Visual search is influenced by 3D spatial layout

Nonie J. Finlayson[1,2] · Philip M. Grove[1]

Abstract Many activities necessitate the deployment of attention to specific distances and directions in our three-dimensional (3D) environment. However, most research on how attention is deployed is conducted with two dimensional (2D) computer displays, leaving a large gap in our understanding about the deployment of attention in 3D space. We report how each of four parameters of 3D visual space influence visual search: 3D display volume, distance in depth, number of depth planes, and relative target position in depth. Using a search task, we find that visual search performance depends on 3D volume, relative target position in depth, and number of depth planes. Our results demonstrate an asymmetrical preference for targets in the front of a display unique to 3D search and show that arranging items into more depth planes reduces search efficiency. Consistent with research using 2D displays, we found slower response times to find targets in displays with larger 3D volumes compared with smaller 3D volumes. Finally, in contrast to the importance of target depth relative to other distractors, target depth relative to the fixation point did not affect response times or search efficiency.

Keywords Visual search · 3D attention · Depth · Spatial attention · Feature-based attention · Search efficiency

✉ Nonie J. Finlayson
   nonie.j@gmail.com

1   The University of Queensland, St Lucia, Australia

2   Department of Psychology, The Ohio State University, Columbus, OH, USA

The deployment of visual attention concentrates an observer's limited visual processing resources on potentially relevant information, while attenuating the influence of other nonrelevant information in the scene. An activity, such as searching for an empty chair in a seminar room, requires the deployment of attention to possible targets, such as chairs in a particular region of the room, and filtering features of that attended region for information that will help guide target selection, such as chair location and availability. Although there is a vast literature examining visual search within 2D displays, such as across a picture of a room (Carrasco, 2011; Cave & Bichot, 1999), there is considerably less research exploring how attention is deployed in 3D space.

Previous studies have reported that it is possible to deploy attention to different depths (Downing & Pinker, 1985; Marrara & Moore, 2000). Downing and Pinker (1985) were the first to demonstrate that responses were slower for targets at a different depth than fixation compared with targets at the same depth as fixation, although both targets' eccentricities were the same. These slower responses found when switching depths illustrates that target localisation does not rely only on 2D coordinates. Marrara and Moore (2000) also found effects of cueing in depth, showing that attention can be shifted in depth without the need for placeholders, objects, or surfaces and is determined based on the perceptual organization of the display. Researchers also have investigated the role of 3D structure in visual search tasks, finding that not only is search for a target uniquely located in depth efficient, but using depth information to segment a display may improve search performance (Finlayson, Retell, Remington, & Grove, 2013; Nakayama & Silverman, 1986).

Although there has been research exploring the effects of depth as a *feature* of search elements (Downing & Pinker, 1985; Marrara & Moore, 2000; Nakayama & Silverman, 1986; de la Rosa, Moraglia, & Schneider, 2008), we know

less about the nature of a 3D search environment, and how visual search is affected by 3D space. Enns and Rensink (1990a; 1990b) demonstrated that scene-based properties improved the efficiency of visual search where those properties directly influence the saliency of the target. Specifically, Enns and Rensink (1990a) showed that adding depth information to the target improves search performance, such that a target is found faster when searching for a 3D box compared with a 2D shaded shape. Aks and Enns (1996) showed that search performance can be modified by adding perspective texture cues to the background of a search display, comparing search in displays perceived as 2D versus 3D. We explore depth as a spatial context within which visual search takes place. Specifically, we investigate four parameters of 3D space and how they affect visual search when irrelevant to the task.

Investigating the deployment of attention in 2D space, Posner, Snyder, and Davidson (1980) first described the allocation of attention as a "spotlight," an analogy where attention resources are concentrated on a particular area, similar to how a spotlight illuminates a selected region of an otherwise dark stage. The zoom lens model (Eriksen & St James, 1986) extended this analogy with the idea of scale adjustment, such that the size of the spotlight can be adjusted according to the task at hand. The attentional spotlight also has been described as a volume extending in depth, based on measured reaction times to targets flanked by distractors at different distances in depth from the target (Andersen, 1990; Andersen & Kramer, 1993). Andersen and Kramer (1993) describe the attentional spotlight as viewer-centered, concentrated between the observer and the point of fixation, with attention dropping off steeply beyond. Parks and Corballis' (2006) data support the idea of viewer-centered attention in depth, recording response times and event-related potentials (ERPs) while participants attended to near or far depths. They found an enhanced ERP signal for attended versus nonattended stimuli for the far depth but not the near depth, consistent with a viewer-centered asymmetry. Our goal is to extend these qualitative descriptions by systematically investigating the influence of four key attributes of the distribution of objects in 3D space on visual search performance. These key attributes are distance in depth, 3D display volume, number of depth planes, and relative target position in depth. We define and discuss each in turn below.

