Feature-location binding in 3D: Feature judgments are biased by 2D location but not position-in-depth

Nonie J. Finlayson *, Julie D. Golomb

Department of Psychology, Center for Cognitive & Brain Sciences, The Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 10 February 2016
Received in revised form 1 July 2016
Accepted 5 July 2016

Keywords:
Object perception
Depth perception
3D space
Spatial congruency bias
Object-location binding
Feature binding

ABSTRACT

A fundamental aspect of human visual perception is the ability to recognize and locate objects in the environment. Importantly, our environment is predominantly three-dimensional (3D), but while there is considerable research exploring the binding of object features and location, it is unknown how depth information interacts with features in the object binding process. A recent paradigm called the spatial congruency bias demonstrated that 2D location is fundamentally bound to object features, such that irrelevant location information biases judgments of object features, but irrelevant feature information does not bias judgments of location or other features. Here, using the spatial congruency bias paradigm, we asked whether depth is processed as another type of location, or more like other features. We initially found that depth cued by binocular disparity biased judgments of object color. However, this result seemed to be driven more by the disparity differences than the depth percept: Depth cued by occlusion and size did not bias color judgments, whereas vertical disparity information (with no depth percept) did bias color judgments. Our results suggest that despite the 3D nature of our visual environment, only 2D location information – not position-in-depth – seems to be automatically bound to object features, with depth information processed more similarly to other features than to 2D location.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

A fundamental aspect of human visual perception is the ability to recognize and locate objects in the environment. Locating and recognizing are often considered separate processes, but the combination of this information is critical for interacting with objects. This idea that different object features or properties need to be integrated has been coined the “binding problem” (Treisman & Gelade, 1980), and can refer to the binding of different object features (e.g. the sun is round and yellow), or the binding of object features to their locations, which might involve different neural mechanisms (Piekema, Rijpkema, Fernández, & Kessels, 2010). Importantly for the latter, in the real world we need to locate objects in a 3D environment. While there is considerable research examining the process of binding object location and features to perceive a coherent object, this has primarily focused on 2D location, with very little understood about how 3D location and features interact for object perception.

In particular, 2D location has often been considered to play a special role in visual perception, above and beyond that seen for other object features (Cave & Pashler, 1995; Chen, 2009; Chen & Wyble, 2015; Golomb, Kupitz, & Thiemann, 2014; Moore, Lanagan-Leitzel, Chen, Haltermann, & Fine, 2007; Pertsov & Husain, 2014; Treisman & Gelade, 1980; Tsal & Lavie, 1988, 1993). However, little research has explored 3D location, specifically position-in-depth, and it is unknown whether the special role seen for 2D location information also extends to depth location.

A recent example of the special role of 2D location information in object perception comes in the form of the spatial congruency bias (Golomb et al., 2014), where two objects are more likely to be judged as the same identity if they appear in the same spatial location. Golomb et al., 2014 showed that despite location information being irrelevant to the task, participants were automatically biased to judge features of two objects as more similar when sharing the same location. Location information has been shown to bias a variety of features, including Gabors, colors, shapes, and faces (Golomb et al., 2014; Shafer-Skelton, Kupitz, & Golomb, Submitted). Moreover, the spatial congruency bias seems to be particular to location, in that identity information does not bias location judgments, nor do features such as color and shape induce a bias of each other (Golomb et al., 2014). This finding of location biasing features is consistent with past research positing a unique role for location during object processing, including the feature integration theory (Treisman & Gelade, 1980), which proposed that spatial attention is required to bind features into a coherent object, as well as more recent work (e.g. Cave & Pashler, 1995; Chen &
Wyble, 2015; Chen, 2009; Moore et al., 2007; Pertzov & Husain, 2014; Tsal & Lavie, 1988, 1993). For example, Tsal and Lavie (1993) found that when instructed to report one of two targets based on the color of a cue, participants were unable to ignore the location of the cue, even though it was irrelevant and detrimental to performance, supporting a unique role of location.

How does depth interact with object features? In particular, does position-in-depth information bias feature judgments in the same way as 2D location information does? A recent study from our lab (Finlayson & Golomb, Under Review) explored the interactions between 2D and depth locations using the spatial congruency effect, finding that 2D locations bias depth judgments, but the reverse is not true: depth information does not bias 2D location judgments. However, while depth may not bias 2D location judgments, it is unknown whether depth biases judgments of other object features. In other words, is depth processed like other features (in which case it should not induce a spatial congruency bias), or is it processed more like another type of location (in which case we would expect depth to bias these other feature judgments)?

