Change blindness in a dynamic scene due to endogenous override of exogenous attentional cues.

Tim J. Smith[1], Peter Lamont[2], & John M. Henderson[3],

[1]Department of Psychological Sciences, Birkbeck, University of London

[2]Psychology, University of Edinburgh

[3]Institute for Mind and Brain, University of South Carolina

Word Count = **1264**

Running Head: Failure to detect dynamic scene changes due to endogenous control.

Key words: Dynamic scene, change blindness, eye movements, visual attention, magic, cuing,

Corresponding Author:

Tim J. Smith

Department of Psychological Sciences, Birkbeck, University of London

Malet Street, London, UK, WC1E 7HX

tj.smith@bbk.ac.uk Phone: +44 (0)131 650 6459 Fax: +44 (0)131 650 3461

ABSTRACT [148 out of 150]

Change blindness is a failure to detect changes if the change occurs during a mask or distraction. Without distraction, it is assumed that the visual transients associated with the change will automatically capture attention (*exogenous* control) leading to detection. However, visual transients are a defining feature of naturalistic dynamic scenes. Are artificial distractions needed to hide changes to a dynamic scene? Do the temporal demands of the scene instead lead to greater *endogenous* control that may result in viewers missing a change in plain sight? In the present study we pitted endogenous and exogenous factors against each other during a card trick. Complete change blindness was demonstrated even when a salient highlight was inserted coincident with the change. These results indicate strong endogenous control of attention during dynamic scene viewing and its ability to override exogenous influences even when it is to the detriment of accurate scene representation.

What we remember of a visual scene is a result of where we attend and which details of attended locations we encode in memory. One of the most striking demonstrations of the interaction between attention and memory is *change blindness*. Change blindness is the failure to detect an obvious change to a scene if the change occurs during a flicker, an eye movement or another onset such as a mudsplash. For example, Rensink and colleagues demonstrated that a disappearing plane engine was hard to identify if flickers were inserted between altered versions of the photograph (Rensink, O'Regan, & Clark, 1997). The flicker masks the visual transients associated with the change that would otherwise automatically (i.e. *exogenously*) capture attention and lead to detection.

Masks or brief periods of occlusion have also been used to hide changes in dynamic scenes (e.g. Levin & Simons, 1997). A coin can be switched with another coin whilst the viewer fixates it as long as the viewer is attending to whether the coin is a head or tail and the change happens during a brief occlusion by the hands (Smith, Lamont, & Henderson, 2012). However, dynamic scenes, by definition contain natural visual transients. Are artificial distractions necessary to hide a change in a dynamic scene? Whilst watching a dynamic scene viewers must decide on how to distribute attention in space and time in order to optimize information uptake. This *endogenous* control of attention must coordinate the viewer's expectations about what is relevant with the demands of the stimulus. Endogenous control has been shown to limit capture by stimulus features in simple displays (Folk, Remington & Johnson, 1992). Does a similar tempering of exogenous control occur during dynamic scenes and can it be used to hide a change in plain sight?

In the present study we asked whether viewers can be made blind to a change in a dynamic scene through endogenous control and whether this can be overridden by exogenous capture of attention. In our study viewers watched a video of a simple card counting task (supplementary video 1). (We encourage readers to view the video now before reading on.) The video depicted a man's hands as he unpacked a deck of blue-backed cards and then dealt them face up on the table. The audio narration instructed viewers to "*count exactly how many red cards are dealt*". In the reveal, the backs of all of the cards were shown to have changed color from blue to red. It is only at this point that participants realise they have been watching a card trick. The secret behind the colour change was simple: Only the first few cards had blue backs, all the rest had red backs, and it is these that the dealer turns over at the end (Figure 1). The critical feature (the change from blue to red; Figure 1a to 1b) was in clear view, was task-relevant (participants were counting "red" faced-cards,) and only 3.4 degrees of visual angle from the attended cards.

In Experiment 1, we showed fifteen participants the video whilst their eye movements were recorded (with an Eyelink 1000). After the video, participants were asked if they had seen when the card backs had changed colour. Various fanciful guesses were offered but none of the fifteen participants reported seeing the cards change colour. Eyetracking data collected during the video showed that all participants fixated the faces of the cards (see supplementary video 2). Participants were shown the video a second time and instructed not to count the cards. During this second presentation of the video most participants (13 out of 15) looked at the backs of the cards (supplementary video 3), and reported seeing the color change after the video finished. This increase in detection across presentations was significant; McNemar exact binomial test(1)=10.083, $p<.001$.