Distance in depth

We define distance in depth to be the z-axis distance between the initial fixation point and a target item. Research with 2D displays suggested that attention takes longer to shift over larger lateral distances (Tsal, 1983), but later evidence demonstrated that attentional shifts are temporally invariant regardless of distance (Remington & Pierce, 1984).

Remington and Pierce (1984) showed that facilitation for a centrally cued peripheral target developed at the same rate for laterally near or far targets. Sperling and Weichselgartner (1995) found similar results and concluded that the duration of a shift of attention is independent of lateral distance and interposed distractors. No research has examined whether or not the distance in depth between the target and fixation affects attention shift duration. However, as lateral shifts were shown to be independent of eye movement characteristics, a similar result could be expected for shifting attention in depth. Therefore, in line with Remington and Pierce (1984), we predict that reaction times and search slopes will not be affected by the targets' z-axis distance from the fixation point.

3D display volume

The 3D display volume refers to the total 3D volume of a display; specifically, the z-axis distance between the nearest and the farthest depth plane in the display. In our manipulation of 3D display values, target, and distractor items maintained the same x and y coordinates regardless of their z-axis (depth) position. In the 2D visual search literature, increasing the size of the search area is associated with a decrease in the processing efficiency of the search array (Castiello & Umilta, 1990; Eriksen & St James, 1986). Eriksen and St James (1986) found that as the 2D size of a cued area increased, reaction time to targets also increased, indicating a reduced concentration of processing resources within the cued area, as fewer resources are dedicated to each location when attention is spread over a larger area. When attention is distributed over a larger region of the visual field, there is a corresponding degradation in processing efficiency for that attended region (Carrasco, 2011). In line with these previous results, we predict that increasing the 3D display volume also will reduce search performance, as measured by response times and search slopes.

Number of depth planes

The number of depth planes refers to the number of discrete depth planes that the target and distractor elements are distributed across within the 3D volume of the search display. Previous literature generally finds that in 2D displays, increased variation in an irrelevant search feature leads to reduced search performance (Farmer & Taylor, 1980; Treisman, 1982). Farmer and Taylor (1980) found that backgrounds with greater color heterogeneity that were irrelevant to the task lead to slower present/absent responses. Treisman (1982) examined 2D spatial layouts and found that participants were slower to find targets when the items in the display were presented in more numerous spatial groups compared to fewer

spatial groups. To explore how increased variation in irrelevant depth information affects search performance, we manipulated the number of depth planes that search elements were distributed across within a given 3D volume. In accordance with previous findings, we predict that increasing depth heterogeneity by spreading the search items over more depth planes will increase both search times and search slope results.

Relative target position in depth

The relative target position in depth refers to the position of a target in depth in reference to other search elements (e.g., in front of or beyond). Previous studies have presented conflicting evidence regarding the effect of target position in depth on search performance. Some studies report asymmetrical results: faster responses to targets the closer they are to the front of the display (O'Toole & Walker, 1997; Reis, Liu, Haivg, & Heft, 2011). Reis et al. (2011) used a real 3D display with fixation initially on the far depth plane, and still participants found targets on the near depth plane faster when searching for a target defined by a single feature. On the other hand, studies using both feature and conjunction search found no response time differences for the manipulation of relative target position in depth (Theeuwes, Atchley, & Kramer, 1998; Dent, Braithwaite, He, & Humphreys, 2012; Finlayson, Remington, Retell, & Grove, 2013). However, previous research often only included target position in depth as a counterbalance measure. In the present report, we systematically examine the effect of target depth position on target selection and predict that if the relative position of the target in depth affects search in the same way as previously reported, we will see faster reaction times as well as shallower search slopes for targets in the front of the search array, compared with those at the back of the array.

Current study

For a complete understanding of how humans deploy attention throughout their environment, it is necessary to understand how attention is deployed in 3D space in addition to how it is deployed across a 2D plane. The latter has been extensively studied while the former has attracted considerably fewer investigations. In this report, we systematically examined target selection in 3D space, exploring two measures for examining the processes of visual attention; reaction times and search slopes. Our experiment employed a visual search paradigm to explore the 3D spatial parameters of distance in depth, 3D display volume, number of depth planes, and relative target position in depth. We measured the response times to find a target and varied distractor set size to measure search slopes. Search slope results reflect the search

efficiency; that is, the rate (ms per item) at which distractors can be rejected and the target is selected, interpreted to occur within the guidance stage of attentional processing (see Wolfe's Guided Search model; Wolfe, 2007). This guidance stage uses input from early stages of visual processing, prioritising salient attributes through a processing bottleneck to later stages of response and decision-making. Response time results reflect processes that occur either before guidance, such as during perception and encoding, or after guidance, during stages of decision making and response (Woodman, Vogel, & Luck, 2001).