A number of studies have demonstrated similarities between 2D and depth effects, such as response priming that is seen for 2D locations (Posner, 1980) also seen for depth (Attchley, Kramer, Andersen, & Theeuwes, 1997; Downing & Pinker, 1985; Finlayson, Remington, Retell, & Grove, 2013; Nakayama & Silverman, 1986), and other findings showing depth as advantageous for object recognition (Caziot & Backus, 2015). On the other hand, several studies have suggested that although depth may play an important role in the visual system, the perceptual and attention effects of depth are weaker or delayed compared to those effects seen for 2D space (Finlayson et al., 2013; Gilinsky, 1951; Kasai, Morotomi, Katayama, & Kumada, 2003; Loomis et al., 2008; Moore, Hein, Grosjean, & Rinkena, 2009). In addition, although neurophysiological and functional neuroimaging research have demonstrated that depth and binocular disparity information is encoded by neurons in much of visual cortex (Backus, Fleet, Parker, & Heeger, 2001; Ban, Preston, Meeson, & Welchman, 2012; Durand, Peeters, Norman, Todd, & Orban, 2009; Neri, Bridge, & Heeger, 2004; Preston, Li, Kourtzi, & Welchman, 2008; Tsao et al., 2003; Welchman, Deubelius, Conrad, Bülthoff, & Kourtzi, 2005), the earliest representations may not be linked with the percept of depth (Preston et al., 2008), and true position-in-depth information may not emerge until later in the visual processing stream compared to 2D location information (Barendregt, Harvey, Rokers, & Dumoulin, 2015; Finlayson & Golomb, Under Review).

Here we explored the interaction between position-in-depth and feature perception using the spatial congruency bias paradigm to ask if depth locations bias feature judgments (object color). We first conducted an experiment testing the influence of irrelevant depth location information on color judgments using binocular disparity to cue depth perception (Experiment 1). We then followed up with two experiments to further probe the role of depth information. First, we re-tested the effects of depth on color judgments using monocular depth cues (occlusion and size: Experiment 2). Second, we dissociated the effects of depth-from-disparity information from pure disparity (eye-specific location) information by testing if vertical disparity, which does not create a depth percept, induces a spatial congruency bias (Experiment 3).

2. Experiment 1

2.1. Method

2.1.1. Participants

Seventeen subjects (9 female; mean age = 19 years; range: 18–21) participated. All participants reported normal or corrected-to-normal color and binocular vision, and were screened for normal stereovision. Informed consent was obtained for all participants, the Ohio State University Behavioral and Social Sciences Institutional Review Board approved the study protocols, and the research was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). All participants were compensated with course credit. Sample size was chosen to match the original spatial congruency experiment reported in Golomb et al. (2014), which had a Cohen’s $d = 1.01$ and statistical power ($1 - \beta$) of 0.96 with $N = 16$; one extra participant was run in Experiment 1 due to over-scheduling. According to criteria set in advance, participants who performed the task with <55% accuracy were excluded from analyses; however, no participants needed to be excluded.

2.1.2. Stimuli

Stimuli were generated with the Psychtoolbox extension (Brainard, 1997) for MATLAB (MathWorks). Depth from binocular disparity was achieved using a Wheatstone stereoscope, with two 24° flat screen LCD monitors facing each other with mirrors set between and reflecting an image from each monitor to each eye of the observer. The viewing distance was 60 cm, with the observer sitting at a chinrest 90° to the monitors. The monitors were color calibrated with a Minolta CS-100 colorimeter.

Stimuli were colored squares on a black background, with size depending on depth location. For the back (far) disparity, stimuli were sized 0.71° × 0.71°, and at the front (close) disparity, stimuli were sized 0.99° × 0.99°. Subjects fixated at the center of the screen on a small 0.27° dot, always presented at the central screen depth (zero disparity). Stimuli were presented peripherally and could vary in horizontal, vertical, and depth location.

2.1.3. Procedure & design

Participants began each trial by fixating in the center for 500 ms, after which the first stimulus appeared in a peripheral location for 250 ms (Fig. 1). This was followed by a blank screen (50 ms) and then a mask (100 ms). Following a random delay period of either 550 ms or 1000 ms, a second stimulus appeared. The second stimulus was presented for the same duration and masked as the first. The first stimulus color was chosen randomly from 180 colors along an isoluminant color wheel (evenly distributed along a circle in CIE $L^*a^*b^*$ color space, centered on $L = 70, a = 20, b = 38$, radius = 60). When the second stimulus differed in color (50% of

Fig. 1. Schematic illustration of the task and stimuli locations for Experiment 1, where the task was to indicate whether the two objects were the same or different colors, while ignoring the irrelevant location information (vertical and depth position). The difference in color between the two objects was subtle (adjusted individually to 75% accuracy threshold). Inset shows sample stimuli for Experiment 2. Experiment 3 was identical to Experiment 1, but the stimuli varied in horizontal position, and vertical disparity was used instead of horizontal disparity (depth).
trials), it was chosen as a small color difference in either direction on the color wheel. The magnitude of the color difference was staircased for each participant during the practice block to converge on 75% accuracy, and was further adjusted between main blocks as necessary (e.g. if >85% or <65% accuracy). The average difference between the two stimuli was 11.5 points on the color wheel.

