Inattentional blindness with same scene at different scales

Henry L. Apfelbaum, Christina Gambacorta, Russell L. Woods, and Eli Peli

Schepens Eye Research Institute and Harvard Medical School, Boston, MA, USA

Abstract

People with severely restricted peripheral visual fields have difficulty walking confidently and safely in the physical environment. Augmented vision devices that we are developing for low-vision rehabilitation implement vision multiplexing, providing two views of the same scene at two different scales (sizes), with a cartooned minified wide view overlaying a natural see-through view. Inattentional blindness may partially limit the utility of these devices as low-vision aids. Inattentional blindness, the apparent inability to notice significant but unexpected events in an unattended scene when attention is fixed on another scene, has classically been demonstrated by overlaying two unrelated game scenes, with unexpected events occurring in one scene while attention is maintained on the other scene by a distractor task. We hypothesized that context like that provided by the related wide view in our devices might mitigate inattentional blindness in a study with two simultaneous views of the same scene shown at different scales. It did not, and unexpected event detection rates were remarkably consistent with our and other mixed-scene studies. Still, detecting about half of the unexpected events bodes well for our use of vision aids that employ vision multiplexing. Without the aids, is it likely that many more events would be missed.

Keywords: Inattentional blindness, augmented vision, low-vision aids, edge filtering, unexpected events, vision multiplexing

Introduction

Many of the devices we develop in our low-vision rehabilitation laboratory employ vision multiplexing – the simultaneous presentation of two different views to one or both eyes (Peli, 2001; Peli *et al.*, 2007). For example, to aid patients with severe peripheral vision loss (with only a 5 to 20º diameter central visual field remaining), our augmented-vision head-mounted device employs spectacles with a small temple-mounted video camera. The wide-angle (e.g. 80º) view of the camera is minified (e.g. 0.2x) and displayed via a beam-splitter (approximately 1 cm^2) imbedded in one of the spectacle's carrier lenses, providing both the natural see-through view and the superimposed augmented view. It is expected that users of these devices will be able to use the augmented view to obtain an improved awareness of their environment and possible hazards, while using the natural see-through view for discriminating detail. Studies in our laboratory have shown that minification does not materially impact the ability to accurately assess potential collisions (Luo *et al.*, 2007; Luo *et al.*, 2009), and the augmented view does improve visual search performance (Luo and Peli, 2006). They have left open questions of the utility of the devices for the important mobility task of detecting unexpected events.

Inattentional blindness (IB) refers to the apparent inability to notice significant but unexpected events (UEs) in an unattended scene when attention is fixed on another scene, even though they both are in the same area of the visual field, as is the case with our vision-multiplexed devices. In a previous study (Apfelbaum *et al.*, 2008), we sought to understand and mitigate the sort of inattentional blindness that low-vision users of our augmented-vision devices might encounter. In the first experiment of that study, we presented subjects with overlapping views of two games (a scene with 3 fellows running in a

Correspondence and reprint requests to: Henry Apfelbaum, Schepens Eye Research Institute, 20 Staniford Street, Boston, MA 02114, USA
Tel: 011 617 912 2588; Fax: 011 617 912 0112
E-mail address: Henry.Apfelbaum@Schepens.Harvard.edu

circle and passing a basketball among them, and a scene showing the hands and forearms of two people playing a hand-slapping game), as was originally done by Neisser and Becklen (1975). We occasionally introduced unexpected events (UEs) in one game, such as a woman with an umbrella walking through the ball game or the hand-slappers pausing to shake hands, while attention was maintained on the other game by a distractor task. In some trials, we filtered one or both of the views to produce cartoon-like images, expecting that the increased transparency of those images and the saliency of the remaining features would alter the IB effect. It did not. About half the UEs remained undetected.

When we examined in detail the scenes in which UEs were detected more frequently, the relationship or intrusiveness of a UE with respect to the attended scene seemed to be significant. This had also been observed by Becklen and Cervone (1983), when their "umbrella woman" by chance seemed to kick the bounced ball and thus was perceived as a part of the attended action.