Method

Participants

Twenty-two volunteer participants were recruited from the first-year participant pool at the University of Queensland. Stereoacuity was measured using the Titmus Stereo test (Stereo Optical Co., Chicago, IL), and all had acuities of 30 arc sec or less at 16 inches, with normal or corrected-to-normal vision.

Apparatus and stimuli

Stimuli were drawn and scripted using Matlab (MathWorks, Natick, MA) and the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997) and presented on two 24" Macintosh Cinema displays (1680 × 1050 pixels) in a mirror stereoscope. The viewing distance was 60 cm.

Participants searched a 3D volume for the presence of a stereoscopically forward tilting bar among backward tilting bars, and we inferred starting position of observers' eyes and attention through a Landolt C task as described below (Flom, Weymouth, & Kahneman, 1963). We manipulated the number of depth planes that the elements were arranged into, the distance between depth planes, and the depth plane occupied by the target.

The search array consisted of 12 or 24 white bars (22.5 arc min × 1.5 arc min) tilted top-backward in depth. The disparity between the top and bottom of each bar was 6 arc min such that the top of the bar had an uncrossed disparity relative to the bottom of the bar. With a set size of 24, the elements were arranged in a matrix consisting of six columns × four rows. The center-to-center vertical separation between any two elements was 2.3°, whereas the horizontal distances varied (2.5° to 2.1°). The set size of 12 consisted of elements arranged in a matrix of four rows × three of the six columns used for set size 24, meaning that the horizontal spacing was doubled. The elements were evenly distributed across two or four depth planes and displayed on a black background. When present, the target (a bar tilted top near) replaced one of the distractors,

leaving 11 distractors with either 2 or 5 in the same depth plane as the target, replacing a backward-tilting bar and equally likely to appear on each depth plane. The fixation point was a white dot (3 arc min radius).

Each item could be presented in one of four stereoscopic depth planes, simulated with binocular disparity. The disparity magnitudes for the four depth planes were 18 arc min crossed disparity (depth plane 1), 6 arc min crossed disparity (depth plane 2), 6 arc min uncrossed disparity (depth plane 3), and 18 arc min uncrossed disparity (depth plane 4), relative to the plane of the display.

In the search task there were four depth plane conditions. First, the *small-separation near* condition consisted of two depth planes situated in front of the plane of the display. These were depth planes 1 and 2 described above, and stimuli were equally divided between the two depth planes. Second, the *small-separation far* condition consisted of two depth planes situated beyond the plane of the display, depth planes 3 and 4 described above Stimuli were equally divided between depth planes 3 and 4. Third, the *large-separation* condition consisted of two depth planes with a large separation between them, depth planes 1 and 4 described above. Stimuli were equally divided between the two depth planes. Fourth, the *four depth planes* condition consisted of stimuli equally divided among depth planes 1, 2, 3, and 4.

The initial position of observers' eyes and attention was controlled with a Landolt C task (Flom et al., 1963). This attention fixation task consisted of a central square frame (6 arc min on each side) with the top or bottom side missing. A circular frame (15 arc min radius) surrounded the square and was flanked top and bottom with two vertical nonius lines (15 arc min long), drawn 18 arc min above and below fixation. All of these elements were presented at 18 arc min uncrossed disparity (depth plane 4). The far depth plane was chosen as the initial fixation point to reduce any advantages for near targets, as has been demonstrated in previous research (O'Toole & Walker, 1997; Reis et al., 2011).

Procedure and design

On each trial (Fig. 1), a fixation dot was displayed at depth plane 4 for 400 ms, followed by the first task, attention fixation. Participants were instructed to ensure that the nonius lines were vertically aligned and then indicate whether the top or bottom side was missing from the square fixation frame with the up and down arrow keys respectively. Once a correct response was provided, the second task, visual search, was presented until the participant responded. Participants were free to make natural eye movements during the search phase of the task. Participants were instructed to search for a bar tilted top near and indicate whether it was present or absent by pressing the left or right arrow key respectively.

