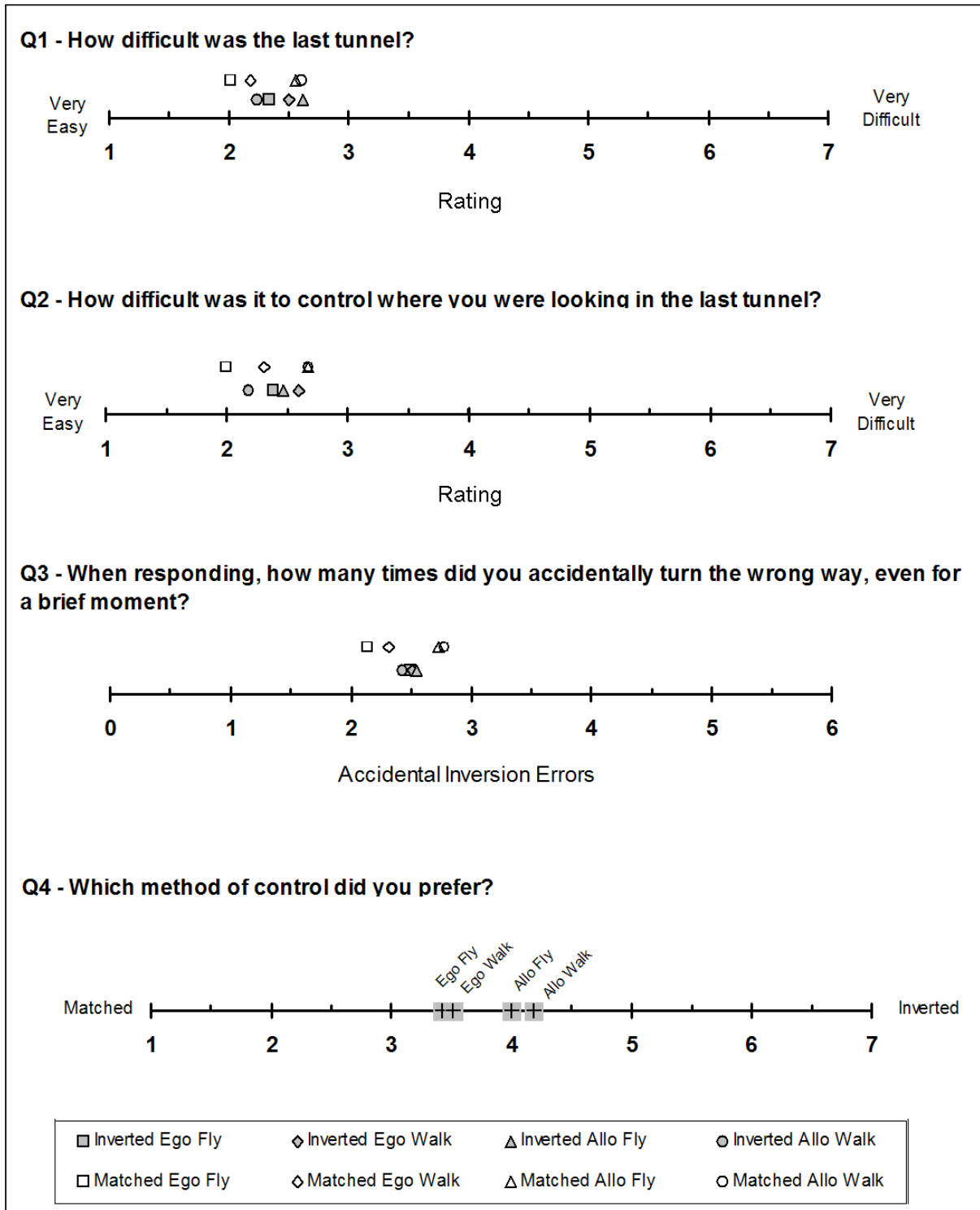


Figure 12: Experiment 2 questionnaire responses.

### *Subjective Measures*

*Task Difficulty.* Context, perspective, and I-R mapping were used as IVs in a repeated measures ANOVA on data from question 1 to test whether these IVs affected the subjectively rated overall difficulty of the task. Results indicate that perspective had a significant effect on the ratings ( $F(1,67) = 12.816, p < .01, \eta_p^2 = .16$ ). Under the allocentric perspective, more participants felt the overall task was more difficult ( $M = 2.53, SD = 1.12$ ) than when using the egocentric perspective ( $M = 2.295, SD = 1.10$ ).

Results also indicated a significant interaction between control-mapping and perspective ( $F(1,67) = 19.50, p < .01, \eta_p^2 = .99$ ). An analysis of the means showed that when using matching control mapping, allocentric ( $M = 2.18, SD = 1.05$ ) was rated more difficult than egocentric ( $M = 2.665, SD = 1.09$ ), but not when using the inverted control method. A paired-sample t-test further indicated that these means were significantly,  $t(133) = -5.13, p < .01$ , different from each other.

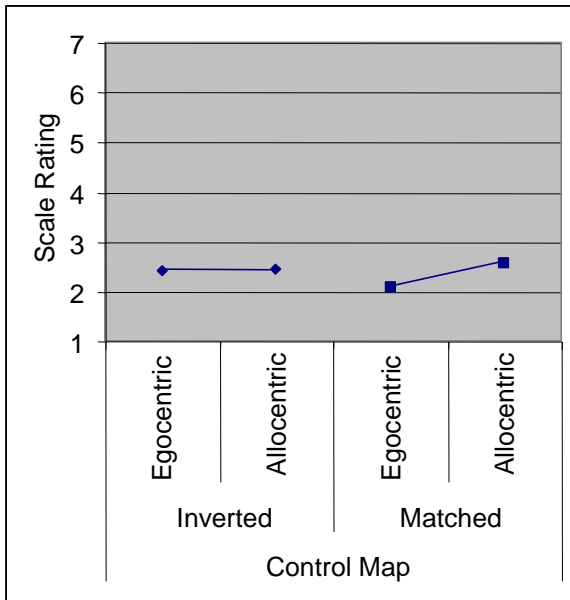


Figure 13: Perspective versus control-map for question 1.

The same repeated measures ANOVA was run on data from the second questionnaire item which pertained to difficulty controlling the view. A main effect was found for perspective ( $F(1, 67) = 12.20, p < .01, \eta_p^2 = .154$ ). An interaction of I-R Mapping and perspective was also found, ( $F(1, 67) = 31.64, p < .01, \eta_p^2 = .32$ ). An analysis of the means showed that when using Matching control mapping, allocentric ( $M = 2.56, SD = 1.08$ ) was rated more difficult than egocentric ( $M = 2.06, SD = 1.11$ ), but not when using the inverted control method. A paired-sample t-test further indicated that these means were significantly different from each other,  $t(135) = -5.732, p < .01$ .

