



Published in final edited form as:

J Vis. ; 10(8): 16. doi:10.1167/10.8.16.

The Advantage of a Ground Surface in the Representation of Visual Scenes

Zheng Bian and George J. Andersen

Department of Psychology, University of California, Riverside, Riverside, CA 92521

Abstract

The present study used change detection tasks to examine whether there is an advantage of a ground surface in representing visual scenes. In 6 experiments a flicker paradigm (Experiment 1 through 4) or a one-shot paradigm (Experiment 5 and 6) was used to examine whether changes on a ground surface were easier to detect than changes on a ceiling surface. Overall, we found that: (1) there was an advantage in detecting changes on a ground surface or changes to objects on a ground surface; (2) this advantage was dependent on the presence of a coherent ground surface; (3) this advantage could propagate to objects connected to the ground surface through “nested” contact relations; (4) this advantage was mainly due to improved encoding rather than improved retrieval and comparison of the ground surface; and (5) this advantage was dependent on the presentation duration of the scene but not the number of objects presented in the scene. Together, these results suggest a unique role of the ground surface in organizing visual scenes.

An important goal of vision is to recover a description of the world from visual images that can be used to guide behavior (Marr, 1982). Our phenomenal experience in perceiving the visual world is that we recover a richly detailed description of the environment. However, recent studies have shown considerable limitations in this description. Studies have found that observers have difficulty detecting significant changes in a scene if a change occurs during a saccade (Henderson & Hollingworth, 1999, 2003), during a blink (O’Regan et al, 2000), during a blank interval inserted between an original and modified scene (Rensink, O’Regan & Clark, 1997; Simons, 1996; Hollingworth, Schrock, & Henderson, 2001), during a film cut (Levin & Simons, 1997; Simons, 1996), or during a “mudsplash” between an original and altered scene (O’Regan, Rensink, & Clark, 1999). This phenomenon, referred to as “change blindness”, has also been demonstrated in real world interactions when the observer’s view of a real scene was temporarily blocked (Simons & Levin, 1998).

The results of change blindness studies demonstrate that observers do not recover a coherent and detailed representation of the visual world. Instead, limited information is encoded and available for further processing. Given this limitation are there any principles used by the visual system to organize the description of the visual world? Previous research has suggested that the ground surface is used by the visual system as a common reference frame to encode the distance of objects on the surface (Gibson, 1950; He & Ooi, 2000). In the present study we assessed this hypothesis in detail by examining whether the ground surface is used as the foundation for organizing a description of the visual world with relational information of objects, object parts, location and distance encoded relative to this foundation.

Background surfaces, including the ground and ceiling surface may be used as the foundation for 3D scene representations because they provide layout information of objects within scenes. For example, Bian, Braunstein and Andersen (2005) found that the perceived depth order of two objects could be altered by optical contact with either a ground or ceiling surface. Studies have also found that many visual tasks are performed in accordance with background surface information. These studies include tasks such as visual search (He & Nakayama, 1992), detection of the direction of apparent motion (He & Nakayama, 1994a), texture segregation (He & Nakayama, 1994b), depth from binocular disparity (He & Ooi, 2000), and the perception of subjective contours (Gillam & Nakayama, 2002). Boundary extension, a phenomenon in which observers tend to report seeing more of the background scene than was originally presented in a picture, was found in pictures with scene layout information but not in pictures with a blank background (Gottesman & Intraub, 2002, 2003). Prior experience with a background scene can have a priming effect on judging the layout information in the scene (Sanocki & Epstein, 1997; Sanocki, 2003). Improved encoding has also been found for information related to the layout of a scene (e.g. the position or the presence/absence of objects in a scene) as compared to information less related to the layout of a scene (e.g. the color of the objects) (Aginsky & Tarr, 2000). In addition, imaging studies using fMRI have found that an area in parahippocampal cortex, referred to as the parahippocampal place area or PPA, responded strongly to layout of 3-D scenes but only weakly to arrays of objects without a coherent background surface (Epstein & Kanwisher, 1998). It was suggested that PPA encodes the spatial layout of the local environment (Epstein, 2005). The results of these studies, considered together, suggest that background surfaces are important for the perception and organization of visual scenes.

However, not all background surfaces have the same ecological importance. Many studies have suggested that the ground surface is the most important background surface among all environmental surfaces. The importance of the ground surface in perceiving 3-D space was discussed as early as 1000 years ago in Alhazen's writings (translation: 1989). In addition, Gibson (1950) argued that the ground surface, compared to other environmental surfaces such as ceilings and side walls, serves a unique role in organizing the visual world. The ground surface supports almost all objects and the locomotion of most land-dwelling animals, including human beings (Gibson, 1950). Objects not in direct contact with the ground are usually supported by the ground surface through a series of "nested contact relations" (Meng & Sedgwick, 2001). In addition, the ground surface is universal whereas other surfaces, in addition to the ground surface, are usually present in artificial environments such as buildings.

The importance of the ground surface, as compared to other environmental surfaces, is evident when considering the information available when viewing the 3D world. An important characteristic of the optical projection of light to the eye is that perspective information (changes in the projected angles of a rigid object as a function of distance) is present when viewing a 3D scene. Perspective information is present for any single object that is visible in a scene. Throughout the scene variations in perspective, for objects located at different distances, defines a structure. This perspective structure can be used to define important properties of the visual world, such as the horizon or the slant of surfaces, from gradients. Perspective structure can also be used to define the layout of the scene including the relative distances of objects and is present for any surface receding in depth such as ground and ceiling surfaces. Although the perspective structure can be identical for ground and ceiling surfaces the usefulness of this information to the observer can vary. Consider, for example, a ground surface extended in depth and visible up to a fixed distance. If the image is rotated 180 deg to produce a ceiling surface the perspective structure for the ground and ceiling surface is identical. However the utility of the perspective structure for determining information important to the observer, such as layout, will not be the same.

When the perspective structure defines a ground surface one can derive egocentric distances of objects in the scene by using eyeheight (the distance from the eye of the observer to the ground). Specifically, absolute distance d can be determined from two alternative calculations. Absolute distance can be specified as

$$d=H/\sin \eta,$$

where d is the absolute distance, H is the eyeheight of the observer and η , is the slant of the ground surface (Ooi, Wu and He, 2001; see also Sedgwick, 1986). Alternatively, absolute distance along a textured ground surface can be determined by

$$d=H \times (\cos\alpha_1/\sin\alpha_2) \times \tan(\beta_1/\beta_2),$$

where H is the eyeheight of the observer, α_1 and α_2 are the projected angles from the observer to two texture elements on the ground surface and β_1 and β_2 are the projected extents of the texture elements (i.e. the calculation of $\tan(\beta_1/\beta_2)$ is the texture gradient of the surface).

However, when the perspective structure defines a ceiling surface egocentric distance cannot be determined by the scaling of eye height because the eye height relative to a ceiling surface is undefined. Although slant information can be determined from texture gradients present in ground and ceiling surfaces (since it does not require eyeheight information, see Howard & Rogers, 2002), absolute distance can only be determined for a ground surface (see also Thompson, Dilda & Creem-Regehr, 2007). We are not ruling out other possibilities that absolute distance can be recovered without the use of eye height scaling. For instance, He, Wu, Ooi, Yarbrough, and Wu (2004) proposed a sequential surface integration process (SSIP) that the visual system uses to judge egocentric distance. According to this process, near distance could be accurately recovered by the visual system through near depth cues such as binocular disparity and vergence. This information is then used as an anchor to perceive distance further away. However, the availability of eye height scaling to recover egocentric distance gives the ground surface a unique advantage over other environmental surfaces, which suggests that the ground surface may serve as the foundation for the perceptual organization of scenes.

Recent studies have shown that perceived egocentric distance is mediated by ground surface information by manipulating optical contact (the contact of images in the 2-D projection) between an object and the ground surface (Meng & Sedgwick, 2001, 2002; Ni, Braunstein & Andersen, 2004, 2005, 2007), by the presence of a discontinuity on the ground (Sinai, Ooi & He, 1998; Feria, Braunstein & Andersen, 2003; He et al, 2004; B. Wu, He & Ooi, 2007), by varying the way the ground surface was scanned (B. Wu, Ooi & He, 2004), or by manipulating the area of the ground surface that was the focus of attention (J. Wu, He & Ooi, 2008).

The importance of the ground surface in the perceptual organization of scenes has also been demonstrated by directly comparing the ground surface with other environmental surfaces. For example, McCarley and He (2000, 2001) used a search task in which objects were arranged to form an implicit ground or ceiling surface. They found faster response times when searching implicit ground versus implicit ceiling surface displays. Bian, Braunstein and Andersen (2005) found that when the ground surface and the ceiling surface provided conflicting information about the relative distance of objects in a scene, observers used information from the ground surface to determine the layout of the scene. When the two surfaces were sidewalls, observers did not show a preference to either surface. They referred

to this result as the *ground dominance effect*. In a follow up study, Bian, Braunstein, and Andersen (2006) varied the relative location of the ground surface and the ceiling surface in the visual field, and found that the ground dominance effect was mainly due to the differences in the projections of ground and ceiling surfaces, with visual field location having a minor effect. Recent research has also found a ground dominance effect for older observers, although the magnitude of the effect was smaller than that found for younger observers (Bian & Andersen, 2008).

One possible reason for the unique role of the ground surface in the perceptual organization of 3-D scenes is evolution (He & Nakayama, 1992, 1994a, 1994b, 1995). An important function of the visual system is to encode features separated in space into representations which serve as the basis for higher-level processing. Patterns that occur repeatedly and that are more relevant to human behavior (e.g., locomotion in the world) may be encoded at a faster speed and in greater detail (McCarley & He, 2000). This would enable the visual system to process information in the environment more efficiently. Since the ground surface provides support, either directly or indirectly, to almost all objects and land-dwelling animals, it is possible that the ground surface serves as a common reference frame against which the locations of objects resting on the ground surface are coded (Gibson, 1950; see also Gibson 1979). He and Ooi (2000) showed that a common surface mediated the judged distance between objects on or close to that surface. More recent research has also found that change detection performance improved when the slant of a receding surface was increased (resulting in a slanted surface more similar to a ground surface; Ozkan & Braunstein, 2009). The results of this research, considered together, suggest that scenes may be organized in a hierarchical fashion with the ground surface used as the foundation for organizing a description of the scene, with objects, object parts, locations and distance information encoded relative to this global description. This hypothesis is consistent with recent research suggesting that the spatial representation of scenes is organized in a hierarchical manner (Rolls, Tromans & Stringer, 2008).

In the current study we used change detection tasks to examine in detail the hypothesis that the ground surface serves as the foundation for organizing a description of the scene with objects, object parts, locations and distance information encoded relative to this global description. If the ground surface is used as an organizing principal for the perception of 3-D scenes, then the ground surface, compared to other environmental surfaces, should be encoded more efficiently and in greater detail. For example, consider two different approaches to organizing a representation of the scene --- a hierarchical structure and a structure based on locally coded information. Furthermore, assume that the unit of information is distance between two points in space in a scene-centered coordinate system. The hierarchical representation could be organized with three levels consisting of a background surface as the top level in the structure, followed by the relative distance or spatial layout of objects, and then followed by the distance of object parts relative to the object. The locally coded representation would be organized as a single level with the distance between all object parts in the scene encoded. The hierarchical representation is more efficient than the locally coded representation because less information is required to encode the scene. For example consider a scene consisting of 5 objects with each object containing 3 parts. In the hierarchical representation the first level would consist of one unit of information (overall depth of the scene), the second level (spatial layout of objects) would consist of 10 units of information (distance between all pairwise sets of objects) and the third level (object parts relative to object) would consist of 15 units of information (3 distances for each object). In contrast, the locally coded representation, based on distance of object parts, would consist of 15 units of information (3 units for each object) and 90 units of information (9 units for all 10 pairwise sets of objects). Thus the hierarchical representation would require that 26 units of information be encoded to describe the scene

whereas the locally coded representation would require 105 units of information to describe the scene. If the observer has limited viewing time then a greater proportion of the scene can be encoded in a hierarchical framework than a locally-coded framework. The purpose of this example is not to argue that scenes are encoded precisely in this manner. Rather, the purpose is to demonstrate how a ground surface may facilitate the encoding of the scene by using a hierarchical representation. A similar account was proposed by He and Ooi (2000) arguing that the visual system may encode the location of objects on a common visual surface using a quasi 2-D coordinate system (X, Y) instead of a 3-D Cartesian coordinate system (X, Y, Z). The advantage of using this strategy is that the visual system could encode the relative distances among objects more efficiently with reduced demand for computation. Our proposal compliments their theory by suggesting a hierarchical structure in the representation of both objects and object parts against a ground surface.

To examine this issue we used change detection tasks in which observers compared a current representation of a scene to a stored representation of a previously presented scene. In 6 experiments a flicker paradigm (Rensink, O'Regan & Clark, 1997) or a one-shot paradigm (Rensink, 2002) was used to compare change detection performance for ground and ceiling surfaces. In Experiments 1 and 2, we examined whether changes to a ground surface were easier to detect than changes to a ceiling surface and whether this effect was due to a preference to focus attention to the ground surface. In Experiment 3, we examined whether changes to objects on a ground surface were easier to detect than changes to objects on a ceiling surface and examined whether disrupting the coherent perspective structure of the background surface could affect the ground surface advantage in change detection. In Experiment 4 we examined if the ground surface advantage in detecting a change would propagate to objects not in direct contact with the ground surface. In Experiment 5, we examined whether the ground surface advantage in change detection was due to improved encoding of the ground surface, or whether improved performance could be due to improved retrieval and comparison of the ground surface. Finally, in Experiment 6 we examined the effect of varying the presentation duration and set size on the ground surface advantage.