Vertical and depth locations of the first stimulus were randomly assigned for each trial, from the following possibilities: vertical location 5.46° above or below fixation, depth position 30 arcmin (0.5°) in front of or behind fixation. Horizontal location was always centered on the screen, in line with fixation. The second stimulus appeared equally likely in one of four locations relative to the first stimulus: same or different depth location by same or different vertical location. These four conditions were counterbalanced and equally likely.

Participants were instructed to judge whether the two colors were the same, and location was irrelevant to the task. Participants responded by keyboard press and were presented with visual feedback (green or red dot) informing them whether or not their response was correct (500 ms). They were also provided with feedback if they broke fixation at any point during the trial, and the trial was aborted and re-run in the block. Participants completed 64 trials per block, comprising 16 trials per each of the four irrelevant-location conditions, in randomized order. Each participant completed one practice block and 5–7 main blocks (two participants were unable to complete the full 7 blocks in time due to slower response times).

Eye position was monitored with an EyeLink 1000 eye-tracking system, recording monocular pupil and corneal reflection position. Fixation was monitored for all experiments. If at any point the participant’s fixation deviated from the central fixation point by greater than 1.5°, the trial was aborted and repeated.

2.1.4. Analyses

Our primary measure for all experiments was the Spatial Congruency Bias (Golomb et al., 2014). For each participant, we first calculated hit and false alarm rates for each location condition. We defined a “hit” as a “same color” response when the stimuli actually were the same color, and a “false alarm” as a “same color” response when the stimuli were actually different colors. We treated this as analogous to a “yes-no” task (Green & Swets, 1966; Macmillan & Creelman, 2004), where participants judged whether the second color was the same as the first reference color. Using the hit rate and false alarm rate, we used signal detection theory to calculate bias (criterion) for each location condition.

For all experiments we focus on the bias measure because our main goal was to assess the spatial congruency bias (Golomb et al., 2014) for position-in-depth. However, as secondary analyses we also report reaction time and d-prime measures to assess whether position-in-depth also results in response facilitation. Values for each of these measures, as well as raw proportion of “same” responses, and alternate ways of calculating bias (normalized c and likelihood ratio β), can be found in Table 1.

Bias (criterion) = \(-z(\text{hit rate}) + z(\text{false alarm rate})\)/2

d' = z(\text{hit rate}) - z(\text{false alarm rate})

Normalized c = bias/d

Likelihood ratio (β) = \(e^{c/2} - \text{hit rate}^2 + \text{false alarm rate}^2\)/2

Values for all measures were averaged separately for each participant and condition and submitted to repeated-measures ANOVAs, with effect size calculated with partial eta squared. Trials on which participants failed to respond, or responded with RTs greater than 2.5 standard deviations of the participant’s mean RT, were excluded.

2.2. Results

2.2.1. Congruency bias

Fig. 2a illustrates the bias of responding “same color” as a function of the different irrelevant location conditions: a negative bias indicates a greater tendency to respond “same color”. We found that irrelevant 2D and depth location information biased color judgments, such that when the two objects were in the same 2D or depth location, participants were more likely to report that the objects were the same color. A two-way repeated-measures