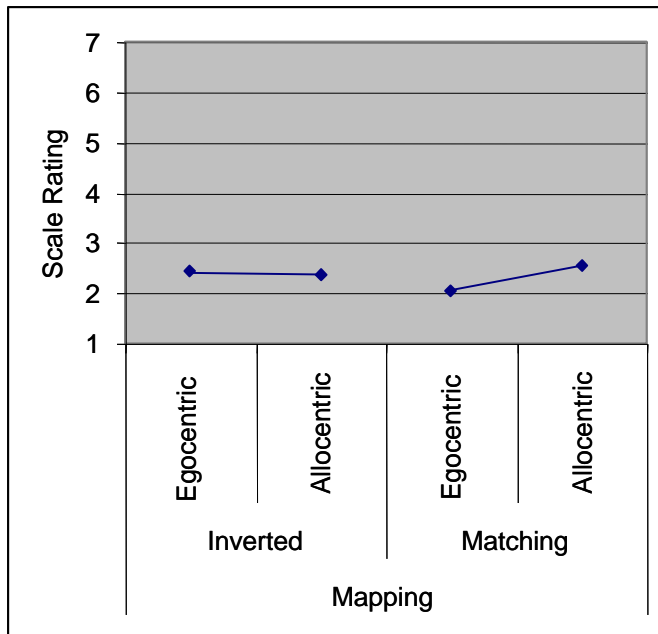


Figure 14: Control-mapping versus perspective for question 2.

*Self-report Performance.* Data for the self-reported number (question 3) of errors was analyzed via the same repeated measures ANOVA used previously. Results indicated an interaction between I-R mapping and perspective,  $F(1, 67) = 18.38, p < .001, \eta_p^2 = .215$ . An analysis of the means showed that when using matching I-R mapping, egocentric ( $M = 3.25, SD = 3.01$ ) reported fewer errors than allocentric ( $M = 2.54, SD = 2.30$ ), but not when using the inverted I-R mapping. A paired-sample t-test further indicated that these means were significantly different,  $t(135) = 3.53, p < .001$ , from each other. It should be noted that the self reported error rate showed a similar interaction however the egocentric error rate was associated with a higher than egocentric.

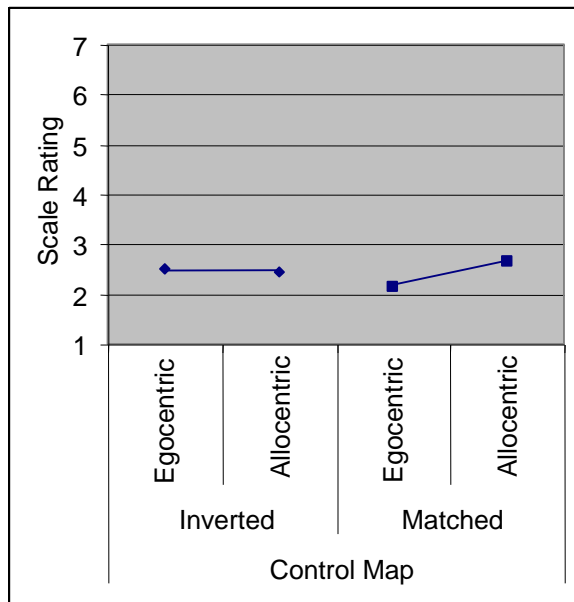


Figure 15: Control-mapping versus perspective for question 3.

*Subjective Preference.* Subjective preference data was analyzed in two ways: a) using a one-sample T-tests comparing the overall mean preference rating as well as ratings broken down by Context and Perspective against the neutral response value, and b) using the same repeated measures ANOVA used for previous analyses.

Table 6 presents the means for the fourth question, subjective preference, across the 4 perspective-context conditions (egocentric-walk, egocentric-fly, allocentric-walk, and allocentric-fly) as well as for the overall mean. Results from the T-test are also provided in Table 6. Only one of the means was significantly different from the neutral scale value of 4: Egocentric ( $t(99) = -2.688, p = .008$ ).



Table 6: Results from question 4 of Experiment 2.

		T-test versus Neutral					
		Value (4)					
		Std.	Std.	T		Mean	
		Dev	Error	Value	Sig.	Difference	
		Mean					
	Overall	3.760	2.431	.172	-1.396	.164	-.240
	Egocentric	3.390	2.269	.227	-2.688	.008*	-.610
	Allocentric	4.130	2.541	.254	.512	.610	.130
	Flying	3.750	2.380	.238	-1.050	.296	-.250
	Walking	3.770	2.494	.249	-.922	.359	-.230
Egocentric	Flying	3.42	1.967	.278	-1.801	.078	-.580
	Walking	3.420	2.278	.322	-1.982	.053	-.640
Allocentric	Flying	3.360	2.284	.323	.230	.819	.080
	Walking	4.080	2.456	.3474	.481	.633	.180

Note: 7-point Likert scale where 1=participant preferred Matching I-R mapping and 7=participant preferred Inverted I-R mapping. \* Indicates significance ( $p < .05$ ).

The ANOVA analysis found a main effect for perspective ( $F(1, 49) = 10.33$ ,  $p < .01$ ,  $\eta_p^2 = .174$ ). The egocentric perspective slightly favored the inverted control method ( $M = 3.465$ ) than the allocentric ( $M = 4.09$ ). Although statistically significant, both means indicated that participants generally felt neutral about the two I-R mappings.

Data from question #5 was not analyzed but was used to categorize participants into a subjective preference group for experiment 3.

## Discussion

Overall, the self-report data did not differ substantially between conditions, despite that numerous differences were statistically significant. The following is a summary of the statistically significant results:

Regarding preference for the I-R mappings (question 4):

The overall preference rating was not different from the neutral value on the preference scale. A slight preference for the matched I-R mapping was found for the egocentric perspective compared to the allocentric perspective.

Regarding the difficulty of the task (questions 1 and 2):

For the matched I-R mapping, the egocentric perspective was rated as slightly easier than the allocentric perspective. There was no difference between perspectives for the inverted I-R mapping.

Regarding the self-reported number of accidental inversion errors:

For the matched I-R mapping, participants reported fewer errors for the egocentric perspective than for the allocentric perspective. There was no difference between the mean number of errors reported for the inverted I-R mapping.

On average, participants did not prefer one I-R mapping over the other. Results from the questionnaire clearly do not suggest a strong subjective stereotype for vertical rotation I-R mapping.

Results from the questionnaire items on task difficulty (questions 1 and 2) indicate that participants found the task more challenging using the allocentric perspective, but only when the matched I-R mapping was being used. The condition for which the task rated as easiest and where the fewest errors were recorded was when

matched I-R mapping was being used during egocentric flying, although the same for egocentric walking was a close second according to results from the first questionnaire item.

One interesting finding was that that while perspective had an effect of self reported task difficulty and I-R preference, it did not have an effect on the "natural" responses that elicited in Experiment 1. In Experiment 1, context affected the frequency of matching and inverted responses but perspective did not have significant effects. However, perspective did have significant effects in Experiment 2 on perceived difficulty, performance, and subjective preference. It should be noted, however, that the differences between the mean ratings were generally quite small.

## EXPERIMENT 3 – PREFERENCE AND INTERFERENCE

### Introduction

The third experiment examines the effects of accidental inversion on training a simple rotation task and has two key research questions. The first research question is whether subjective preference for a particular I-R mapping affects performance after having trained with the non-preferred input-response mapping. The second research question is, after having practiced with a particular I-R mapping, whether performance is affected by short-term, unexpected exposure to the opposite I-R method. This experiment, like the second, focuses exclusively on vertical rotations.

The first rationale behind this experiment builds upon results from the first two experiments. Experiment 1 indicated there is not a strong, single stereotype response for vertical rotation. Experiment 2 indicated that there is also not a strong, single subjective preference for one I-R method for controlling vertical rotation. Results from the second experiment indicated that nearly equal numbers of participants subjectively preferred both I-R mapping. This third experiment further examines whether performance by operators using their subjective preferred I-R mapping follows principles established by research in stimulus-response theory. Thus, this experiment effectively asks the question if operators can self select the I-R mapping for which they have the strongest stereotype response.