Experiment 1

The purpose of the first experiment was to examine whether a change on a ground surface was easier to detect than a change on a ceiling surface. The displays simulated a ground and ceiling surface defined by a random checkerboard pattern. A flicker paradigm (Rensink, O'Regan & Clark, 1997) was used to present the original scene (A) and a modified scene (A') in a sequence of A, A, A', A'. The modified scene was produced by changing the luminance of one square in the original scene. If there is a ground surface advantage in detecting a change in the scene then detection performance should be faster and more accurate when the change is on a ground as compared to a ceiling surface.

Method

Observers—The observers were 9 undergraduate students (4 male and 5 female) from the University of California, Riverside. All observers were paid for their participation, were naive regarding the purpose of the experiment, and had normal or corrected-to-normal visual acuity.

Stimuli—The stimuli were computer generated 3-D scenes composed of a ground surface and ceiling surface, each with a 6×6 random black-white square texture. Each square was measured as $238 \text{ cm} \times 238 \text{ cm}$. The average luminance of the stimulus was 60.8 cd/m^2 . Examples of the stimuli are shown in Figure 1. The simulated distances from the observer to the near and far ends of the plane were 571 cm and 2000 cm, respectively. (The calculation of the scene dimensions was based on an eye-height of 120 cm.)

Design—Two independent variables were manipulated: (1) the surface in which a change occurred (ground or ceiling), and (2) the inter-stimulus-interval (ISI) between two consecutive scenes (80 ms, 160 ms, or 240 ms). For “change” trials, 18 squares in the center of each surface were randomly selected to be the candidate targets that changed luminance across consecutive scenes (on each change trial only 1 of the 18 squares changed luminance). This manipulation produced 108 change trials. We also included 36 no-change trials (6 trials for each of 6 combinations) to ensure that observers followed instructions. A total of 144 trials were evenly divided into two blocks. Eight practice trials (6 change trials and 2 no-change trials) were inserted at the beginning of each block. The two experimental blocks were preceded by a 48-trial practice block composed of 36 change trials and 12 no-change trials. The order of the trials for each observer in each block was randomized.

Apparatus—The displays were presented on a 21-inch (53 cm) flat screen CRT monitor with a pixel resolution of 1280 by 1024, controlled by a Windows XP Professional Operating System on a Dell Dimension XPS workstation. The dimensions of the display on the monitor were 40.0 cm (W) × 30.0 cm (H), subtending a visual angle of $31.3^\circ \times 23.7^\circ$. A black viewing hood was placed in front of the monitor to cover the edges of the screen. A 19-cm diameter glass collimating lens, which magnified the images by approximately 19%, was located between the observer and the monitor. The purpose of the collimating lens was to remove accommodation as a flatness cue and thus increase the perceived depth of the 3-D scenes. The distance between the eyes and the collimating lens was approximately 10 cm and the distance from the eyes to the monitor was 85 cm. A chin rest was mounted at a position appropriate to this viewing distance. An optical mouse was used by the observers to initiate each trial and to respond.

Procedure—The experiment was run in a dark room. The observers viewed the display binocularly through the collimating lens with their head position fixed by a chin rest. On each trial, a white cross first appeared in the center of the screen. The observers were instructed to fixate the cross and press the left button of the mouse to initiate the trial. The cross then disappeared and a scene composed of two surfaces was presented. The initial scene (A) and the modified scene (A') were presented for 250 ms each in the sequence of A, A, A', A', with a gray screen presented for various ISI (80 ms, 160 ms, or 240 ms) after each scene (see Figure 2). The purpose of presenting each scene twice was to create temporal uncertainty about when the change occurred (Rensink, O'Regan & Clark, 1997). The initial and modified scenes were thus alternated every 660 ms, 820 ms, or 980 ms, depending on the ISI. The task of the observers was to observe the scenes carefully and continuously and detect the square that changed luminance between two successive scenes. The observers were informed that the target square was likely to appear with equal probability on the ground or ceiling surface. Observers were shown examples of a change on a ground and ceiling surface. They were allowed to move their eyes once a trial began. They were instructed to respond as soon as they found a change by pressing the left button of the mouse, although they were not instructed that the response time was recorded. Twenty-five percent of the trials contained no change in the scene. If observers did not find the target square, they were instructed to continue viewing the scenes. On each trial, the sequence was repeated 10 times or until the observer responded. The average number of alterations that an observer needed to detect a change was recorded. Feedback was not provided during the practice trials or the experiment.

Results and discussion

Due to the relative small number of no-change trials (6 trials for each of the 6 combinations), we did not use the false alarm rate and the hit rate to calculate the sensitivity score (d'). Instead, the hit rate and the number of alterations needed to detect a change were

measured as dependent variables. No-change trials served as “catch trials” to ensure that observers followed instructions. We established an exclusion criterion of a false-alarm rate of 10% or greater. All observers showed a false alarm rate less than 10%.

The hit rate (proportion of change trials detected) was calculated for each subject in each condition and analyzed in a 2 (surface in which a change occurred) by 3 (ISI) analysis of variance (ANOVA). The main effect of surface type was significant ($F(1, 8) = 6.99, p < .05$). The hit rate was 95.1% when the change occurred on a ground surface and 88.5% when the change occurred on a ceiling surface. The main effect of ISI was also significant ($F(2, 16) = 3.86, p < .05$). Post hoc comparisons (Tukey HSD Test) indicated a significant difference ($p < .05$) between the 80 ms (93.8%) and 240 ms (88.9%) ISI conditions. No other pairwise comparisons were significant ($p > .05$). Although the interaction between surface type and ISI was not significant, ($F(2, 16) = 2.10, p = .15$, see Figure 3) there was a trend for the hit rate to decline with greater ISIs for the ceiling surface.

A two-way repeated-measures ANOVA was also conducted on the mean number of alterations required to detect a change. Data from change trials in which no response occurred (“miss” trials) were not included in the analysis. The main effect of surface type was significant, ($F(1, 8) = 10.18, p < .05$, see Figure 4). The mean number of alterations for the ground and ceiling surface was 5.07 and 6.18, respectively. The main effect of ISI ($F(2,16) = 0.12$) and the interaction between ISI and the surface type ($F(2,16) = 0.74$) were not significant ($p > .05$).

Overall the results indicate better performance in detecting a change on a ground surface than on a ceiling surface. Although the accuracy rate was similar between the two surfaces when the ISI was 80 ms, observers required an average of 1.5 more alterations, or 32% more time, to detect a change on a ceiling as compared to ground surface.

Experiment 2

Overall, the results of Experiment 1 were consistent with the hypothesis that a ground surface, compared to other environmental surfaces, serves a special role in the perceptual organization of scenes. We believe that this effect is due to the unique projection of the ground surface. That is, the ground surface recedes in depth from bottom to top of its projected image, whereas the ceiling surface recedes in depth from top to bottom of its projected image. Since a ground surface exists almost everywhere and is utilized more frequently in everyday behavior, our visual system should be more adapted to the projection of a ground surface than to the projection of a ceiling surface. However, this effect may be due to the location of the ground surface in the visual field. Previous studies have found improved performance for processing visual information in the lower, as compared to upper, visual field. These studies include tasks such as visual search (He, Cavanagh, & Intriligator, 1996; Ellison & Walsh, 2000), figure-ground segregation (Rubin, Nakayama & Shapley, 1996; Vecera, Vogel & Woodman, 2002), and visually-guided actions (Danckert & Goodale, 2001). Although we did not restrict eye movements and thus the ground surface may vary location in the visual field during a trial, the ground surface was always located in the lower part of the display whereas the ceiling surface was always located in the upper part of the display.

Another possible explanation for the results of Experiment 1 is that observers were focusing attention more to the ground surface than the ceiling surface. Previous research on change detection has found results suggesting that attention is an important factor in detecting changes in a scene. For example, Rensink, O'Regan & Clark (1997) found that objects rated with higher interest were easier to detect than objects rated with marginal interest.

In Experiment 2 we modified the displays in order to investigate the importance of location in the visual field and attention. The first modification was that one surface (a ground or ceiling) was presented on each trial. This modification ensured that observers were not preferentially attending to one surface over another surface within a trial. If the ground surface advantage observed in Experiment 1 was due to preferential attention to the ground surface, then performance should be similar between the two surfaces. However, if the results of Experiment 1 were due to a ground surface advantage then greater change detection should occur for displays with ground as compared to ceiling surfaces. The second modification was that the surface was presented in the bottom, middle, or top of the display, similar to the stimuli examined in Bian, Braunstein and Andersen (2006). If the ground surface advantage was due to the location of the ground surface in the lower visual field, then performance should be the same for ground and ceiling surfaces. However, if the ground surface advantage obtained in Experiment 1 was due to more efficient encoding of the visual representation, then greater accuracy and faster responses in detecting a change should occur for scenes with a ground surface regardless of the location in the display.

Method

Observers—The observers were 10 undergraduate students (6 male and 4 female) from the University of California, Riverside. All observers were paid for their participation, were naive regarding the purpose of the experiment and had normal or corrected-to-normal visual acuity. Four of the observers had participated in Experiment 1, and their performance was not significantly different from the performance of the remaining observers in Experiment 1.

Stimuli—The stimuli were similar to that used in Experiment 1 with the following exceptions. Only one surface was presented on each trial. The surface was located either in the bottom, middle or the top region of display (the center of the surface was 10.6 cm or 300 pixel below or above the center of the display in the bottom and top condition, respectively). Each surface was composed of 16×8 squares (each square was $89 \text{ cm} \times 179 \text{ cm}$) in order to increase task difficulty. Finally, an 80 ms ISI was used. Examples of the stimuli are shown in Figure 5.

Design—Two independent variables were manipulated: (1) surface type (ground or ceiling), and (2) location of the surface (bottom, middle, or top of display). For “change” trials, 36 out of the 72 squares (12×6) in the center of each surface were randomly selected to be the candidate targets that changed luminance across consecutive scenes. This resulted in 216 “change” trials. We also included 72 “no change” trials (12 replications of each of the 6 combinations) to ensure that observers followed instructions. A total of 288 trials were evenly divided into 3 blocks. Eight practice trials (6 change trials and 2 no-change trials) were inserted at the beginning of each block. The experimental blocks were preceded by a 48-trial practice block composed of 36 change trials and 12 no-change trials. The order of trials for each observer in each block was randomized.

Apparatus and Procedure—The apparatus and the procedure were the same as in Experiment 1. The observers did not know the type of surface or its location in the display before each trial. Feedback was not provided during the practice trials or the experiment.

Results and discussion

Data from one observer (with a false alarm rate greater than 10%) were excluded from the analysis. The hit rate and mean number of alterations in each condition were analyzed in a 2 (surface type) by 3 (location) ANOVA. The main effect of surface type was significant, ($F(1, 8) = 6.31, p < .05$). According to this result the hit rate was greater for the ground surface (mean hit rate of 89.6%) as compared to the ceiling surface (mean hit rate of 86.6%).

The main effect of location ($F(2, 16) = 0.43$) and the interaction of surface type and location ($F(2, 16) = 0.33$; see Figure 6) were not significant ($p > .05$).

A similar pattern of results occurred for the number of alterations needed to detect a change. The main effect of surface type was significant ($F(1, 8) = 8.21, p < .05$). According to this result the mean number of alterations needed to detect a change was greater for the ceiling surface (8.79 alterations or 5806 ms) as compared to the ground surface (8.28 alterations or 5464 ms). The main effect of location was not significant ($F(2, 16) = 0.06, p = .95$). In addition, the interaction between surface type and location was not significant ($F(2, 16) = 1.33, p = .29$), although there was a trend for the ground surface advantage to decrease as the location varied from the bottom region to the top region (see Figure 7). These results, combined with the results from the hit rate analysis, are consistent with the results obtained in Experiment 1 and suggest a ground surface advantage independent of its location in display.

Experiment 3

In Experiments 1 and 2 we found that changes on a ground surface were easier to detect than changes on a ceiling surface and that this effect could not be accounted for by the location of the ground surface in the display or by a preference to allocate more attention to the ground surface. These results suggest that this effect was mainly due to the unique projection of the ground surface. In the current experiment, we examined whether changes to objects supported by a ground surface were easier to detect than changes to objects attached to a ceiling surface, and examined whether this effect was dependent on the presence of a coherent ground surface. Observers were presented with a ground or ceiling surface scene in which eight objects were randomly positioned in the scene. Observers were instructed to detect a change in the location of one of the objects in the scene. If the ground surface serves as a foundation for organizing a description of the visual world, then the properties of objects resting on a ground surface should be encoded by the visual system more efficiently than the properties of objects attached to a ceiling surface. Such properties may include the layout of the scene, the location of objects, the presence of objects, and the distance between objects (Aginsky & Tarr, 2000). Specifically, if there is greater efficiency in encoding object properties when ground surfaces are present then changes to these properties, such as location, should be easier to detect for ground as compared to ceiling surfaces.

Previous research (Varakin & Levin, 2008, Yokosawa & Mitsumatsu, 2003) has found that altering the coherent structure of a scene, by randomly repositioning regions of the scene, results in decreased performance in detecting a change in the scene. In Experiment 3 we examined this issue by randomly repositioning regions of ground and ceiling surfaces and thus altering the perspective structure of the scene. If the ground surface advantage is dependent on a coherent perspective structure of the scene, then we should find a decline in the ground surface advantage as a function of the degree to which the coherent perspective structure of the surface has been disrupted.