Participants completed 640 trials in total, with a rest halfway through. Depth plane condition (small-separation near, small-separation far, large-separation, four depth planes), target depth (1 nearest, 2, 3, 4 farthest), set size (12, 24), and target presence (present, absent) were all randomized within each experiment. There were equal numbers of trials in each depth plane condition and the target had an equal probability of appearing at each depth within a given condition. That is, the target was presented on each possible depth plane on 50% of target present trials for the two depth plane conditions, and 25% of target present trials for the four depth plane condition. The target was present on 75% of trials, with target-absent trials acting as catch trials, and the distractors were equally divided among possible depth planes. This resulted in a minimum of 15 trials per condition, with double that in the conditions that only had two depth planes, and took participants from 30-60 minutes to complete.

Results

Data from one participant were removed due to accuracy rates of less than 85% in the search task. Average accuracy for the remaining participants was 94% with no sign of a speed-accuracy trade-off. All error trials were removed from the data set before analysis. Less than 1.1% of all trials had incorrect responses to the Landolt C task, and these were also removed before analysis. We calculated mean response times for target present trials in each condition.

Figure 2 illustrates response times to find a target as a function of depth plane condition and target depth. Although the small-separation near and small-separation far conditions are plotted separately in the figure, in order to investigate the effect of target depth (i.e., the depth plane the target was presented on) across 3D volume and number of depth planes, we use depth plane 1 from the "small-separation near" condition and depth plane 4 from the "small-separation far" condition. This allows us to compare search results for small and large 3D volumes containing the same number of depth planes (two), by comparing the response times to the target in the large-separation depth plane 1 (front) and 4 (back) directly to the response times in the small-separation conditions at depth plane 1 (front depth plane in small-separation near condition) and depth plane 4 (back depth plane in small-separation far condition).

The two small-separation conditions were presented at different distances relative to the fixation point; however, a three-way ANOVA with set size (12, 24), target depth (front, back), and depth condition (near, far) found that mean target-present response times averaged across set size and target depth did not differ significantly between the small-separation near ($M = 1990$ ms) and small-separation far ($M = 2046$ ms) conditions ($F_{1,20} = 2.16$, $p = 0.157$).
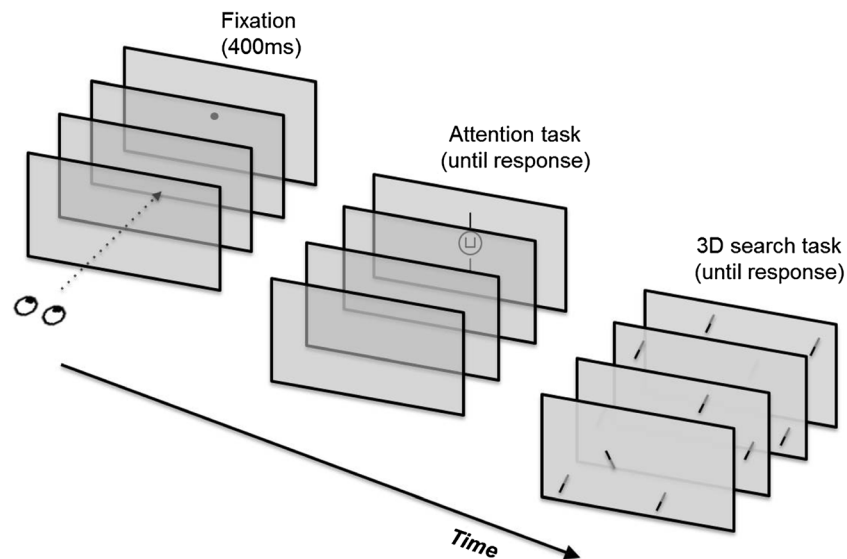
Fig. 1 Cartoon diagram of a typical trial sequence. The participant fixated for 400 ms, then completed a Landolt C task to draw attention to depth plane 4, and then completed the search task, illustrated here as a four-depth plane search condition, set size 12 (target present on the front depth plane)

Search slopes also did not differ significantly between the small-separation near ($M = 69$ ms/item) and small-separation far ($M = 72$ ms/item) conditions ($F_{1,20} = 0.21$, $p = 0.649$), and there was no three-way interaction ($F_{1,20} = 0.31$, $p = 0.586$). Therefore, for subsequent analyses the small-separation near and far conditions are referred to as *small-separation*. The reason we do not collapse across depth planes into small separation front (depth planes 1 & 3) and small separation back (depth planes 2 & 4) is to control for possible effects of absolute disparity (i.e., distance in depth from the initial fixation point). However, we note that averaging across the conditions

in this way leads to the exact same pattern of results for main ANOVA and follow-up analyses.