The second rationale for this experiment is that even a small number of errors due to accidental inversion can have serious consequences for some human-machine systems. Given that I-R mapping may be configured by the operator, it is reasonable to expect that the operator may on occasion accidentally use their non-preferred I-R

mapping. Anecdotal evidence indicates that such circumstances already occur in some human-machine systems. In such circumstances, the operator would likely switch the I-R mapping to their preferred mapping at their earliest possible opportunity. Usage scenarios that may force an operator to use their non-preferred I-R mapping include: a) when the human-computer interface does not permit the mapping to be changed, b) the operator does not know how to change the mapping, c) the operator selects a different mapping than desired due to human error or poor labeling, or d) the operator is not provided an indication of the current I-R mapping. Any of the above scenarios may have caused the operator to have been exposed to an I-R mapping that is opposite to the one that they prefer, even if only for brief period of duration.

### *Summary of Experimental Design*

#### *Groups*

Two groups were formed defined by which participants used their preferred and which used their non-preferred I-R mappings. The two I-R preference groups are described below:

#### I-R preference (two groups)

1. Preferred – participants used the I-R mapping that they subjectively preferred.
2. Non-preferred – participants used the I-R mapping that was opposite to the mapping that they subjectively preferred.

### *Independent Variables*

Two within-subjects independent variables were manipulated in this experiment, speed and phase. Speed defined how quickly the participant had to respond to critical stimuli and was defined as follows:

Speed (three levels, within-subjects)

1. Slow – the task progressed at the slowest speed which allowed for the most time to react when a critical stimulus appeared.
2. Medium – the task progressed at twice the slow speed which allowed for less time to react to a critical stimulus than the slow speed, but less than the fast speed.
3. Fast – the task progressed at three times the slow speed which allowed for the least time to react when a critical stimulus appeared.

Phase defined the speed and I-R preference conditions for each trial and is outlined below:

Phase (four levels, within-subjects)

1. Training – consisted of the first 18 trials. Of these, the first 6 trials were at a slow speed, the second 6 trials were at a medium speed, and the last 6 trials were at fast speed. Participants in the preferred group used their preferred I-R mapping and participants in the non-preferred group used their non-preferred I-R mapping.
2. Baseline – consisted of the next 6 trials after the training phase and was at a fast speed. Groups used the same I-R preference used during the training phase.

3. I-R mapping switch phase (switch phase) – consisted of the next 6 trials after the baseline phase and was at a fast speed. Groups used the I-R preference opposite to that used during the training and baseline phases.
4. I-R mapping return phase (return phase) – consisted of the next 6 trials after the switch phase and was at a fast speed. Groups returned to using the same I-R preference used during the training and baseline phases.

#### *Dependant Variables*

The primary dependant variable measured in this experiment was number of accidental inversion errors. An accidental inversion occurred when a participant responds to a target by rotating their avatar in the direction opposite the target.

#### *Hypotheses*

1. Preferred group will perform better (i.e., make fewer errors) than the non-preferred group during the training phase.
2. The performance of the non-preferred group will be comparable to the preferred group at the baseline phase.
3. Performance of both the preferred and non-preferred groups will be worse at the switch phase than the baseline phase.
4. Performance of the non-preferred group will be better than the preferred group in switch phase. This hypothesis is based on the prediction that performance of the preferred group, having successfully selected the I-R mapping that is more consistent with their stereotype response than the other

- I-R mapping option, will be worse than performance of the non-preferred group, which, according to stimulus response theory, should be less affected by using a less compatible I-R mapping.
5. Performance of both the non-preferred and the preferred groups will not achieve the same level of performance in the return phase as was observed in the baseline session.
  6. Performance of the preferred group will be better (i.e., fewer errors) than the non-preferred group in the return phase.

## Method

### *Sampling Pool*

The same 50 participants that took part in Experiment 2 also took part in this experiment.

### *Participant Assignment*

Participants were assigned to one of two groups, congruent or opposite, based on results from experiment 2. The congruent group used the I-R mapping that they preferred at the end of experiment 2 and the opposite group used the I-R mapping that they did not prefer. Both groups had 25 participants.

### *Stimuli and Apparatus*

The same virtual environment (VE) and avatar from experiment 2 was used. A computer mouse was used to control the avatar's line of sight. The VE was fixed at an egocentric perspective and the avatar only walked down the corridors.



## *Procedure*

### *Instructions*

Participants were instructed to rotate their line of sight to gaze directly at blue targets as when they appeared as the avatar moved down the corridor. A tone would play once a participant passed a target (regardless if they were gazing at it or not). Participants were instructed to return the line of sight to the center of the facing wall at the end of the corridor when the tone was heard.

### *Task*

The task was the same used in experiment 2 with the exception of the speed of the avatar's movement through the VE tunnels during the training session. For the first set of six training trials, the avatar moved down the tunnel at the slowest speed. For the second set of six training trials, the avatar moved down the tunnels at the medium speed. For the third and final set of six training phase trials, the avatar moved down the tunnel at the fastest speed. The avatar moved down the corridors at the fastest speed for trials in the baseline, transfer, and Return phases.

### *Test Session*

Each corridor represented a single trial. Participants completed a total of 48 trials consisting of 18 training trials, 6 baseline trials, 6 interference trials, and 6 return trials (see Table 7). Each trial consisted of 6 targets. The order of the trials and placement of the target panels were randomized.

Table 7: Phases in Experiment 2

Training Phase (T)	$T_{S1}, T_{S2}, T_{S3}, T_{S4}, T_{S5}, T_{S6}$
	$T_{M1}, T_{M2}, T_{M3}, T_{M4}, T_{M5}, T_{M6}$
	$T_{F1}, T_{F2}, T_{F3}, T_{F4}, T_{F5}, T_{F6}$
Baseline Phase (B)	$B_{F1}, B_{F2}, B_{F3}, B_{F4}, B_{F5}, B_{F6}$
Switch Phase (S)	$S_{F1}, S_{F2}, S_{F3}, S_{F4}, S_{F5}, S_{F6}$
Return Phase (R)	$R_{F1}, R_{F2}, R_{F3}, R_{F4}, R_{F5}, R_{F6}$

---

Speed of movement: S=Slow, M=Medium, F=Fast

## Results

Results are presented below by phase and used an alpha level of .05.

### *Training Phase*

Data on the number accidental inversion errors was first analyzed separately for each training phase. One way ANOVAs were conducted on I-R mapping preference group (preferred versus non-preferred) against each set of the training phase plus the baseline phase (training set 1, training set 2, training set 3, baseline set). A main effect was found for the first training set ( $F(1,48) = 4.86, p = 0.03$ ). Analysis of the means indicate that participants in the non-preferred conditions made more errors in the first training phase ( $M$  non-preferred group = 3.16,  $SD=2.85$ ,  $M$  preferred group = 1.64,  $SD=1.93$ ). It is also important to point out that the mean number of errors between both preferred group and non-preferred groups were not statistically significant during the third training set and baseline phase.