<table>
<thead>
<tr>
<th>Expt</th>
<th>Location</th>
<th>Color</th>
<th>p(“Same”)</th>
<th>RT (ms)</th>
<th>Bias (c)</th>
<th>d-prime</th>
<th>Normalized c</th>
<th>Likelihood ratio (β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Same Y Same Z</td>
<td>Same</td>
<td>0.91</td>
<td>799</td>
<td>-0.47</td>
<td>1.92</td>
<td>-0.24</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.31</td>
<td>824</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Same Y Diff Z</td>
<td>Same</td>
<td>0.86</td>
<td>856</td>
<td>-0.31</td>
<td>1.74</td>
<td>-0.17</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.29</td>
<td>852</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diff Y Same Z</td>
<td>Same</td>
<td>0.85</td>
<td>826</td>
<td>-0.27</td>
<td>1.57</td>
<td>-0.19</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.31</td>
<td>842</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diff Y Diff Z</td>
<td>Same</td>
<td>0.85</td>
<td>846</td>
<td>-0.29</td>
<td>1.60</td>
<td>-0.18</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.31</td>
<td>855</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Same XY Same Z</td>
<td>Same</td>
<td>0.94</td>
<td>745</td>
<td>-0.44</td>
<td>2.37</td>
<td>-0.21</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.25</td>
<td>768</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Same XY Diff Z</td>
<td>Same</td>
<td>0.92</td>
<td>778</td>
<td>-0.55</td>
<td>2.03</td>
<td>-0.28</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.33</td>
<td>797</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diff XY Same Z</td>
<td>Same</td>
<td>0.85</td>
<td>791</td>
<td>-0.33</td>
<td>1.58</td>
<td>-0.23</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.33</td>
<td>791</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diff XY Diff Z</td>
<td>Same</td>
<td>0.86</td>
<td>783</td>
<td>-0.30</td>
<td>1.72</td>
<td>-0.21</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.32</td>
<td>819</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Same X Same Y</td>
<td>Same</td>
<td>0.89</td>
<td>664</td>
<td>-0.45</td>
<td>1.81</td>
<td>-0.24</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.33</td>
<td>722</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Same X Diff Y</td>
<td>Same</td>
<td>0.86</td>
<td>691</td>
<td>-0.24</td>
<td>1.91</td>
<td>-0.13</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.26</td>
<td>723</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diff X Same Y</td>
<td>Same</td>
<td>0.83</td>
<td>683</td>
<td>-0.20</td>
<td>1.56</td>
<td>-0.13</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.29</td>
<td>718</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diff X Diff Y</td>
<td>Same</td>
<td>0.78</td>
<td>683</td>
<td>-0.12</td>
<td>1.39</td>
<td>-0.09</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diff</td>
<td>0.29</td>
<td>711</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ANOVA with factors vertical location (same/different) and depth location (same/different) revealed that although in the right direction, neither of the main effects reached significance (Y: $F_{1,16} = 3.37$, $p = 0.085$, $\eta^2_p = 0.17$; Z: $F_{1,16} = 4.38$, $p = 0.053$, $\eta^2_p = 0.22$ respectively), but there was a significant two-way interaction ($F_{1,16} = 4.95$, $p = 0.041$, $\eta^2_p = 0.24$). Follow-up t-tests showed a significant bias for same compared to different depth location when vertical location was held the same ($t_{16} = -2.59$, $p = 0.020$, $d = 0.73$), and a significant bias for same compared to different vertical location when depth was held the same ($t_{16} = -2.42$, $p = 0.028$, $d = 0.54$).

2.2.2. Other effects

As noted above, our primary measure of interest was the congruency bias. However, other measures are listed in Table 1. There was a significant influence of vertical location on $d'$ ($F_{1,16} = 7.82$, $p = 0.013$, $\eta^2_p = 0.33$), with no main effect for depth location ($F_{1,16} = 1.28$, $p = 0.274$, $\eta^2_p = 0.07$), and no interaction ($F_{1,16} = 2.51$, $p = 0.133$, $\eta^2_p = 0.14$). RT priming was significant for depth location ($F_{1,16} = 11.66$, $p = 0.004$, $\eta^2_p = 0.42$) but not for vertical location ($F_{1,16} = 1.48$, $p = 0.242$, $\eta^2_p = 0.09$), with no interaction ($F_{1,16} = 2.33$, $p = 0.146$, $\eta^2_p = 0.13$).

2.3. Discussion

The 2D spatial congruency effect was primarily replicated: when depth location was the same, there was a bias for same versus different vertical location. Our main question was whether irrelevant depth location information also induced a spatial congruency bias, such that stimulus colors were judged as more similar when they appeared at the same depth. We found a spatial congruency bias for depth location, although this interacted with vertical location, such that we only saw a significant biasing of color judgments when both vertical and depth locations were the same.

However, there is an important potential confound here: In this experiment binocular (horizontal) disparity was used to cue depth perception. Binocular disparity is achieved by small horizontal differences in opposite directions for each eye, and 2D location information is known to induce a strong spatial congruency bias, even for very small differences in location. Thus, it is unclear if the bias we found here for “depth” location was truly due to the difference in perceived depth position, or if may have actually been due to these small 2D location differences for each eye. In Experiments 2 and 3, we attempt to dissociate these possibilities in two ways. In Experiment 2, we test whether the bias for depth-from-disparity generalizes to other, non-disparity depth cues, in particular occlusion and size. In Experiment 3, we then test whether disparity differences alone induce a bias, by varying vertical disparity between the eyes (which does not create a depth percept).

