2020

Discriminating Targets among Distractors in a Virtual Shopping Environment with Different Rack Orientations: Testing a Model of Visibility

Tyler Sinclair Whitlock
Wright State University

DISCRIMINATING TARGETS AMONG DISTRACTORS IN A VIRTUAL SHOPPING ENVIRONMENT WITH DIFFERENT RACK ORIENTATIONS: TESTING A MODEL OF VISIBILITY

A Thesis submitted in partial fulfillment of the
requirements for the degree of
Master of Science

by

TYLER SINCLAIR WHITLOCK

B.S., Ohio State University, 2016

2020

Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

JULY 28, 2020

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Tyler Sinclair Whitlock ENTITLED Discriminating Targets among Distractors in a Virtual Shopping Environment with Different Rack Orientations: Testing a Model of Visibility BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

Scott Watamaniuk, Ph.D.

Thesis Director

Debra Steele-Johnson, Ph.D.

Chair, Department of Psychology

Committee on

Final Examination

John Flach, Ph.D.

Robert Gilkey, Ph.D.

Pratik Parikh, Ph.D.

Barry Milligan, Ph.D.

Interim Dean, Graduate School

Whitlock, Tyler Sinclair. M.S., Department of Psychology, Wright State University, 2020. Discriminating Targets among Distractors in a Virtual Shopping Environment with Different Rack Orientations: Testing a Model of Visibility.

Abstract

Objective: This study measured observers' abilities to identify letter targets distributed among number distractors in a virtual shopping environment. Head-turning behavior was also continuously recorded throughout each trial. The data were then used to test whether a model's prediction for the duration of visibility needed for target detection in a virtual shopping environment (Parikh & Mowrey, 2014) generalize to the more realistic shopping task of identifying a target on a shelf. Currently, the model predicts the visibility of the locations of targets in traditional racks oriented 90° to the aisle (perpendicular) as well as racks oriented at 30°, 45, 135°, and 150° to the central aisle. **Background:** Exposure (whether a portion of the rack is seen) and intensity (how long that rack portion is seen) are the two variables of interest in the model. According to the analytical and computational models developed by Parikh and Mowrey (2014), traditional 90° racks in retail shopping environments result in lower exposure and intensity than racks at other angles. A previous study confirmed these model predictions with a simple target detection task (small red targets on empty grey racks) in a virtual environment. However, discriminating a target on a stocked shelf requires more time and is more representative of typical shopping behavior. **Methods:** The 24 participants completed 10 target discrimination trials as they were moved through a virtual shopping environment. **Hypothesis:** We hypothesized and found a significant effect of orientation on discrimination performance. Additionally, we hypothesized that the percentage of total targets correctly identified would be lower than the simple detection rate in Parikh and Mowrey (2014) but found mixed results. Model fit was first assessed via a d' metric. The d' values were generally low, but they were best at intensities higher than that needed for detection due to the additional time needed to identify the targets among distractors. However, the observed non-normal distributions of hits and false alarms make the d' analysis difficult to interpret. Subsequently, a chi-square analysis was done. The chi-square analysis also showed evidence for higher intensities needed for discrimination than for detection in the 30°, 45°, and 90° rack orientations. Limitations and modifications needed for the model to achieve a better match to human discrimination performance are discussed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

The present study evaluates the performance of the model of location visibility proposed by Parikh and Mowrey (2014) using a discrimination task in a simple virtual shopping environment. The model was originally created based on theoretical human parameters to examine the impact of the layout of a shopping environment on rack visibility. Specifically, how product rack angle (30°, 45°, 90°, 135°, & 150°) and head rotation influence the visibility of specific product rack locations, and consequently target detection. The model from Parikh and Mowrey (2014) was tested using an easy task of detecting a small red target on a grey background. A key parameter in the model is how much time a person needs to have a target in their field of view in order to detect it, referred to as "exposure time" or "intensity". The present study further tests the model by investigating whether it can generalize to behavior that is a closer approximation to more naturalistic shopping behavior. Specifically, we test whether the exposure times determined from the detection task will be adequate to accomplish a more difficult discrimination task.

The impact and arrangement of space is important in many different areas such as medical (e.g., the ICU) and retail shopping environments. Space is an important factor in retail environments due to the competing considerations of costs and customer comfort. Retailers are particularly interested in space because of the importance of product placement demands of suppliers to allow for higher visibility of promotional products. Space is also important in medical environments. Medical environments often require staff to monitor and physically check on several patients which can be a time-consuming process if the staff must visit each patient's individual room. However, the environment could be designed to maximize visibility of patients from a single vantage point allowing staff to quickly check on multiple patients with minimal movement. In this way productivity could be improved allowing more time to attend to other crucial patient needs.

Background

Several scientists have considered the role of exposure in retail environments to be a major factor in product purchases. Cairns (1962) believed sales were directly related to the number of people who were able to see, or be exposed to, the target. Similarly, Drèze, Hoch, & Purk, (1994) posited that purchase likelihood was influenced by product visibility. Logically, this makes sense - a product not seen cannot be purchased. Although exposure is unlikely to be the only factor involved in purchasing a product, it should be thought of as a necessary component. Another critical component is how long that product is exposed to someone, referred to as intensity.

The layout or usage of retail space is the result of several different factors and goals. One of the primary factors in layout is the efficient utilization of space. There is often a trade-off between how space is utilized and visibility. For example, a layout that maximizes utilization of the space may prevent some of the items or racks in that space from being easily seen (exposed) due to obstruction. However, for most retail spaces, the goal is to strike a balance between exposure and efficient space utilization. Traditional product racks oriented 90° relative to the walk aisle have remained a staple because they are quite efficient in terms of utilization of space. However, these traditional 90° racks may not provide the best balance in terms of exposure and/or intensity. Parikh and Mowrey (2014) define exposure as whether a portion of the rack can be seen. If a portion of the rack can be seen at any time, it is treated as exposed. Exposure can also be thought of ‘holistically’ – that is what proportion of the total rack area could be seen after navigating through the environment. Intensity is defined as the duration that a portion of the rack can be seen while traversing the environment (Parikh & Mowrey, 2014). These experiments by Parikh and Mowrey to empirically determine the advantages and disadvantages of different rack orientations on target detection can provide new and potentially better ways of utilizing space.

Models

Parikh and Mowrey (2014) developed analytical and algorithmic models to provide evidence of the benefits of different rack orientations. Their analytical model calculates rack exposure while the algorithmic model simulates exposure and intensity of a person going through a central aisle with varying rack orientations surrounding them (see Figure 1). Their models assume a positive correlation between exposure/intensity and target detection; thus, an increase in exposure and intensity would result in an increased probability of target detection. Their models provide an estimate of what can potentially be seen, such as targets on a rack, but human performance may not reflect the same behavior. Human performance may differ from the theoretical performance due to different variables such as lighting, sound, and smell that affect where and when attention is directed. For example, the model could predict that a target at a particular location is seen, but if the human is not paying attention to that portion of the rack even though it is in their field of view then the target may not be seen. A familiar scent from a specialty food, a sound from another aisle or person, or even different lighting meant to highlight a new product, could redirect attention away from a target meaning it could be missed.

Parikh and Mowrey's models provide a picture of what could potentially be seen but because humans are impacted by many variables not in the models, such as attention and distractions, their performance may be quite different.

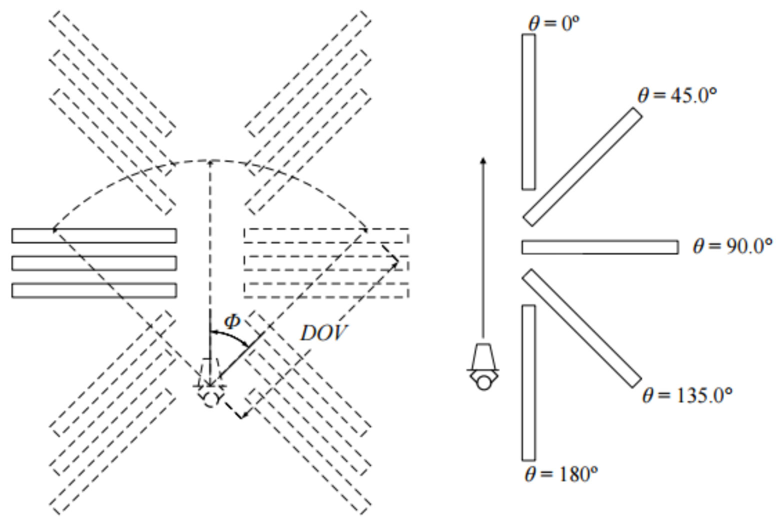

Figure 1. Representation of a participant in a virtual environment with various different rack orientations. Phi (ϕ) = head turn angle, Theta (θ) = rack orientation, and DOV = Depth of focused Vision. Adapted from Parikh & Mowrey (2014), Figures 1 and 2.

Parikh and Mowrey created a slicing algorithm (see Figure 2) to determine when a target would be exposed on a rack and how intense that exposure would be. The slicing algorithm creates an individual's functional field of regard based on depth of view (assumed to be 25 ft) and head turn angle (Φ), in their example assumed to be 45 degrees, but can be any amount. A Φ of 45 degrees was assumed because for most people it is a more comfortable head turn than larger turn angles (e.g., 90°) which would be more physically demanding and strenuous. Their model also assumes a maximum eye rotation of 15° which, when combined with the head turn, results in a 60° field of view on each side of straight ahead for a total 120° field of view. Similarly, the algorithm assumes a maximum 25-foot depth of focused vision. This value was arrived at by considering that a 0.5” x 0.5” target, viewed from a distance of 25 feet, would subtend approximately the same size (in visual angle) as the smallest letter identifiable by someone with 20/20 vision (0.095° or 5.7 min arc).

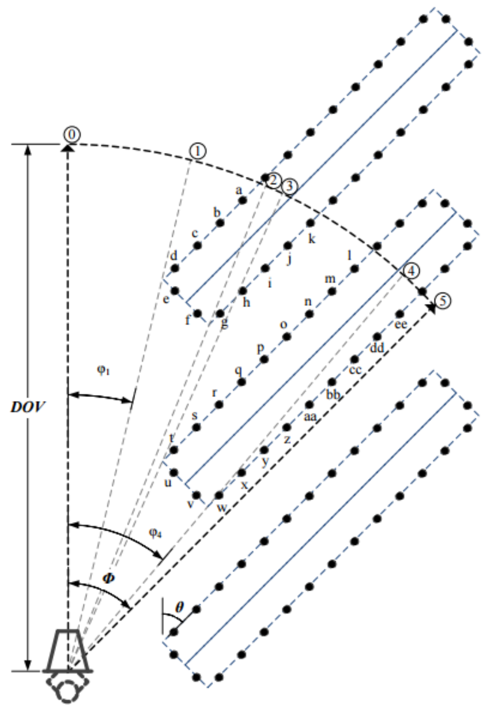

Figure 2. Visual representation of the slicing algorithm used by Parikh and Mowrey to determine intensity and exposure for portions of each rack. Phi (ϕ) = head turn angle, Theta (θ) = rack orientation, and DOV = Depth of focused Vision. Numbers 0 through 5 represent different portions of the rack and the viewing path to that portion. Adapted from Parikh & Mowrey (2014), Figure 6.

However, it is not enough to simply know how far a person can see clearly because the way the person views the world is also important. In this regard, Parikh and Mowrey’s (2014) algorithm assumes a horizontal scan pattern based on the results of

Ebster and Garaus (2015) who showed that individuals in shopping situations typically scan from side to side (horizontally) instead of top to bottom (vertically). Utilizing these variables, 0.5 inch or smaller targets that are either occluded by other adjacent racks or beyond the 25 ft depth of view are considered to be not seen.

Parikh and Mowrey (2014) then utilized the slicing algorithm and computed exposure and intensity values for simulations of a person walking down a central aisle surrounded by aisles on the left and right with five different aisle orientations (30°, 45°, 90°, 135°, & 150°). Their research suggests that acute (e.g., 30°) and obtuse angles (e.g., 150°) both have advantages when compared to a traditional 90° rack in terms of exposure or intensity. In a follow-up paper, Mowrey (2016) calculated that there would be a 150-250% increase in holistic exposure for racks oriented at 30° compared to traditional 90° racks depending on whether the flow of foot traffic was unidirectional or bidirectional. This result is depicted in Figure 3, taken from Mowrey (2016). Notice that as rack orientation varies the change in exposure (*E* - the total proportion of a rack that falls within the model's field of view) relative to the 90° rack orientation also varies. The abrupt changes in relative exposure at 45° and 135° are due to several factors including the obstruction of rack locations from preceding racks and the alignment of rack faces with an observer's head angle as an observer moves passed racks. These factors change differentially with rack orientation and their combined effects get compounded when the calculations are made as an observer moves down an aisle and passes more racks. A more complete explanation appears in Mowrey (2016). However, similar to Parikh and

Mowrey’s (2014) study, Mowrey (2016) posited that rack orientations with higher exposure do not result in higher intensity values. Layouts with high exposure can contain portions of the rack that are seen only for a brief period of time - leading to higher exposure but lower intensity values. So, any rack orientation maximizing intensity will limit exposure assuming a constant travel rate/pattern. Thus, there is no perfect orientation when considering both exposure and intensity. Which rack orientation is best for a retail store is likely to depend on traffic flow, scanning patterns, and what the store wants to prioritize in terms of exposure, intensity, or efficient space utilization.