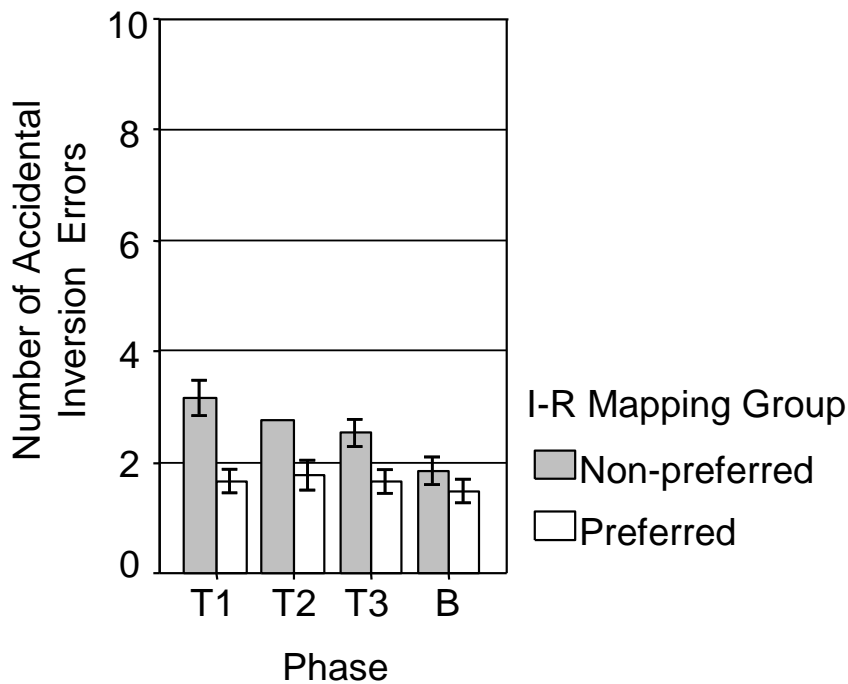


Figure 16: Accidental inversion errors training & baseline phases by group.

### *Switch and Return Phases*

A 3 (critical phase) x 2 (I-R mapping group) mixed ANOVA was also conducted on the number of accidental inversion error data. Critical phase (baseline, switch, and return) was a within subjects variable and I-R mapping group (preferred versus non-preferred) was a between-subjects variable. A main effect was for phase ( $F(2, 96) = 56.47, p < .01, \eta_p^2 = .541$ ). An analysis of the means using t-tests indicates significant differences between the baseline ( $M = 1.66, SD = .26$ ) and the switch phases ( $M = 7.18, SD = .73$ ) ( $t(49) = -7.93, p < .01$ ), the switch and the return phases ( $M = 2.28, SD = .31$ );  $t(49) = 7.12, p < .01$ ), as well as between the baseline and the return phases ( $t(49) = -2.12, p = .04$ ).

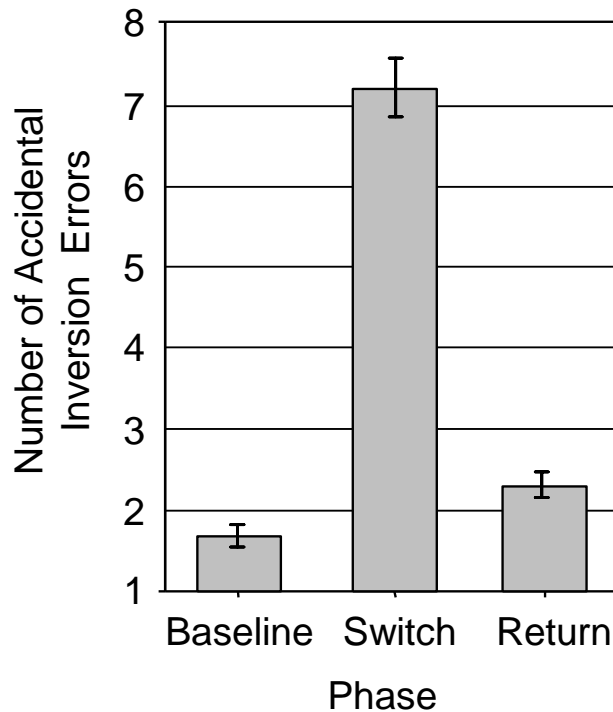


Figure 17: Overall means for accidental inversion errors for critical phases.

More importantly, an interaction was found between phase and I-R mapping group ( $F(2, 96) = 4.626, p = .012, \eta_p^2 = .088$ ). This result was further analyzed using separate one-way ANOVAs on I-R mapping preference group (preferred versus non-preferred) against each critical phase (baseline, switch, and return). A main effect was found for the switch set ( $F(1, 48) = 4.06, p = 0.05, \eta_p^2 = ?$ ). Analysis of the means indicate that participants in the non-preferred conditions made fewer errors in the switch phase ( $M$  non-preferred group = 5.72,  $SD=5.07, M$  preferred group = 8.64,  $SD=5.18$ ).

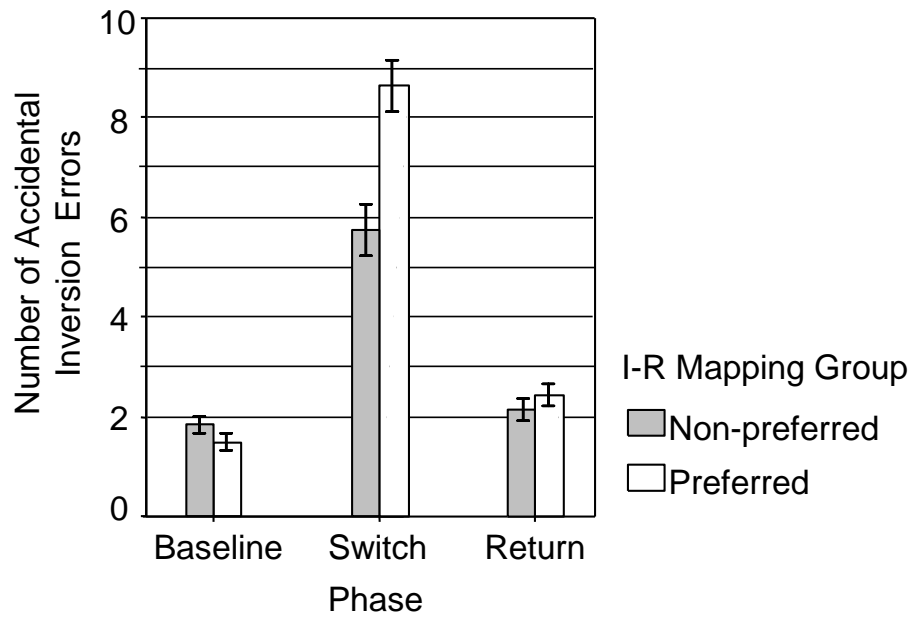


Figure 18: Accidental inversion errors for critical phases by I-R mapping group.

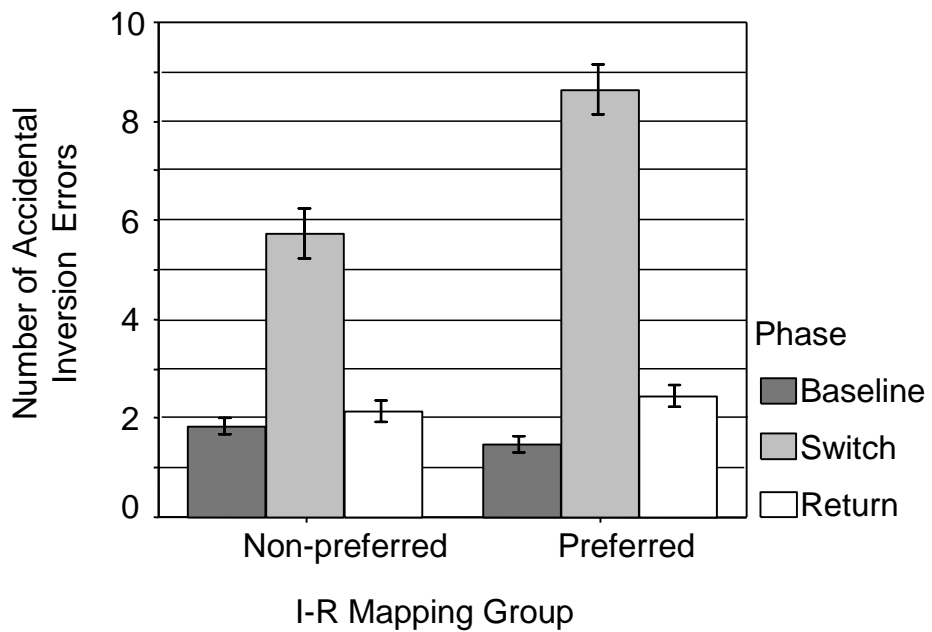


Figure 19: Correspondence between I-R mapping groups.