3. Experiment 2

3.1. Method

3.1.1. Participants

Sixteen subjects (9 female; mean age = 22 years; range: 19–36) participated. None were excluded.

3.1.2. Stimuli, procedure & design

Stimuli were similar to Experiment 1, with the following exceptions: The colored squares were presented on a mid-gray background (40% contrast), with size either $0.90^\circ \times 0.90^\circ$ (back) or $1.1^\circ \times 1.1^\circ$ (front). An $8^\times 8^\prime$ square filled with random noise (light and dark gray colored pixels: 24% and 55% maximum luminance of the display, respectively) was always present on the screen, centered on fixation at the central screen depth (see inset in Fig. 1). Stimuli were presented to appear either in front of (front) or behind (back) this square. Front versus back depth differences were cued both by size differences (larger for front) and occlusion cues; “front” stimuli were presented such that they occluded part of the square, while “back” stimuli were partially occluded by the square (one-quarter of the stimulus). All depth cues were monocular in this experiment; there were no disparity differences between stimuli.

Stimuli were presented peripherally in one of eight locations. The horizontal and vertical locations were 4° above or below and to the left or right of fixation, centered on the four corners of the occlusion square. Each of these four positions could be presented as “front” or “back” depth. Position was assigned randomly for the first stimulus, and the second stimulus could appear in one of four conditions relative to the first stimulus: same or different 2D location by same or different depth location. When 2D location was different, it differed in both horizontal and vertical position (i.e., diagonally opposite corner of occlusion square). These four conditions were counterbalanced and equally likely. As in Experiment 1, location was always irrelevant to the task, which was to judge same/different color. Timing was the same as Experiment 1, except that stimuli were presented for 500 ms instead of 250 ms.

3.2. Results

3.2.1. Congruency bias

Fig. 3 illustrates the bias to respond “same color” as a function of the irrelevant location conditions. Depth location information cued using occlusion and size did not bias color judgments. A two-way repeated-measures ANOVA with factors 2D location (same/different) and depth location (same/different) revealed a main effect of 2D location ($F_{1,16} = 8.13$, $p = 0.012$, $\eta^2_p = 0.35$), but no effect of depth location ($F_{1,16} = 0.46$, $p = 0.510$, $\eta^2_p = 0.03$), and no significant interaction ($F_{1,16} = 1.61$, $p = 0.223$, $\eta^2_p = 0.10$).

3.2.2. Other effects

For $d'$ (Table 1), there was a significant influence of 2D location ($F_{1,16} = 17.82$, $p = 0.001$, $\eta^2_p = 0.54$), but no effect for depth location ($F_{1,16} = 1.06$, $p = 0.320$, $\eta^2_p = 0.07$), with no interaction ($F_{1,16} = 4.24$, $p = 0.057$, $\eta^2_p = 0.22$). RT priming was significant for both 2D and depth location (2D: $F_{1,16} = 9.61$, $p = 0.007$, $\eta^2_p = 0.39$; Z: $F_{1,16} = 15.08$, $p = 0.001$, $\eta^2_p = 0.50$), with a significant interaction ($F_{1,16} = 11.37$, $p = 0.004$, $\eta^2_p = 0.43$).
In terms of the spatial congruency bias, in Experiment 2 we again replicated the 2D location bias found in Golomb et al. (2014) and in Experiment 1, but here we found that depth location using monocular cues did not result in a significant spatial congruency bias. It is worth noting that both experiments showed significant reaction time priming for same versus different depth, as seen in previous research (Atchley et al., 1997; Downing & Pinker, 1985), indicating that participants were sensitive to the depth information from both binocular and monocular cues. However, only the binocular disparity cue depth information resulted in a spatial congruency bias.

The finding that depth from binocular disparity biases color judgments but depth from occlusion and size does not do so suggests that the spatial congruency bias is not a generalizable phenomenon common to all depth cues. This leads to the question: was the congruency bias seen in Experiment 1 driven by depth location information at all? In other words, does depth-from-disparity bias color judgments, or was the effect due to low-level disparity differences in 2D location between the eyes?

To test this question, in Experiment 3 we asked if vertical disparity information biases color judgments. Vertical disparity stimuli involve displaying items at slightly different vertical locations in each eye, with no associated depth percept. If the results from Experiment 1 were due to the depth percept, we would not expect vertical disparity to exhibit a spatial congruency bias, whereas if it were low-level disparity differences producing this bias, then vertical disparity should also bias color judgments.