Method

Observers—The observers were 20 undergraduate students (8 male and 12 female) from the University of California, Riverside. All observers were paid for their participation, were naive regarding the purpose of the experiment and had normal or corrected-to-normal visual acuity. None of the observers had participated in any of the previous experiments.

Stimuli—In the normal surface condition, the stimuli were computer generated 3-D scenes with either a ground surface or a ceiling surface with a random black and white checkerboard texture. There were also 6 cubes (colored in red, green, and blue with a

simulated size of 42 cm × 94 cm × 18 cm) and 6 pyramids (colored in red and green with a simulated area of 3845 cm² in the bottom triangle and a simulated height of 25 cm) attached to the surfaces. In the jumbled surface conditions, the surfaces were evenly divided into either a 12-section grid (6 × 2), a 48-section grid (12 × 4), or a 196-section grid (24 × 8). Each individual grid was randomly rotated in 0°, 90°, 180°, or 270° and then randomly repositioned (see Figure 8). Since there were no coherent global perspective cues in the jumbled background surfaces, a jumbled ground surface could not be distinguished from a jumbled ceiling surface. However, the objects were not jumbled and thus maintained a layout simulating either an implicit ground surface or an implicit ceiling surface. That is, for jumbled ground-surface displays, the projected size of the objects decreased systematically with increased height in the display. If the jumbled background surface was removed the objects were positioned as if supported by an implicit ground surface. We will refer to this condition as a “ground-like” layout. Similarly, for conditions in which a ceiling surface was jumbled the projected size of the objects increased systematically with increased height in the display. If the jumbled background surface was removed, the objects were positioned as if attached to an implicit ceiling surface. We will refer to this condition as a “ceiling-like” layout (see McCarley & He, 2001).

Design—Two independent variables were manipulated: (1) the configuration of the background surface (normal, jumbled with 6 × 2 grid, 12 × 4 grid or 24 × 8 grid), and (2) the type of layout (ground-like or ceiling-like). For each of the 8 combinations, each of the 12 objects changed its location, producing 96 “change” trials. The direction of the location change was randomly chosen between 0° to 359° relative to the implicit surface, and the amount of change was randomly chosen between 100% and 200% of the length of a square in the normal background surface condition. We also included 32 “no change” trials (4 trials for each of the 8 combinations) to ensure that observers followed instructions. A total of 128 trials were evenly divided into 2 blocks. Eight practice trials (6 change trials and 2 no-change trials) were inserted at the beginning of each block, resulting in 72 trials per block. The experimental blocks were preceded by a 72-trial practice block composed of 54 change trials and 18 no-change trials. The order of trials for each observer in each block was randomized.

Apparatus and Procedure—The apparatus and procedure were the same as that used in Experiments 1 and Experiment 2. Feedback was not provided during the practice trials or the experiment.

Results and discussion

The false alarm rate for all observers was less than 5%. The mean hit rate and response time was analyzed for each subject in each condition and analyzed in a 2 (type of layout) by 4 (configuration) ANOVA. For the hit rate, the main effect of type of layout was not significant, ($F(1, 19) = 1.06, p = .32$). The main effect for the configuration of the background surface was significant ($F(3, 57) = 3.18, p < .05$). Post hoc comparisons (Tukey HSD test) indicated significant differences between the 24 × 8 grid configuration and all other 3 conditions ($p < .05$). The interaction between the configuration and type of layout was not significant ($F(3, 57) = 1.01, p = .40$).

The overall results for response time are presented in Figure 9. The main effect of type of layout was significant, ($F(1, 19) = 6.02, p < .05$). Observers responded faster to ground-like layout displays (5055 ms) as compared to ceiling-like layout displays (5341 ms). The main effect of configuration ($F(3, 57) = 0.20$) and the interaction of configuration and type of layout ($F(3, 57) = 0.70$) was not significant.

A planned-comparison between ground-like and ceiling-like layout was conducted for each level of surface configuration to further examine whether there was a ground-like layout advantage in change detection. With the normal surface configuration, we found that changes to objects on a ground surface were significantly faster to detect than changes to objects on a ceiling surface ($F(1, 19) = 6.22, p < .05$). However, when the perspective structure of the surface was disrupted by jumbling, there was no ground surface advantage in any of the three jumbled conditions ($F(1, 19) = 1.96, p = 0.18$ for 6×2 grid, $F(1, 19) = 0.13, p = 0.72$ for 12×4 grid, and $F(1, 19) = 0.35, p = 0.56$ for 24×8 grid, respectively). These results are consistent with our hypothesis that the ground surface advantage in change detection is due to the perspective structure of the ground surface in organizing the representation of 3-D scenes.

Experiment 4

In the previous experiments we demonstrated that it was easier to detect changes on a ground surface or changes to the objects resting on a ground surface as compared to those on a ceiling surface. However, most 3-D scenes consist of more complicated spatial configurations than those we have investigated so far. Many objects are not directly supported by the ground surface. For example, a book may lie on a table, or a pillow may sit on a chair. Since both the table and the chair are resting on the ground surface, the book and the pillow are connected to each other and ultimately to the ground surface through “nested” contact relations (Meng & Sedgwick, 2001). It has been shown that “nested” contact relations are used in egocentric distance judgments (Meng & Sedgwick, 2001). In the current experiment, we examined whether the ground surface advantage in change detection could propagate to objects connected to the ground surface through “nested” contact relations. If the ground surface is used by the visual system as a common reference frame in organizing 3-D scenes, then objects indirectly supported by the ground surface should also be encoded more efficiently as compared to objects indirectly connected to the ceiling surface.

Method

Observers—The observers were 15 undergraduate students (6 male and 9 female) from the University of California, Riverside. All observers were paid for their participation, were naive regarding the purpose of the experiment and had normal or corrected-to-normal visual acuity. None of the observers had participated in any of the previous experiments.

Stimuli—The stimuli were computer generated 3-D scenes with either a ground surface or a ceiling surface defined by a random black and white checkerboard texture. There were also 8 green cubes, 8 blue slabs, and 8 red posts on the surface. In condition 1, each green cube was attached to a blue slab, which in turn was attached to a red post and positioned on the background surface (see Figure 10). Thus, the green cube was connected to the background surface through two levels of “nested” contact relations. In condition 2, the red post was positioned adjacent to the slab, so that the slab was optically (but not physically) connected to the background surface. Previous studies have found that observers perceive objects as physically attached to a background surface due to optical contact (Gibson, 1950; Meng & Sedgwick, 2001). Thus, in this condition the cube was connected to the surface through one level of “nested” contact relation. In the third condition, both the slab and the post were positioned adjacent to the cube, such that optically the cube was in direct contact with the background surface. The simulated distance of the cubes was the same across these three conditions. As a result the magnitude of change was identical (in visual angle) regardless of the nested contact condition.

Design—Two independent variables were manipulated: (1) type of surface (ground or ceiling), and (2) the level of “nested” contact relations between the cube and the background surfaces (0, 1, or 2). For each of the 6 combinations, each of the 8 cubes changed its location in a randomized direction 3 times, producing 144 “change” trials. We also included 54 “no change” trials (3 replications for each of the 6 combinations) to ensure that observers followed instructions. A total of 198 trials were evenly divided into 3 blocks. Eight practice trials (6 change trials and 2 no-change trials) were inserted at the beginning of each block resulting in 74 trials per block. The experimental blocks were preceded by a 22-trial practice block composed of 16 change trials and 6 no-change trials. The order of trials for each observer in each block was randomized.

Apparatus and Procedure—The apparatus and procedure were similar to that used in Experiments 1, 2 and 3. A flicker paradigm was used. On each trial, one of the cubes might change location from one scene to the next. The direction and amount of location change was similar to that examined in Experiment 3. In addition, when a cube that was positioned on a slab changed location the position change was confined to the surface area of the slab. The observers were instructed to respond as soon as they detected a change by pressing the left button of the mouse. Observers were instructed that the only possible change on each trial was the location of one of the cubes, not the slabs or posts. Observers were shown examples of all three types of nested contact conditions during instructions. All other properties of the scene remained the same across different scenes. Feedback was not provided during the practice trials or the experiment.

Results and discussion

The hit rate and mean response time were derived for each observer in each condition and analyzed in a 2 (type of surface) by 3 (level of nested relations) ANOVA. With regard to accuracy, all observers were very accurate with a hit rate over 95% and false alarm rate less than 5% in all conditions. The ANOVA for hit rate indicated no significant differences ($p > .05$) were found for the main effect of surface type ($F(1, 14) = 0.01$), the main effect of nested relations ($F(2, 28) = 1.18$), and the interaction of surface type and nested relations ($F(2, 28) = 0.57$).

With regard to response time the main effect of surface type was significant, ($F(1, 14) = 12.68, p < .01$). The mean response time for the ground and ceiling surface conditions were 4461 ms and 4750 ms, respectively. The main effect for nested relation ($F(2, 28) = 0.88, p = .43$) was not significant. The interaction between surface type and nested relation ($F(2, 28) = 1.72, p = .20$) was not significant, either (see Figure 11). Planned comparisons were conducted to examine the effects of the ground surface at each level of “nested” contact relations. When the objects were connected directly to the background surface or through one level of “nested” contact relation, there was a significant ground surface advantage ($F(1, 14) = 6.67, p < .05$, and $F(1, 14) = 10.26, p < .01$, respectively). There was no significant difference between detection performance for the two surfaces ($F(1, 14) = 0.02, p = .90$) when the objects were connected to the background surface through two levels of “nested” contact relations. This result, consistent with our hypothesis, suggests that objects not in direct contact with the ground surface can benefit from improved encoding of the ground surface by the visual system. However, this benefit is limited to the first level of a “nested” contact relation.

Experiment 5

In the previous experiments we found that changes on a ground surface or changes to the objects on a ground surface were easier to detect than those on a ceiling surface. These results suggest that the ground surface and the objects on a ground surface were encoded

more efficiently by the visual system. However, this effect could also be due to greater efficiency in retrieving and comparing information when a ground surface is present. Recent studies have shown that change blindness could be due, in part, to a failure of retrieval and comparison processes (Hollingworth, 2003). In the current experiment we examined whether a difference in retrieval and comparison processes contributed to the ground surface advantage in change detection using a one-shot paradigm (Rensink, 2002). On each trial, an initial scene was presented for 20 sec, followed by a brief mask and then followed by a test scene. The observers were instructed to judge whether the two scenes were identical or different (i.e., a change in location of one object). A postcue manipulation was used to examine whether the ground surface advantage was the result of retrieval and comparison processes (Hollingworth, 2003). For half the trials a cue was presented in the test scene. For this condition observers were instructed to judge whether a change in location had occurred for the cued object. If the ground surface advantage is due to more efficient retrieval and comparison then the ground surface advantage should decrease when a post-cue is presented because of decreased memory load for the test scene.

Method

Observers—The observers were 11 undergraduate students (5 male and 6 female) from the University of California, Riverside. All observers were paid for their participation, were naive regarding the purpose of the experiment and had normal or corrected-to-normal visual acuity. None of the observers had participated in any of the previous experiments.

Stimuli—The stimuli were computer generated 3-D scenes with either a ground surface or a ceiling surface, each with a 32×16 random black and white checkerboard texture. Each texture element measured $44.6 \text{ cm} \times 89.2 \text{ cm}$. The simulated distance from the observer to the near and far ends of the plane were 571 cm and 2000 cm, respectively (based on an eye-height of 120 cm). Eight objects (cubes and pyramids colored in blue and green) were randomly located on the surface. The locations of the objects were randomized across trials. Examples of the stimuli are shown in Figure 12.

Design—Two independent variables were manipulated: (1) surface type (ground or ceiling), and (2) postcue (present or absent). A total of 64 “change” trials were presented (16 replications for each of the four combinations). We also included 64 “no change” trials (16 replications for each of the 4 combinations) to ensure that observers followed instructions. A total of 128 trials were divided into 2 blocks. Eight practice trials (4 change trials and 4 no-change trials) were inserted at the beginning of each block resulting in 72 trials per block. The experimental blocks were preceded by a 32-trial practice block composed of 16 change trials and 16 no-change trials. The order of trials for each observer in each block was randomized.

Apparatus—The apparatus was the same as that used in previous experiments.

Procedure—On each trial, a white cross first appeared in the center of a black background. The observers were instructed to fixate the cross and press the left button of the mouse to initiate the trial. The cross then disappeared and a scene composed of a surface with 8 objects was presented. The initial scene was displayed for 20 sec during which time the observers were instructed to remember the location of each object and their spatial configuration. The initial scene was replaced by a mask for 200 ms followed by the test scene. The test scene was either identical to the initial scene, or differed from the initial scene in the location of one object. The task of the observer was to determine if there was a change in the test scene as compared to the initial scene. The observers were instructed that the only change in the scene would be a change in the location of one object. If they detected

a change they were instructed to click the left mouse button. If they did not detect a change they were instructed to click the right mouse button. In the postcue present condition, a red arrow was presented in the test scene that pointed to one of the objects (see Figure 13). For this condition observers were instructed to determine whether the object that was cued had changed location. Feedback was not provided during the practice trials or the experiment.