## Discussion

Results from the training phase supported the first two hypotheses. As expected, the non-preferred group took more practice trials to achieve the same level of performance as the preferred group. Moreover, both groups were able to train to nearly the same level of performance by the last set of trials in the training phase given the amount of practice that was provided. Thus, results indicate that, for the simple task used in the experiment, the participants that were forced to use their non-preferred mapping initially performed significantly worse than participants using their preferred mapping, but were able to achieve equivalent level of performance after only a brief period of practice.

Results from the switch and return phases indicate that there indeed are differences between the two groups despite the equivalent levels of performance observed in the baseline phase. The key finding is that, while both preferred and non-preferred groups made significantly more errors in the baseline phase than in the switch phase (when forced to flip their I-R mapping), the group trained using their preferred I-R mapping performed significantly worse than the group trained using their non-preferred I-R mapping. Compared to performance in the baseline phase, the non-preferred group made about 3 times as many errors and the preferred group about 6 times.

Performance of both groups was not significantly different once the I-R mapping was changed back to the original state (i.e., the mapping used during training) in the return phase. Performance in the return phase was worse for both groups compared to the baseline phase.

Interestingly, performance of the non-preferred I-R mapping group during the switch phase was not equivalent to performance of the preferred group during the baseline phase, which suggests that the experience of using their non-preferred I-R mapping affected their performance in at least the short term. These results suggest that recovery from a brief and unexpected exposure to another I-R mapping while conducting a task is rapid under the conditions of this experiment, although should not be expected to equal pre-interruption performance.

The pattern of results from the training, switch, and return phases suggest the following characteristics regarding accidental inversions errors for the task used in experiment 3:

1. Initial performance is better when an operator is allowed to utilize the I-R mapping that they prefer rather than when forced to use their non-preferred I-R mapping.
2. Performance using the non-preferred mapping improves consistently over a relatively short period of constant practice.
3. Performance markedly decreases when operators are suddenly and unexpectedly forced to use the I-R mapping opposite to that used in training.
4. Performance decreases when suddenly and unexpectedly forced to switch I-R mappings. Switch from the non-preferred mapping to the preferred mapping and vice versa were both negatively affected. However, performance was significantly worse for operators going from using their preferred I-R mapping to their non-preferred than operators going from their non-preferred mapping to their preferred.

5. Performance was not significantly affected after a brief period using the I-R mapping opposite to that used during training (after returning again to the I-R mapping used during training). Moreover, differences in performance vanish between operators who used their preferred I-R method and operators who did not.

## CONCLUSIONS AND RECOMMENDATIONS

### Conclusions

The research in this dissertation sought answers to three sets of research questions. The first set of research questions was: a) is there a strong population stereotype for 3D rotational control using a 2-DOF interface? b) what factors mediate whether an individual's stereotype response? Experiment 1 provided evidence for a strong stereotype response for horizontal rotations and a weak stereotype response for vertical rotations. The split between the two logical response options to a horizontal rotation was 90% / 10%, whereas the split for vertical rotations was 60% / 40%. None of the factors assessed in this study affected the proportion of response types for horizontal rotations. The type of response made to vertical rotations, however, was affected by whether the context of the stimuli was "flying" or "walking". Results showed that only under the flying condition was the proportion of response types affected. Thus, a key conclusion from the first experiment is that vertical rotations in the context of flight are associated with a stronger stereotype response compared to vertical rotations in the context of walking.

The second set of research questions was: a) do operators subjectively prefer one input-response method over another for rotational control? b) what factors mediate



an operator's preference? Results from Experiment 2 indicated that subjective preference was fairly neutral on average although a slight preference for one input-response mapping was found for the egocentric perspective compared to the allocentric perspective. Participants rated their performance as better (i.e., fewer accidental inversion errors) in the egocentric versus the allocentric perspective. Interestingly, while context appeared to have a significant effect on stereotype response, as evidenced in Experiment 1, it did not have a significant effect on subjective preference as measured in Experiment 2.

The third set of research questions was: a) does subjective preference affect the rate at which accidental inversion errors occur? b) does brief exposure to a different input-response control mapping affect the rate at which accidental inversion errors occur? C) how recoverable is performance after a brief exposure to another I-R control mapping? Results from Experiment 3 indicated that one's subjective preference for an I-R mapping affected initial performance (i.e., fewer accidental inversion errors) such that more errors occurred when forced to use one's non-preferred I-R mapping. For the task used in the experiment, the group using their non-preferred I-R mapping achieved equivalent performance by the end of the training phase and during the baseline base. Both groups made more errors during the switch phase (which required the group trained using their non-preferred I-R mapping to use their preferred mapping and the group trained using their preferred mapping to use their non-preferred mapping). However, the group trained using their preferred mapping made more errors than the group trained using their non-preferred mapping. In the return phase both groups once again used the I-R mapping they were trained with and, despite that performance

between the two groups improved markedly and were comparable, both made significantly more errors than made in the baseline phase.

Conclusions that may be drawn from Experiment 3 are that being able to use the I-R method that one subjectively prefers may affect performance for some tasks. Moreover, despite that performance using a non-preferred I-R method may be, after practice, equivalent to the level of performance using one's preferred I-R mapping, differences may still exist in terms of resilience to interference or confusion caused by a temporary exposure to another I-R mapping. Lastly, while a temporary exposure to another I-R mapping may have a substantial impact, performance appears to rapidly return to previous levels, before the interruption.

#### Theoretical Implications

Theoretical accounts may be made for the two primary findings in this research: a) that vertical rotational control stereotypes are, at least in the context examined, weaker than stereotypes for horizontal control, and b) that usage of one's subjectively preferred input-response relationship for a vertical rotation task can affect performance, especially when the input-response mapping is suddenly reversed.

The findings on population stereotypes for vertical versus horizontal rotation are similar to results by Hoffman (1997) that indicate a relationship between stereotype strength and performance for horizontal but not vertical rotational control. Hoffman concluded that principles used to drive response stereotypes were of greater strength for horizontal rotations than for vertical and that input-response relationships for the two types of arrangements were associated with different sets of expectancies. Hoffman suspected that this result may have been due in part to familiarity given the

commonality of devices with horizontally arranged displays and controls. Another component of this familiarity may be the fact that the human visual field is larger horizontally than vertically due to the horizontal arrangement of the eyes.

Theoretical implications may also be drawn from the findings on the interaction between subjective preference for input-response mapping and performance. Classic studies on stimulus response theory (e.g., Fitts & Deininger, 1954, and Fitts & Seeger, 1953) predict that violation of a strong stereotype response may cause a more severe degradation of performance than violation of a weak stereotype response. Findings from experiment 3 are consistent with this pattern but also emphasizes that, in cases where the population stereotype is weak, the stereotype responses biases of the individual should be considered.