4. Experiment 3

4.1. Method

4.1.1. Participants

Sixteen subjects (9 female; mean age = 20 years; range: 18–24) participated. None were excluded.

4.1.2. Stimuli, procedure & design

The stimuli and procedure were the same as Experiment 1, except that instead of testing horizontal disparity, Experiment 3 tested vertical disparity. Vertical disparity was the same magnitude as horizontal disparity in Experiment 1: 30 arcmin (0.5°), half in one direction (randomly up or down) for one eye, and the other half in the opposite direction for the opposite eye. Thus, on trials where vertical disparity was different, the two stimuli covered the same vertical positions, but the eye-specific position reversed between stimuli. Likewise, whereas Experiment 1 varied 2D location using vertical differences of 5.46° above or below fixation, Experiment 3 varied 2D location using horizontal differences of 5.46° left or right of fixation (with vertical location aligned with fixation). The four irrelevant location conditions (same/different 2D horizontal location × same/different vertical disparity) were counterbalanced and equally likely.

4.2. Results

4.2.1. Congruency bias

Fig. 4 illustrates the bias to respond “same color” as a function of the irrelevant location conditions. Irrelevant vertical disparity location information biased color judgments, such that participants were more likely to report that the two stimuli were the same color when they had the same vertical disparity information, compared to different vertical disparity (eye-specific location) information. A two-way repeated-measures ANOVA with factors vertical disparity (same/different) and horizontal location (same/different) revealed significant main effects of both horizontal location and vertical disparity (X: $F_{1,15} = 9.12, p = 0.009, \eta^2_p = 0.38$, Y: $F_{1,15} = 16.55, p = 0.001, \eta^2_p = 0.535$ respectively), with no significant interaction ($F_{1,15} = 3.55, p = 0.079, \eta^2_p = 0.19$).

4.2.2. Other effects

For d' (Table 1), there was a significant influence of horizontal location ($F_{1,15} = 29.30, p < 0.001, \eta^2_p = 0.66$), but none for vertical disparity ($F_{1,15} = 0.40, p = 0.539, \eta^2_p = 0.03$), with no interaction ($F_{1,15} = 2.92, p = 0.108, \eta^2_p = 0.16$). For RT priming there were no significant main effects (X: $F_{1,15} = 0.03, p = 0.871, \eta^2_p < 0.01$, Y: $F_{1,15} = 1.37, p = 0.260, \eta^2_p = 0.08$) or interaction ($F_{1,15} = 2.93, p = 0.107, \eta^2_p = 0.16$).

4.3. Discussion

As in the previous experiments, 2D location information resulted in a strong spatial congruency bias. Surprisingly, vertical disparity also resulted in a significant spatial congruency bias for color judgments. Because vertical disparity entails different eye-specific location information without producing any depth percept, this finding suggests that the results from Experiment 1 were likely due to the low-level disparity differences for each eye rather than depth information.
5. General discussion

We investigated the effect of depth location information on color judgments across three experiments. Previous work reported a spatial congruency bias, where two objects were more likely to be judged as having the same features when they appeared in the same 2D location (Golomb et al., 2014). In the current paper we replicated this prior result and tested whether depth location also biases feature judgments. While we found initial evidence that depth from binocular disparity seemed to bias color judgments in Experiment 1, the results from Experiments 2 and 3 suggest that depth location does not in fact bias color judgments, indicating that the spatial congruency bias does not extend to 3D location.

Before further discussion, it is important to note that the congruency bias reflects a different type of effect than response facilitation measured by reaction time or sensitivity. Both RT and d’ measure facilitation; that is, an increase in performance when an irrelevant dimension is repeated. The congruency bias, on the other hand, does not necessarily improve performance, but rather results in a shift in the responses, and has been argued to reflect something more fundamental about the role of location in object perception (Golomb et al., 2014). In this sense the congruency bias could be seen as similar to the Simon or Stroop tasks (Lu & Proctor, 1995; Simon, 1990; Stroop, 1935), such that when the location is the same, participants might be unable to suppress a response to that property, even though it is task irrelevant. However, while the Simon and Stroop tasks are typically understood as response interference effects, Golomb et al. (2014) argued that the congruency bias reflects more of a perceptual-level shift. While the bias (criterion) measure is traditionally associated with changes in response, bias effects can in fact result from either perceptual or response processes (Mack, Richler, Gauthier, & Palmeri, 2011; Wixted & Stretch, 2000), and may reflect a perceptual-level effect even when there is no effect on d-prime/sensitivity (Morgan, Hole, & Glennerster, 1990; Witt, Taylor, Sugovic, & Wixted, 2015).