In the second experiment of our previous study (Apfelbaum *et al.*, 2008), we showed that inattention, not the nature of the unexpected events, was responsible for IB. We simply showed some of the same videos, with overlays just has had been done in the first experiment, but in some trials we had the subjects respond to game events in the scenes that had the UEs. When the UEs were in the attended scene they were always detected. Thus it was the contextual relationship between the UEs and the attended scene that mitigated IB.

IB has proven to be extremely robust, both with natural-scenes (like the ball and hand games) and synthetic stimuli (such as crosses and x's, as in Mack and Rock, 1998). For our purposes, studies with natural scenes are of most interest. Historically, those studies have superimposed two independent scenes, such as the ball game and hand game of Neisser and Becklen's original study, or, different instances of the same game with no correlation between the two scenes (Becklen and Cervone, 1983; Simons and Chabris, 1999; Most *et al.*, 2000; Most *et al.*, 2001; Koivisto *et al.*, 2004). We refer to these as "mixed scene" studies. As noted above, when UEs seemed to become part of the attended scene context they were more readily detected, and when they were, in fact, part of the attended scene, they were always detected. We hypothesized that the context provided when viewing the same scene action at two scales, as is the case with our augmented-vision device, might serve to mitigate IB. Here we report on a study that tested this hypothesis. We call this a "same-scene" study. A finding that same-scene context mitigates IB would be important, given IB's insensitivity to other manipulations, while a negative finding would nonetheless identify yet another dimension of IB's robustness.

Methods

Sidewalk Shell Game

Our augmented-vision device presents patients with severe peripheral vision loss a cartoon-like display from a camera with a wide field of view (e.g. 80º), minified to subtend about 15º in the device. To test our hypothesis with normally-sighted subjects, it was necessary to construct a scenario in which attention would remain focussed on a narrow view of a scene while UEs occurred in the superimposed wider view of the same scene, with both views filling the same display area.

We chose to use a simulated sidewalk shell game, as a subject's attention could be directed to the limited area needed for following the shell game, while UEs could occur in the larger context of the scene. One actor (the "performer") shuffled three shells with a marker placed under one of them, and the subject was asked to follow the shell action and determine the final position of the marker. Two actors as game players (the "observers") stood at either end of the table, and two actors (the "passersby") entered, paused to watch, and exited during the game (Fig. 1a). In some trials, the passersby performed UEs, such as swapping unusual hats with the performer.

We used two cameras, one of which had a narrow view of the scene (1 m wide) while the other, whose output was cartooned and superimposed, provided a wider (1.6 m) view. The narrow, full colour, view showed little more than the table top, the actors' torsos and faces, the shells, and the performer's arms and hands (Fig. 1b), while the wider and taller view included the area above the actors' heads in which most of the UE activity occurred (Fig. 1a). The narrow view provided just enough view between the observers and the performer to include the torso-level motion of the passersby entering and leaving, thus providing an additional contextual relationship between the activity in the narrow and wide views.

Videos and Video Processing

A Canon ZR10 miniDV camcorder (Canon Inc., Tokyo, Japan) was used to record the narrow view, and a Canon ZR85 miniDV camcorder was used to record the wide shots.

Since cartooning did not affect IB in the previous study, we chose in this study to always cartoon the wide view, as that is the configuration best approximating our augmented-vision device. Adobe Premiere Pro2.0 (Adobe Systems, Inc., San Jose, CA, USA) was used to adjust the colour balance of the wide-view videos such that the shells and table covering were approximately isoluminant, ensuring that the shells were not visible when edge filtered, since the filter used only the S-video luminance channel. Thus, the subjects would have to attend to the narrow view to accurately track the shell with the marker. We processed the videos with a modified ValueVision edge-filtering device similar to the one developed for the previous study (DigiVision, Inc., San Diego, CA, USA). In this study, however, the resulting edge images were unipolar (containing white lines only), as this, too, better represented our augmented-vision device.