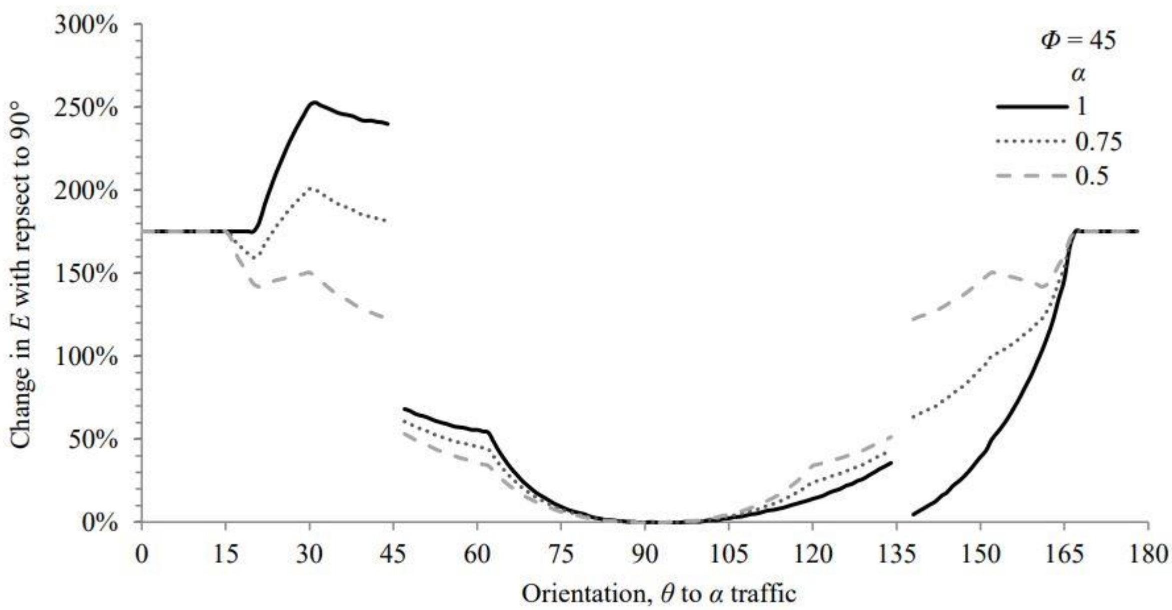

Figure 3. Change in exposure values (E) with respect to 90° as a function of rack orientation. Alpha (α) values represent shopper traffic flow with 1 representing unidirectional traffic, 0.75 representing mixed traffic, and 0.5 representing bidirectional traffic. Data are based on an environment comprised of a central aisle with 5 racks on either side and assuming a 45° head rotation. Adapted from Mowrey (2016), Doctoral Dissertation, Figure 19.

An unpublished study (Parikh, personal communication) was conducted to test the predictions made by the Parikh and Mowrey (2014) model. The 20 participants in the study performed an easy target detection task with no distractors. Participants were grouped into four orientation conditions in which each participant completed 3 trials with the 90° rack orientation and then 6 trials of their assigned orientation condition, either 30°, 45°, 90°, or 135°. This allowed for more trials of each of the more unusual rack orientations to help compensate for the lack of experience with them. Participants viewed a virtual environment in which they were moved along a central aisle and tasked with detecting 0.5” x 0.5” red squares distributed amongst gray-colored rack faces with no distractors. The results supported the model predictions in that observers detected the most targets with racks oriented at 45° and 135° (47% and 45%, respectively), they performed poorest for racks oriented at 90° (37%), and performance for racks oriented at 30° and 150° fell in between these two extremes (44% and 43% respectively). Moreover, the model was able to match human performance when the intensity threshold was 0.1 sec.

The current study examines how well the model fits human performance when applied to a target discrimination task amid distractors instead of a simple detection task. It tests model fit by determining the intensity threshold that best matches the human performance observed in the study. In terms of model fit, I hypothesize that the intensity threshold will need to be higher than in the detection study to match the increase in processing time needed to identify the targets amongst distractors. In terms of human performance, there is a main hypothesis and three sub-hypotheses. The main hypothesis is that there will be a significant effect of orientation on discrimination performance. The first sub-hypothesis is that the 45° rack orientation will significantly outperform the other orientations. The second sub-hypothesis is that the 30° rack orientation will significantly outperform the other orientations except for the 45° rack orientation. Finally, all rack orientations will produce better performance than the 90° rack orientation. Additionally, we expected that discrimination performance in the current study will be lower than detection performance in the original unpublished study.

Method

Participants

There were 27 participants recruited for the study. Three were excluded (two due to not following instructions during the study, one due to not passing the vision requirements) resulting in a sample of 24 participants (7 male/17 female) between the ages of 18-25 years (*M* = 19.04, *SD* = 1.65). Of the participants, 22 reported being right-handed, one left-handed, and the other ambidextrous. No participants reported having issues turning their head or having any form of color-blindness. Participants were recruited via the SONA system at Wright State University. The SONA system is a research participation system for introduction to psychology students to fulfill a portion of the course requirement. All participants were administered a vision test (Snellen chart following standard acuity testing procedures, Marsden, Stevens, & Ebri, 2014) and had normal or corrected to normal visual acuity (minimally 20/25). Additionally, participants completed a pre-test training where they needed to identify at least 4 out of 10 targets or they would be dismissed from the study - all participants met this criterion.

Apparatus

Experiments were performed at the Appenzeller Visualization laboratory at Wright State University. The Display Infrastructure for Virtual Environments (DIVE) configuration consists of 27 55-inch full-HD (1920x1080, 120 Hz, 4000:1 contrast ratio, 8ms response time, 450 cd/m^2 brightness) LED-backlit LCD displays with 3mm bezels (Samsung UE55A) mounted to form three walls (see Figure 4). Each wall is 87 inches in length and 144 inches wide and consists of nine displays in a 3x3 arrangement. Eleven cameras mounted above the screens in the environment allow for horizontal (x-axis) and vertical measurements (z-axis) of head position to be computed by monitoring the position of small localizer spheres (tracked at a 60 Hz rate) attached to the polarized glasses worn by the participants (see Figure 5). The virtual environment provided a 12 x 12 ft walkable space, large enough for participants to feel comfortable and able to examine their surroundings in a naturalistic way. This was important because many smaller virtual environments may make the participant feel constrained, which distracts from the immersion of the virtual environment and could thus affect their behavior. In the center of the floor there was a black-outlined square with an “X” in the center to identify where the participant was to stand during the experiment.

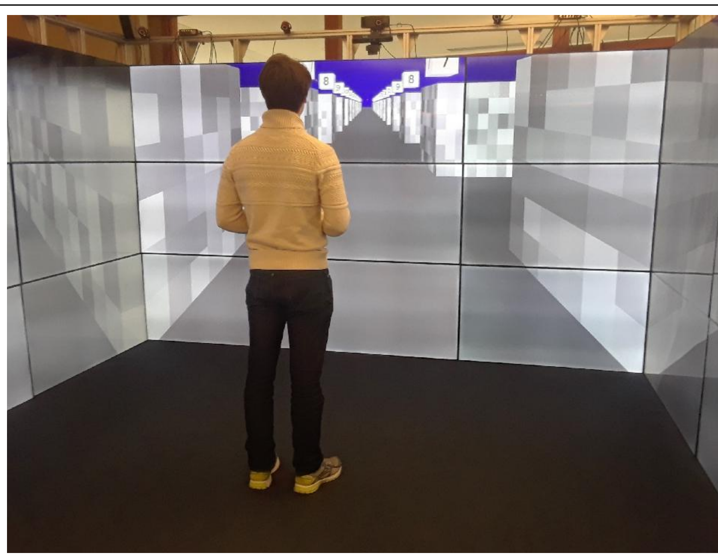

Figure 4. The author in the Appenzeller Visualization laboratory viewing a virtual environment with a 90° rack orientation.

The virtual shopping environment was rendered in 3D and was created utilizing the DIVE environment with OpenSceneGraph software. In order to create the 3D effect, two 'half-images' were generated for every display frame. Each half image was viewed by one eye, through the wearing of polarized glasses (Figure 5) and each half image was updated at 60 Hz (interleaved). The 3D environment was produced by introducing appropriate disparity between these half-images. Participants who normally wore eyeglasses to correct their acuity were required to wear their eyeglasses during the experiment and were offered velcro ties to fix the temples of the polarized glasses to their eyeglasses for improved stability.

Figure 5. Image of polarized glasses with localizer spheres attached.

Task

Within the virtual environment, participants experienced being moved along a straight central aisle of a simulated store flanked by 15 display racks on each side. The width of the central and side aisles was 5 ft. The racks had a height of 6.5 ft tall and each rack was 30 ft long. Six alpha-numeric characters appeared on each rack face at random locations and on a randomly selected 10 rack faces, one of the characters was a target. Participants were to detect and identify the targets as they were moved through the environment.

Independent Variables

The independent variable in this experiment is rack orientation relative to the central aisle. All participants experienced 5 different rack orientations (30°, 45°, 90°, 135°, & 150° - see Figure 6) and completed two trials with each rack orientation. Each trial had a different distribution of targets so that in every trial participants needed to examine each rack face for a possible target.

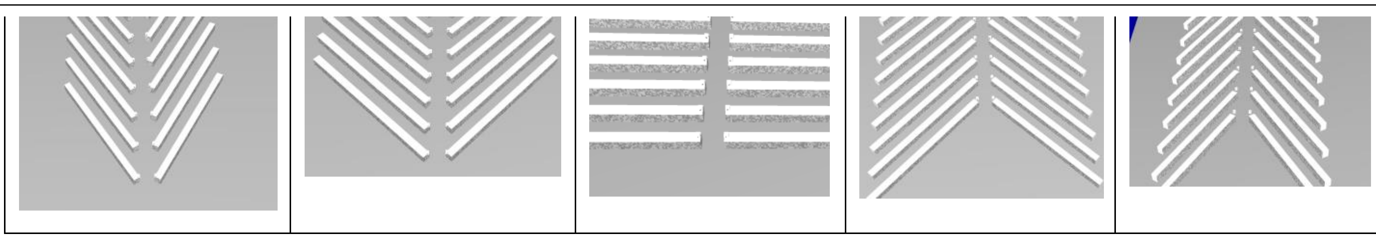

Figure 6. Representations of the 5 rack orientations (from left to right: 30°, 45°, 90°, 135°, & 150°) used in this study.

Dependent Measures

There were two dependent measures: target discrimination performance and head rotation behavior. Participants in the virtual environment were given the task of identifying letter targets presented at a simulated size of 8” x 5” distributed among all the rack faces. From the participant’s perspective in the central aisle, these targets would subtend on average 5.0 x 3.2 deg of visual angle at their nearest location (on the rack face adjacent to the central aisle and the participant directly in front of the rack face) to 1.0 x 0.64 deg at their most distant location (on the rack face at the opposite end from the central aisle). In each trial there were a total of 10 targets. All rack faces contained six alpha-numeric characters, but on only 10 rack faces one of those characters was a target. Targets never appeared on the first two front and last two back rack faces because these could be viewed for longer than all of the other rack faces due to the traveled path's starting and ending points. Targets were letters randomly chosen with replacement from among these: B, C, D, E, F, H, J, K, L, M, T. Since a letter could appear more than once within the same trial, this prevented participants from inferring a target based on a process of elimination given previous letters they had identified. Distractors were single digit numbers excluding the numbers 1 and 0 due to their similarity to letters such as upper-case I, lower-case L, and upper-case O (see Figure 7). Both targets and distractors were presented in the same font and black color while rack faces were colored grey, orange, blue, green, red, or yellow (presented at near physical equiluminance) such that no two rack faces forming a single lateral aisle were the same color (see Figure 8). Given that the rack faces had a luminance ranging from 6.11-7.4 cd/m^2 (measured with a Minolta 1° luminance meter) and the alpha-numeric characters had a luminance of 0.05

cd/m^2, targets and distractors had contrasts ranging from 98.38-98.66%.

Figure 7. Example of a rack face (orange color) that a participant may have seen during the current

Target discrimination judgements were recorded manually by an experimenter using a scoring sheet with a list of all targets and the color of each rack face (see Figure 8). The experimenter was seated outside of the test environment approximately 15 feet directly behind the participant to minimize being in the participant's field of view. This allowed the participant's attention to be focused on the task instead of

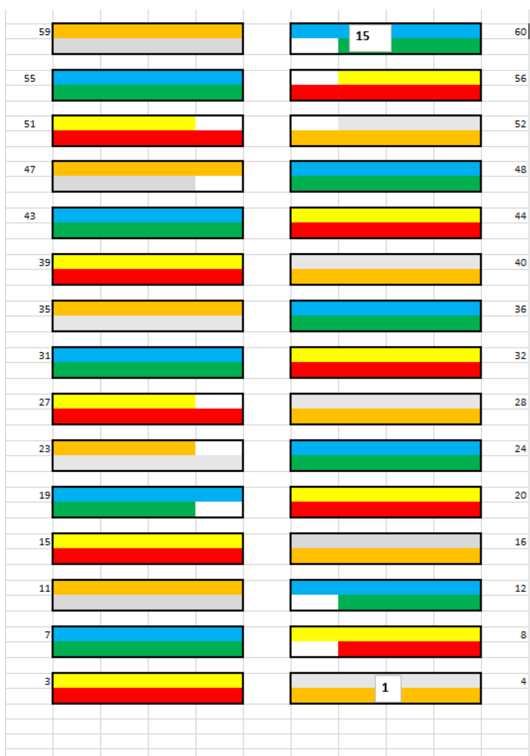

Figure 8. Example scoring sheet used to mark and record participant responses. White spaces represent that a target is placed on that rack.

being drawn to the experimenter.

Discrimination data was coded by target location (front/back face of rack, left/right side aisle) so that target location variables could be further analyzed along with the more general percent of targets discriminated for each rack orientation.

Participant head movements in the virtual environment were examined by looking at the number of times a participant turned their head crossing the straight-ahead position (center crosses), bias towards one side (percentage of time looking to the left side versus the right side), average amount of head rotation in degrees (Φ), and average angular speed of rotation.

Procedure

Design

The present study involved a five condition (30°, 45°, 90°, 135°, & 150° rack orientation) within-subjects design. To familiarize participants with the virtual environment and task, all participants went through a basic virtual environment consisting of a central aisle with 15 surrounding racks on both sides (see Figure 9). Participants were virtually moved down the central aisle at a constant speed of 4.06 ft/s. This speed was chosen based on participant preference during pilot testing. The only difference between the five rack conditions was the angle of the racks relative to the central aisle. However, due to some rack orientations taking up more space than others (obtuse orientations, ie. 135° & 150°), the time required to traverse from the start to the end point of the virtual environment was not constant. Thus, a trial for the 90° racks took approximately 73 seconds while that for the 150° racks took approximately 119 seconds.