In the original spatial congruency bias report, Golomb et al. (2014) reported that even when judgments were made using a sliding scale that eliminated the response conflict, participants were more likely to rate two objects as more similar when location was the same, and that this effect was only present for perceptually difficult discriminations (Golomb et al., 2014).

Thus, the spatial congruency bias may carry different theoretical implications than a sensitivity effect, even though both may be perceptual in nature. Moreover, it is possible for the two effects to co-exist, such that location information may both bias and improve feature judgments. In previous reports of the spatial congruency bias (Golomb et al., 2014; Shafer-Skelton et al., Submitted), the bias was sometimes accompanied by a sensitivity effect, as we found here for 2D information, but in several of the original experiments there was only a bias effect and no change in sensitivity, suggesting that observers give more “same” responses even when sensitivity is equalized. Thus, the congruency bias seems to tap into something fundamental about the object-location binding process, where object location (at least in 2D) is automatically incorporated into perception of object features.

Here we investigated the role of depth location information in this process, asking if depth is processed like another type of location (in which case we would expect depth to induce a spatial congruency bias and bias color judgments); or if depth is more like other features (in which case it should not induce a congruency bias). Note that in the original spatial congruency bias report (Golomb et al., 2014), features such as color and shape did not induce a congruency bias, even when the differences were highly salient, whereas even small, near-threshold differences in 2D location biased feature judgments. In Experiments 1 and 2, we tested large, salient differences in depth location, finding that binocular – but not monocular – depth cues biased color judgments.

The results from Experiments 1 and 2 suggest that the influence of depth is tenuous at best, and is not generalizable across different depth cues. In Experiment 3 we further probed the role of depth information, dissociating the effects of depth-from-disparity information from pure disparity (eye-specific location) information. As the same vertical disparity and different vertical disparity conditions both have the same amount of overlap between the two eyes, any effect of color perception or fusion should be the same across these conditions. Combined with our results from Experiment 2, the finding that vertical disparity biased color judgments even though it does not create a depth percept suggests that depth location does not bias object perception. However, we cannot rule out an alternative interpretation that the congruency bias seen in Experiment 1 was due to combined depth and disparity information, and that binocular depth but not monocular depth biases the perception of object features. Binocular disparity is arguably one of the stronger and more realistic cues for depth perception (Finlayson, Remington, & Grove, 2012; McKee & Taylor, 2010), and Finlayson et al. (2012) demonstrated that there is variation in the perception of motion in depth depending on the cues used to simulate depth. That said, in both Experiments 1 and 2 we found RT priming effects for depth, as expected based on previous 3D attention literature (Arcthy et al., 1997; Downing & Pinker, 1985), indicating that participants perceived and were sensitive to depth information in both cases. Therefore we believe it is unlikely that these results are due to a difference in depth cue, and more likely due to position-in-depth information not biasing color judgments, with the binocular disparity results from Experiment 1 reflecting disparity and not depth effects.

Our results indicate that depth does not induce a spatial congruency bias, similar to the lack of bias induced by other features such as shape and color (Golomb et al., 2014). We therefore propose that – at least in this context – depth information is treated more similarly to other types of object features, rather than as an aspect of object location. Our findings provide support for a special role of location in the binding process, but only for 2D location. An account of location as a privileged feature proposes that irrelevant location information is automatically encoded with other object features, biasing their perceptual judgments. In this account, 2D location serves as an index to group or bind features of an object together, an important cue for object recognition (Golomb et al., 2014; Kahneman, Treisman, & Gibbs, 1992). The fact that this special role does not seem to extend to position-in-depth information is consistent with studies suggesting weakened or delayed effects of depth compared to 2D location (Finlayson et al., 2013; Gilinsky, 1951; Kasai et al., 2003; Loomis et al., 2008; Moore et al., 2009), and suggests that other research finding similar perceptual and attentional responses for 2D and depth effects (Arcthy et al., 1997; Downing & Pinker, 1985; Finlayson et al., 2013; Nakayama & Silverman, 1986) may instead reflect later processes or feedback effects less involved in the binding process.