Wide and close views were combined using Premiere so that the final clips used in the study had the same action occurring simultaneously from two different overlaid perspectives. Since the DigiVision filter employed interlaced NTSC format, captured and processed video was maintained in uncompressed, interlaced, DV format to avoid deinterlacing artefacts that would be particularly noticeable in the cartooned views. The camcorders were used to convert between NTSC and DV format, to avoid deinterlacing by the computer's S-video capture device. Figures 1c and 2 have examples of the overlaid images shown to subjects, and the supplemental files with the on-line version of this paper provide samples of the video clips.

Sequence

Through the course of each trial, actions were performed at specifically-designated times, as detailed in Table 1. In all trials, the performer first showed the marker, and then covered it with one of the 3 shells. After 30 seconds of shuffling, the performer placed the shells in an evenly spaced row. The video was paused, and the subject was instructed to guess which shell he or she believed to be covering the silver marker. After this was noted, the video was then restarted, and observer 1 pointed to the shell she believed was covering the marker. This shell was lifted by the performer. This was repeated with observer 2, if observer 1 guessed incorrectly. Finally, the performer revealed the correct location of the marker if both observers 1 and 2 guessed incorrectly. No sleight of hand was used, as we wanted the subjects to be able to track the shell with the marker accurately if paying careful attention. Pilot tests identified a shuffle technique and rate that permitted accuracies of about 95%.

Unexpected Events

UEs were introduced in the wide view of some of the video clips. In one event (Fig. 2a), the second passerby entered wearing a wizard's hat. The hat had a cone-shaped outline with star and moon details that were clearly visible in the cartooned outline. The hat was placed on the performer's head by passerby 2, and then removed by passerby 1, who placed the hat on his own head before leaving. In two other UE trials, the event action was similar, but a small flat sailor's hat (Fig. 2b) or a wide-brimmed safari hat (Fig. 2c) was used. In another event (Fig. 2d), the performer wore the safari hat at the outset. Passerby 2 entered wearing the sailor's hat, and switched that hat for the safari hat. Passerby 1 then took the sailor's hat off of the performer's head, and left wearing it. In another event clip (Fig. 2e), no hats were swapped; instead, the second passerby was dressed in a harlequin costume. The bold costume patterns were readily visible in the cartooned view and at torso level in the narrow view. The harlequin mask and hat provided unusual features in the cartooned view.

Trials

Each trial began with a game video clip, followed by a pause for the subject to guess the marker's location, then the video revelation of the actual position, and finally, the questions to discern if an unexpected event was noticed. The overlaid wide and narrow views were shown during game play, but only the narrow view was shown during the revelation segments. This was done so that the trials in which the performer initially wore a hat that was subsequently removed would not alert the subject to the hat changes when not engaged by the distractor task.

In each subject session, 9 videos were presented. The first two trials were used to introduce the shell game task and overlaid video treatment. They did not include UEs, and were not included in subsequent analyses. The UE order of the remaining trials was balanced across subjects, based on a pair of 5x5 digram-balanced Latin squares (Keppel and Wickens, 2004), giving 10 different presentation orders. The five different UE clips were presented at trial positions 3, 5, 7, 8, and 9, and two clips presented as catch trials, with no UEs, were placed at positions 4 and 6. Each of the 10 possible order presentations was used 3 times, for a total of 30 subject sessions.

Subjects

A total of 32 subjects participated in the study, with two excluded due to poor visual acuity (binocular visual acuity less than 20/40). The remaining 30 subjects (8 males) met the inclusion criteria, being between 18 and 40 years old (18-35 years actual), with normal or corrected to normal eyesight (binocular visual acuity 20/40 or better, 20/15 – 20/40 actual). All subjects were recruited by craigslist advertisements (http://boston.craigslist.com/vol/) and signed Schepens Institutional Review Board-approved consent forms. Subjects were paid $10 for their session and completed the study in an hour or less. They were reimbursed up to $4 for local transportation costs.

Physical Setup

The subject was asked to sit in front of a 20” CRT monitor, at a viewing distance that was comfortable, generally about 1m. Videos were presented using Windows Media Player Version 10 in full screen mode. The experimenter sat to the side of the subject, in order to pause the video at the appropriate time and ask questions at the conclusion of each game.