### Recommendations

The following are recommendations for the design of human-machine systems meant to control 3D rotation using a 2-DOF input device:

- Operators should be able to select what I-R mapping to use to affect both horizontal and vertical rotations.
- Operators should be allowed to experience all I-R mapping options before using the system for the actual task. Enough time should be allowed in order to form a subjective preference. Operators should be able to switch I-R mapping during these test trials.
- After selecting a particular I-R mapping, operators should be able to practice using it for several minutes, preferably until a level of performance is

achieved. Feedback on performance is recommended. This is especially important after having selected or changed the I-R mapping.

- Operators should be able to change which I-R mapping is used at any time while using the system.
- At any time, an operator should be able to test out an I-R mapping in a such a way that feedback is only provided on the behavior of the I-R mapping and does not affect the actual system. This is especially important after having selected or changed the I-R mapping.
- The capability to change the I-R mapping should be easily accessible both in terms of time and steps required.
- The capability to change the I-R mapping should be guarded against accidental activation.
- I-R mappings should be labeled in such a way that each mapping option is easily distinguished from the others. Terms such as “Normal” and “Inverted” should be avoided.
- The currently selected I-R mapping setting should be always displayed at times before the operator is able to affect the system.
- An operator should be provided with a way to easily remember which I-R mapping they have previously indicated that they prefer.

For horizontal rotation (i.e., along the Y axis):

- A reasonable default setting is such that to rotate an object to the left, a left directional input is used and to rotate an object to the right, a right directional

input is used. It may be assumed that 9 out of 10 users will expect that the system will behave this way.

For vertical rotation (i.e., along the X axis):

- Control of vertical rotation via 2-DOF device does not have a strong stereotype response. Instead of a default mapping, the system should make the operator aware that he or she must select a mapping and is free to try out the different options.
- Operators should be made aware of when or under what conditions they can change the I-R mapping for vertical rotation.

#### Practical Implications

The recommendations from this research have implications for any human-machine system where rotation occurs. Three examples of domains where this research is relevant are human-in-the-loop video-based surveillance systems, entertainment systems, and camera-based or virtual surgery systems.

Many surveillance systems allow a human operator to take remote control of a camera that can be rotated. Examples include military platforms such as the U.S. Navy P3 Orion and the P8-A Poseidon aircraft, both of which support surveillance missions. Multiple surveillance operators are working together on P3 or P8-A aircraft. As such, these operators may have different subjective preferences for vertical rotation control of imaging devices. As with any military operation, it is critical to prosecute the mission as effectively as possible and minimize human error in regard to capturing accurate surveillance data.

This research may also benefit surgical systems where the surgeon and medical personnel observe procedures via a display of a laparoscopic camera or virtual simulation. One particular application that is relevant to this research is robot-assisted laparoscopic surgery, which involves a human controlling an input device which physically affects the patient. Accidental inversion errors in this context could lead to moving a surgical apparatus in the wrong direction. For example, the surgeon may rotate a laser to the opposite side of the target area that was intended. Of particular concern is the practice of some surgeons of training dexterity by playing video games (e.g., Rosser, et al., 2007; Morris, 2004; see also Reiting, Schmalstieg, Bornik, & Beichel, 2006; Arnold & Farrell, 2002). The concern is that the video games used for training may utilize an input-response mapping opposite to that of the surgical system or trainer.

#### Recommendations for Future Research

One valuable opportunity for future research is to observe accidental inversion as it occurs on by a trained human-operator on skilled task in a longitudinal study. One of the challenges of measuring accidental inversion outside of an artificial task in a controlled experiment is recognition. Experiments 2 and 3 used tasks which clearly had a correct and incorrect response. Thus, reviewing recordings of camera control during laparoscopic surgery, for example, does not necessarily indicate when a camera rotation was the intention of the surgeon or a correction to an accidental rotation in the wrong direction. Nonetheless, real world data on accidental inversion would be useful in understanding the incidence of the phenomena on a per task basis.

Along the same vein as the above research, another opportunity is to longitudinally examine the effects of random brief exposure to input-response mappings opposite to that currently being. One aspect of such research could examine the number of inputs required to return to previous levels of performance before the interruption occurred.

Thirdly, future research could focus on whether individual difference variables interact with the occurrence of accidental inversion. One benefit of research on individual differences is that screening can be used to identify users that are suitable for certain tasks or control arrangements. Additionally, users could potentially be identified to use either default input-response mapping or to go through a sequence where one is able to safely experience multiple mappings before making a selection.

## APPENDIX A: EXPERIMENT MATERIALS



## Experiment 1

### *Informed Consent Form*

#### Informed Consent Form

This research study examines reaction time to moving pictures. As part of the study you will be asked to use a control device to respond to images that appear on the computer screen.

There are no anticipated risks to you for taking part in this study, except those associated with normal computer use. You are free to withdraw your consent to participate and may discontinue your participation in the interview at any time without consequence.

If you have any questions about this research project, please contact me at \_\_\_\_\_. My faculty supervisor is Dr. Valerie Sims. Questions or concerns about research participants' rights may be directed to the UCFIRB office, University of Central Florida Office of Research, Orlando Tech Center, 12443 Research Parkway, Suite 207, Orlando, FL 32826. The phone number is (407) 823-2901.

Sincerely,  
Derek Diaz

\_\_\_\_\_ I have read the procedure described above for this study.

\_\_\_\_\_ I voluntarily agree to participate in this study.

\_\_\_\_\_  
Participant

### *Instructions*

#### Spoken:

“In this study you will be using the \_\_\_\_\_ (mouse/joystick) to respond to movement that will appear on the monitor. This part will last about 18 minutes followed by a few questionnaires.”

#### Presented on screen before the start of the study:

“You will be using the mouse / joystick to respond in this study.

When the experiment starts you will appear to be moving through a corridor.

You will constantly appear to be moving forward.

However, at random times, the screen will turn.

As soon as this happens, move the mouse / joystick

AS IF \*YOU CAUSED\* the turn to happen.

Respond as quickly as possible when you see a turn.

--- Please wait for the signal to begin ---”



**Set - B**

How difficult was the last tunnel?

-----	-----	-----	-----	-----	-----	
1	2	3	4	5	6	7
Very			Neutral			Very
Easy						Difficult

How would you rate your ability to control where you were looking in the last tunnel?

-----	-----	-----	-----	-----	-----	
1	2	3	4	5	6	7
Very			Neutral			Very
Easy						Difficult

How frequently did you accidentally turn the wrong way in the last tunnel?

-----	-----	-----	-----	-----	-----	
1	2	3	4	5	6	7
Very			Neutral			Very
Easy						Difficult

Which method of control do you prefer?

-----	-----	-----	-----	-----	-----	
1	2	3	4	5	6	7
This			Neutral			The previous
Method						Method

If you had to choose a method of control to use for now on, which would you prefer?