A two (set size: 12, 24) by two (target depth: nearest, farthest) by three (depth plane condition: small-separation, large-separation, four depth planes) repeated-measures ANOVA conducted on the target-present response times revealed a main effect of set size ($F_{1,20} = 66.62$, $p < 0.001$, $\eta_p^2 = 0.77$), such that targets were responded to faster when in the smaller set size ($M = 1705$ ms) compared with the larger set size ($M = 2581$ ms). We found a main effect of target depth ($F_{1,20} = 29.90$, $p < 0.001$, $\eta_p^2 = 0.60$), such that targets on the nearest
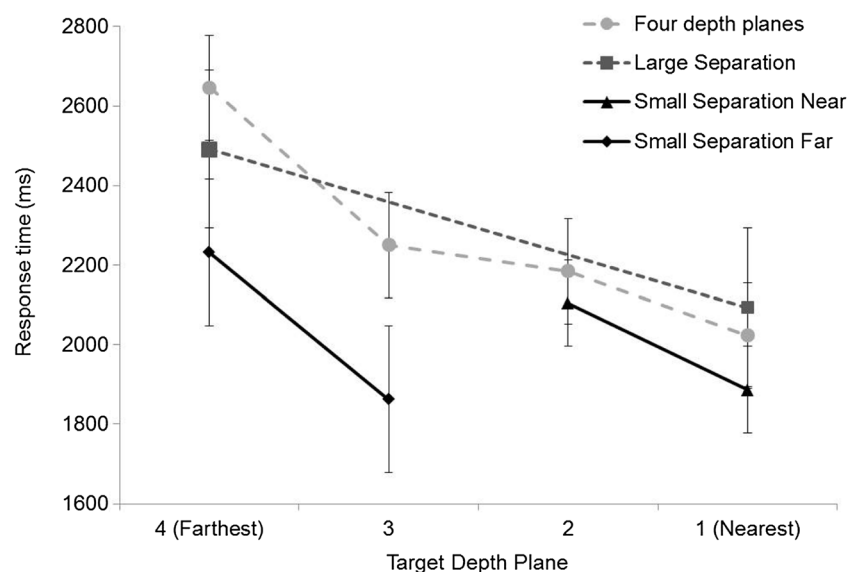
Fig. 2 Mean response times for target present data with target depth plane on the x-axis and depth plane conditions in separate lines, collapsed across set size. Error bars show standard error

depth plane (M = 1993 ms) were responded to faster than targets on the farthest depth plane (M = 2420 ms).

We found a significant main effect of depth plane condition ($F_{2,40}$ = 16.88, p < 0.001, η_p^2 = 0.46) and a significant interaction between target depth and depth plane condition ($F_{2,40}$ = 4.73, p = 0.014, η_p^2 = 0.19). Following up the main effect of depth plane with paired t tests Bonferroni corrected for three comparisons (α = 0.017), the small-separation condition (M = 2065 ms) was significantly faster than either the large-separation condition (M = 2288 ms, p < 0.001) or the four depth plane condition (M = 2319 ms, p < 0.001), with no significant difference between large-separation and four depth plane conditions (p > 0.05). This same pattern of results was reflected in the interaction, which averaged response times across set size. Unpacking the interaction with ANOVAs for each target depth, we found significant differences across depth plane conditions for both depth plane 1 ($F_{2,40}$ = 5.87, p = 0.006, η_p^2 = 0.23) and depth plane 4 ($F_{2,40}$ = 13.09, p < 0.001, η_p^2 = 0.40). Follow-up comparisons showed that targets in the small-separation condition were found to be significantly faster than either the large-separation (t_{20} = −3.21, p = 0.004 & t_{20} = −3.67, p = 0.002 depth plane 1 and 4 respectively) or the four depth plane condition (t_{20} = −2.37, p = 0.028 & t_{20} = −4.29, p < 0.001, depth plane 1 and 4 respectively). Response time results replicated a prior pilot study.[1]

Furthermore, we found significant interactions between target depth and set size ($F_{2,40}$ = 5.79, p = 0.006, η_p^2 = 0.22), and depth plane condition and set size ($F_{1,20}$ = 7.17, p = 0.014, η_p^2 = 0.26), but no three-way interaction (p = 0.210). We followed up these set size interactions by comparing the search slopes for the target depth conditions and the depth plane conditions (Fig. 3). To calculate search slopes for each participant, response times for set size 12 were subtracted