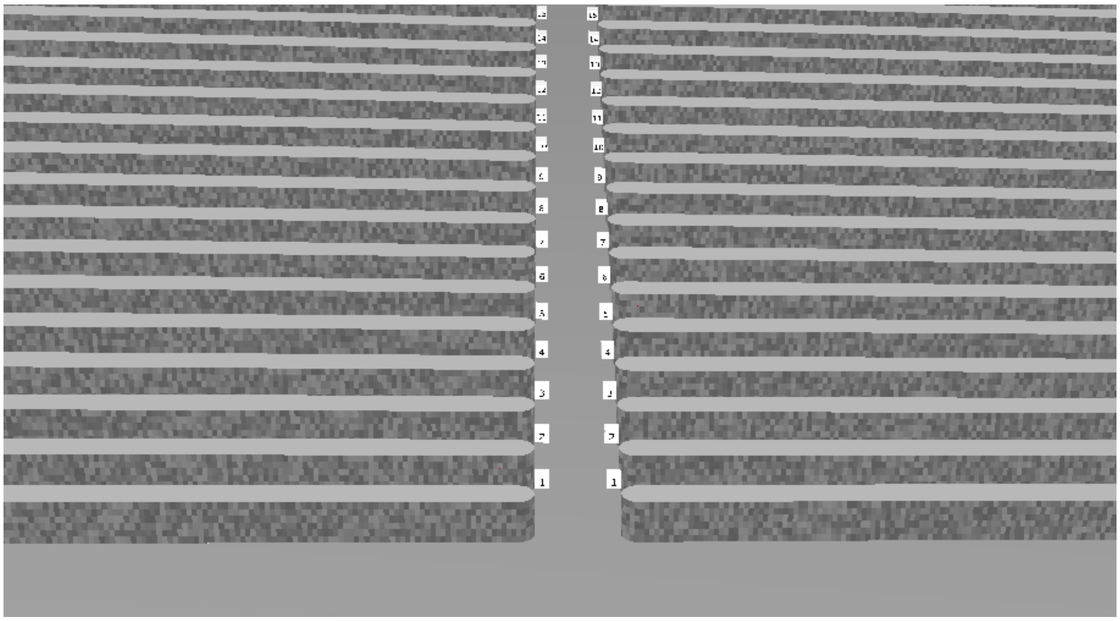

Figure 9. Schematic representation of a top view of the virtual shopping environment with 15 racks on either side of the central aisle oriented at 90°.

First participants were introduced to the study and given an overview, then given an informed consent document to sign. Participants who did not consent were thanked for their time and dismissed, but this did not occur. Once participants consented to participate, they were given a visual acuity test using a Snellen chart. If they met the acuity requirement (20/25 or better), participants then completed an electronic pre-survey providing their demographic information and shopping habits. Participants then were given a simulator sickness questionnaire (Kennedy, Lane, Berbaum, & Lilienthal, 1993). If a participant reported having half of the symptoms on the simulator sickness questionnaire at a moderate or severe level they were not allowed to participate, but this never occurred. Participants were then administered a postural stability test (consisting of the sharpened Romberg and stand on one leg eyes closed tests from Hamilton, Kantor, & Magee; 1989) as a second component to assess simulator sickness. The postural stability test was also completed after the experiment. Simulator sickness was then assessed by comparing the pre- and post-experiment measurements of postural stability for significant deviations. If a participant was found to have a significant deviation between their pre and post measures, they were asked to remain in the laboratory until transportation was arranged for them to return home. However, no subjects had this deviation.

Once participants were cleared to participate, they were led to the virtual environment. Participants were instructed to stand in the middle of the floor on the black-outlined box with an "X" in it. Participants were then given the 3D glasses and controller to hold. The controller was used to start the trial by pressing a button when ready. Once the participants donned the 3D glasses and reported that they were comfortable, the task was then explained to participants. Once participants understood the task, they were then asked to examine the different rack colors and confirm a short one- to two-syllable color name to reduce possible reporting confusion. Participants' head-turning ability was then assessed by having them turn their head as far to the right, left, up, and down as felt comfortable, returning to the center after each direction. Following this, and after answering any task-related questions, participants completed a practice trial (90° rack orientation) to confirm they understood the task. Participants who satisfactorily completed the practice trial by correctly identifying at least four of the 10 targets and responding in the required manner (stating the rack face color followed by the letter name) were given the option to perform a second practice trial or continue onto the main test. Participants who did not satisfactorily discriminate the number of tasks or stay inside the square during the the task were required to complete a second practice trial. Any participant not able to perform the task satisfactorily in the practice trials was dismissed and the session terminated.

Once a participant completed two practice trials or said they wanted to move forward after the first satisfactory practice trial, the experiment session began. All participants went through 10 trials, two trials for each of the five rack orientations. In each trial the targets were distributed differently among the 56 possible rack faces so no participant saw the same distribution of targets twice. A python script was used to create these target distributions. The script first randomly chose 10 rack faces (five front and five rear rack faces) out of the 56 possible rack faces to contain a target. The possible rack faces for targets was only 56 rather than 60 because the first two and the last two rack faces were excluded as they are not obscured to the same extent by other racks and participants would have had more time to view them. All rack faces had 6 distractor numbers placed in randomly selected locations on the rack face. If the rack face was to have a target, one of the distractors was replaced with a randomly chosen target letter. This process was then repeated independently for all ten experimental trials. A practice trial was created using the same process, but target locations were restricted to the half of the rack face closest to the central aisle in order to prioritize understanding of the task. The randomization of target and distractor location was implemented to require participants to consider the entirety of a rack face instead of quickly scanning just one portion of the rack face. To keep task difficulty level constant, all participants completed the same 10 trials (and thus experienced the same target distributions) but in different orders. Participants were assigned to one of four orders via the random number generator from random.org (Haar, n.d.) before they started participating. Participants started with either the 30° or the 150° rack orientation which then increased or decreased respectively with each subsequent trial. Once participants completed the first five trials, the pattern was either repeated or reversed depending upon which order the participant was assigned.

Results

Discrimination Performance.

Target discrimination performance for both front and back rack faces for each of the 5 rack orientations are shown in Figure 10. Values are averaged across all participants and trials to create an aggregate score for each rack face for each rack orientation. A two-way repeated measures analysis of variance revealed a significant effect of rack face ($F(1,47) = 531.77$, $p < .05$), orientation ($F(4,188) = 62.38$, $p < .05$), and an interaction between rack face and orientation ($F(4,188) = 314.87$, $p < .05$). A follow-up pairwise analysis found a significant difference between front and back targets hits for all conditions (see Table 1) with the 90°, 135°, and 150° rack orientations producing better performance for front face target hits and 30° and 45° producing more target hits on the back face.

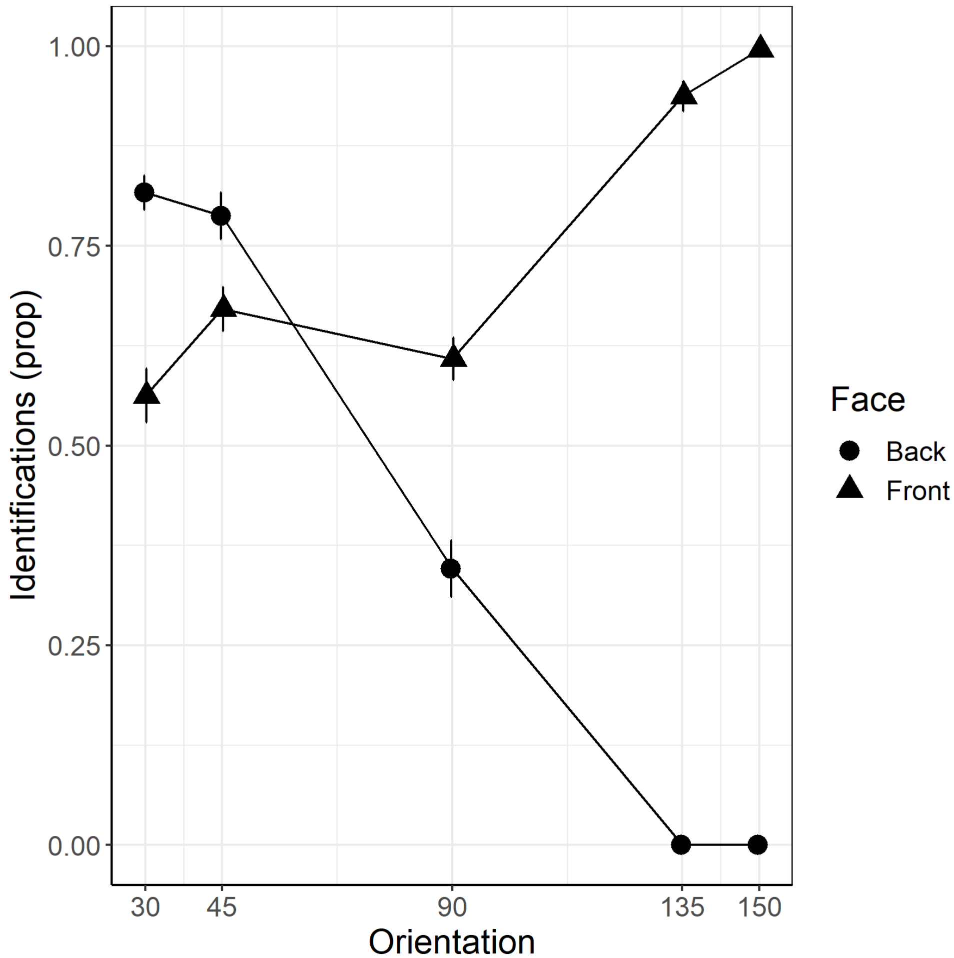

Figure 10. Target identification for front (filled triangles) and back (filled circles) rack faces by rack orientation. Error bars represent ± 1 standard error of the mean.

Table 1.

Pairwise comparisons (t-tests) for front versus back rack face target identifications. Level of significance is indicated by asterisks: **** - p < .0001, *** - p < .001, ** - p < .01, * - p < .05, ns - p >= .05. The p value was adjusted (p.adj) using the greenhouse-geisser correction.

	Orientation	.y.	group1	group2	n1	n2	statistic	df	p	p.adj	p.adj.signif
1	135	Identifications	Back	Front	48	48	-49.40	47.00	0.00	0.00	****
2	150	Identifications	Back	Front	48	48	-239.00	47.00	0.00	0.00	****
3	30	Identifications	Back	Front	48	48	6.38	47.00	0.00	0.00	****
4	45	Identifications	Back	Front	48	48	3.10	47.00	0.00	0.00	**
5	90	Identifications	Back	Front	48	48	-5.64	47.00	0.00	0.00	****

I had hypothesized that target discrimination performance would be poorer than target detections as measured in the previous unpublished study (Parikh, personal communication). However, this was not confirmed in our initial analysis. In fact, discrimination performance was better than detection for all rack orientations (see Table 2, Raw Percent Correct). This was unexpected and puzzling. One thought was that perhaps the contrast of the red target in the detection experiment was lower than the letters in the current study making it more difficult to detect. However, we were not able to find a measure of the contrast of the red target to justify this explanation. However, a closer look at the detection experiment provided a more likely explanation. In the detection study, 50-60% of the targets were purposely positioned on rack locations that the model of visibility, using a higher estimated 90° for average head rotation (Φ), predicted would not be visible. This made sense as the researchers wanted to test if the model predictions would be close to human performance. However, in the present study targets were simply randomly positioned and not purposely put on rack positions expected not to be seen. Thus, the two studies are not comparable in terms of the proportion of targets that one would consider even possible targets, at least according to the Parihk and Mowrey model. To correct for this, the percent of targets detected in the unpublished Parihk study and the percent of targets discriminated in the current study were recalculated using the Parihk and Mowrey (2014) model's predicted number of targets that would be visible given a 90° average head rotation (Φ). This data corrected for target positioning also appears in Table 2. Using these corrected data, my hypothesis that discrimination performance would be poorer than detection performance is supported for all but one rack orientation, 150°.

Table 2.

Average target detections for each rack orientation for the unpublished study (Parikh, personal communication) and target identifications for the current study. Data are provided as the raw performance (computed relative to all possible targets) and corrected for target positioning (computed relative to the model-predicted number of visible targets)

	Raw Percent Correct		Corrected For Target Positic	
Rack Orientation	Detection	Identification	Detection	Identificatio
30°	44	69	87	73
45°	47	73	105	73
90°	37	48	106	95
135°	45	47	106	82
150°	43	50	88	100

The better discrimination than detection performance seen for the 150° orientation was not expected and it is not entirely obvious why this would be so after correcting the data for target positioning. One potential reason could be a difference in level of engagement by participants in the two studies. Participants tasked with identifying targets amongst distractors may have attempted to meet the increased difficulty or treated it as more of a game than those participants who simply were detecting a red target against a grey background. However, since engagement was not measured in either study this cannot be verified. Moreover, it is difficult to imagine that a difference in the level of engagement would only occur for the 150° rack orientation condition. Thus, the reason for the better discrimination than detection performance for the 150° orientation condition is unknown.

Head Rotation Behavior.

To examine whether participants searched for targets differently when rack orientation differed, participants head rotation behavior was analyzed. An algorithm determined head turns by identifying a start point (time and position) and end point (time and position) for each head movement in the following way. If the head did not move for 0.5 sec, the algorithm identified its current position and time as a starting point. The starting point was updated every 0.5 sec if the head did not move during that time. When the head moved more than 2 degrees in 0.5 sec, the direction of the head movement (left (positive)/right (negative)) was determined and the algorithm continued to check head position in 0.5 sec intervals until the largest position deviation from the starting point was found and this was identified as the endpoint of that head movement. Head rotations less than 10 degrees were discarded to ensure that recorded head turns were not artifacts of a participant repositioning themselves. From these measurements we computed the number of head turns crossing the center position, the bias towards searching one side, as well as the average head rotation magnitude and their average angular velocity.

First, the average head rotation magnitude is plotted as a function of rack orientation in Figure 11. Overall, the largest head turns were made in the 135° and 90° rack orientation conditions, followed by the 45°, 150°, and 30° orientations. A one-way repeated measures analysis of variance revealed a significant effect of orientation on head rotation magnitude (*F(4,188)* = 123.349 *p* < .001). Table 3 shows the results of all pairwise comparisons. Head turn magnitude for each rack orientation was significantly different from that for all other rack orientations except between 135° and 90°.