## APPENDIX B: CORRELATION MATRIX FROM EXPERIMENT 2

Table 8: Correlation Matrices for Experiment 2, Questionnaire Items 1 - 3

		Q1 I1F	Q2 I1F	Q3 I1F	Q1 I1W	Q2 I1W	Q3 I1W	Q1 N1F	Q2 N1F	Q3 N1F
Q1 I1F	Corr. Sig.									
Q2 I1F	Corr. Sig.	.852(**) .000								
Q3 I1F	Corr. Sig.	.580(**) .000	.586(**) .000							
Q1 I1W	Corr. Sig.	.442(**) .001	.417(**) .003	.127 .381						
Q2 I1W	Corr. Sig.	.355(*) .011	.409(**) .003	.123 .395	.892(**) .000					
Q3 I1W	Corr. Sig.	.327(*) .020	.269 .059	.437(**) .002	.584(**) .000	.511(**) .000				
Q1 N1F	Corr. Sig.	.181 .208	.104 .474	-.027 .854	.161 .264	.145 .316	.101 .486			
Q2 N1F	Corr. Sig.	.160 .268	.130 .370	-.071 .625	.222 .122	.185 .197	.167 .245	.926(**) .000		
Q3 N1F	Corr. Sig.	-.056 .700	-.136 .347	.003 .984	-.075 .604	-.099 .496	.149 .301	.826(**) .000	.779(**) .000	

\*\* Pearson Correlation is significant at the 0.01 level (2-tailed).

\* Pearson Correlation is significant at the 0.05 level (2-tailed).

		Q1 N1W	Q2 N1W	Q3 N1W	Q1 I3F	Q2 I3F	Q3 I3F	Q1 I3W	Q2 I3W	Q3 I3W
Q1 N1W	Corr. Sig.									
Q2 N1W	Corr. Sig.	.804(**) .000								
Q3 N1W	Corr. Sig.	.660(**) .000	.799(**) .000							
Q1 I3F	Corr. Sig.	.209 .146	.065 .656	.034 .816						
Q2 I3F	Corr. Sig.	.138 .339	-.011 .940	-.011 .939	.885(**) .000					
Q3 I3F	Corr. Sig.	.189 .189	.176 .221	.296(*) .037	.666(**) .000	.647(**) .000				
Q1 I3W	Corr. Sig.	.234 .102	.264 .064	.290(*) .041	.471(**) .001	.398(**) .004	.492(**) .000			
Q2 I3W	Corr. Sig.	.241 .092	.200 .163	.208 .147	.354(*) .012	.338(*) .016	.420(**) .002	.869(**) .000		
Q3 I3W	Corr. Sig.	.070 .627	.190 .187	.281(*) .048	.190 .186	.182 .207	.503(**) .000	.702(**) .000	.719(**) .000	

\*\* Pearson Correlation is significant at the 0.01 level (2-tailed).

\* Pearson Correlation is significant at the 0.05 level (2-tailed).

		Q1 N3F	Q2 N3F	Q3 N3F	Q1 N3W	Q2 N3W	Q3 N3W
Q1 N3F	Corr. Sig.						
Q2 N3F	Corr. Sig.	.917(**) .000					
Q3 N3F	Corr. Sig.	.807(**) .000	.796(**) .000				
Q1 N3W	Corr. Sig.	.442(**) .001	.354(*) .012	.154 .286			
Q2 N3W	Corr. Sig.	.487(**) .000	.419(**) .002	.308(*) .030	.815(**) .000		
Q3 N3W	Corr. Sig.	.520(**) .000	.512(**) .000	.446(**) .001	.643(**) .000	.678(**) .000	

\*\* Pearson Correlation is significant at the 0.01 level (2-tailed).

\* Pearson Correlation is significant at the 0.05 level (2-tailed).

## APPENDIX C: IRB HUMAN SUBJECTS PERMISSION LETTERS





Office of Research & Commercialization

Valerie Sims, Ph.D.  
Department of Psychology  
College of Arts and Sciences  
University of Central Florida  
4000 Central Florida Boulevard  
Orlando, Florida 32816

Dear Dr. Sims:

With reference to your protocol entitled, "Axis Inversion Study 1," I am enclosing for your records the approved, executed document of the UCFIRB Form you had submitted to our office.

Please be advised that this approval is given for one year. Should there be any addendums or administrative changes to the already approved protocol, they must also be submitted to the Board. Changes should not be initiated until written IRB approval is received. Adverse events should be reported to the IRB as they occur. Further, should there be a need to extend this protocol, a renewal form must be submitted for approval at least one month prior to the anniversary date of the most recent approval and is the responsibility of the investigator (UCF).

Should you have any questions, please do not hesitate to call me at 823-2901.

Please accept our best wishes for the success of your endeavors.

Cordially,

A handwritten signature in black ink that reads 'Chris Grayson'.

Chris Grayson  
Institutional Review Board (IRB)

Copies: Mr. Derek Diaz  
Dr. Richard Tucker  
IRB File



Office of Research & Commercialization

Valerie Sims, Ph.D.  
Department of Psychology  
College of Arts and Sciences  
University of Central Florida  
4000 Central Florida Boulevard  
Orlando, Florida 32816

Dear Dr. Sims:

With reference to your protocol entitled, "Axis Inversion Study 2," I am enclosing for your records the approved, executed document of the UCFIRB Form you had submitted to our office.

Please be advised that this approval is given for one year. Should there be any addendums or administrative changes to the already approved protocol, they must also be submitted to the Board. Changes should not be initiated until written IRB approval is received. Adverse events should be reported to the IRB as they occur. Further, should there be a need to extend this protocol, a renewal form must be submitted for approval at least one month prior to the anniversary date of the most recent approval and is the responsibility of the investigator (UCF).

Should you have any questions, please do not hesitate to call me at 823-2901.

Please accept our best wishes for the success of your endeavors.

Cordially,

A handwritten signature in cursive script that reads "Chris Grayson".

Chris Grayson  
Institutional Review Board (IRB)

Copies: Mr. Derek Diaz  
Dr. Richard Tucker  
IRB File



Office of Research & Commercialization

Valerie Sims, Ph.D.  
Department of Psychology  
College of Arts and Sciences  
University of Central Florida  
4000 Central Florida Boulevard  
Orlando, Florida 32816

Dear Dr. Sims:

With reference to your protocol entitled, "Axis Inversion Study 3," I am enclosing for your records the approved, executed document of the UCFIRB Form you had submitted to our office.

Please be advised that this approval is given for one year. Should there be any addendums or administrative changes to the already approved protocol, they must also be submitted to the Board. Changes should not be initiated until written IRB approval is received. Adverse events should be reported to the IRB as they occur. Further, should there be a need to extend this protocol, a renewal form must be submitted for approval at least one month prior to the anniversary date of the most recent approval and is the responsibility of the investigator (UCF).

Should you have any questions, please do not hesitate to call me at 823-2901.

Please accept our best wishes for the success of your endeavors.

Cordially,



Chris Grayson  
Institutional Review Board (IRB)

Copies: Mr. Derek Diaz  
Dr. Richard Tucker  
IRB File

## REFERENCES

- Alluisi, E.A., (1961). Toward optimizing man's tactile communication. *Perceptual and Motor Skills*, 12, 235-245.
- Alluisi, E.A., Warm, J.S. (1990). Things that go together: A review of stimulus-response compatibility and related effects. In *Stimulus-Response Compatibility*. Proctor, R.W., & Reeve T.G. (Eds). Elsevier Science Publishers B.V. (North-Holland).
- Arnold P., & Farrell, M.J. (2002). Can virtual reality be used to measure and train surgical skills? *Ergonomics*, 45, 362 – 379.
- Bertelson, P. (1963). S-R relationships and reaction times to new versus repeated signals in a serial task. *Journal of Experimental Psychology*, 65, 478-484.
- Bowman, D., Kruijff, E., Laviola, J., & Poupyrev, I. (2004). *3D User Interfaces: Theory and Practice*. Addison-Wesley, Boston.
- Brainard, R.W., Irby, T.S., Fitts, P.M., & Allusisi, E.A., (1962). Some variables influencing the rate of gain of information. *Journal of Experimental Psychology*, 63, 105-110.
- Brebner, J. & Sandow, B. (1976) The effect of scale side on population stereotype. *Ergonomics*, 19(5), 571-580.
- Chan W.H., & Alan H.S. (2003). Movement compatibility for rotary control and circular display—Computer Simulated Test and real Hardware Test. *Applied Ergonomics*, 34, 61–71.
- Chan W.H., Chan. A.H.S. (2007). Strength and reversibility of movement stereotypes for lever control and circular display. *International Journal of Industrial Ergonomics*, 37, 233–244.