How robust is participant belief that only the card faces are relevant (i.e. their endogenous control of attention)? Can attention be exogenously drawn to the location of the change? To investigate this question, we replicated the experiment with a new set of fifteen viewers. In experiment 2 a sudden colour onset (a 240ms bright pink outline to the deck of cards) was used to try to "pull" attention to the deck of cards just before the color change. Similar unexpected onsets have increased change detection in static flicker paradigms (Scholl, 2000). Even in this condition, eyetracking revealed that viewers continued to fixate the card faces, ignoring the visual transient (supplementary video 4). Once again, none of the fifteen viewers reported seeing the card backs change color. As in the first experiment, most of the viewers (12 out of 15) noticed the changing color during second viewing and looked directly at the card backs (a significant increase from the first viewing; McNemar exact binomial test(1)=11.077, p<.001).

Due to the dynamic nature of the stimuli and the predictable trajectory of the cards as they are dealt, it is possible that participants selectively filter out irrelevant transients such as the salient outline. In Experiment 3, we endeavored to remove this filter by instructing participants both to count the number of red cards and to report a "pink flash" via button press. Nine of the fifteen participants reported seeing the flash with four participants even saccading towards the deck of cards in response to the onset (supplementary video 5). Nevertheless, none of the participants noticed the color change. All fifteen participants noticed the change during second viewing (a significant increase from the first viewing; McNemar exact binomial test(1)=13.067, p<.001).

In summary, 45 out of 45 participants failed to notice a color change that took place close to fixation during first viewing. The majority of participants (40 out of 45) were able to detect the change during a second viewing (no significant difference across experiments; Fisher's Exact $\chi^2(2) = 3.120$, p=.343, n.s.). Our results demonstrate that expectations about the relevant features and locations within a dynamic scene can override stimulus factors competing for attention and even limit detection of an otherwise salient visual change near fixation. This study extends previous findings of failure to detect changes in an edited dynamic scene (Levin & Simons, 1997), the presence of an object in a dynamic scene (i.e. *inattentional blindness*; e.g. Simons & Chabris, 1999) and the use of misdirection to hide small changes, such as a dropped cigarette (Kuhn & Tatler, 2005) by demonstrating that large changes to a scene do not need to be artificially masked or hidden within a complex event if the viewer chooses to look away from the change.   Control of attention during dynamic visual scenes is much more complex than observed in static scenes and closely related to our spatiotemporal expectations about the events depicted.

References

Folk, C. L., Remington, R. W., & Johnstone, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance , 18*, 1030-1044.

Kuhn, G., & Tatler, B. W. (2005). Magic and fixation: Now you don't see it, now you do. *Perception*, 34, 1153-1161.

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review, 4*, 501-506.

Rensink, R.A., O'Regan J.K., & Clark, J.J. To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*, 8:368-373. 1997

Scholl, B. J. (2000). Attenuated change blindness for exogenously attended items in a flicker paradigm. Visual Cognition, 7(1/2/3), 377-396.

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

Smith T. J., Lamont, P., & Henderson, J. M (2012). The penny drops: Change blindness at fixation. *Perception* 41(4) 489 – 492

Acknowledgments

This project was funded by the Leverhulme Trust (Ref F/00-158/BZ) awarded to JMH and a University of London starter grant awarded to TJS. Thanks to Rachael Bedford for statistical advice and Parag K. Mital for use of CARPE for dynamic heatmap generation. The trick was created by Peter Lamont and Richard Wiseman, as part of a "Psychology of magic" project funded by the Wellcome Trust in 2001, and subsequently published in MAGIC magazine (The deja vu colour changing deck, *MAGIC*, 11, 2003, 87-8).