Results and discussion

We calculated the proportion of hits and false alarms and derived a sensitivity score (d') for each observer in each condition. The d' scores were subsequently analyzed in a 2 (surface type) by 2 (postcue) ANOVA with repeated measures. The main effect of surface type was significant ($F(1, 10) = 20.24, p < .01$). According to this result sensitivity was higher for ground (mean d' of 1.73) as compared to ceiling (mean d' of 1.25) surfaces. The main effect of postcue was significant ($F(1, 10) = 5.58, p < .05$). According to this result sensitivity was higher for the postcue present (mean d' of 1.65) as compared to postcue absent (mean d' of 1.33) conditions. This result is consistent with previous research that found that change blindness was partially due to a failure of retrieval and comparison processes (Hollingworth, 2003). The interaction between surface type and postcue was not significant ($F(1, 10) = 0.78, p = .40$, see Figure 14), suggesting that the ground surface advantage is not due to retrieval and comparison processes. Instead, the results suggest that the ground surface advantage is likely the result of more efficient encoding of the scene.

Experiment 6

The results of Experiment 5 failed to find evidence that the ground surface advantage in change detection was the result of improved retrieval and comparison processes and suggest that the ground surface advantage could be due to greater efficiency in encoding the scene. In Experiment 6 we directly tested this hypothesis. If the scene is encoded with greater efficiency when a ground surface as compared to a ceiling surface is present, then the ground surface advantage should vary as a function of the time available for encoding the scene. Specifically, the background surface may not be completely encoded when the display duration is short. As a result, there should be little or no advantage of a ground surface. As the presentation time is increased, the representation of the ground surface scene should be more complete if the encoding of the scene is more efficient. Thus, a larger ground surface advantage should occur with longer display durations.

An additional issue that has not been addressed is whether there is an optimal number of objects that can be encoded when a ground surface is present. Studies on visual memory have found that observers optimally code 4 objects in a scene (Luck & Vogel, 1997). In Experiment 3 there were 12 objects present in the display. In Experiment 4 and 5 there were 8 objects present in the display. In the current experiment, we examined whether the ground surface advantage varied with the number of objects present in the scene.

In the current experiment we examined these issues using a one-shot change detection paradigm similar to the paradigm used in Experiment 5. Specifically, an initial scene was presented for a period of time, followed by a mask and then followed by a second scene. Observers had to indicate whether the second scene was the same or different from the initial scene. We systematically varied the presentation duration of the initial scene (from 100 ms to 5000 ms) and the number of objects presented (from 3 to 6).

Method

Observers—The observers were 12 undergraduate students (5 male and 7 female) from the University of California, Riverside. All observers were paid for their participation, were

naive regarding the purpose of the experiment and had normal or corrected-to-normal visual acuity. None of the observers had participated in any of the previous experiments.

Stimuli—The stimuli were similar to those in Experiment 5, with the exception that no post cue was presented.

Design—Three independent variables were manipulated: (1) surface type (ground or ceiling), (2) the presentation duration of the initial scene (100 ms, 250 ms, 500 ms, 1 sec, 2.5 sec, or 5 sec), and (3) the number of objects present (3, 4, 5, or 6 objects). For each of the 12 combinations of the surface presented and the presentation duration, each of the objects in the 4 set size conditions changed location in a randomized direction for 8, 6, 5, and 4 times, respectively. This produced 1164 “change” trials. An equal number of “no change” trials were also included. A total of 2328 trials were evenly divided into 6 blocks, with 394 trials in each block containing one level of presentation duration. The order of the presentation duration was counterbalanced across observers with a Latin Square design. Two practice trials (1 change trial and 1 no-change trial) were inserted at the beginning of each block for a total of 396 trials per block. Each participant finished the experiment either in 2 sessions with 1.5 hours per session or in 3 sessions with 1 hour per session, depending on the order of the presentation durations. The experimental blocks were preceded by a 32-trial practice block composed of 16 change trials and 16 no-change trials. The order of trials for each observer in each block was randomized.

Apparatus—The apparatus was the same as that used in Experiments 1 - 5.

Procedure—The procedure was similar to that used in Experiment 5. Observers had to determine if there was a change in the test scene as compared to the initial scene and to respond by clicking one of two mouse buttons. The observers were instructed that the only change in the scene was the location of an object. Feedback was not provided during the practice trials or the experiment.

Results and discussion

We calculated the proportion of hits and false alarms and derived a sensitivity score (d') for each observer in each condition. The d' scores were subsequently analyzed in a 2 (surface type) by 6 (duration) by 4 (number of objects) ANOVA. The main effect of surface type was significant ($F(1, 11) = 18.24, p < .01$). According to this result, sensitivity to detect a change was greater when the scene was a ground (mean d' of 1.65) as compared to ceiling surface (mean d' of 1.46). The main effect of duration was significant ($F(5, 11) = 21.25, p < .01$). The mean d' for the 100 ms, 250 ms, 500 ms, 1 sec, 2.5 sec, and 5 sec duration conditions were 0.87, 1.34, 1.60, 1.75, 1.83, and 1.94, respectively. Post hoc comparisons (Tukey HSD test) indicated significant differences between the 100 ms duration and all other durations examined, and between the 250 ms and the 1, 2.5, and 5 sec duration conditions. These results indicate that performance increased with an increase in the presentation duration of the initial scene. The main effect of set size was significant ($F(3,30) = 25.7, p < .05$). The mean d' for the 3, 4, 5, and 6 object conditions were 1.75, 1.79, 1.44, and 1.25, respectively. Post hoc comparisons (Tukey HSD test) indicated significant differences between the 3 object and the 5 and 6 object conditions, and between the 4 object and the 5 and 6 object conditions. These results indicate that change detection performance declined when the number of objects was greater than 4.

The interaction between the surface type and the presentation duration was significant ($F(5, 55) = 2.94, p < .05$, see Figure 15). An examination of the simple effects for presentation duration indicated that there was no significant effect ($p > .05$) of surface type for durations

of 100 ms ($F(1, 11) = 1.87$), 250 ms ($F(1, 11) = 0.17$), and 500 ms ($F(1, 11) = 0.02$) conditions. In contrast, there was a significant effect ($p < .05$) of surface type for the 1000 ms ($F(1, 11) = 22.75$), 2500 ms ($F(1, 11) = 12.92$), and 5000ms ($F(1, 11) = 14.18$) presentation durations. These results indicate that the ground surface advantage occurred for durations greater than 500 ms and thus the time required to encode the surface and the relative locations of objects in the scene is greater than 500 ms.

Although the main effect of set size was also significant the interaction of set size and surface type was not significant ($F(3, 33) < 1$, see Figure 16). These results, considered together, suggest that the ground surface advantage is not dependent on the amount of information to be encoded, at least for the conditions examined in the current experiment. No other significant effects were found.

General Discussion

In the present study we tested the hypothesis that the ground surface is used as an organizing principle for a description of the visual world. In Experiment 1 we presented observers with ground and ceiling surfaces and found a significant advantage for both hit rate and response time when detecting changes on a ground surface. In Experiment 2, we presented observers with only one surface on each trial and varied the location of the surface in the display. The results indicated a significant advantage in hit rate and response time for detecting changes on the ground surface similar to the results of Experiment 1. One interesting trend in the RT results occurred when the location of the surface in the display was systematically varied in this experiment. Although the interaction between the surface and the location was not significant, the fastest RTs for the ceiling surface occurred when the surface was in the top part of the display. In a similar fashion the fastest RT for the ground surface occurred when the surface was in the bottom of the display. This suggests that there may be a top-down influence on how environmental surfaces are used for the organization of visual scenes. An important issue for future research will be to examine the role of top-down processing on the perception of scenes with environmental surfaces.

In Experiment 3 we demonstrated that changes to objects on a ground surface were easier to detect than those on a ceiling surface, and that this advantage was dependent on a coherent ground surface. In Experiment 4, we demonstrated that the ground surface advantage in change detection was not limited to the objects in direct contact with the ground surface. The benefit of a ground surface could propagate to objects connected to the surface through “nested” contact relations. The results of Experiment 4 indicate that the ground surface advantage could propagate to one level of nested relations but not to additional levels of nested relations. One explanation for this result is that the 2 level nested relation stimuli (which consisted the object positioned on a slab which was located on a post) may not have provided as stable a structure as the one level nested relation (which consisted of an object positioned on a slab). An important issue for future research will be to examine whether stability of nested relations is important for the propagation of the ground surface effect.

In Experiment 5 we examined whether differences in retrieval and comparison contributed to the ground surface advantage in change detection. We found a similar benefit for both the ground and ceiling surface when a postcue was presented, suggesting that the superior performance in detecting changes to a scene with a ground surface was due to improved encoding, rather than improved processing for comparison and retrieval, of the scene by the visual system.

In Experiment 6 we manipulated the presentation duration of the initial scene to directly examine whether the ground surface advantage was due to more efficient encoding. In

addition, we examined whether the ground surface advantage was dependent on the number of objects in the scene. The results indicate that detection performance was greater for ground surfaces, as compared to ceiling surfaces, for presentation durations greater than 500 ms. These results suggest that the visual system requires more than 500 ms to derive a representation of the scene sufficient for a ground surface advantage in change detection. In addition, presentation durations greater than 500 ms resulted in greater sensitivity for a ground surface but not for a ceiling surface.

These results extend the findings from previous studies demonstrating the importance of background surfaces in organizing 3-D scenes (Gottesman & Intraub, 2002, 2003; He & Nakayama, 1992, 1994a, 1994b, 1995; He & Ooi, 2000; Sanocki & Epstein, 1997; Sanocki, 2003), and demonstrate that the ground surface serves a foundational role in organizing a description of the visual world. In addition, the results demonstrate that scenes with a ground surface are encoded more efficiently than scenes with a ceiling surface. The results are also consistent with previous studies demonstrating the preferential processing of the information on a ground surface over a ceiling surface in layout judgments (Bian, Braunstein, & Andersen, 2005, 2006; Bian & Andersen, 2008) and in visual search (McCarley & He, 2000, 2001). Our results suggest that the representation of 3-D scenes is encoded more efficiently when a ground surface is present as compared to when a ceiling surface is present. The greater efficiency for encoding scenes when a ground surface is present may be due to the type of representation used. For instance, it is possible that the ground surface is encoded using a hierarchical structure in which the ground surface is encoded at the top level of the structure with objects and object parts encoded at more subordinate levels of the representation. Scenes with a ceiling surface, on the other hand, may be encoded in a single level with local information of objects and parts of objects encoded at the same level. As discussed earlier a hierarchical structure requires less information to encode the scene and thus can be encoded with greater efficiency than a locally coded representation. An important topic for future research will be to examine the type of the representation used for perception of scenes and how it is encoded when different environmental surfaces are present.

For an object positioned on the ground surface, the height of its projected image increases with increasing distance from the observer. If the object is attached to a ceiling surface, however, the height of its projected image decreases as it is located further away from an observer. That is, the height in the picture cue is consistent with the layout of objects on a ground surface but not with the layout of objects on a ceiling surface. Hence, ground and ceiling surfaces have a symmetrically reversed optical pattern. Despite such similarity in geometry, the two surfaces are quite different ecologically. A ground surface is universal and supports most objects in the environment. In addition, ground surfaces are used for locomotion of most animals. A ceiling surface, however, has much less importance in everyday life. Our visual system, being highly adaptive, may have learned to encode more behaviorally relevant optical patterns with a higher efficiency in order to achieve maximum economy in processing visual information (McCarley & He, 2000). One demonstration This suggests that ground surface information may be used as a foundation for organizing scenes and the relations of components of a scene (objects, object parts, locations and distance). In addition, this suggests that, when looking at a scene, the perceptual organization of the scene may be optimal when ground surface information is present. The results from this study, considered in light of other studies showing a ground surface advantage over other environmental surfaces in various visual tasks (Bian, Braunstein & Andersen, 2005, 2006; McCarley & He, 2000, 2001), suggest that the ground surface serves a critical role in the perceptual organization of 3-D scenes.

In summary, the results of the present study indicate an advantage of the ground surface over the ceiling surface in detecting a change in the visual world. This advantage was not due to the location of the ground surface in the display. Our results, considered together with the results of previous studies examining the importance of the ground surface in perceiving 3-D scenes (Meng & Sedgwick, 2001, 2002; McCarley & He, 2000, 2001; Ni, Braunstein & Andersen, 2004, 2005, 2007; Sinai, Ooi & He, 1998; Feria, Braunstein & Andersen, 2003; He et al, 2004; Wu, Ooi & He, 2004), indicate a unique role of the ground surface in the perceptual organization of visual scenes.

Acknowledgments

This research was supported by NIH AG031941 and EY18334. We thank three anonymous reviewers for comments on an earlier draft of the paper.