Moreover, the idea that 2D – but not 3D – location may serve as a fundamental index or cue for object binding carries interesting neural implications for representations of object location and identity. One fundamental question is to what extent “what” and “where” information is processed separately in the brain. While original accounts suggested a strict dichotomy of two separate visual streams (Goodale & Milner, 1992; Mishkin, Ungerleider, & Macko, 1983; Ungerleider & Mishkin, 1982), recent evidence supports a more nuanced story (Carlson, Hogendoorn, Fontejn, & Verstraten, 2011; Cichy, Chen, & Haynes, 2011; DiCarlo & Maunsell, 2003; Golomb & Kanwisher, 2012; Kravitz, Kriegeskorte, & Baker, 2010; Op De Beeck & Vogels, 2000; Schwarzlose, Swisher, Dang, & Kanwisher, 2008). As proposed in...
Golomb et al. (2014), the spatial congruency bias suggests that object identity may never be represented fully independently of location, and our findings are consistent with this proposal, with the caveat that we only see this advantage for 2D location, not depth locations. Depth information has been reported in both dorsal and ventral stream visual areas, including known object- and feature-processing areas such as LOC, MT, and V4 (DeAngelis & Newsome, 1999; Finlayson, Zhang, & Golomb, Under Review; Hubel & Wiesel, 1970; Neri, 2004; Parker, 2007; Preston et al., 2008; Tanabe, Doi, Umeda, & Fujita, 2005; Tsao et al., 2003; Welchman et al., 2005), raising the interesting possibility that 2D location information may be more integrated than depth information with the object feature information in these regions. Another possibility is that this aspect of the binding process occurs earlier in visual processing, perhaps before depth information is fully represented. Binocular disparity information is found in early visual cortex (Hubel & Wiesel, 1970; Skalicky, 2016), but perceptually relevant position-in-depth information may not emerge until later visual areas (Barendregt et al., 2015; Finlayson et al., Under Review; Preston et al., 2008).

Regardless, it appears that 2D location plays a more fundamental role in the binding process than does position-in-depth. However, it is unclear if the spatial congruency bias is only seen for 2D location, with all other features treated equally, or if it perhaps reflects a hierarchy (e.g. Felleman & van Essen, 1991; van Essen & Zeki, 1978) where features processed earlier in the visual processing stream might bias the judgments of more complex features processed later. For example, depth might not influence low-level features like color or orientation, but perhaps might influence a more complex judgment such as face perception. Other research from our group has shown that 2D location biases features regardless of complexity (e.g., exerting similar effects on Gabors and faces; Shafer-Skelton et al., Submitted), and that neither color nor shape bias one another (Golomb et al., 2014), but the possibility of a hierarchy of 3D location processing cannot be ruled out in our current study.

Finally, a surprising and important result we uncovered was the robust effect of eye-specific location information on object perception; i.e., that color judgments were biased by 2D location information that was only present as a relative difference between the two eyes. For example, if Object 1 was centered 0.5° above the midline in the left eye and 0.5° below the midline in the right eye, participants were more likely to judge Object 2 as being the same color if it maintained this exact disparity information, compared to a subtle swap between eyes (0.5° below the midline in the left eye and 0.5° above the midline in the right eye), even though the average position across eyes was identical in both cases. This indicates a very low-level and early effect of 2D location, before location information from each eye is combined and averaged, suggesting that the spatial congruency bias occurs very early in visual processing. This would be consistent with research showing that the spatial congruency bias is present in low-level retinotopic (eye-centered) coordinates rather than the more ecological spatiotopic (world-centered) coordinates across eye movements (Shafer-Skelton et al., Submitted). However, this comparison is particularly interesting in light of evidence showing that depth information is represented explicitly in visual cortex (Bridge & Parker, 2007; Finlayson et al., Under Review; Hubel & Wiesel, 1970), whereas spatiotopic representations are not (Gardner, Merriam, Movshon, & Heeger, 2008; Golomb & Kanwisher, 2012). Thus it is interesting that neither depth nor spatiotopic position seem to influence the congruency bias, yet tiny differences in eye-specific position can cause substantial influences on judgments of features such as color. Although some theories of object-location binding have suggested that binding occurs later in visual processing (e.g., Riesenhuber & Poggio, 1999), perhaps even in medial temporal lobe or prefrontal cortex (e.g., Hannula & Ranganath, 2008; Mitchell, Johnson, Raye, & D’Esposito, 2000; Rao, Rainer, & Miller, 1997), this eye-specific finding suggests that at least certain aspects of object-location binding occur much earlier, relying solely on low-level 2D location cues.

Acknowledgments

This work was supported by research grants from the National Institutes of Health (R01-EY025648) and Alfred P. Sloan Foundation (BR-2014-098). We thank members of the Golomb Lab for assistance with data collection.

References


Finlayson, N. J., & Golomb, J. D. (Under Review). 3D spatial vision: 2D location biases depth judgments but not vice versa.


Golomb, J. D., & Kanwisher, N. (2012). Higher level visual cortex represents retinotopic, not spatiotopic, object location. Cerebral Cortex, 22(12), 2794–2810.