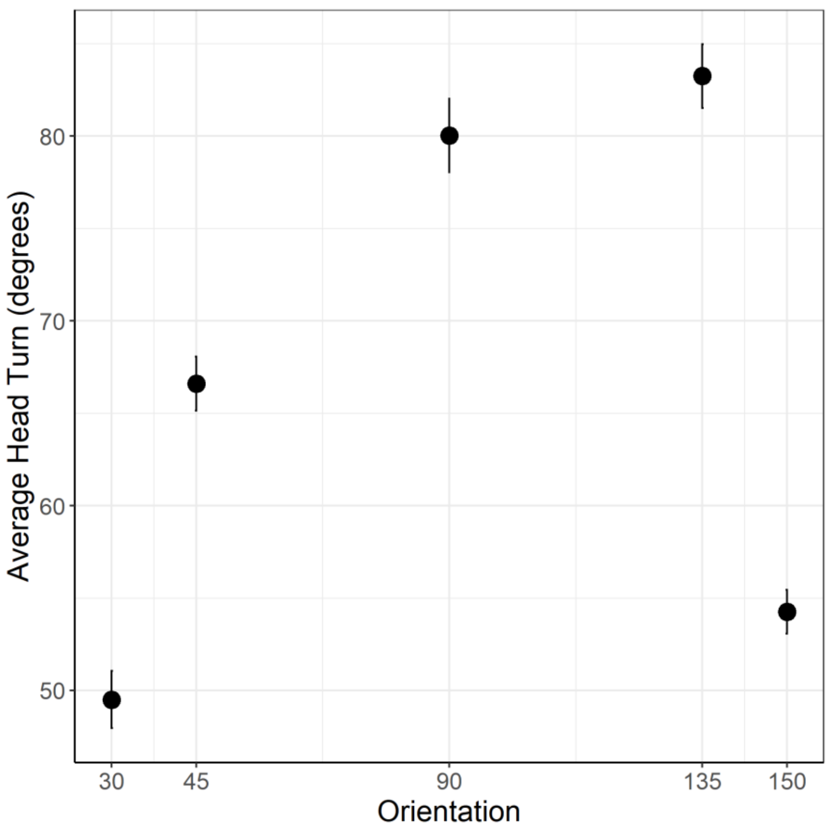

Figure 11. Average head rotation magnitude plotted as a function of rack orientation. Error bars represent ± 1 standard error of the mean.

Table 3.

Pairwise comparisons (t-tests) for average head turn angle. Level of significance is indicated by asterisks: **** - p < .0001, *** - p < .001, ** - p < .01, * - p < .05, ns - p >= .05. The p value was adjusted (p.adj) using the greenhouse-geisser correction.

	.y.	group1	group2	n1	n2	statistic	df	p	p.adj	p.adj.sig
1	Average Head Turn	135	150	48	48	15.55	47.00	0.00	0.00	****
2	Average Head Turn	135	30	48	48	15.29	47.00	0.00	0.00	****
3	Average Head Turn	135	45	48	48	9.33	47.00	0.00	0.00	****
4	Average Head Turn	135	90	48	48	1.50	47.00	0.14	1.00	ns
5	Average Head Turn	150	30	48	48	3.55	47.00	0.00	0.01	**
6	Average Head Turn	150	45	48	48	-9.38	47.00	0.00	0.00	****
7	Average Head Turn	150	90	48	48	-12.10	47.00	0.00	0.00	****
8	Average Head Turn	30	45	48	48	-12.24	47.00	0.00	0.00	****
9	Average Head Turn	30	90	48	48	-11.91	47.00	0.00	0.00	****
10	Average Head Turn	45	90	48	48	-6.83	47.00	0.00	0.00	****

Additionally, average head turn angular velocity followed a similar pattern as mean head turn magnitude, but here the 90° rack orientation elicited the fastest head turns followed by the 135°, then 45°, 150°, and 30° orientations (see Figure 12). A one-way repeated measures analysis of variance revealed a significant effect of orientation on average angular velocity ($F(4,188) = 90.69$, $p < .001$). The results of follow-up pairwise comparisons are shown in Table 4. Participants in the 135° rack orientation tended to turn their heads significantly faster than for the other rack orientations, except in comparison to 90°. Participants tended to have the next fastest head turns in the 45° rack orientation, then 150°, and the slowest head turns in the 30° rack orientation. Considering this data in relation to the average head turn magnitude shows that when observers produced larger head turns they also tended to move their heads faster, whereas smaller head turns tended to be slower.

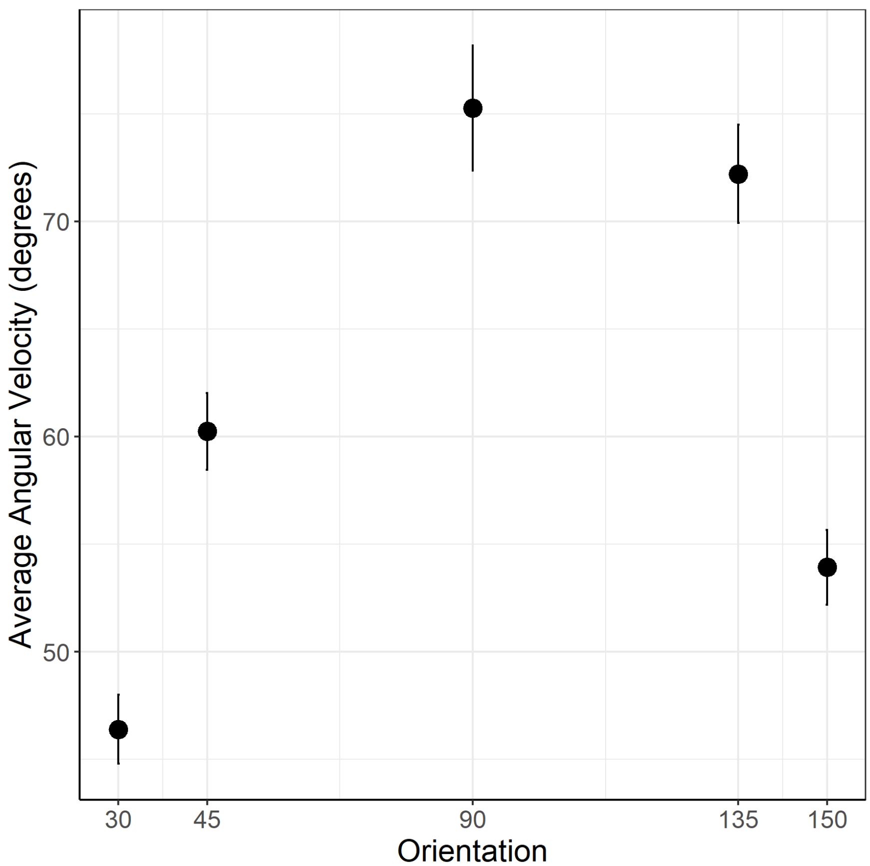

Figure 12. Average head turn angular velocity plotted as a function of rack orientation. Error bars represent ± 1 standard error of the mean.

Table 4.

Pairwise comparisons (t-tests) for average angular velocity of head turns. Level of significance is indicated by asterisks: **** - p < .0001, *** - p < .001, ** - p < .01, * - p < .05, ns - p >= .05. The p value was adjusted (p.adj) using the greenhouse-geisser correction.

	.y.	group1	group2	n1	n2	statistic	df	p	p.adj	p.a
1	Average Angular Velocity.	135	150	48	48	12.33	47.00	0.00	0.00	***
2	Average Angular Velocity.	135	30	48	48	14.53	47.00	0.00	0.00	***
3	Average Angular Velocity.	135	45	48	48	8.49	47.00	0.00	0.00	***
4	Average Angular Velocity.	135	90	48	48	-1.76	47.00	0.08	0.84	ns
5	Average Angular Velocity.	150	30	48	48	5.57	47.00	0.00	0.00	***
6	Average Angular Velocity.	150	45	48	48	-4.59	47.00	0.00	0.00	***
7	Average Angular Velocity.	150	90	48	48	-9.06	47.00	0.00	0.00	***
8	Average Angular Velocity.	30	45	48	48	-10.24	47.00	0.00	0.00	***
9	Average Angular Velocity.	30	90	48	48	-11.28	47.00	0.00	0.00	***
10	Average Angular Velocity.	45	90	48	48	-7.01	47.00	0.00	0.00	***

When examining the average number of head turns per orientation condition a different pattern in the data emerged. As can be seen in Figure 13, the 30° and 150° rack orientations produced the most head turns, followed by 45° and 135° and then 90°. A one-way repeated measures analysis of variance revealed a significant effect of rack orientation on number of head turns ($F(4,188)$ = 93.43, p < .001). The results of follow-up pairwise comparisons are shown in Table 5. The most extreme rack orientations, 30° and 150°, produced a significantly higher number of head turns than the other rack orientations and were not significantly different from each other. Examining the next less extreme rack orientations of 45° and 135°, participants in these rack orientations had significantly more head turns than they did in the 90° rack orientation, but 45° and 135° were not distinguishable from each other. Finally, participants in the 90° rack orientation tended to turn their heads the least number of times.

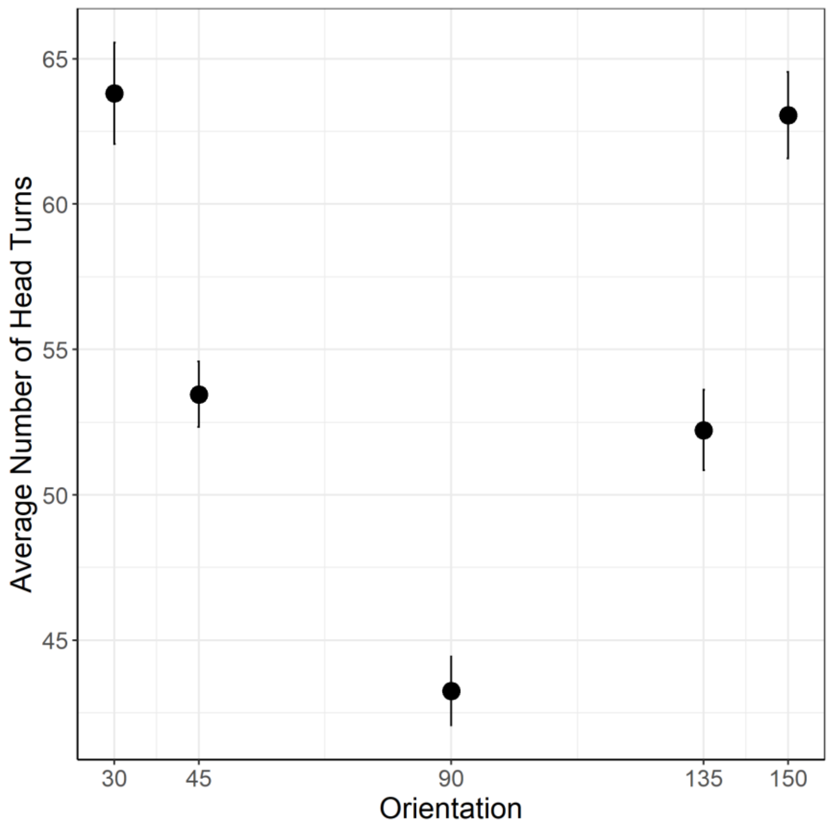

Figure 13. Average number of head turns plotted as a function of rack orientation. Error bars represent ± 1 standard error of the mean.

Table 5.

Pairwise comparisons (t-tests) for average number of head turns. Level of significance is indicated by asterisks: **** - p < .0001, *** - p < .001, ** - p < .01, * - p < .05, ns - p >= .05. The p value was adjusted (p.adj) using the greenhouse-geisser correction.

	.y.	group1	group2	n1	n2	statistic	df	p	p.adj	p.adj.
1	Head Turns	135	150	48	48	-10.45	47.00	0.00	0.00	****
2	Head Turns	135	30	48	48	-7.95	47.00	0.00	0.00	****
3	Head Turns	135	45	48	48	-1.08	47.00	0.29	1.00	ns
4	Head Turns	135	90	48	48	6.22	47.00	0.00	0.00	****
5	Head Turns	150	30	48	48	-0.61	47.00	0.54	1.00	ns
6	Head Turns	150	45	48	48	9.58	47.00	0.00	0.00	****
7	Head Turns	150	90	48	48	15.61	47.00	0.00	0.00	****
8	Head Turns	30	45	48	48	7.95	47.00	0.00	0.00	****
9	Head Turns	30	90	48	48	13.18	47.00	0.00	0.00	****
10	Head Turns	45	90	48	48	12.02	47.00	0.00	0.00	****

When one considers the number of head turns along with the previous two measures, average head turn velocity and magnitude, an interesting pattern can be seen. Specifically, rack orientations of 90° and 135° elicited fewer head turns but they were larger and faster. However, rack orientations of 30° and 150° elicited more head turns but they were smaller and slower.

Another variable related to head turn behavior was bias (percent of time looking to the left versus right side, head orientation had to be greater than 10 degrees to be counted as not straight ahead). A one-way ANOVA found a significant effect of rack orientation on bias ($F(4,188) = 16.412$, $p < .001$). As can be seen in Figure 14, participants in each orientation tended to spend more time oriented towards the left side than the right side. The 90° rack orientation had the largest bias towards the left side at 15.25 percent more than the right followed by the 135° rack orientation at 13.16 percent more than the right side. The results of a followup pairwise analysis t-test are shown in Table 6. An interesting pattern emerged in which participants in the 30° and 150° rack orientations were significantly less biased to the left side than the 45°, 90°, and 135° rack orientations. Both the 30° and 150° orientations and 45°, 90°, and 135° orientations were not significantly different from each other.