Session Procedures

To mask the true nature of the experiment, subjects were told that we were testing out a video-overlapping technique used in some of the devices in our lab, to determine if presenting more than one view of the same scene interfered with the subject’s ability to follow the scene. They were told that they would be scored on how many times they could correctly guess the location of the marker at the end of each trial, and no mention was made of UEs. Questions asked after each game clip was shown were designed to avoid alerting the subjects to the existence of the UEs.

At the beginning of the session, the subject was instructed to follow the movement of the shell covering the silver marker. After viewing each video clip, the subject was asked a series of questions. The first question was, “How difficult would you say that was? Very easy, easy, somewhat difficult, very difficult, or practically impossible?” This was followed by, “Was there anything particular that made this video easier or more difficult to follow?” If a subject responded that there was interference or a distraction, the experimenter asked what that was, and tried to judge if the response indicated that the subject had noticed one of the UEs. For example, if a subject said that the movement of the hands on the shells was difficult to follow, this was not scored as a detection, but if a subject was distracted by movement of the people in the background, either mentioning a hat or the touching of the performer’s head, the experimenter scored it as a detection. Subjects were told the study was in pilot phases, and that any information about the difficulty of the videos would be very helpful. The experimenter acted very interested in all feedback to the difficulty question, carefully noting all comments made by the subject, to encourage the subject to mention all details they noticed, even if the subject did not feel they were important.

Measures

For each subject, each of the 5 UEs was scored as detected or undetected. Accuracy of correct guess of the marker location was also recorded, as an indicator of attention to the distractor task during the UE trials. Trials with incorrect identification of the marker position were included in the UE detection analyses.

Results

UEs in same-scene overlaid videos were detected less frequently than the events in the different-scene overlaid videos used in our previous study (Apfelbaum *et al.*, 2008). UEs were detected in 68 (45%) of the trials in which they were shown, as compared to the 123 total detections (57%) in the previous study. The difference is statistically significant ($p = 0.029$, by difference of proportions, method 10 of Newcombe, 1998). Thus, attention was not captured readily by events in the unattended superimposed view, even though the views were contextually related and the events were synchronized and co-located.

Detection rates for each of the five UEs ranged from 33% for detection of the hat swap to 67% for detection of the wizard's hat (Fig. 3). That range was smaller than in the previous study (19% to 86%).

Accuracy of the distractor task, identifying the shell with the marker, was high (95%), with 8 subjects guessing wrong just once, and no misses by the remaining subjects. UEs were detected on half of the trials with wrong guesses. The safari-hat trial accounted for half of the wrong guesses, while the harlequin trial accounted for 3 and the wizard hat trial accounted for one. There was a non- significant trend for there to be more marker errors on the UE trials that were detected more frequently (Spearman correlation, $r_4 = 0.56$, $p=0.32$). This apparent trend may be misleading, as the errors could have come from small differences in the difficulty of following the marker in the UE trials, despite our attempts to make them equally difficult. The perceived difficulty of the distractor task in this (same-scene) study was less than our previous study (Apfelbaum *et al.*, 2008, Mann-Whitney, $z340=3.29$, $p=0.001$). Overall, this is evidence against the possibility that a harder distractor task is the explanation for the lower UE detection rate in this study compared to our previous study (Apfelbaum *et al.*, 2008).

Discussion

Our subjects viewed videos of a sidewalk shell game, with a cartoon-like representation of the full scene superimposed on a narrow view of the game play. The subjects tried to follow the location of the marker placed under one shell at the start of a trial and identify its location when the play stopped. With their attention thus engaged, subjects should have been able to track the marker accurately, and indeed they did. During the game play, in some trials, UEs occurred that were only visible in the wide view. We hypothesized that subjects would notice the UEs more frequently than they did in previous studies in which the overlaid scenes were unrelated, but that was not the case.