Cho, Y.S., & Proctor, R.W. (2003). Stimulus and response representations underlying orthogonal stimulus-response compatibility effects. *Psychonomic Bulletin & Review*, 10, 45-73.

Chua R., Weeks D.J., Ricker L. K,& Poon P. (2001). Influence of operator orientation on relative organizational mapping and spatial compatibility. *Ergonomics*, 44, 751–765.

Davis, R., Morray, N., & Treisman, A. (1961). Imitating responses and the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 13, 78-89.

Dutta, A., & Proctor, R. W. (1992). Persistence of stimulus-response compatibility effects with extended practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 801-809.

Eimer, M. (1995) Stimulus-response compatibility and automatic response activation: evidence from psychophysiological studies. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 837-854.

Ellis R., & Tucker M. (2000). Micro-affordance: The potentiation of components of action by seen objects. *British Journal of Psychology*, 91, 451-471.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement, *Experimental Psychology*, 47(6), 381-391.

Fitts, P. M. (1959). *Human information handling in speeded tasks* (Report No. RC-109). Yorktown Heights, NY: IBM Research Center.

Fitts, P.M., & Deininger, R.L. (1954). S-R compatibility: Correspondence among paired elements within stimulus and response compatible codes. *Journal of Experimental Psychology*, 48, 483-492.

Fitts, P.M. and Seeger, C.M. (1953). S-R compatibility: spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, 46, 199-210.

Garvey W.D., & Knowles W.B. (1954). Response time patterns associated with various display-control relationships. *Journal of Experimental Psychology* 47(5),315-322.

Hawkins H.L., & Underhill J.R. (1971). S-R compatibility and the relative frequency effect in choice reaction time. *Journal of Experimental Psychology* 91(2), 280-286.

Hinckley K., Paush R., Proffitt D. (1997). Attention and visual feedback: the bimanual frame of reference, *Proceedings of the 1997 symposium on Interactive 3D graphics*, (pp.121-126). New York: ACM Press.

Hommel, B., & Prinz, W. (Eds.). (1997). *Theoretical issues in stimulus-response compatibility*. Amsterdam: North-Holland.

Hotta, A., Takahashi, T., Takashi, K., & Kogi, K. (1981). Relations between direction-of-motion stereotypes for controls in living space. *Journal of Human Ergology*, 10, 73-82.

Jeannerod, M. (1981). Intersegmental coordination during reaching at natural objects. In: Long J, Baddeley AD (eds) *Attention and performance IX*. Erlbaum, Hillsdale, 153–169.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. *Lecture Notes in Computer Science*, 1404, 1-18.

Kornblum, S., Hasbroucq, T. & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility – A model and taxonomy, *Psychological Review*, 97, 253-270.

Kornblum, S. & Lee, J-H. (1995). Stimulus-Response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response, *Journal of Experimental Psychology: Human Perception and Performance*, 21, 855-875.

Loveless, N.E. (1962). Direction-of-motion stereotypes: a review, *Ergonomics*, 5, 357-383.

MacKenzie, C. L., & Iberall, T. (1994). *The grasping hand*. Amsterdam: Elsevier Science.

Megaw, E. D. (1972). Directional errors and their correction in a discrete tracking task, *Ergonomics*, 15, 633-643.

Morrin, R.E., & Grant, D.A. (1955). Learning and performance of a keypressing task as a function of the degree of spatial stimulus-response correspondence. *Journal of Experimental Psychology*, 49, 39-47.

Morris, B.R. (2004, February 26). Pre-Op Warm-Up: A Few Video Games. *The New York Times*. Retrieved November 21, 2007, from <http://query.nytimes.com/gst/fullpage.html?res=9C05E1DA123CF935A15751C0A9629C8B63&sec=&spon=&pagewanted=1>.

Norman, D.A. (1988). *The Psychology of Everyday Things*. New York: Bantam Books.

Poupyrev I., & Ichikawa T. (1999). Manipulating objects in virtual worlds: categorization and empirical evaluation of interaction techniques. *Journal of Visual Languages and Computing*, 10(1), 19-35.

Proctor, R. W., & Reeve, T.G. (1990). Research on Stimulus-Response compatibility: toward a comprehensive account. In Proctor, Robert W. & Reeve, T. Gilmour (Eds.), *Stimulus-Response Compatibility*, Amsterdam, NL: Elsevier Science, 483-494.

Proctor, R.W., & Vu, K.-P.L. (2005). Roles of task-defined associations and reference frames in spatial stimulus-response compatibility. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 141-153). Washington, DC: American Psychological Association.

Proctor, R.W., Wang, D.-Y.D., & Pick, D.F. (2004). Stimulus-response compatibility with wheel-rotation responses: Will an incompatible response coding be used when a compatible coding is possible? *Psychonomic Bulletin & Review*, 11, 841-847.

Reeve, T.G., & Proctor, R.W. (1984). On the advance preparation of discrete finger responses. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 541-553.

Reitinger, B., Schmalstieg, D., Bornik, A., Beichel, R. (2006). Spatial Analysis Tools for Virtual Reality-based Surgical Planning. *IEEE Symposium on 3D User Interfaces*, 37-44.



Rosser, J.C., Lynch, P.J.; Cuddihy, L., Gentile, D.A., Klonsky, J., Merrell, R. (2007). The Impact of Video Games on Training Surgeons in the 21st Century. *Archives of Surgery*, 142, 181-186.

Sanders, A.F. (1980). Stage analysis of reaction processes. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 331-354). Amsterdam: North-Holland.

Scali, S., Wright, M., & Shillito, A. (2003). 3D Modeling Is Not for WIMPs. *Proceedings of HCI International*. Lawrence Erlbaum Associates, Vol.1.

Vu, K.P. L, & Proctor, R. W. (2001). Stimulus-response compatibility in interface design. *Proceedings of the Ninth Conference on Human-Computer Interaction*, pp 1368-1372.

Vu, K.P. L, & Proctor, R. W. (2003). Naïve and experienced judgments of stimulus-response compatibility: Implications for interface design. *Ergonomics*, 46, 169-187.

Vu, K.P.L, & Proctor, R.W. (2004). Mixing Compatible and incompatible mappings: Elimination, reduction, and enhancement of spatial compatibility effects. *Quarterly Journal of Experimental Psychology*, 57A, 539-556.

Wallace, R.J. (1971). S-R compatibility and the idea of a response code. *Journal of Experimental Psychology*, 88, 354-360.

Wickens, C.D. (1987). Information processing, decision making, and cognition. In G. Salvendy (Ed.), *Handbook of human factors* (pp 72-107). New York: John Wiley & Sons.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance* (2nd ed.). New York: HarperCollins Publishers Inc.

Woodson, W.E., & Conover, D.W. (1970). *Human engineering guide for equipment designers* (2nd ed.). Berkeley, CA: University of California Press.