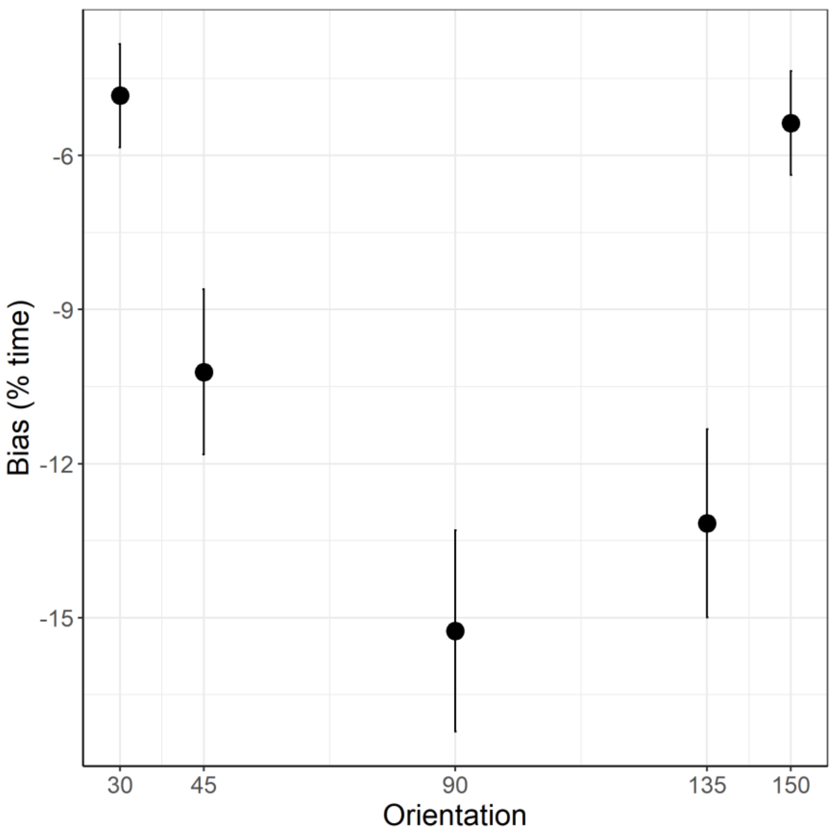

Figure 14. Bias towards one side (negative values represent the left side; positive values represent the right side) plotted as a function of rack orientation. Error bars represent ± 1 standard error of the mean.

Table 6.

Pairwise comparisons (t-tests) for bias, percentage of time spent looking towards left side (negative) vs right side (positive). Level of significance is indicated by asterisks: **** - p < .0001, *** - p < .001, ** - p < .01, * - p < .05, ns - p >= .05. The p value was adjusted (p.adj) using the greenhouse-geisser correction.

	.y.	group1	group2	n1	n2	statistic	df	p	p.adj	p.adj.signif
1	Bias	135	150	48	48	-5.57	47.00	0.00	0.00	****
2	Bias	135	30	48	48	-5.20	47.00	0.00	0.00	****
3	Bias	135	45	48	48	-2.29	47.00	0.03	0.26	ns
4	Bias	135	90	48	48	0.98	47.00	0.33	1.00	ns
5	Bias	150	30	48	48	-0.48	47.00	0.63	1.00	ns
6	Bias	150	45	48	48	3.85	47.00	0.00	0.00	**
7	Bias	150	90	48	48	5.83	47.00	0.00	0.00	****
8	Bias	30	45	48	48	3.62	47.00	0.00	0.01	**
9	Bias	30	90	48	48	5.54	47.00	0.00	0.00	****
10	Bias	45	90	48	48	2.57	47.00	0.01	0.13	ns

A final measure of head-turn behavior was the average number of times a head turn was made across the center (straight ahead) per orientation condition. The results are plotted in Figure 15 and show that rack orientations of 30°, 45°, 135°, and 150° produced a similar number of head turn center crosses, but the 90° rack orientation produced the fewest head turn center crosses. A one-way repeated measures ANOVA revealed a significant effect of orientation on center crosses ($F(4,188)$ =14.389, p < .001). Results of a follow-up pairwise analysis are shown in Table 7. Participants in the 90° rack orientation made significantly fewer head turn center crosses than in the other rack orientations, which were not statistically different from each other.

Overall, the head turn data provides evidence that participants seem to be adapting their behavior to the rack orientation. Specifically, an interesting relationship between number of head turns, average head turn in degrees, and angular velocity emerges. When participants were in an environment with rack orientations such as 90° or 135°, they made larger head turns that were faster, but they made fewer of them. However, when participants were in an environment with rack orientations such as 30° or 150°, they made smaller head turns that were slower and they made more of them. Moreover, when observers showed more of a bias to look to one side (such as for the 90° rack orientation), they also tended to make fewer center crosses.

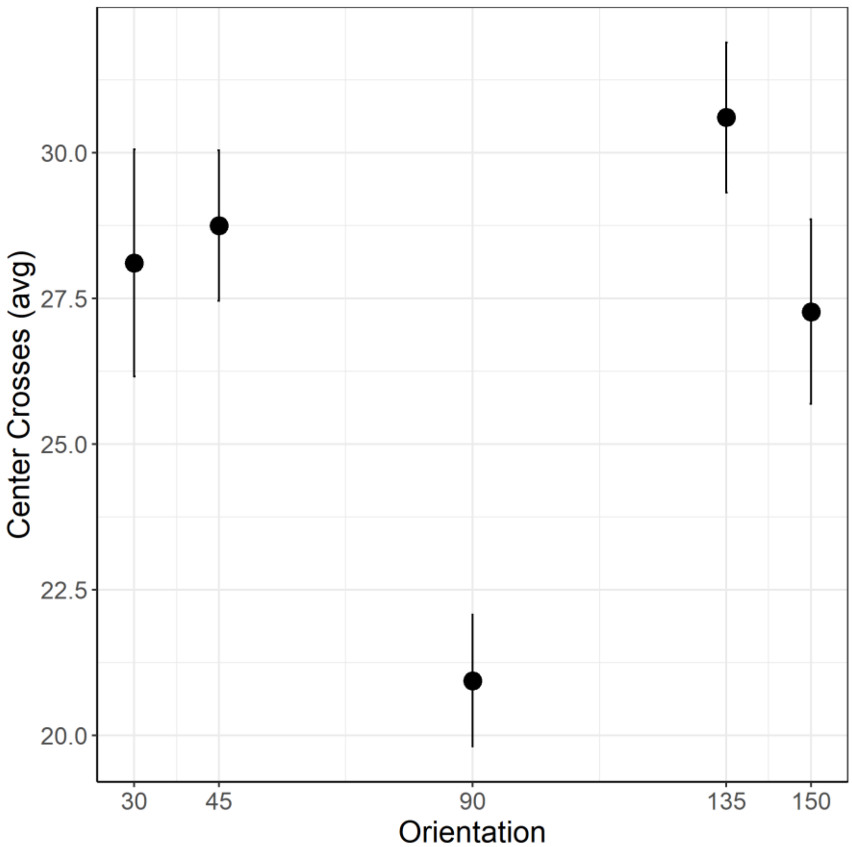

Figure 15. Average number of center crosses plotted as a function of rack orientation. Error bars represent ± 1 standard error of the mean.

Table 7.

Pairwise comparisons (t-tests) for average number of head turns crossing the center (straight ahead). Level of significance is indicated by asterisks: **** - p < .0001, *** - p < .001, ** - p < .01, * - p < .05, ns - p >= .05. The p value was adjusted (p.adj) using the greenhouse-geisser correction.

	.y.	group1	group2	n1	n2	statistic	df	p	p.adj	p.adj.sig
1	Center Crosses	135	150	48	48	2.81	47.00	0.01	0.07	ns
2	Center Crosses	135	30	48	48	1.68	47.00	0.10	1.00	ns
3	Center Crosses	135	45	48	48	2.00	47.00	0.05	0.52	ns
4	Center Crosses	135	90	48	48	10.55	47.00	0.00	0.00	****
5	Center Crosses	150	30	48	48	-0.54	47.00	0.59	1.00	ns
6	Center Crosses	150	45	48	48	-1.16	47.00	0.25	1.00	ns
7	Center Crosses	150	90	48	48	4.27	47.00	0.00	0.00	***
8	Center Crosses	30	45	48	48	-0.41	47.00	0.68	1.00	ns
9	Center Crosses	30	90	48	48	3.80	47.00	0.00	0.00	**
10	Center Crosses	45	90	48	48	6.98	47.00	0.00	0.00	****

Model Performance

In this section I will attempt to 'fit' the model of Parikh and Mowrey (2014) to the collected discrimination data. Recall that this model was built to help assess what a shopper would see as they walked through the central aisle of a store, rotating their head as they walk, with product racks extending out from each side of the central aisle. Also recall that what the model determines is whether any portion of a display rack falls within an observer's field of regard (exposure), assuming a 30° stationary field of view, and for how long (intensity). While the model was fit to data from a simple detection task (Parikh, personal communication), it is important to realize that the model does not perform any task. For that previous detection task, the model 'fit' produced an intensity value (or threshold) that resulted in the same number of targets exceeding that value in the model as the number of targets detected by the human observers. So, in the model, if a target falls within the prescribed observer's field of view for long enough, then the model assumes that the target is 'seen'. It is important to note that the model can make mistakes in only two ways: 1) the model does not 'see' a target that the human saw because it does not have an intensity that is equal to or greater than the requisite value (threshold) set by the experimenter, or 2) it 'sees' a target (intensity at that rack location exceeded the threshold) that the human observer did not.

The model has three main variables of interest to the current study – step size, average phi, and an intensity threshold. The model algorithm determines intensity values for each target location based on the step size and average phi. The average phi value is the magnitude of the head rotation for the simulated participant going through the central aisle. The step size represents how long it takes the simulated person to be transported through the virtual environment, passed all of the racks. Setting step size to a value lower than 1.0 basically reduces the speed with which the person is moved through the environment. This would increase the duration someone would see the racks and increase the intensity values for the targets, distractors, and all locations on the racks. In the current study we assume a step size of 1.0 to be consistent with the previous detection study (Pratikh, personal communication) but used participants' average maximum head-turn angle to each side for phi (head-turn angles: 37°, 50°, 66°, 56°, and 40° for rack orientations of 30°, 45°, 90°, 135°, and 150° respectively). Finally, the intensity threshold refers to how long a person needs to see a target for it to be detected, or in this study, discriminated. Recall our initial hypothesis is that the model will require a longer time for target discrimination than detection, so we hypothesize that the model intensity threshold will need to be at a higher value to account for the current discrimination data.

In the detection experiment (Pratikh, personal communication), the model was fit to the data by setting step size = 1 and phi = 90°. Then the intensity threshold was progressively increased until the number of target locations that met or exceeded the threshold was equal to the average number of targets detected by the human observers. Note that when the intensity threshold is very low (ie. 0.01 sec), every location on a rack would meet or exceed the threshold (unless a location never appeared in the assumed field of view), such that the model would predict that every target in the field of view would be seen. Therefore, to reduce the number of target locations predicted to be seen by the model, intensity threshold is increased. It is important to note that this fitting procedure does not guarantee that the target locations that meet or exceed the model's intensity threshold were the same ones that the human observers correctly detected.

For the current discrimination data, the model fit was initially determined by computing the d' value for each orientation across observers. For this d' calculation, human discrimination performance was used as the ground truth. That is, only targets correctly identified by the human observer were considered as "target present" (or signal plus noise) trials, targets not correctly identified by the human observer were considered as "target absent" (noise alone) trials. For each value of intensity threshold tested (the model was run for all intensity threshold values from 0 to 32 in increments of 0.1), the proportion of "hits" (how many targets exceed the intensity threshold that were also correctly discriminated by the observer) and "false alarms" (how many targets exceed the intensity threshold that were not correctly discriminated by the observer) the model predicted were determined. I then used a standard formula to compute (d' = z(False Alarms) – z(Hits). Since the z-score for a proportion of 1.0 or 0.0 cannot be determined (it is plus or minus infinity), a correction was used to enable d' to be computed: if the proportion was 1.0, the value was changed to 1-1/(2N), if the proportion was 0.0, the value was changed to 1/(2N) based on recommendations from the *Signal Detection Theory User's Manual* (Macmillan & Creelman, 2004). This analysis was done across observers for each rack orientation to get an overall model score for each rack orientation. These orientation-specific scores were then averaged together to get a holistic score for the model. Recall that the intensity threshold represents how long a target needs to be within the defined field of view for a target to be considered seen. For example, with an intensity threshold of 8, if a target's intensity was determined to be 7.9 per the model algorithm, the model would not predict it to be seen. Here we are determining the maximum intensity needed to produce the best model fit as indicated by the highest d' value. The maximum intensity value was chosen as it is a more conservative estimate of the amount of time needed for discrimination.

Figure 16 shows the model's performance (d') plotted as a function of intensity threshold. Holistically, the model had mixed results with the present target discrimination task. Averaging across all orientations, best model performance was observed when the intensity threshold value was .006 sec, resulting in a d' = 1.398. While that seems like good performance, a closer look at the model performance for individual rack orientations reveals an interesting phenomenon. Looking at Figure 17, notice that the d' values for the 135° and 150° orientations are much higher than the scores for most of the other orientation and in particular have high scores even for the shortest intensity threshold tested where every target that falls in the model's assumed field of view would be predicted to be seen. is unusual and the reason is related to how the human observers' target hits were distributed between the front and back faces of the racks (see Figure 10). For the two obtuse rack orientations (135° & 150°) the observers correctly discriminated virtually all of the targets on the front face of the racks and no targets on the back face of the racks. This occurred because observers apparently adopted a task strategy in which they did not even attempt to look at the back face of the racks, likely because it was extremely difficult to inspect the entire front rack face and then inspect the back rack face. As a result, observers did not move their head a large enough angle for that back-rack face to even enter into their field of view. Since the model's field of view was computed assuming the observers' average head rotation (phi), the model also did not 'look' at the back-rack face, so the model essentially does not make any false alarms, making for high d' values. Thus, for the 135° and 150° rack orientations, the model produces intensity values for the front rack face that are quite high, whereas the back-rack face values are 0. Therefore, the obtuse orientations have a different score from the other orientations primarily because observers have adopted a different search strategy in those conditions. Due to this difference, another average was calculated utilizing only the 30°, 45°, and 90° rack orientations. Examining Figure 16 shows that excluding the 135° and 150° rack orientations from the average lowers d', with the largest drop seen at lower intensity thresholds. Interestingly, the best model fit is now found at an intensity threshold of (1.306 seconds), but at a d' = 0.841 This is also the highest peak point for the grand average for all intensity thresholds larger than 1.306 seconds.