We interpret these results to mean that the added context provided by the same-scene simultaneous action in this study did not mitigate IB. An alternative interpretation could be that the context did improve the ability to notice UEs, but the events used in this study were inherently more difficult to detect than those in the previous studies. We believe that to be unlikely. Overall detection rates in this study (45%) were actually slightly lower than those of our previous study (57%, Experiment 1, Apfelbaum *et al.*, 2008), although the difference, while statistically significant, is within the range of variability of experiments of this sort and certainly not indicative of mitigation. Results were more consistent across the range of UEs than in the earlier study. They are also comparable to results others have found. Simons and Chabris (1999) reported an average 54% detection rate in the transparent overlay conditions most comparable to ours, although the rate varied considerably with difficulty of the distractor tasks. Becklen and Cervone reported a 34% average detection rate, while Neisser and Dube (unpublished, but summarized in Neisser, 1979) found that a 21% rate more than doubled to 48% if the subjects were first given (non-UE) practice trials, as was the case in our studies.

If unconvinced that UE difficulty did not mask any effect of same-scene context, it may be possible to do additional experiments to disambiguate these alternatives. For example, we considered overlaying the wide and narrow views from different tapings of the game, to eliminate the direct relationship between action in the views. If detection rates drop, we would have a potential indication that (lack of) context does affect detection, although even in this case, the effect could be more related to interference between the views than lack of context, and lack of a drop would still leave the question of difficulty ambiguous.

While the presentations used in this study correspond directly to the relationships encountered when using our augmented-vision device - two views of the same scene at two scales, with the wide view cartooned – there are significant differences. Since we tested normally-sighted subjects instead of patients with peripheral visual field loss, we presented the cartooned wide view on the full screen, rather than as an inset sized roughly to correspond to the patient’s field of view. The zoom factor of 1.6 was less than the typical minification factors of 3-5 used in our augmented-vision device. As such, we had closer contextual overlap than in the device. If context did not mitigate IB in this study, it is unlikely to help more with the device. Teasing out subtler effects is unlikely to help us in the design of such devices. Since UEs *were* detected about half of the time, we are nonetheless encouraged that such devices will improve the likelihood of detecting many events outside of the wearer’s narrow visual field that would otherwise go unnoticed.

Since it was impossible to follow the shell action accurately in the wide view, we did not have a distractor task equivalent to that used in the second experiment of our previous study. Hence, we could not simply change instructions so that the subject would perform the distractor task while attending to the view containing the UEs. Thus we could not perform an equivalent experiment to show in this case that attention, rather than overlaying, is the critical issue, and attended UEs would always be detected. However, it was clear in pilot tests that the UEs were easily noticed if attention was directed to the wide view.

Errors in guessing the marker location could have been an indication of lack of attention to the distractor task, but that seems not to be the case. Rather, they clustered on two of the scenes, and UEs were detected half of the time, which is a result similar to detections when the marker location was identified correctly. Thus the difficulty of following the shell with the marker in those scenes was likely higher. This was borne out by close investigation of the videos, and we concluded that attention would have remained captured. This is consistent with the observation (Apfelbaum *et al.*, 2008), that UE detections seem to be affected strongly by interactions between the events in the two scenes.

It is reasonable to ask if demonstrations of inattentional blindness found with overlapping scenes in laboratory studies relate to any real-world phenomenon. There are certainly common anecdotal accounts of missed UEs in daily life, but controlled studies are lacking. Jovancevic *et al.* (2006) reported that walking subjects did not fixate virtual pedestrians on a collision course (seen in a head-mounted display) about 40% of the time. Failure to fixate increased to about 60% if the subjects were tasked with following a virtual leader. That study closely approximated real-world conditions, and no scene overlaps were involved. They have since found comparable results when the potentially-colliding pedestrians were real people (Jovancevic-Misic and Hayhoe, 2009). These findings could be interpreted as supporting IB as a real problem for normally-sighted people when walking. However, pedestrian collisions do not occur as frequently as would be implied by an inattentional blindness interpretation of those results. Fixation on an object on a collision course is not necessary for identification of a potential collision and taking an appropriate avoidance action (Stoffregen and Riccio, 1990; Royden and Hildreth, 1999). It is likely that many of those non-fixated potentially-colliding pedestrians were monitored with peripheral vision. People with severely restricted visual fields do not have that luxury, hence the value of rehabilitation aids such as our augmented-vision device. Since our subjects were normally-sighted, the detection rates achieved may be different for patients with peripheral field loss. However, our purpose was to see if context would mitigate IB. If it did not have that effect for normally-sighted subjects it is not likely to have a stronger effect for patients. In addition, since the vision multiplexing of our augment-vision HMDs does involve superimposed views, the laboratory studies of IB are all the more relevant.