References

- Aginsky V, Tarr MJ. How are different properties of a scene encoded in visual memory? *Visual Cognition*, 7(1/2/3). 2000:147–162.
- Alhazen, I.; Sabra, AI. *The optics of Ibn-Haytham*. Vol. 1. Warburg Institute; London: University of London: 1989. Book of optics. (Date of original work unknown, approx. 11th century)
- Bian Z, Andersen GJ. Aging and the perceptual organization of 3-D scenes. *Psychology & Aging*. 2008; 23(2):342–352. [PubMed: 18573008]
- Bian Z, Braunstein ML, Andersen GJ. The ground dominance effect in the perception of 3-D layout. *Perception & Psychophysics*. 2005; 67(5):815–828.
- Bian Z, Braunstein ML, Andersen GJ. The ground dominance effect in the perception of relative distance in 3-D scenes is mainly due to characteristics of the ground surface. *Perception & Psychophysics*. 2006; 68(8):1297–1309. [PubMed: 17378416]
- Danckert J, Goodale MA. Superior performance for visually guided pointing in the lower visual field. *Experimental Brain Research*. 2001; 137:303–308.
- Ellison A, Wash V. Visual field asymmetries in attention and learning. *Spatial Vision*. 2000; 14:3–9. [PubMed: 11334180]
- Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature*. 1998; 592:598–601. [PubMed: 9560155]
- Epstein R. The cortical basis of visual scene processing. *Visual Cognition*. 2005; 12(6):954–978.
- Feria CS, Braunstein ML, Andersen GJ. Judging distance across texture discontinuities. *Perception*. 2003; 32:1423–1440. [PubMed: 14992421]
- Gibson, JJ. *The perception of the visual world*. Houghton-Mifflin; Boston: 1950.
- Gibson, JJ. *The ecological approach to visual perception*. Houghton-Mifflin; Boston: 1979.
- Gillam B, Nakayama K. Subjective contours at line terminations depend on scene layout analysis, not image processing. *Journal of Experimental Psychology: Human Perception and Performance*. 2002; 28(1):43–53.
- Gottesman CV, Intraub H. Surface construal and the mental representation of scenes. *Journal of Experimental Psychology: Human Perception and Performance*. 2002; 28(3):589–599. [PubMed: 12075890]
- Gottesman CV, Intraub H. Constraints on spatial extrapolation in the mental representation of scenes: view-boundaries vs. object-boundaries. *Visual Cognition*. 2003; 10(7):875–893.
- He S, Cavanagh P, Intrilligator J. Attentional resolution and the locus of visual awareness. *Nature*. 1996; 383:334–337. [PubMed: 8848045]
- He ZJ, Ooi TL. Perceiving binocular depth with reference to a common surface. *Perception*. 2000; 29:1313–1334. [PubMed: 11219987]
- He ZJ, Nakayama K. Surfaces versus features in visual search. *Nature*. 1992; 359:231–233. [PubMed: 1528263]

- He ZJ, Nakayama K. Apparent motion determined by surface layout not by disparity or three-dimensional distance. *Nature*. 1994a; 367:173–175. [PubMed: 8114913]
- He ZJ, Nakayama K. Perceiving textures: beyond filtering. *Vision Research*. 1994b; 34(2):151–162. [PubMed: 8116275]
- He ZJ, Nakayama K. Visual attention to surfaces in 3-D space. *Proceedings of the National Academy of Sciences*. 1995; 92:11155–11159.
- He ZJ, Wu B, Ooi TL, Yarbrough G, Wu J. Judging egocentric distance on the ground: Occlusion and surface integration. *Perception*. 2004; 33:789–806. [PubMed: 15460507]
- Henderson JM, Hollingworth A. The role of fixation position in detecting scene changes across saccades. *Psychological Science*. 1999; 10(5):438–443.
- Henderson JM, Hollingworth A. Global transsaccadic change blindness during scene perception. *Psychological Science*. 2003; 14(5):493–497. [PubMed: 12930482]
- Hollingworth A. Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*. 2003; 29(2):388–403. [PubMed: 12760623]
- Hollingworth A, Schrock G, Henderson JM. Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition*. 2001; 29(2):296–304. [PubMed: 11352212]
- Howard, IP.; Rogers, BJ. Seeing in depth. Vol. 2. University of Toronto Press; 2002. depth perception.
- Levin DT, Simons DJ. Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*. 1997; 4(4):501–506.
- Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. *Nature*. 1997; 390:279–281. [PubMed: 9384378]
- Marr, D. *Vision*. W. H. Freeman; New York: 1982.
- McCarley JS, He ZJ. Asymmetry in 3-D perceptual organization: Ground-like surface superior to ceiling-like surface. *Perception & Psychophysics*. 2000; 62:540–549. [PubMed: 10909244]
- McCarley JS, He ZJ. Sequential priming of 3-D perceptual organization. *Perception & Psychophysics*. 2001; 63:195–208. [PubMed: 11281096]
- Meng JC, Sedgwick HA. Distance perception mediated through nested contact relations among surfaces. *Perception & Psychophysics*. 2001; 63:1–15. [PubMed: 11304007]
- Meng JC, Sedgwick HA. Distance perception across spatial discontinuities. *Perception & Psychophysics*. 2002; 64:1–14. [PubMed: 11916293]
- Ni R, Braunstein ML, Andersen GJ. Perception of scene layout from optical contact, shadows, and motion. *Perception*. 2004; 33:1305–1318. [PubMed: 15693673]
- Ni R, Braunstein ML, Andersen GJ. Distance perception from motion parallax and ground contact. *Visual Cognition*. 2005; 12(6):1235–1254.
- Ni R, Braunstein ML, Andersen GJ. Perception of scene layout from ground contact, occlusion and motion parallax. *Visual Cognition*. 2007; 15(1):48–68.
- O'Regan JK, Deubel H, Clark JJ, Rensink RA. Picture changes during blinks: looking without seeing and seeing without looking. *Visual Cognition*. 2000; 7:191–211.
- O'Regan JK, Rensink RA, Clark JJ. Change blindness as a result of 'mud-splashes.'. *Nature*. 1999; 398:34–34. [PubMed: 10078528]
- Ooi TL, Wu B, He ZJ. Distance determined by the angular declination below the horizon. *Nature*. 2001; 414:197–200. [PubMed: 11700556]
- Ozkan K, Braunstein ML. Change detection for objects on surfaces slanted in depth [Abstract]. *Journal of Vision*. 2009; 9(8):916–916a. <http://journalofvision.org/9/8/916/>, doi:10.1167/9.8.916.
- Rensink RA. Change detection. *Annual Review of Psychology*. 2002; 53:245–277.
- Rensink RA, O'Regan JK, Clark JJ. To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*. 1997; 8(5):368–373.
- Rolls ET, Tromans JM, Stringer SM. Spatial scene representations formed by self-organizing learning in a hippocampal extension of the ventral visual system. *European Journal of Neuroscience*. 2008; 28(10):2116–2127. [PubMed: 19046392]

- Rubin N, Nakayama K, Shapley R. Enhanced perception of illusory contours in the lower versus upper visual hemifields. *Science*. 1996; 271:651–653. [PubMed: 8571128]
- Sanocki T, Epstein W. Priming spatial layout of scenes. *Psychological Science*. 1997; 8(5):374–378.
- Sanocki T. Representation and perception of scenic layout. *Cognitive Psychology*. 2003; 47:43–86. [PubMed: 12852935]
- Sedgwick, HA. Space perception. In: Boff, KR.; Kaufman, L.; Thomas, JP., editors. *Handbook of perception and human performance*. Wiley; New York: 1986. p. 21-1-21-57.
- Simons DJ. In sight, out of mind: when object representations fail. *Psychological Science*. 1996; 7(5): 301–305.
- Simons DJ, Levin DT. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*. 1998; 5(4):644–649.
- Sinai MJ, Ooi TL, He ZJ. Terrain influences the accurate judgment of distance. *Nature*. 1998; 395:497–500. [PubMed: 9774104]
- Thompson WB, Dilda V, Creem-Regehr SH. Absolute distance perception to locations off the ground plane. *Perception*. 2007; 36:1559–1571. [PubMed: 18265839]
- Varakin DA, Levin DT. Scene structure enhances change detection. *The Quarterly Journal of Experimental Psychology*. 2008; 61(4):543–551. [PubMed: 18300186]
- Vecera SP, Vogel EK, Woodman GF. Lower region: a new cue for figure-ground assignment. *Journal of Experimental Psychology: General*. 2002; 131:194–205. [PubMed: 12049239]
- Wu B, He ZJ, Ooi TL. In accurate representation of the ground surface beyond a teture boundary. *Perception*. 2007; 36:703–721. [PubMed: 17624117]
- Wu B, Ooi TL, He ZJ. Perceiving distance accurately by a directional process of integrating ground information. *Nature*. 2004; 428:73–77. [PubMed: 14999282]
- Wu J, He ZJ, Ooi TL. Perceived relative distance on the ground affected by the selection of depth information. *Perception & Psychophysics*. 2008; 70(4):707–713. [PubMed: 18556932]
- Yokosawa K, Mitsumatsu H. Does disruption of a scene impaire change detection? *Journal of Vision*. 2003; 3:41–48. [PubMed: 12678624]

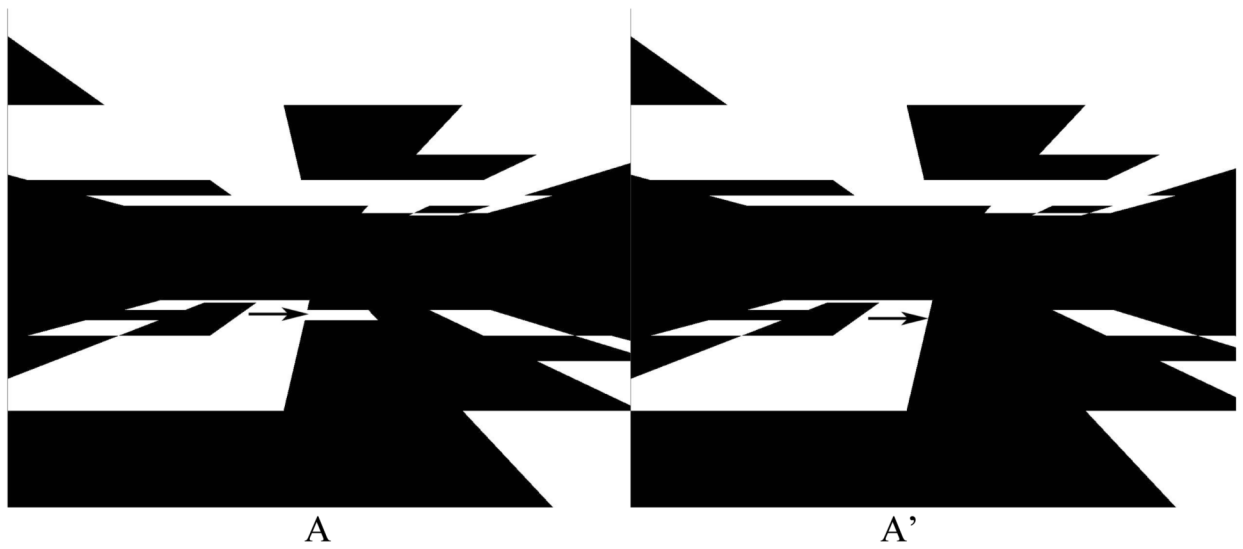


Figure 1. Example of the stimuli used in Experiment 1. (A) is the original scene and (A') is the modified scene. The black arrows point to the square that changed luminance between two scenes and were not present in the displays.

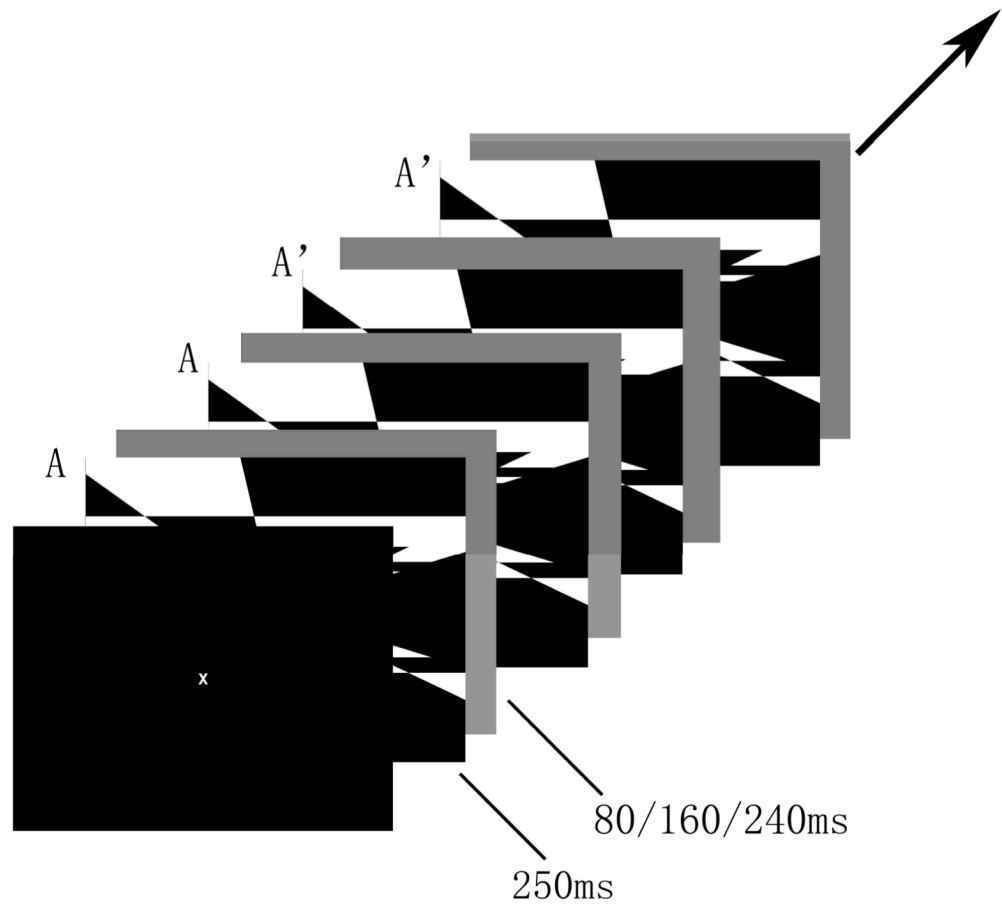


Figure 2. Example of the sequence of each trial used in Experiment 1. A trial continues for 20 alterations (40 scenes) or until the subject responds.