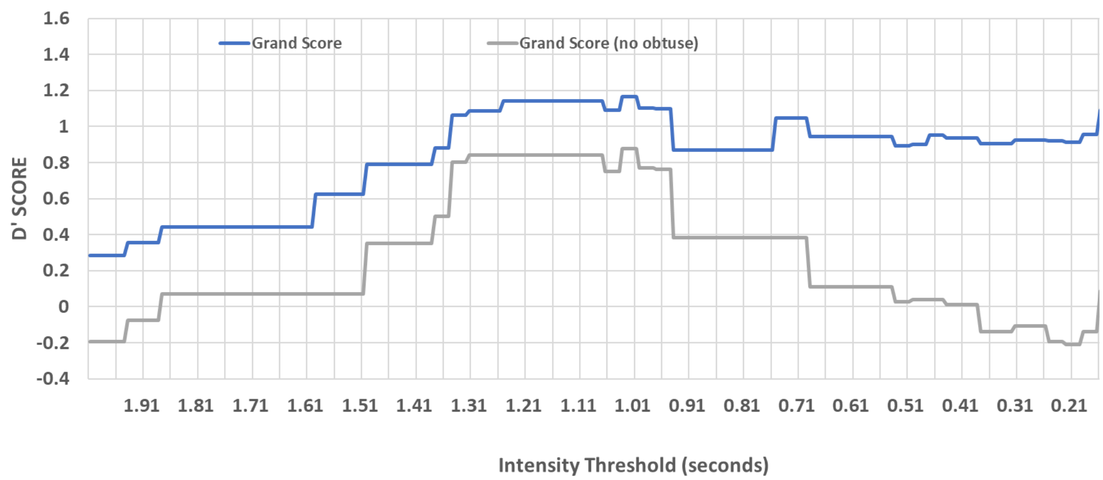

Figure 16. Average d' model scores as a function of intensity threshold. The blue line represents the model score averaged across all orientations and the silver line represents the average model score for just the 30°, 45°, and 90° orientations.

Due to the differences in model fit seen in Figure 16, it can be useful to compare model performance among orientations to find insights into the strengths and weaknesses of the model. Examining Figure 17 shows that for a rack orientation of 45° the model starts initially with a d' = 0.818 at a threshold of 0.006 sec, but, improves to a d' = 1.022 with a peak intensity threshold of 1.306 sec. For the 30° rack orientation, the starting d' is similar equal to 0.161 at a threshold of 0.06 and improves to a d' of 0.570 at an intensity threshold of 1.369 sec. Examining Figure 18 you can see that the reason for the peak is due to the model reducing its false alarms to zero. For the 90° rack orientation, d' starts higher at 0.793 and improves to a d' of 1.316 at an intensity threshold of 1.025 sec. The 90° rack orientation is interesting because the model's best performance occurs at a much lower intensity threshold than the 30° and 45° orientations.

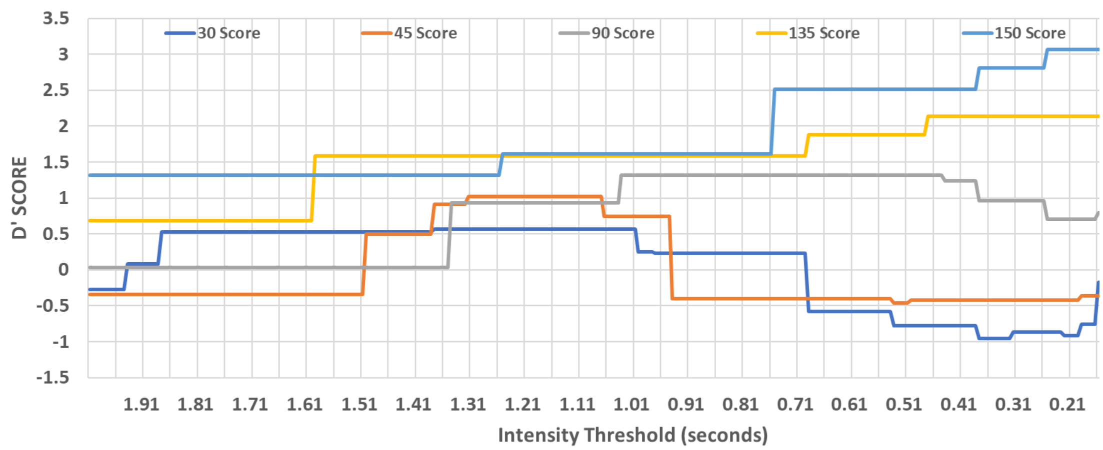

Figure 17. d' model scores as a function of intensity threshold for individual rack orientations on loglinear coordinates. Each score represents the average across observers and across the two runs per orientation completed by each observer. The x-axis is reversed to show the peak intensity threshold necessary.

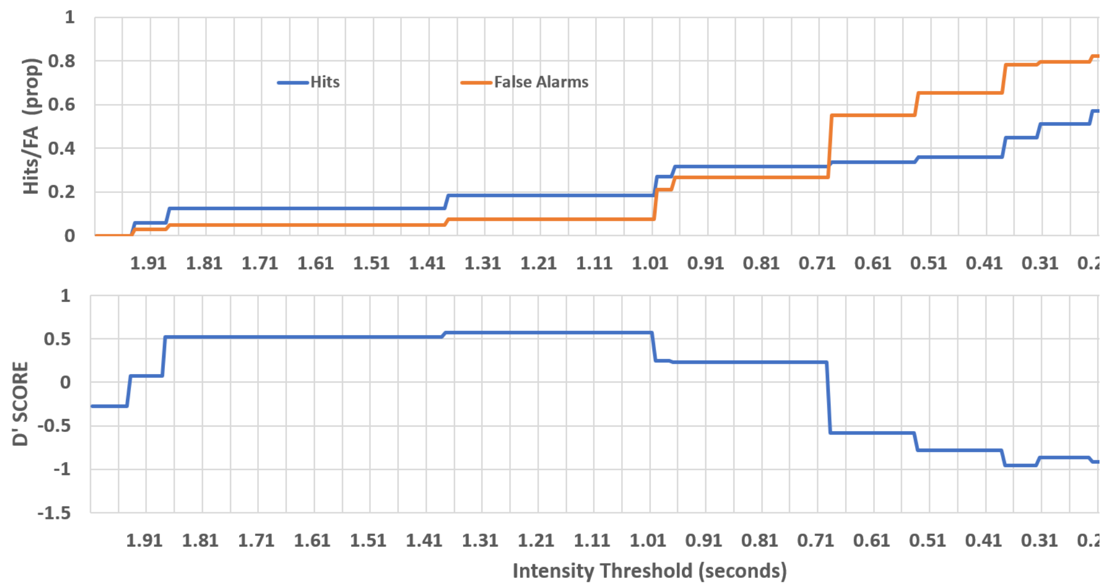

Figure 18. The top graph shows the average proportion of hits (TP) and false alarms (FP) for the 30° rack orientation as a function of intensity threshold. The bottom graph shows the d' score for the 30° rack orientation. In both graphs, the x-axis is reversed.

However, as seen in Figure 19, the highest d' score still corresponds to the point where the model quits making false alarms, just like in the 30° orientation condition (see Figure 18). This curve is more akin to the curves seen in the obtuse orientations, which spike earlier due to observers having more difficulty in seeing the back face and summarily ignoring it, a behavior mimicked by the model. The lower amount of back face target discrimination (see Figure 10) make sense with that increased difficulty. Model performance for the 135° and 150° rack orientations shows a dramatic difference from the 30°, 45°, and 90° rack orientations – it starts much higher and outperforms all orientations even at the highest intensity thresholds. For both the 135° and 150° rack orientation, the model slowly improves as intensity threshold is decreased, unlike the other orientations which peak and then generally decrease. Figure 20 shows this dramatic difference in the 150° rack orientation, where the false alarm errors drop off immediately at low intensity thresholds, and then model fit slowly degrades as intensity threshold is increased due to target exposures falling below the intensity threshold. Model performance for the 135° and 150° rack orientations reaches a maximum d’ of 2.156 (at an intensity threshold of 0.006 sec) and 3.062 (at an intensity threshold of 0.006 sec) respectively.

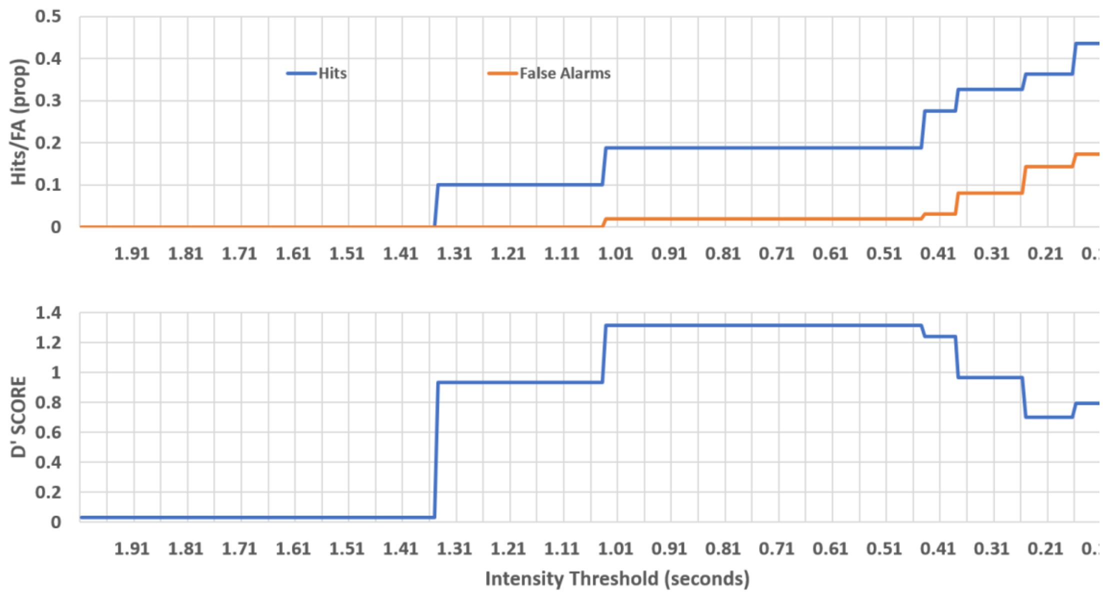

Figure 19. The top graph shows the average proportion of hits (TP) and false alarms (FP) for the 90° rack orientation as a function of intensity threshold. The bottom graph shows the d' score for the 90° rack orientation. In both graphs, the x-axis is reversed.

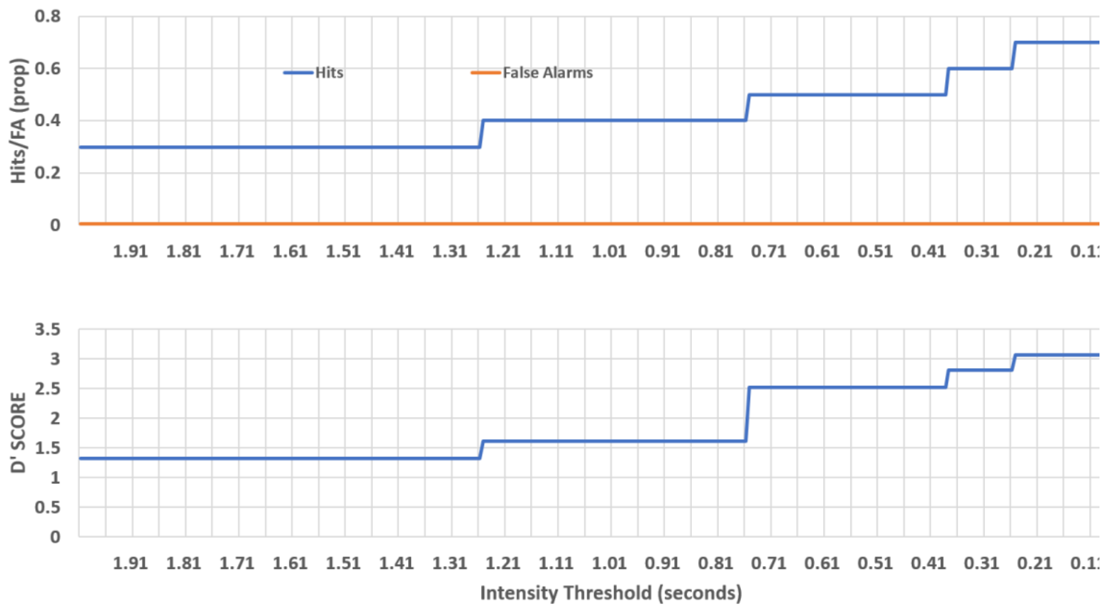

Figure 20. The top graph shows the average proportion of hits (TP) and false alarms (FP) for the 150° rack orientation as a function of intensity threshold. The bottom graph shows the d' score for the 150° rack orientation. In both graphs, the x-axis is reversed.

Initially, fitting the model using a d' method seemed to make sense and provided valuable insights into the model. The first insight is that the unique nature of the obtuse orientations, which nearly prevents participants from seeing the back face of the racks, enables the model to perform well. The second is that for the 30° and 45° rack orientation conditions, the model performed poorly at low intensity thresholds but improved as intensity threshold increased and both peaked at similar thresholds. Finally, amongst the 30°, 45°, and 90° orientations, all benefitted from utilizing a higher intensity threshold than that of the earlier target detection study, lending some evidence to the increased time necessary for discrimination as we hypothesized. However, the unusual curves of d’ as a function of intensity threshold led us to question their underlying cause.

To get a better understanding of the d’ model fits, we examined the underlying noise and signal+noise distributions. An ROC curve is a useful way to visualize this by plotting the hit and false alarm rates as intensity threshold changes. If the underlying distributions have a normal shape and equal variance, we should get a prototypical ROC curve like that seen in Figure 21. This would provide evidence that the usage of signal detection theory and d’ is appropriate.