Conclusions

We conclude that the context provided by a cartoon-like wide view superimposed on the sort of narrow view of a scene available to a person with severely restricted peripheral vision did not mitigate inattentional blindness, as the detection rate of unexpected events was remarkably similar to that found in studies with no contextual relationship between superimposed views. However, we note that detection rates on the order of 50% are significantly better than almost no detection at all, as would likely be the case without the augmented view, so we continue to be encouraged by the potential of these devices to aid people with restricted peripheral vision.

Acknowledgements

This research was supported in part by National Institutes of Health grant EY-12890 and Department of Defense grant W81XWH.

Commercial relationships: Eli Peli has patent rights to the cartooned augmented-vision display.

References

Apfelbaum, H. L., Apfelbaum, D. H., Woods, R. L. and Peli, E. (2008) Inattentional blindness and augmented-vision displays: Effects of cartoon-like filtering and attended scene. Ophthal. Physiol. Opt. 28, 204-217.

Becklen, R. and Cervone, D. (1983) Selective looking and the noticing of unexpected events. Mem. Cognit. 11, 601-608.

Jovancevic-Misic, J. and Hayhoe, M. (2009) Adaptive Gaze Control in Natural Environments. J. Neurosci. 29, 6234-6238.

Jovancevic, J., Sullivan, B. and Hayhoe, M. (2006) Control of attention and gaze in complex environments. J. Vis. 6, 1431-1450.

Keppel, G. and Wickens, T. D. (2004) Design and Analysis, A Researcher's Handbook. Pearson Prentice Hall, Upper Saddle River, NJ.

Koivisto, M., Hyona, J. and Revonsuo, A. (2004) The effects of eye movements, spatial attention, and stimulus features on inattentional blindness. Vision Res. 44, 3211-3221.

Luo, G., Lichtenstein, L. and Peli, E. (2007) Estimating collision risk from minified images through a HMD when walking and standing (abstract). Optom. Vis. Sci. 84, E-abstract 070007.

Luo, G. and Peli, E. (2006) Use of an augmented-vision device for visual search by patients with tunnel vision. Invest. Ophthalmol. Vis. Sci. 47, 4152-4159.

Luo, G., Woods, R. L. and Peli, E. (2009) Collision judgment when using an augmented-vision head- mounted display device. Invest. Ophthalmol. Vis. Sci. 50, 4509-4515.

Mack, A. and Rock, I. (1998) Inattentional Blindness. Cambridge, MA, US: MIT Press. (1998). xiv, 273pp.,

Most, S. B., Simons, D. J., Scholl, B. J. and Chabris, C. F. (2000) Sustained inattentional blindness: The role of location in the detection of unexpected dynamic events. Psyche 6, article 14.

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E. R. and Chabris, C. F. (2001) How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. Psychol. Sci. 12, 9-17.

Neisser, U. (1979) The control of information pickup in selective looking. In: Perception and its development: a tribute to Eleanor J. Gibson A. D. Pick ed, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 201-219.

Neisser, U. and Becklen, R. (1975) Selective looking: Attending to visually specified events. Cognit. Psychol. 7, 480-494.

Newcombe, R. G. (1998) Interval estimation for the difference between independent proportions: Comparison of eleven methods. Stat. Med. 17, 873-890.

Peli, E. (2001) Vision multiplexing: An engineering approach to vision rehabilitation device development. Optom. Vis. Sci. 78, 304-315.