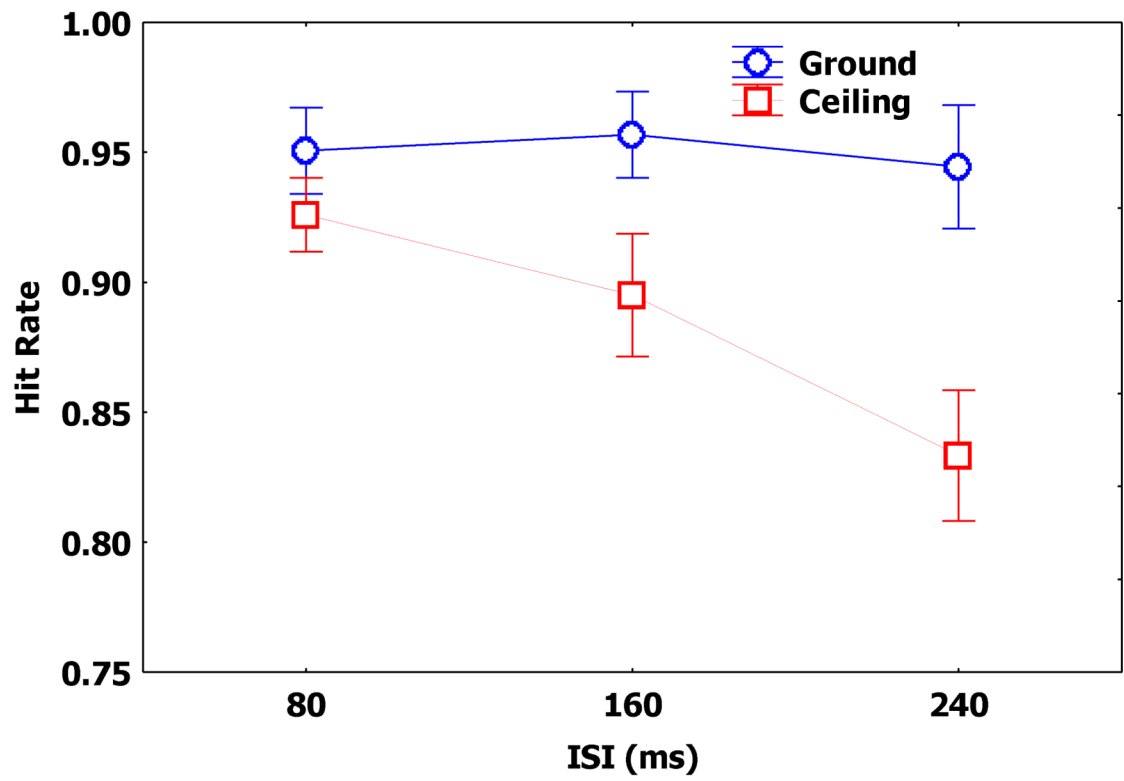


Figure 3. Hit rate as a function of ISI and surface type from Experiment 1. Error bars represent ± 1 standard error.

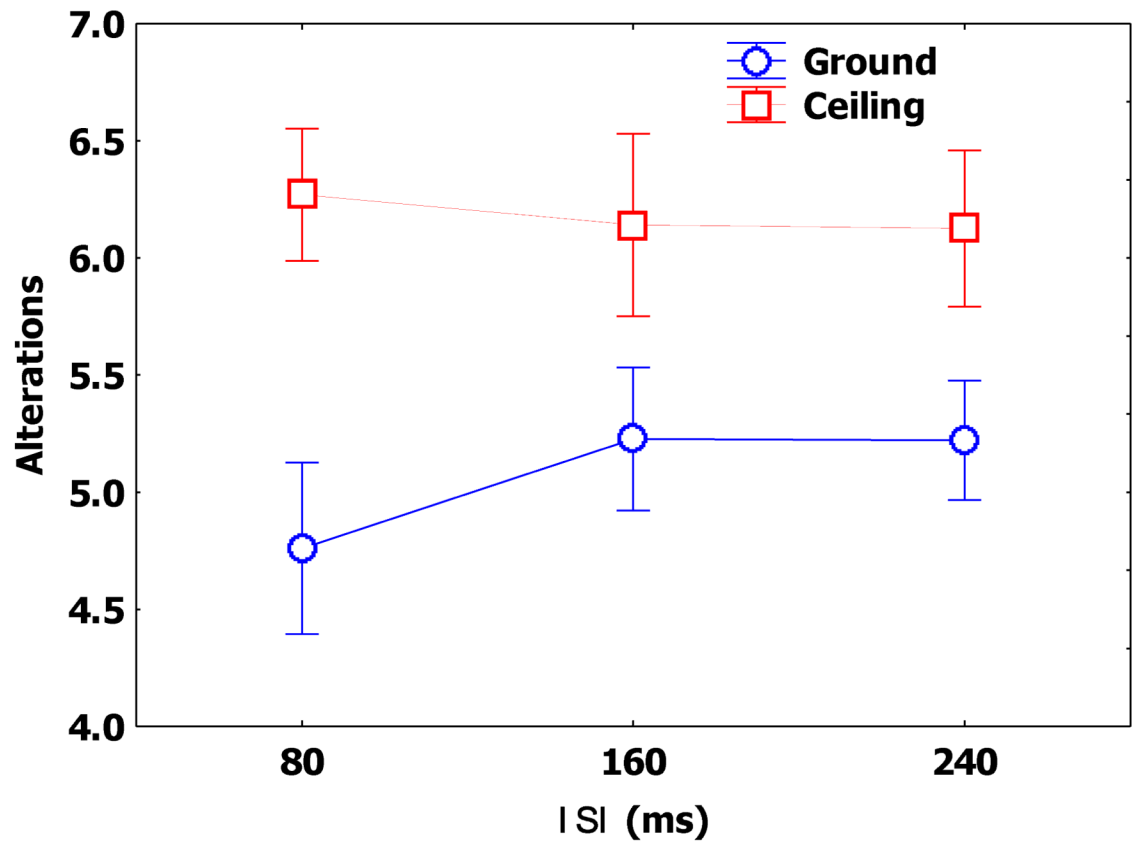


Figure 4. The number of alterations needed for change detection as a function of ISI and surface type from Experiment 1. Error bars represent ± 1 standard error.

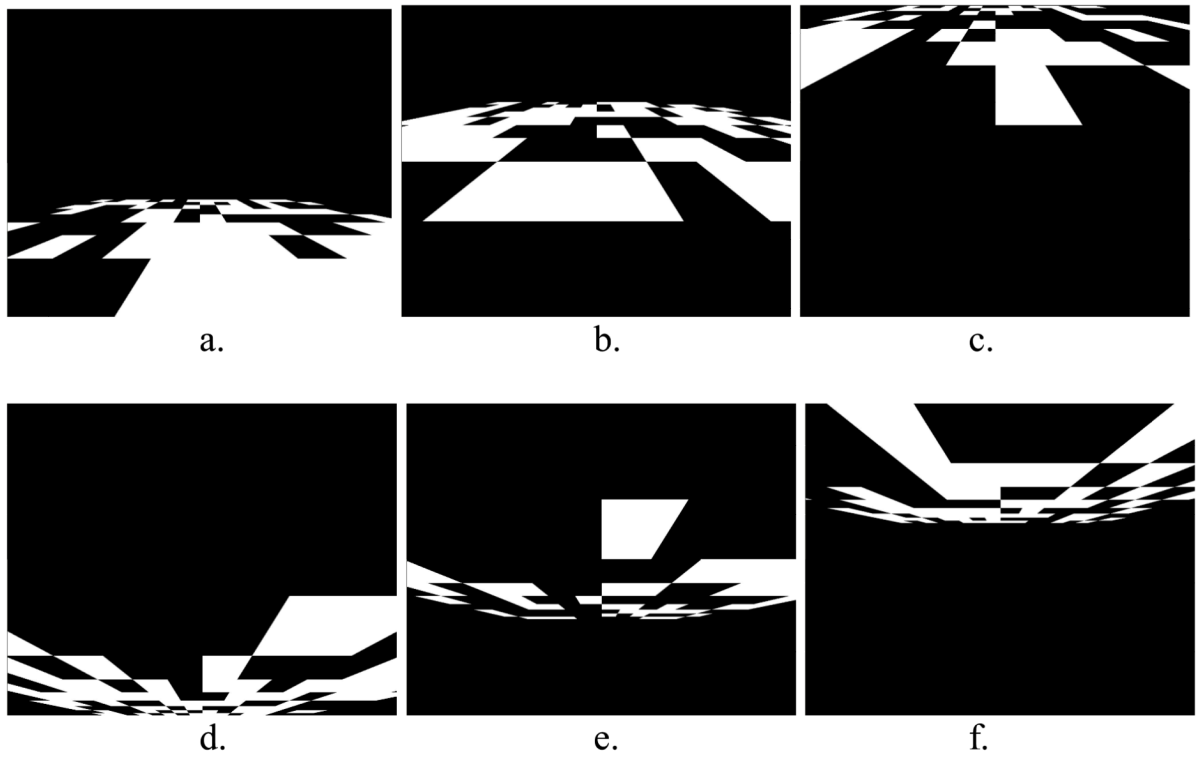


Figure 5. Examples of the stimuli used in Experiment 2. (a) – (c) ground surface at bottom, middle and top of display; (d - f) ceiling surface at bottom, middle and top of display.

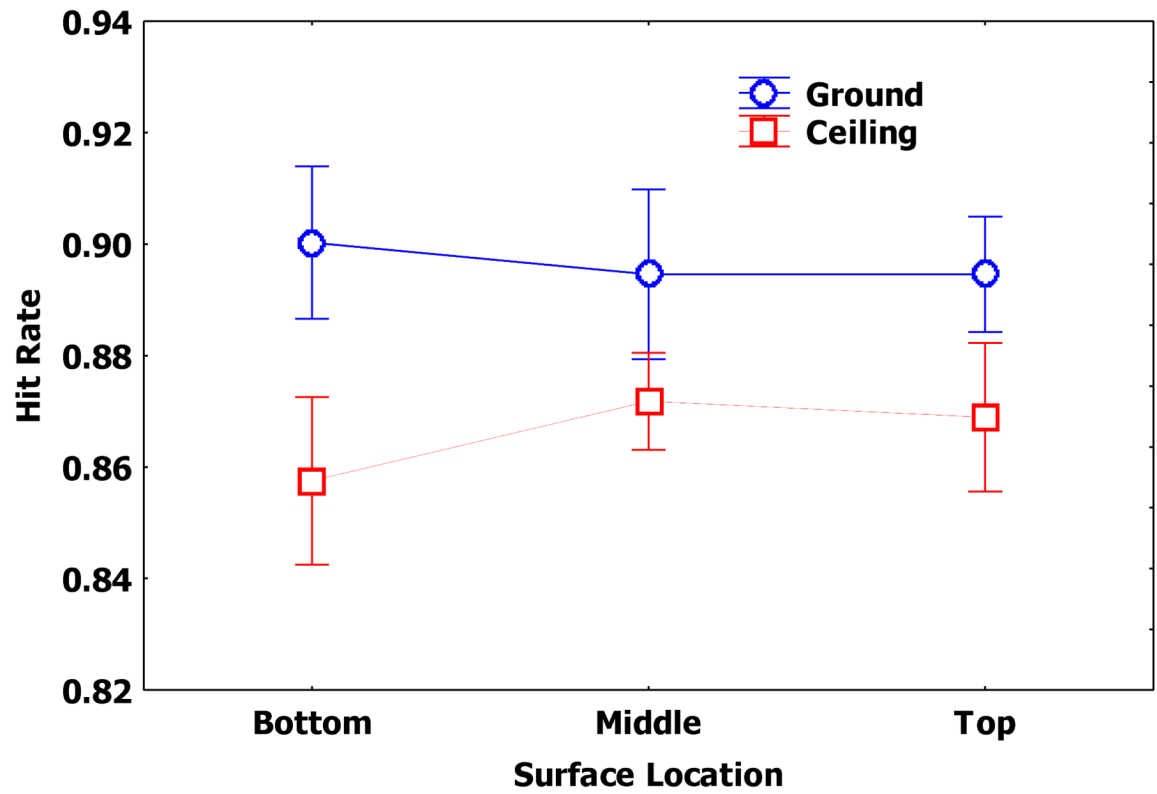


Figure 6.
The hit rate as a function of type of surface presented and display location from Experiment 2. Error bars represent ± 1 standard error.