---

[1] In pilot work, 22 naïve participants were tested, reporting normal or corrected to normal binocular vision. The design and procedure was exactly as in the experiment described above, except set size was not manipulated, with only 12 elements in each search array. A two-way ANOVA (depth plane condition by target depth) on response times revealed a similar pattern of results as the current study. A main effect of target depth (F[1,18] = 29.20, p < .001, η_p^2 = .62) showed that targets on the nearest depth plane (M = 1446 ms) were responded to faster than targets on the farthest depth plane (M = 1758 ms), and a main effect of depth plane condition (F[2,36] = 6.06, p = .005, η_p^2 = .25) showed that targets in the small-separation condition were found significantly faster than either the large-separation (p < .001) or the four depth planes condition (p = .037), with no significant difference between the large-separation and four depth planes conditions (p = .543).

from those at set size 24 and divided by the difference in number of items (24 – 12 = 12). Paired-sample t tests, Bonferroni corrected for three comparisons (α = 0.017), revealed that for the target depth condition, the search slope for the nearest targets (61 ms/item) was significantly shallower than the search slope for the farthest targets (87 ms/item) (t_{20} = −2.68, p = 0.014). For the depth plane variable, both the small-separation (69 ms/item) and the large-separation (63 ms/item) search slopes were significantly shallower than the four depth planes condition (90 ms/item) (t_{20} = −2.67, p = 0.015 and t_{20} = −2.64, p = 0.016, respectively) but did not differ from each other (t_{20} = 0.88, p = 0.390).
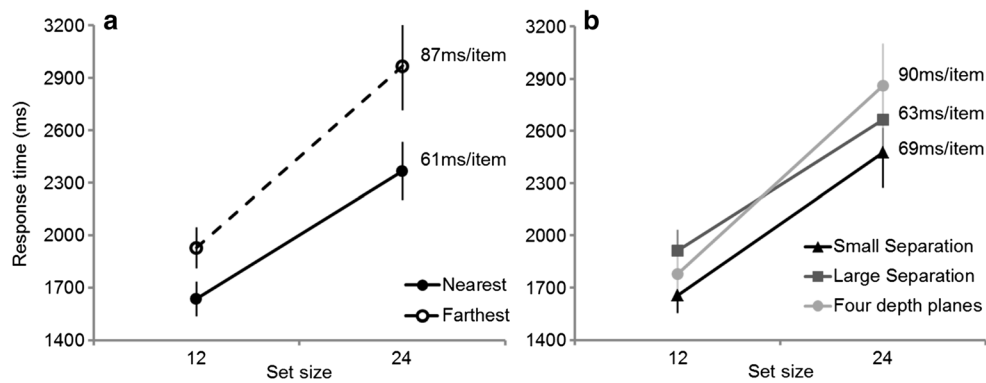
Discussion

This report is the first systematic attempt to characterise target selection in 3D space as a function of the target irrelevant characteristics: distance in depth, 3D display volume, number of depth planes, and relative target position in depth. The differential response times to detect targets in each of the 3D display conditions indicate that attentional deployment is influenced by the 3D spatial layout of the search array. We manipulated the 3D display parameters not directly related to the search elements and discuss the effects of each of the four chosen 3D space parameters on visual search performance in turn below. Revisiting the theoretical interpretations of the measures, variations in the response time to find the target reflect time to shift attention around the 3D array as affected by the different parameters, and search slopes reflect the efficiency of these attention shifts, in this case considered to be affecting the guidance of search items (Wolfe, 2007).

Our first prediction, relating to distance in depth, was that reaction times and search slopes would be unaffected by the target's proximity in depth to the fixation point. Consistent with this hypothesis, we found no difference in response times or search slopes between the near and far small-separation conditions, indicating that the relative position of the display in relation to initial fixation did not affect search performance. We infer that attention shifts are invariant across depths, consistent with previous research in 2D displays (Remington & Pierce, 1984; Sperling & Weichselgartner, 1995).

Next, we predicted that increasing the 3D display volume would reduce search performance, as measured by response times and search slopes. We found that response times to find the target were faster on average in the small-separation condition compared with the large-separation condition, which compared the 3D volume of the search display after controlling for number of depth planes. We also found that when targets were on depth plane 1, responses were faster in the small-separation condition compared with the large-separation condition. Because participants' attention was controlled to start each search trial at the back depth plane, when

Fig. 3 Search slope results shown for mean response times for target present data with set size on the x-axis, for (**a**) Target Depth, where the target appeared in the nearest (i.e., depth plane 1: filled black circles) versus the farthest (i.e., depth plane 4: unfilled white circles) depth planes, and for (**b**) Depth Plane condition, for small-separation (triangle), large-separation (square) and four depth plane (circle) conditions. Search slope values are reported next to each slope as increase in ms per item added to the display

the target was on depth plane 1 the participants were shifting their attention over the same distance in depth. However, when targets on depth plane 1 were presented in a larger search volume (i.e., large-separation condition compared with small-separation condition), their search performance was reduced (i.e., longer response times). We see the same pattern for targets on depth plane 4.