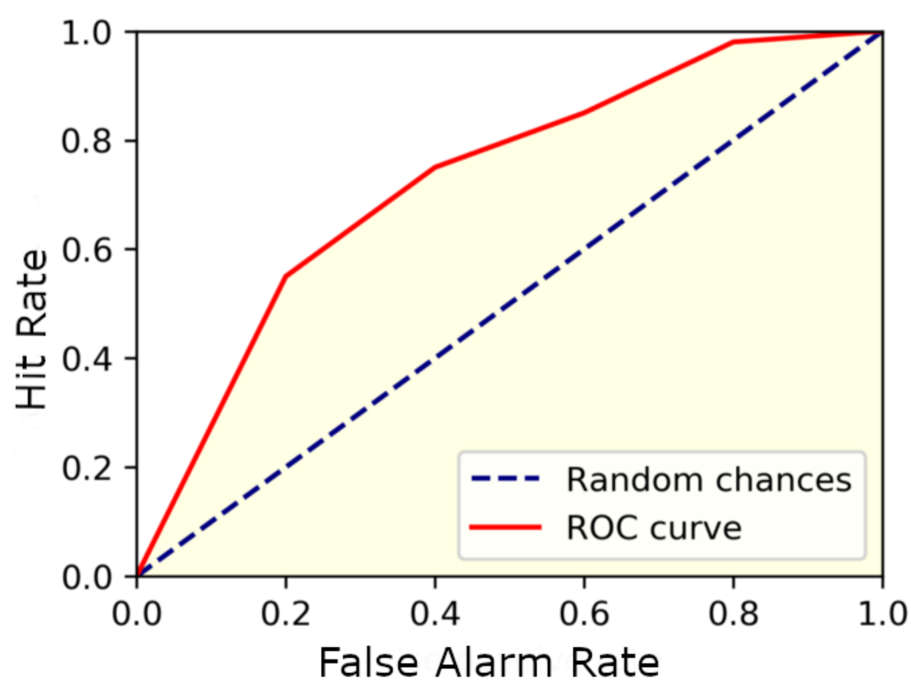

Figure 21. An example ROC curve. Random chance is represented by the dotted blue line whereas the solid red line represents a more standard ROC curve. adapted from Bui (2020).

Unfortunately, when plotting the hit and false alarm rate as intensity threshold increased, we do not see standard-looking ROC curves (see Figures 22-24). Examining Figure 22, as the intensity threshold (the criterion) changes d' also changes. This is inconsistent with signal detection theory. The fact that the underlying distributions of noise alone and signal+noise trials are not equivariance normal and d' changes as the criterion changes indicates that assessing model fit using our innovative d' method is not appropriate. Figure 23 shows the closest curve to a normal ROC curve providing some potential evidence for the d' method. Figure 24 however presents a very abnormal ROC curve which is a nearly vertical line that jumps to the far-right corner instead of showing any curve. Due to these abnormal ROC curves, we conclude that the d' analysis, while valuable in terms of understanding participant behavior, is not appropriate to evaluate the fit of the model to the data.

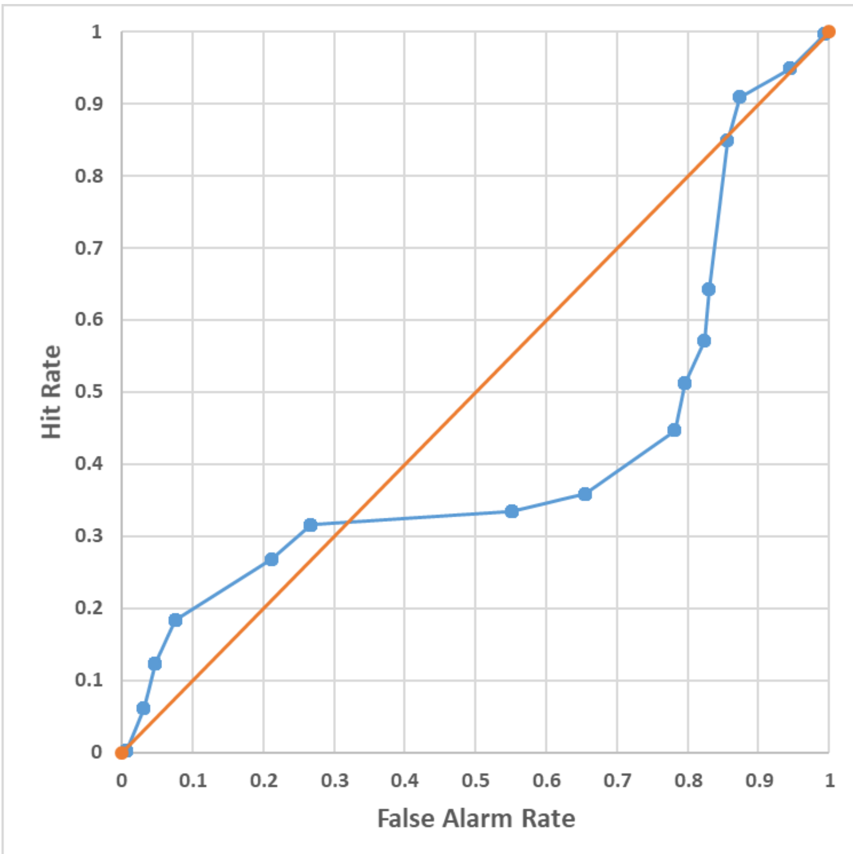

Figure 22. Receiver operating characteristic curve for the 30° rack orientation created by varying intensity threshold.

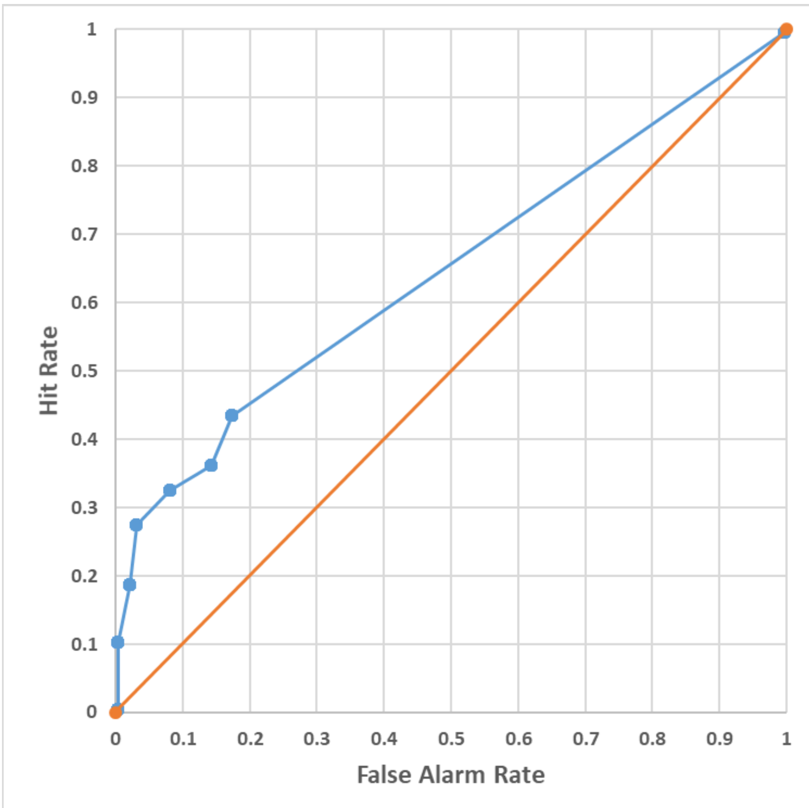

Figure 23. Receiver operating characteristic curve for the 90° rack orientation created by varying intensity threshold.

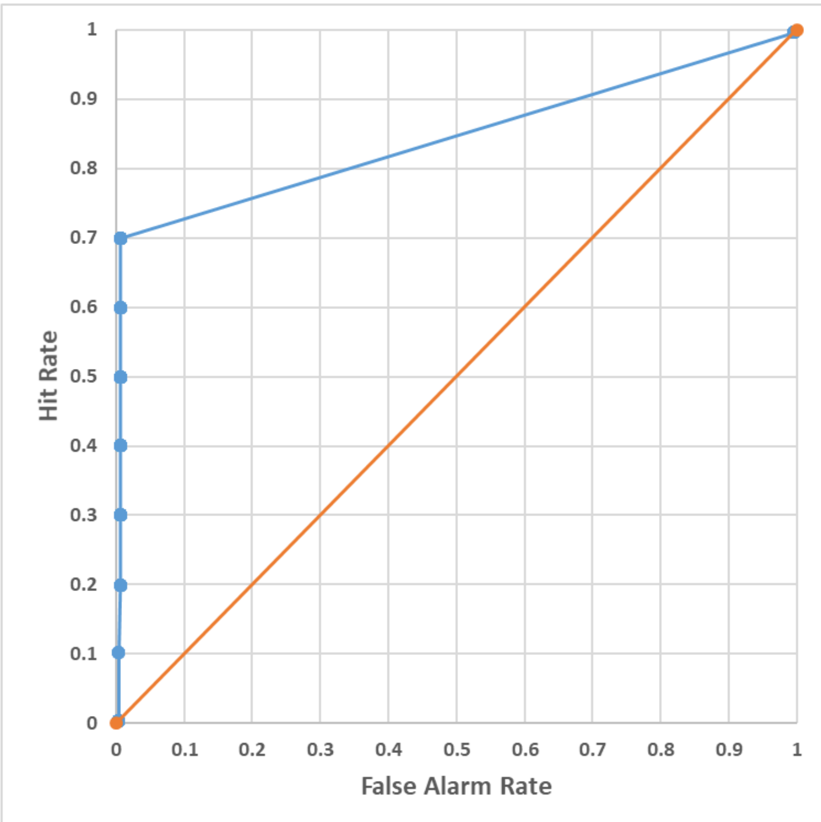

Figure 24. Receiver operating characteristic curve for the 150° rack orientation created by varying intensity threshold.

The next attempt to fit the model to the data was through a 2x2 chi-square analysis. Utilizing a 2x2 chi-square, the accuracy of the model in terms of matching human observer target hits and misses can be determined. In order to visualize how intensity plays a role, the chi-square value for intensity from 0–50 in steps of 0.5 is plotted in Figure 25. In the chi-square plots, a peak indicates a point of best model fit in terms of intensity. Overall, when looking at the curves for each orientation in Figure 25 we find mixed results in terms of the suitability of the chi-square method (finding a peak) for fitting the model to the data. Examining the 135° and 150° rack orientations in Figure 25, the chi-square values start at their highest score with an intensity threshold of 0 and steadily decrease. This is a similar pattern to the d' model graphs for the same rack orientations. Unfortunately, this behavior suggests that the chi-square analysis reaches a degenerate state for the 135° and 150° rack orientations as there is no peak. Examining the 30° rack orientation in Figure 25 we do find a brief peak at an intensity threshold of 5.5 seconds with a significant chi-square value ($\chi(1,1) = 47.60$, $p < .01$). The 45° rack orientation shows interesting behavior with its highest peak at the start following a drop and then a second peak-plateau at the 20.5 seconds. The starting peak at 0 is significant ($\chi(1,1) = 34.70$, $p < .01$) and the second peak-plateau at 20.5 is also significant ($\chi(1,1) = 19.88$, $p < .01$) This suggests some evidence for the suitability of the chi-square analysis fitting the model to the data, but finding two peaks in one function prevents more confidence. Finally, examining the 90° rack orientation a significant peak at 6.5 seconds ($\chi(1,1) = 57.58$, $p < .01$) is found. It is unclear why there is a difference in the intensity values where the model fit peaks for the 30°, 90°, and 45° rack orientations. It could be that targets appearing at different angles relative to the observer (since the targets appear flat on the rack face) vary in how easily they are discriminated. However, the current data cannot resolve this question. The pattern seen in the 90° rack orientation is interesting because it starts at a high value as seen in the 135° and 150° rack orientations but does improve later on. A likely reason for this behavior is again due to a difficulty of seeing the back rack faces between the ease of the acute orientations and the high difficulty in the obtuse orientations.

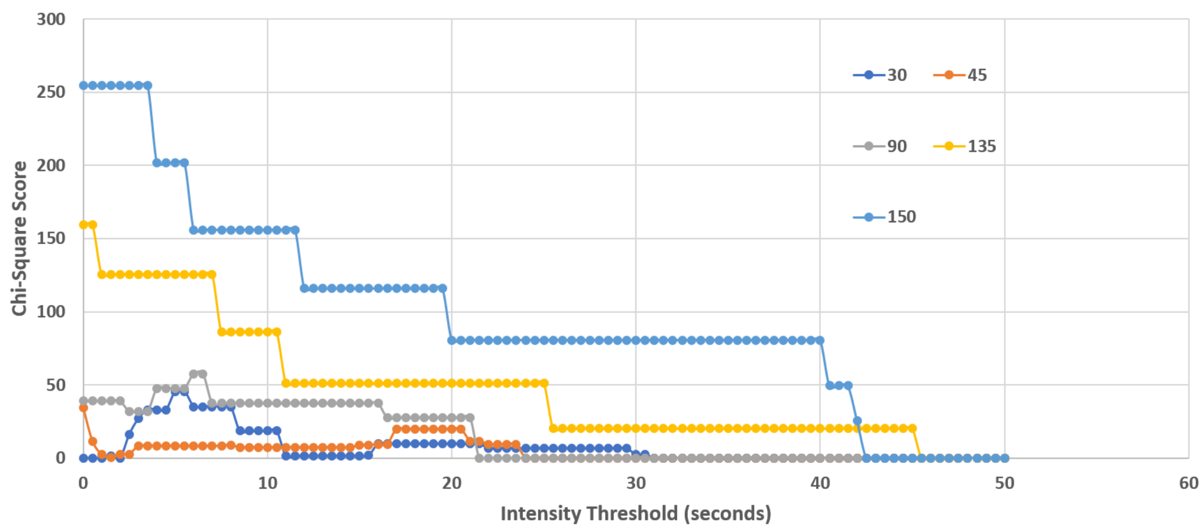

Figure 25. Chi-square statistic scores as a function of intensity thresholds for all rack orientations.