Peli, E., Luo, G., Bowers, A. and Rensing, N. (2007) Applications of augmented vision head-mounted systems in vision rehabilitation. J.SID 15, 1037-1045.

Royden, C. S. and Hildreth, E. C. (1999) Differential effects of shared attention on perception of heading and 3-D object motion. Percept Psychophys 61, 120-33.

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

Stoffregen, T. A. and Riccio, G. E. (1990) Responses to optical looming in the retinal center and periphery. Ecologic Psych 2, 251.

(a) (b) (c)

Figure 1: (a) Wide view of the game scene, showing the performer, two stationary observers, and one of the two passersby. (b) Narrow view of the shell action. (c) Narrow view and cartooned wide view superimposed, as presented to the subject. The shells can't be followed using the wide view, which is principally where unexpected events occur.

TABLE 1. Action sequence for each trial.

Time (sec)	Action
0-5	The performer lifts the shell covering the silver marker. The location is displayed for 5 seconds.
6	The performer covers the marker with a shell and begins interchanging the shells in a random fashion.
11-15	The first passerby enters from the left and stops left of the performer.
16-20	The second passerby enters from the right and stops right of the performer.
21-24	The second passerby exits to the left, passing behind the performer.
25-29	The first passerby exits to the right, passing behind the performer.
30-36	The performer stops moving shells. Cartooning disappears. The video is paused and the subject asked to guess the location of the marker.
37-40	Motion resumes. Observer one guesses.
41-44	Observer two guesses (if necessary).
45	The performer reveals the correct location of the marker (if necessary).

ACTION	UE SCENE
(a) Wizard's hat trial: The second passerby enters wearing a wizard's hat and places it on the performer's head. He pauses to watch the game and then exits. The first passerby removes the hat and exits wearing it.	
(b) Sailor's hat trial: Same as (a) except for use of a sailor's hat.	
(c) Safari hat trial: Same as (a) except for use of a safari hat.	
(d) Hat swap trial: The performer initially wears the safari hat. The second passerby wears the sailor's hat and swaps hats with the performer. He pauses to watch the game, and then exits wearing the safari hat. The first passerby removes the performer's hat and exits wearing it.	
(e) Harlequin trial: The second passerby wears a harlequin costume and hat, and a white mask with a large clown nose.	

Figure 2. Unexpected event scenes. The photos illustrate frames captured during each of the 5 unexpected events. Video samples are available in the supplemental material with the on-line version of this paper.

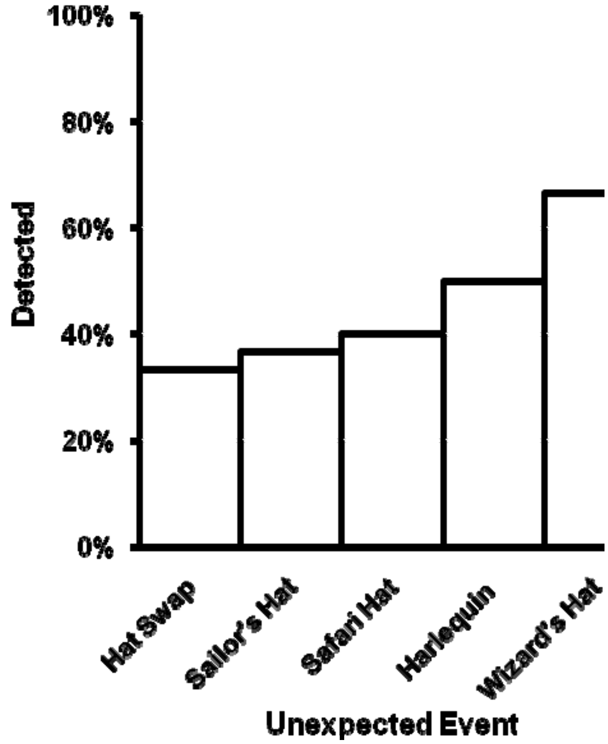

Figure 3. Detection frequency of unexpected events.