Looking at our search slope findings for this second prediction, we found no differences in search efficiency results between the small-separation and large-separation conditions. Our search slope findings indicate no set size effects for different 3D volumes of the search array. However, we demonstrate that search within a larger 3D volume (such as the large-separation compared with small-separation condition) requires more time per time find the target. The response time advantage for smaller 3D volumes is consistent with previous research with 2D displays (Castiello & Umilta, 1990; Eriksen & St James, 1986), which found that increasing the 2D search display size increased response times. However, because we found no difference in search slope results between the small and large search volumes, we infer that the 3D search volume affects search performance outside of the guidance mechanisms. It is possible that searching through a larger compared with smaller 3D space adds a perceptual constant when viewing the display. Alternatively, the 3D volume of the display may affect the response and decision making process. We cannot specify whether the increased time to find targets occurs before or after guidance. However, we speculate that a larger 3D volume reduces search performance by adding perceptual noise into the search process, as expected with more visual input.

Third, we predicted that increasing depth heterogeneity by spreading the search items over more depth planes would increase both search times and search slope results. However, we found no response time differences when searching an array separated into two (i.e., large-separation

condition) compared with four depth planes (equal 3D volumes). We did find shallower search slopes for the large-separation condition compared with the four depth planes condition, after controlling for display volume and target depth, indicating reduced search efficiency for increased depth heterogeneity (as indexed by number of depth planes). As such, it appears that search arrays arranged into fewer depth groups have an advantage over arrays with more depth groups in target selection. We found that the number of depth planes affects search efficiency, suggesting that depth heterogeneity affects the guidance stage of attentional processing, likely through bottom-up salience of the search items. Our findings are broadly similar to previous studies, which found that variation in color (Farmer & Taylor, 1980) as well as 2D spatial groups (Treisman, 1982) both resulted in slower target selection compared with homogenous displays. These studies indicate that when search array items are heterogeneous in an unrelated feature or spatial grouping, search times increase, whereas our study only found reduced search efficiency for increasing depth heterogeneity. Whereas there also is evidence opposing the negative effects of irrelevant distractor heterogeneity (Pashler, 1988), Duncan and Humphreys (1989) proposed that the important factor was target-distractor discriminability, such that distractor heterogeneity only affects search times when targets are similar to distractors. Our findings of reduced search performance with increased depth heterogeneity are consistent with this argument, as we see fairly steep search slopes in our results indicating low target-distractor discriminability. It should be noted that with only two and four depth planes we have limited predictive power, with insufficient heterogeneity to fully explore response times for 3D search array organization, and with greater depth heterogeneity it is possible that we also would see an effect on search times. We do see a trend towards this with a larger set size (Fig. 3b).

Finally, we confirmed our prediction of faster response items and shallower search slopes for targets in the front of the search array compared with those at the back of the array. Notably, looking at the small-separation near and far conditions, we see equivalent response times for targets at the relative "front" of each display, despite the "front" depth plane being depth plane 1 for the *small-separation near* condition and depth plane 3 for the *small-separation far* condition, and we found no significant difference between these two conditions. These data indicate the importance of relative position over distance from the initial fixation depth plane, demonstrating a target selection advantage for elements at the relative front of a 3D display. In contrast and as described above, we found that distance of a target from initial fixation does not appear to play an important role in search performance, at least not under the limited distances used in our study.

Past research exploring target depth has yielded varied results, finding no differences between front and back targets (Theeuwes, Atchley, & Kramer, 1998; Dent, Braithwaite, He, & Humphreys, 2012; Finlayson, Remington, Retell, & Grove, 2013), but also results consistent with ours showing an advantage for front targets (O'Toole & Walker, 1997; Reis et al., 2011). However these investigations have tended to employ target depth as a counterbalance measure, and often the search was restricted to one of the multiple depth planes present. Our results suggest that searching the display from front to back could be an automatic strategy that participants adopt, despite their attention starting at the back of the display. Additionally, elements closer to an observer may attract attention first due to being more immediately relevant to their behaviour, as these front objects are more likely to require an avoidance action than objects further away. This follows from the behavioural urgency hypothesis (Franconeri & Simons, 2003; Finlayson, Remington, & Grove, 2012), which predicts attentional capture by stimuli that indicate the potential need for immediate action, and in this case enable a perceptual processing primacy for near objects through the attentional selection system. More specifically, the improved search efficiency we found for front targets suggests guidance via bottom-up salience prioritisation and enhancement of the signal of objects closer to an observer.