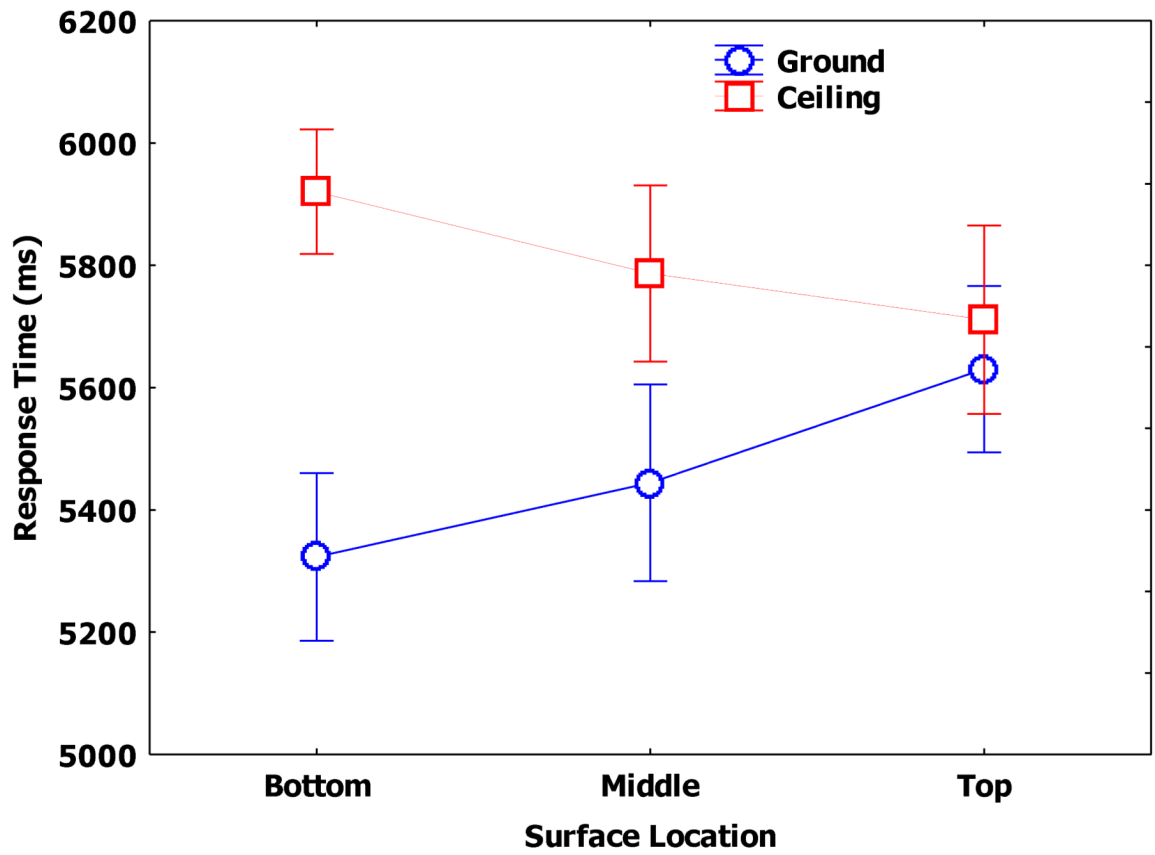


Figure 7. Mean response time as a function of the type of surface presented and location in the display from Experiment 2. Error bars represent ± 1 standard error.

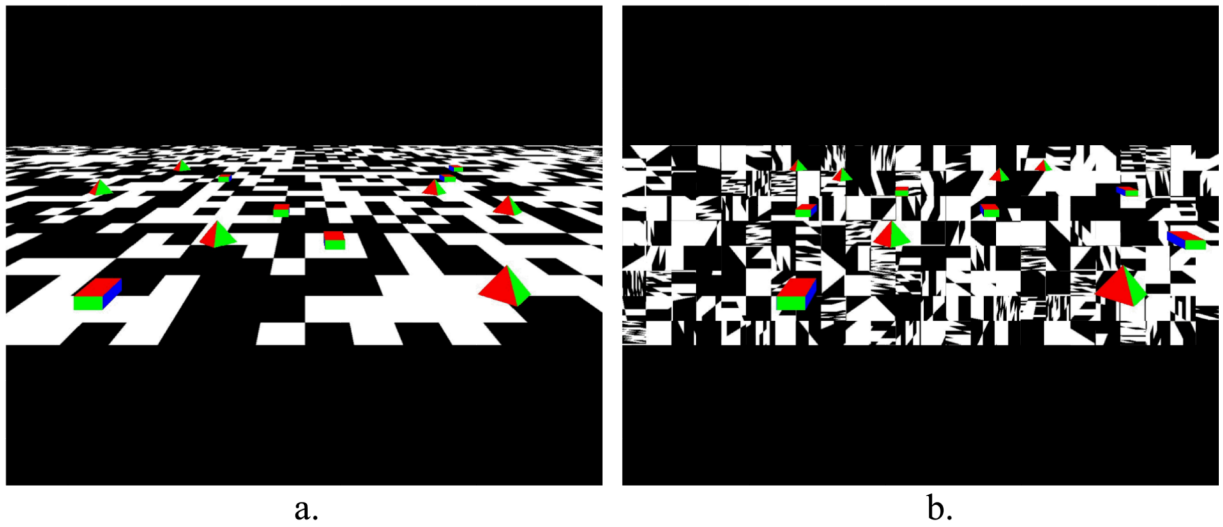


Figure 8. Examples of the stimuli used in Experiment 3. (a) normal ground surface; (b) ground surface with 24×8 grid. The ceiling surface displays were inverted images of the ground surface displays.

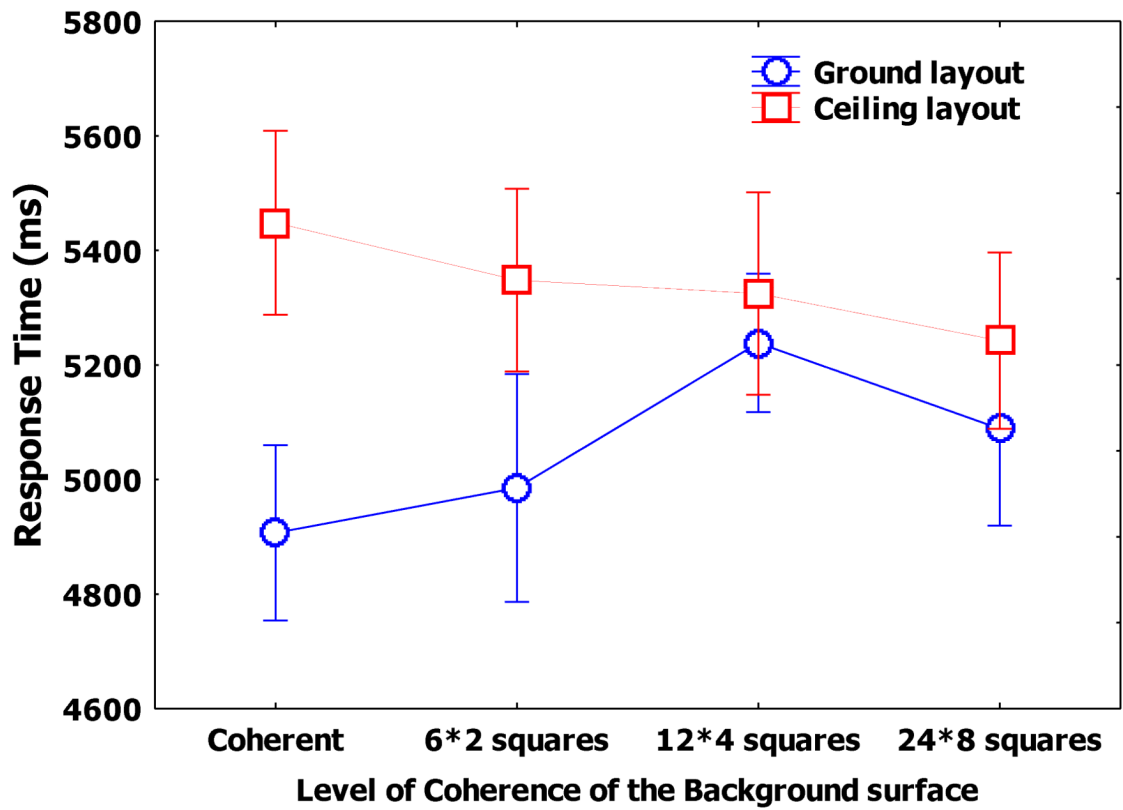
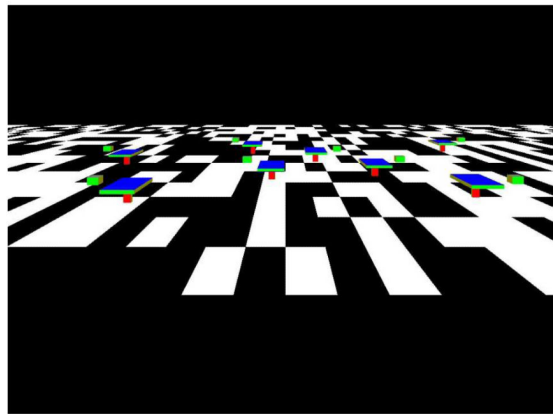
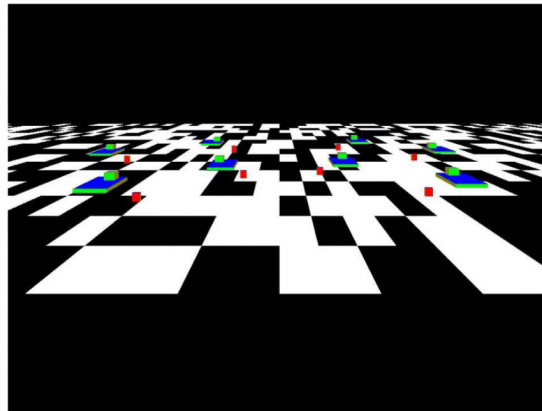


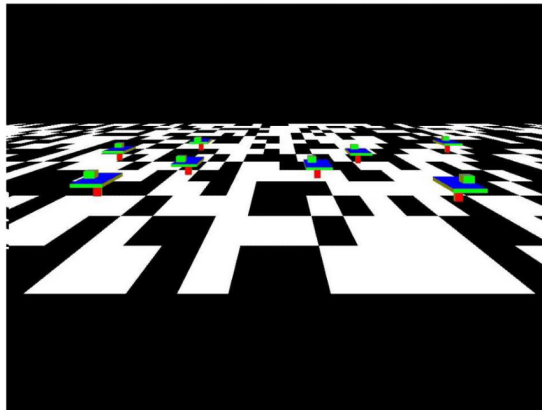
Figure 9. Mean response time as a function of surface configuration and layout of the objects from Experiment 3. Error bars represent ± 1 standard error.



a.



b.



c.

Figure 10.

Examples of the stimuli used in Experiment 4. (a) green cubes directly connected to the ground surface; (b) green cubes connected to the ground surface through 1 level of “nested” contact relation (a slab); (c) green cubes connected to the ground surface through 2 levels of “nested” contact relations (a slab and a post). The ceiling surface displays were inverted images of the ground surface displays.

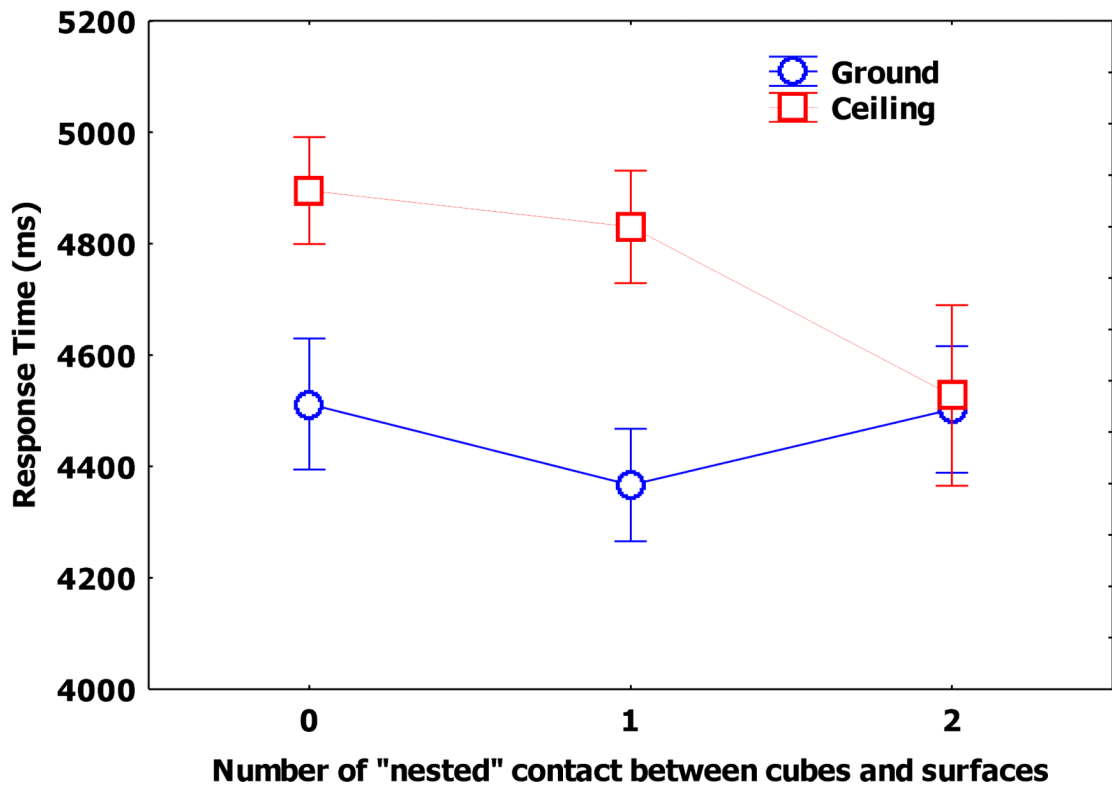


Figure 11. Mean response time as a function of type of surface presented and the level of “nested” contact relations from Experiment 4. Error bars represent ± 1 standard error.