Overall, the data is appropriate for the 2x2 chi-square analysis and does not violate important assumptions making it a more suitable method to fit the model to the data than the d' analysis. While the pattern observed does suggest the analysis reaches a degenerate case for the obtuse orientations, observers adopting a search strategy that essentially ignores the back faces of racks for these rack orientations is likely the cause. The intensity thresholds for the 30° and 90° rack orientations for discrimination are higher than those measured for detection. This increased intensity for discrimination is also true for the 45° rack orientation if the initial peak in the chi-square function is disregarded.

Discussion

This paper presents a test of the model of target visibility (Parikh & Mowrey, 2014) using a discrimination task (target amongst distractors) more akin to a naturalistic shopping environment. In terms of target discrimination performance, the main hypothesis that rack orientation would have a significant effect on target performance was supported. Only one sub-hypothesis was not supported – that all rack orientations would produce better discrimination performance than the 90° rack orientation. The first sub-hypothesis that the 45° rack orientation would result in the best performance was supported. The second sub-hypothesis that performance in the 30° rack orientation would be better than all other rack orientations except 45° was also supported. The 30° and 45° rack orientations were found to produce significantly higher target discriminations than the 90° orientation. This gives retailers and others designing environments which have the flexibility to experiment with rack orientation (e.g., libraries) evidence to try one of these orientations. The obtuse rack orientations did not produce significantly different target identifications than the 90° orientation. The difficulty of seeing targets on the back faces lead participants to adopt a different search strategy which gave a low prioritization to targets on the back face. Participants in the other rack orientations were able to identify targets on the back face so they did not de-prioritize it as highly. The 30° and 45° orientations had a similar ratio of target hits on the front and back rack face. The obtuse orientations would see improved performance in a bi-directional study where the back face would become a front face when the environment is traversed from the opposite direction. This would mean that for the obtuse rack orientations, both rack faces down any aisle would have high intensity values. However, it is important to note that to see high performance of 80-100% you would need to treat these forward and reverse trials as either one combined trial or have these trials occur simultaneously by allowing enough time for the observer to turn around completely rather than just turn their head.

Now that the human behavior has been summarized, the analysis of how well the model fit the human behavior can be discussed. Overall, the model had mixed results when extending it to the present discrimination task. Interestingly, finding the appropriate way of fitting the model ended up being a point of interest. Initially, utilizing signal detection theory and d’ seemed a useful method of evaluating model fit. Using this analysis, the model performed well for the 135° and 150° orientation conditions, but poorly for the 30°, 45°, and 90° orientation conditions. This was due to the model (and observers) ignoring the back rack faces in the 135° and 150° rack orientations and thus making no false alarms, a behavior not seen with the other orientations. Model fit was always best whenever the model quit making false alarms, which in the case of the 30°, 45°, and 90° rack orientations were at higher intensity thresholds. When examining the underlying distributions of signal and signal+noise trials, it was observed that the distributions were not Gaussian in shape nor were they of equal variance. This violation of signal detection theory assumptions made computing d’ an inappropriate way to determine model fit. Utilizing a 2x2 chi-square approach across intensity thresholds, a more consistent pattern was found. The model provided the best fit to the data for the

135° and 150° rack orientations at the minimum intensity, but the model fit improved for the 30°, 45°, and 90° rack orientations at larger intensity values. Since the data did not violate the underlying assumptions necessary for the chi-square analysis, we have more confidence in these results.

This led to mixed results for our initial hypothesis of increased processing time for a person to be able to discriminate a target amongst distractors instead of simple detection. However, the 30°, 45°, and 90° rack orientation conditions did see large improvements in model fit with increased processing time (as seen through higher intensity thresholds). One reason for the mixed results is the difference in complexity for the model between the 135° and 150° conditions and the others due to a shift in search strategy. This shift in search strategy consisted of observers giving a low priority to scanning the back-rack face due to the difficulty of perceiving it versus the ease of examining the front rack face. Participants likely thought the increased difficulty in examining the back face was not worth the risk of missing a target on the front face. Since participants did not have the same search strategy between conditions (or at least had a more equal search priority between rack faces for the non-obtuse rack orientations) we see this difference in model fit. Without that constraint in the 30/45/90 rack orientations, we see a poorer model fit overall. Examining Figure 25, the chi-square statistic decreases and then increases at higher intensity values. One way this would be possible is that participants had a different scanning pattern than the model does allowing those lower intensity targets to be discriminated more often than the model expects.

People tend to adapt their behavior to the task environment and task demands, but since the model has no aspect of learning or adapting, it cannot make the same changes.

The obtuse orientations highlight some limitations of the model particularly well. Performance in these orientations is poor (only 50%), but the chi-square model fit is good due to targets on the back-rack face never being predicted to be seen. This is dissimilar to the other orientations since it means the model starts off with an inherent advantage of being correct for half of the targets no matter the intensity threshold. The other rack orientations see improvement by increasing the intensity threshold. This removes targets that are only present in the field of regard for a brief period of time and which are likely not to be discriminated correctly by the human observers. However we do not see this same pattern of improvement in the obtuse orientations. Targets on the front rack face are in the field of regard for a long period of time, meaning chi-square model fit only slowly decreases over time as the high intensity targets are eventually not considered exposed.

One important overall point about the model is that it is a model of rack location visibility being applied to a discrimination task. The model itself can be thought of as a video camera with a specific field of view constantly rotating left/right at a consistent pace (phi) while being moved down a central aisle. Akin to a video camera, all points in this field of view are considered equally visible. For a computer or surveillance camera, this idea of equivalence is fine because the video can be examined post-hoc in a pixel-by-pixel and frame-by-frame manner, but this is not representative of vision in people.

People do not equally see all points in the visual field – the fovea (approximately the central 2°, which is smaller than the model's field of view) provides the most detail and acuity decreases drastically from there to the far periphery (Fairchild, 2013). In addition, attention plays a significant role such that what and where people are attending tends to determine what people see. If this was not the case, issues of not perceiving significant events such as a deer running onto the road, or unusual events such as a gorilla walking through a basketball court during a game (Simons & Chabris, 1999) would not occur. This difference in representation of vision between a person and the model is an important distinction because it shows that people are different from the model. Since the model does not have an attentional component it will continually maintain its scanning pattern and all locations in it field of view are equally visible. Whereas a person could see a target and have their attention shifted towards it, disrupting and or shifting their scanning pattern as well as making all other locations in their field of view temporarily less visible.

Another important difference between the model and human observers is that the model is not doing a task, it simply determines what locations on a rack are visible and for how long. This means that for the model, if a target falls within its visual field for the pre-selected intensity the target is predicted to be seen (detected or discriminated). Whereas humans can make mistakes due to misperceiving or misreporting, the model makes no such errors. If these types of errors happen only rarely, like misreporting the character identity even though it was clearly perceived, this would not be too problematic as it would simply add some small amount of measurement error to the data. However, there are many things that impair human visual perception. One is crowding. Crowding occurs when a target in the periphery appears close to other similar objects (distractors). When the spacing between the target and distractors is less than about one-half of the eccentricity of the target, discrimination of the target is reduced to near chance (Bouma, 1970). The model is unaffected by such crowding effects and thus it makes no errors based on them.

A final difference between the model and human behavior is the search or scanning pattern. The model assumes a smooth and continuous left-right scanning pattern, which is not akin to human behavior. Human behavior involves attention, which can be grabbed and shift/disrupt their scanning pattern (or lead to a complete change) which cannot happen in the model. This is highlighted well in the chi-square plots where for some rack orientation conditions there is no smooth curve up to a peak and then a decrease to zero. A relationship between intensity (how long a target is seen) and discrimination is theoretically reasonable in that targets in view for longer should have a higher percentage chance of being discriminated. The lack of a clear rise to peak and then decline in the chi-square results suggests participants are utilizing different scanning patterns depending upon the condition. For example, if a person detects or identifies a target on a shelf before fully turning their head, they would not continue scanning that rack (knowing that there is only one target per rack face), but instead move to examining another rack. This is an adaptation that people make naturally but is one that the model in its current form cannot make as it always makes complete scans to the left and right at the set head turn magnitude.

These results lead to the important question of "what is the best store layout?" as it will inevitably be asked by retailers and designers. One piece of this question involves the target performance while the other piece is the user experience of these orientations. The evidence for the target performance points towards using the 30°/45° rack orientations as they produced the highest discrimination performance. The user experience can relate to the idea of effort in finding and identifying targets on each rack orientation, but also includes how successful (aka the number of target discriminations) the user was in achieving their goal. If you go into a store but only can find a few of the items on your shopping list, even if you did not work hard to find them, the user experience would be poor. Although outside the scope of this thesis, I considered that one may try to better quantify user experience in the present virtual shopping conditions by determining an effort-performance tradeoff measure. The effort portion would be a combination of three variables: head turn count, average head turn magnitude, and average angular velocity of head turns. The performance would be the proportion of targets discriminated. While such a composite user experience measure could be helpful and reduce the data to a single numerical value, arriving at a formulation that generalizes across environments and tasks would be challenging.

This user experience measure based on an effort-performance tradeoff would be of the most value to retailers and others interested in implementing different rack orientations due to the importance placed on performance, which is necessary for sales. Recently, there has been an increase in options such as grocery delivery or curbside pickup. Grocery delivery and curbside pickup can offer advantages over traditional shopping such as shorter time (convenience) and obtaining all the items on your list. While there can still be frustrations with these options (not receiving item(s) or receiving substitutions), the increased convenience brought about by these options points towards a needed emphasis on the user experience in stores. This new emphasis on the user experience in stores is important to keep people shopping in person. The best method to create this user experience measure is not explicitly clear and is ripe for exploration in future work. For example, the weighting of the different components of the effort portion could be done in several ways. One line of reasoning would be to more heavily weight larger and faster head turns as these may intuitively be thought of as more strenuous or effortful. However, a fast head turn may be performed so quickly that a person may not experience them as more effortful. The idea of effort in a head turn is also not well conceptualized – after all, turning the head is naturally done many times throughout the day and often reflexively as part of the orienting response. A well conceptualized effort variable will likely require input from ergonomics and additional experiments to examine this relationship between head turns and effort. There is a possibility that the suggested effort-performance tradeoff may look like the well-studied speed-accuracy tradeoff which could present an interesting parallel.

Conclusions

Overall, the results in this study found mixed viability for the Parikh and Mowrey (2014) model of rack exposure to account for target discrimination in a virtual shopping environment. For all rack orientations except 135° and 150°, intensity thresholds were longer than those needed for a target detection task to achieve the best model fits. Additionally, the 30° and 45° orientations performed the best in terms of target discriminations while significantly outperforming the 90° orientation. These results provide evidence for the viability of the 30° and 45° rack orientations as competitors to the traditional 90° rack orientation. Retailers and other environments such as intensive care units should consider utilizing 30° and 45° orientations of item such as racks or counters to improve visibility of important items such as passengers or products.

An important conclusion from the study is the finding that people align and adjust their behavior to the constraints imposed upon them by the task and environment. Participants in the obtuse orientations had a very minimal priority on identifying targets on the back rack face due to the inherent difficulty in seeing that back rack face. Participants in the 30° and 45° rack orientations more equally prioritized the front and back rack faces resulting in a more equal distribution of those target discriminations. Interestingly, the 90° rack orientation condition presented a middle point where observers seemed to give the back rack face a lower priority, but not to the extremes as seen in the 135° and 150° orientation conditions. Another pattern of behavior found is the relationship between the size (magnitude) of head-turns and their speed and frequency.

When people were making larger head turns (such as in the 90° and 135° rack orientation conditions), these head turns were faster and fewer overall. If the head turns were smaller such as in the 30° and 150° rack orientation conditions, they tended to be slower but there were also more of them. There is clearly a relationship between human scanning pattern and the constraints imposed upon them by the rack orientation.

Finally, the interesting patterns found in the chi-square model fits taken along with the relationships between patterns of scanning behavior suggest the scanning pattern of the model does not represent human behavior well. The smooth scanning pattern represented in the model does not have the adaptability that people embody. This makes it likely that the intensity of a location for some people will not be the same as in the model. In order for the model to better fit human data, variables such as attention, crowding, and a variable scanning pattern should be considered for implementation.

References

Anderson, E. E. (1979). An analysis of retail display space: theory and methods. *Journal of Business*, 103-118.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178.

Bui, H. (2020). ROC Curve Transforms the Way We Look at a Classification Problem. Retrieved August 20, 2020, from https://towardsdatascience.com/a-simple-explanation-of-the-roc-curve-and-auc-64db32d75541

Cairns, J. P. (1962). Suppliers, retailers, and shelf space. *The Journal of Marketing*, *26*, 34-36.

Drèze, X., Hoch, S. J., & Purk, M. E. (1994). Shelf management and space elasticity. *Journal of Retailing, 70*(4), 301-326. doi:10.1016/0022-4359(94)90002-7

Ebster, C., & Garaus, M. (2015). *Store design and visual merchandising: Creating store space that encourages buying*. New York, NY: Business Expert Press.

Fairchild, M. D. (2013). *Color appearance models*. John Wiley & Sons.

Haahr, M. (n.d.). True Random Number Service. Retrieved from https://www.random.org/

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator

Sickness. *International Journal of Aviation Psychology*, *3*(3), 203. https://doi-org.ezproxy.libraries.wright.edu/10.1207/s15327108ijap0303_3

Marsden, J., Stevens, S., & Ebri, A. (2014). How to measure distance visual acuity. *Community eye health*, *27*(85), 16.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

Mowrey, C. H. (2016). *Retail Facility Design Considering Product Exposure* (Doctoral dissertation). Retrieved from OhioLink. (wright1472487958)

Parikh, P. J., & Mowrey, C. (2014). Designing retail facilities to improve product exposure. *Progress in Material Handling Research*. Charlotte, NC: Material Handling Institute.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28*, 1059-1074.

Wischgoll, T. (2005, August 26). *Advanced Visual Data Analysis*. Retrieved from http://avida.cs.wright.edu/appenzeller.html

Appendix