Conclusions

Our data demonstrate that target selection in a 3D search array is predicted by 3D display volume, number of depth planes, and relative target position in depth, but not by distance in depth from the initial fixation plane. We demonstrate that the depth of the target relative to the other distractors is more important for search tasks than its depth relative to the fixation point, with a target at the relative "front" of a display being easier to find than one at the "back." Furthermore, we show that search is faster in smaller 3D volumes and that spreading search items across more depth planes reduces search efficiency but does not appear to affect response times. This systematic investigation of four parameters specifying object layout in 3D space provides a foundation for future 3D visual attention research.

References

Aks, D. J., & Enns, J. T. (1996). Visual search for size is influence by a background texture gradient. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1467–1481.

Andersen, G. J. (1990). Focused attention in three-dimensional space. *Perception & Psychophysics, 47,* 112–120.

Andersen, G. J., & Kramer, A. F. (1993). Limits of focused attention in three-dimensional space. *Perception & Psychophysics, 53*(6), 658–667.

Brainard, D. H. (1997). The Psychophsics Toolbox. *Spatial Vision, 10,* 433–436.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51,* 1484–1525.

Castiello, U., & Umilta, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica, 73,* 195–209.

Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review, 6*(2), 204–223.

de la Rosa, S., Moraglia, G., & Schneider, B. A. (2008). The magnitude of binocular disparity modulates search time for targets defined by a conjunction of depth and colour. *Canadian Journal of Experimental Psychology, 62*(3), 150–155. doi:10.1037/1196-1961.62.3.150

Dent, K., Braithwaite, J. J., He, X., & Humphreys, G. W. (2012). Integrating space and time in visual search: How the preview benefit is modulated by stereoscopic depth. *Vision Research, 65,* 45–61.

Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. P. O. Martin (Ed.), *Attention and Performance XI* (pp. 171–187). Hillsdale: Erlbaum.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological review, 96*(3), 433.

Enns, J. T., & Rensink, R. A. (1990a). Influence of scene-based properties on visual search. *Science, 247,* 721–723.

Enns, J. T., & Rensink, R. A. (1990b). Sensitivity to three-dimensional orientation in visual search. *Psychological Science, 1*(5), 323–326.

Eriksen, C. W., & St James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40*(4), 225–240.

Farmer, E. W., & Taylor, R. M. (1980). Visual search through color displays: Effects of target-background similarity and background uniformity. *Perception & Psychophysics, 27*(3), 267–272.

Finlayson, N. J., Remington, R. W., & Grove, P. M. (2012). The role of presentation method and depth singletons in visual search for objects moving in depth. *Journal of Vision, 12*(8), 13. **1-9**.

Finlayson, N. J., Remington, R. W., Retell, J. D., & Grove, P. M. (2013). Segmentation by depth does not always facilitate visual search. *Journal of Vision, 13*(8), 11.

Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963). Visual resolution and contour interaction. *Journal of the optical society of America, 53*(9), 1026–1032.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics, 65*(7), 999–1010.

Marrara, M. T., & Moore, C. M. (2000). Role of perceptual organization while attending in depth. *Perception and Psychophysics, 62*(4), 786–799.

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320,* 264–265.

O'Toole, A. J., & Walker, C. L. (1997). On the preattentive accessibility of stereoscopic disparity: Evidence from visual search. *Perception & Psychophysics, 59*(2), 202–218.

Parks, N. A., & Corballis, P. M. (2006). Attending to depth: Electrophysiological evidence for a viewer-centered asymmetry. *NeuroReport, 17,* 643–647.

Pashler, H. (1988). Cross-dimensional interaction and texture segregation. *Perception & Psychophysics, 43*(4), 307–318.

Pelli, D. G. (1997). The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 427–442.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109,* 160–174.

Reis, G., Liu, Y., Havig, P., & Heft, E. (2011). The effects of target location and target distinction on visual search in a depth display. *Journal of intelligent Manufacturing, 22,* 29–41.

Remington, R. W., & Pierce, L. (1984). Moving attention: Evidence for time-invariant shifts of visual selective attention. *Attention, Perception & Psychophysics, 35*(4), 393–399.

Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review, 102*(3), 503–532.

Theeuwes, J., Atchley, P., & Kramer, A. F. (1998). Attentional control within 3-D space. *Journal of Experimental Psychology: Human Perception and Performance, 24*(5), 1476–1485.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance, 8*(2), 194–214.

Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 9*(4), 523–530.

Wolfe, J. M. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York: Oxford University Press.

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science, 12*(3), 219–224.