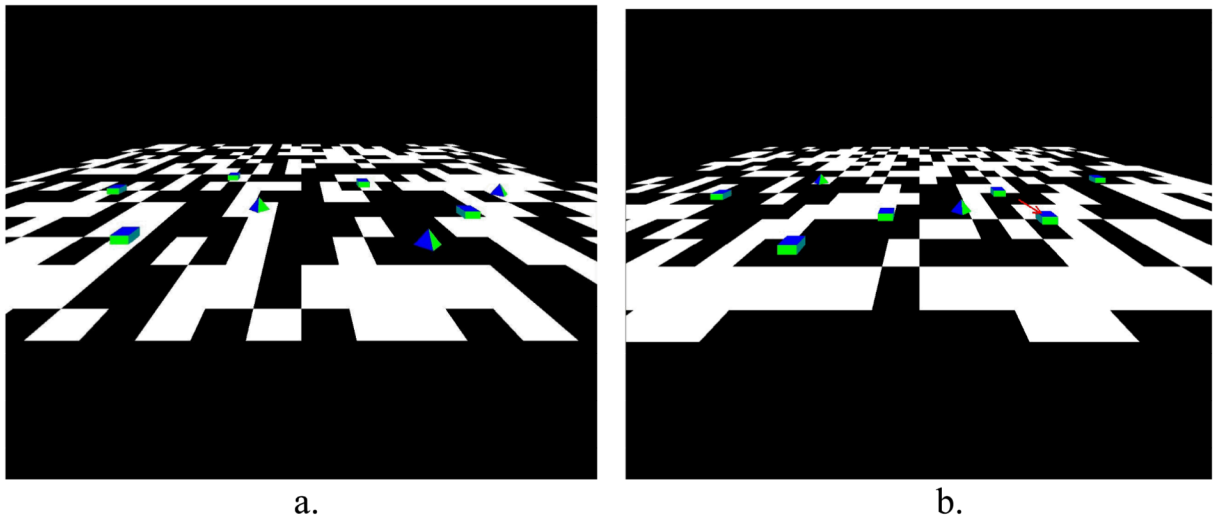


Figure 12. Examples of the stimuli used in Experiment 5. (a) test scene of a ground surface without postcue; (b) test scene of a ground surface with postcue. The ceiling surface displays were inverted images of the ground surface displays.

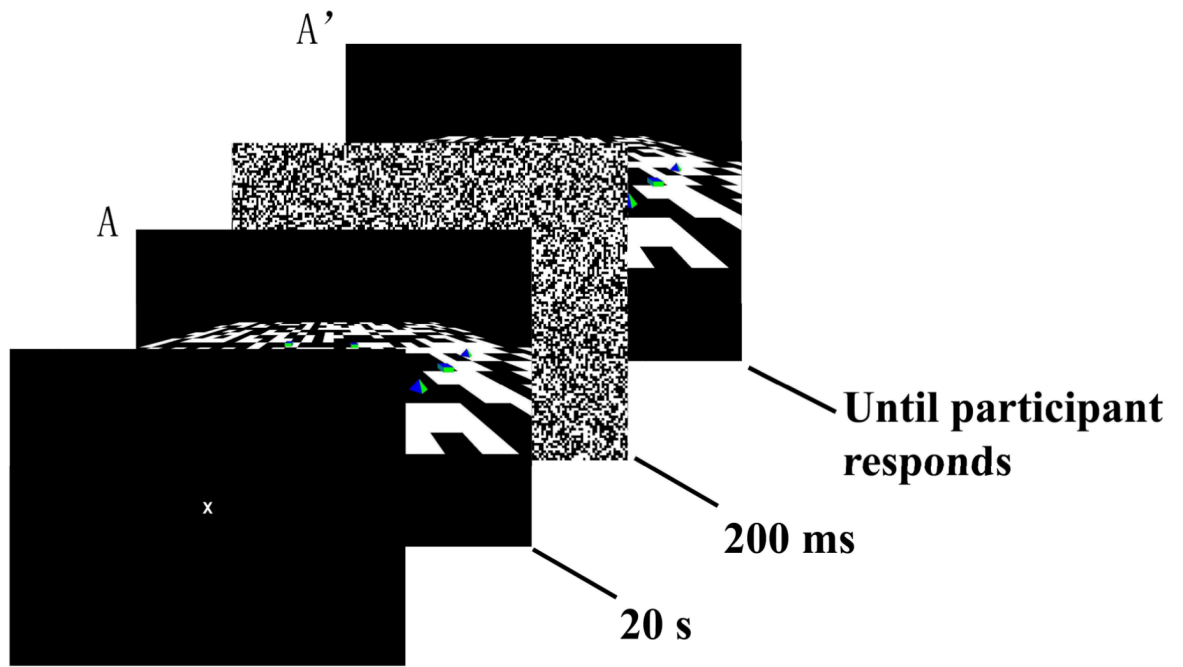


Figure 13.
 Example of the sequence of each trial used in Experiment 5.

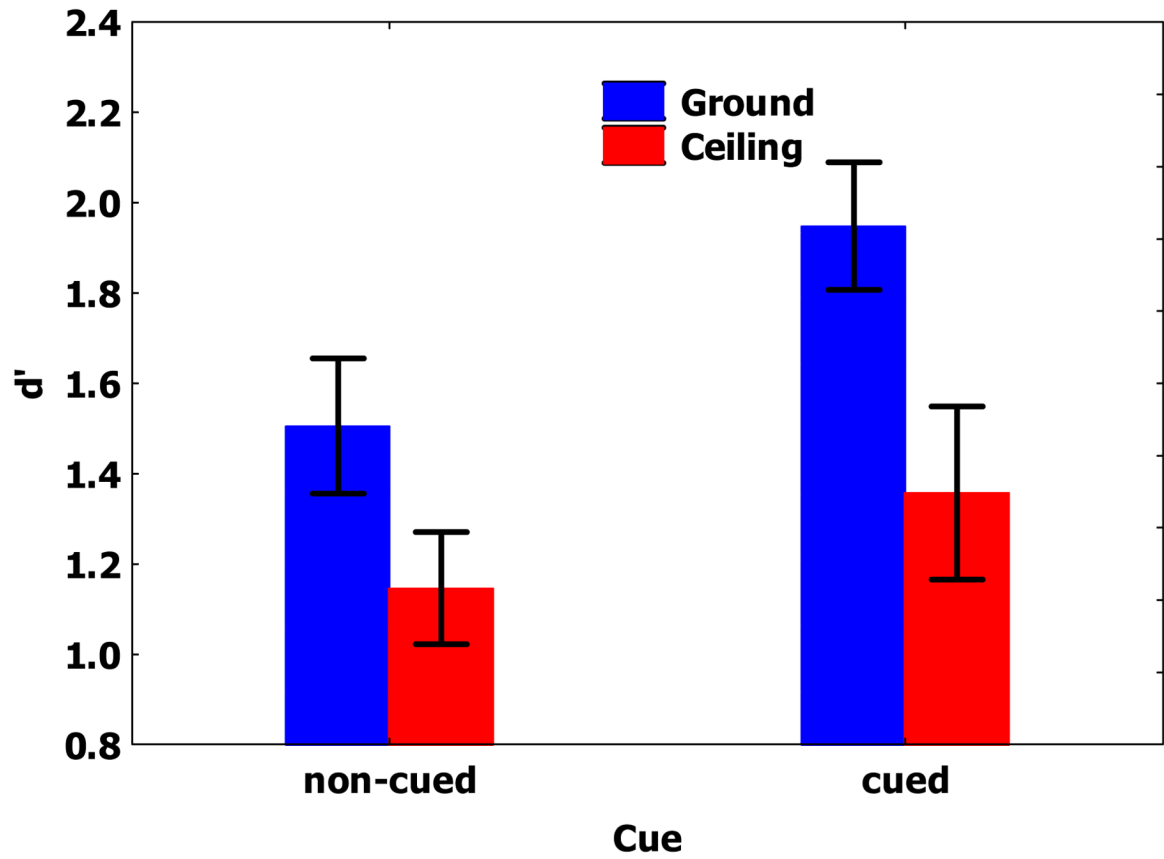


Figure 14. Mean sensitivity (d') as a function of the type of surface presented and the presence/absence of a postcue from Experiment 5. Error bars represent ± 1 standard error.

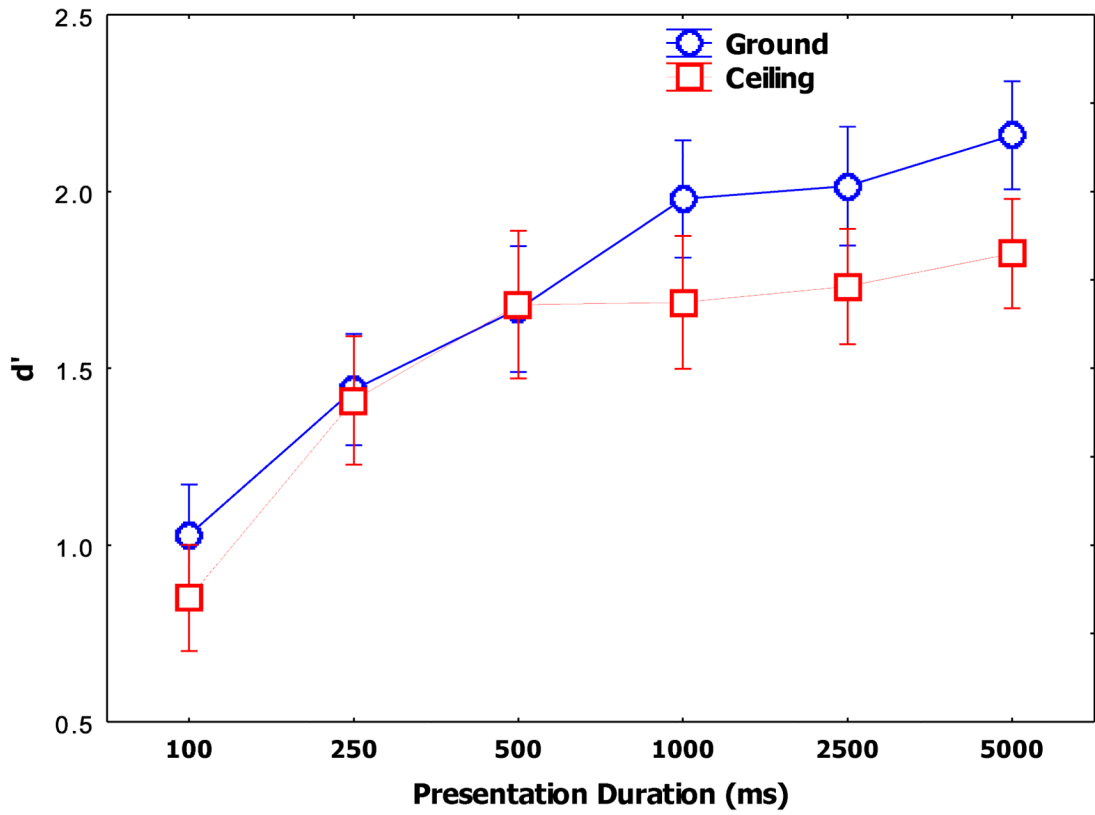


Figure 15. Mean sensitivity (d') as a function of type of surface presented and presentation duration from Experiment 6. Error bars represent ± 1 standard error.

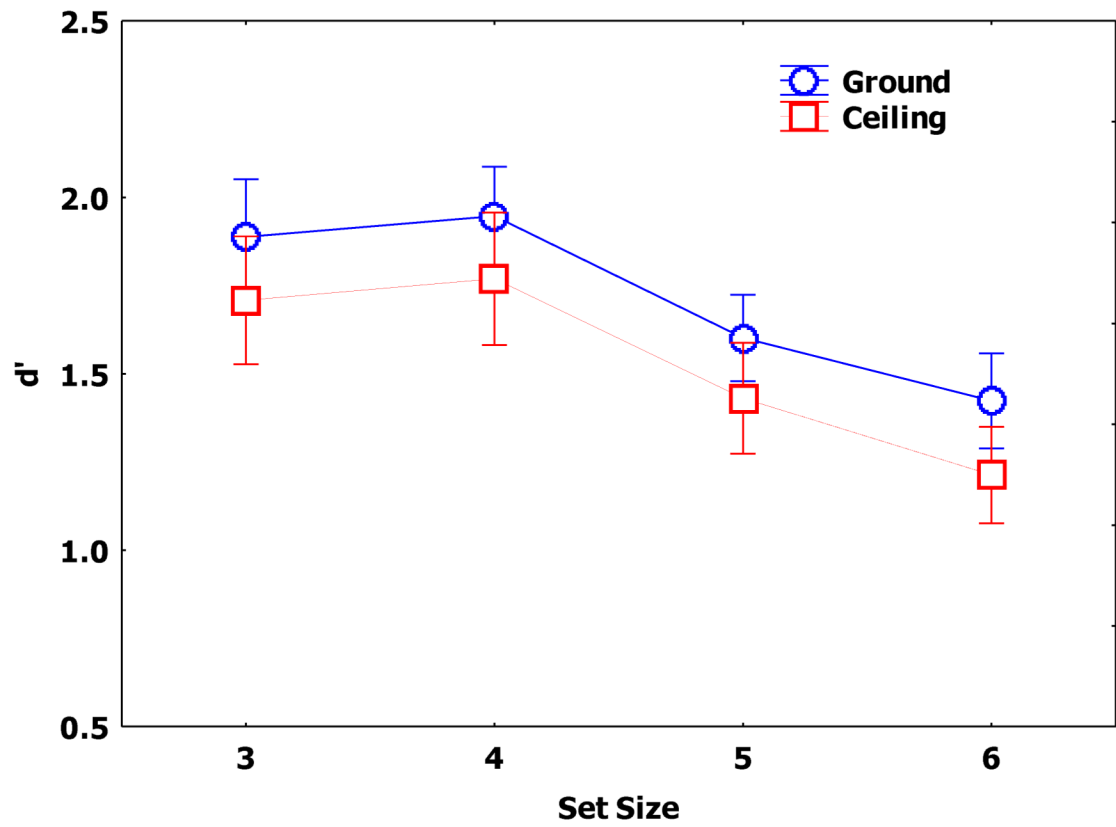


Figure 16. Mean sensitivity (d') as a function of type of surface presented and set size from Experiment 6. Error bars represent ± 1 standard error.