Figure Caption

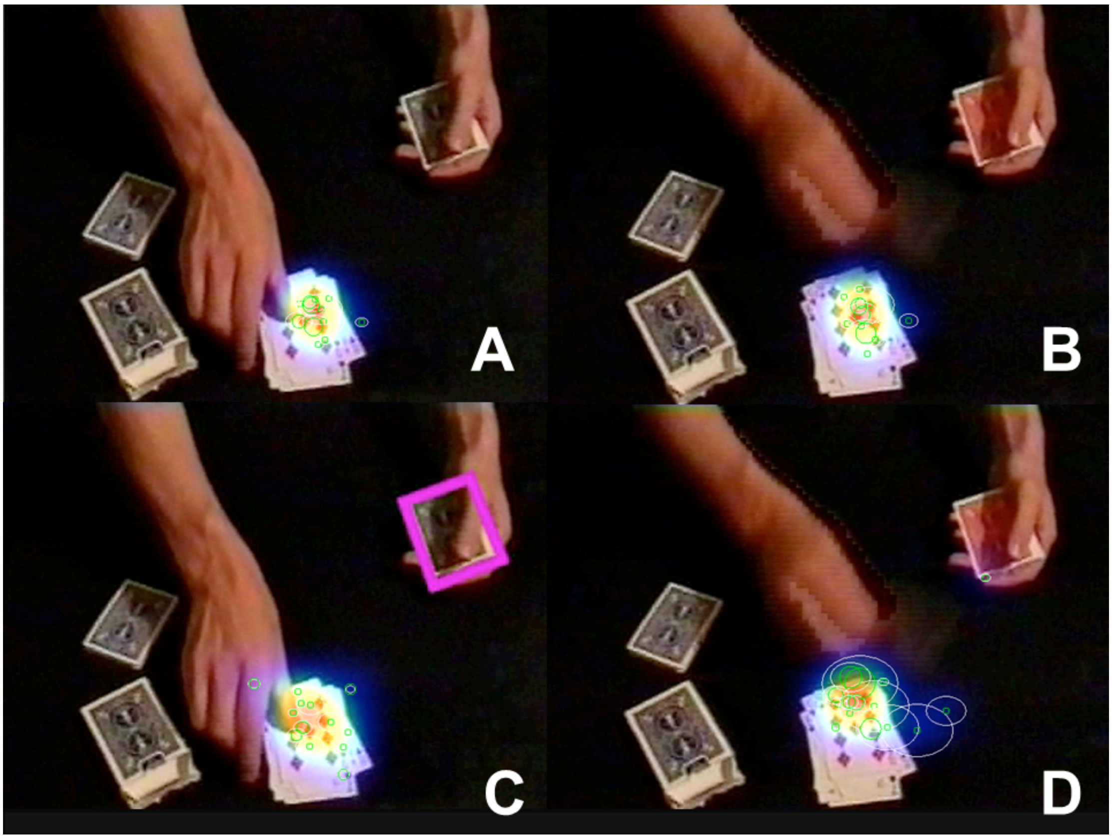

Figure 1: Screen shots from the video used in all experiments. Ellipses represent gaze locations of fifteen participants from each experiment. Diameter of ellipse indicates fixation duration and heatmap indicates degree of gaze clustering (more clustered = hotter colour). A) Experiment 1: 480 ms preceding the change of the backs of the cards from blue to red; B) Experiment 1: 240 ms following the color change; C) Experiment 3: The pink outline of the cards in the dealer's hand presented in Experiments 2 and 3; D) Experiment 3: Following the onset and 240 ms after the color change.

Supplementary On-Line Materials

Materials and Methods

Participants

Forty five members of the Psychology Department subject pool participated for course credit or pay: Experiment 1 (N=15, 10 female), Experiment 2 (N=15, 9 female), and Experiment 3 (N=15, 8 female). All participants had normal or corrected to normal vision, were naïve with respect to the purposes of the study and were informed of their rights of participation according to the British Psychological Society's ethical guidelines.

Apparatus

Eye movements were monitored by an SR Research Eyelink 1000 eyetracker. Fixation position was sampled at 1000Hz and saccades were detected using a 17-sample saccade detection model with a velocity threshold of 30°/sec, an acceleration threshold of 8000°/sec^2, and a minimum amplitude of 0.5°. Viewing was binocular, but only the right eye was tracked. The videos were presented on a 21 inch CRT monitor at a viewing distance of 90 cm with a refresh rate of 140 Hz. The experiment was controlled with SR Research Experiment Builder software.

Stimuli**.**

A single video was presented to participants (full-color, 720x576 pixels, 25fps, xvid codec, 1 min 5sec length) on a 800 x 600 pixel screen (subtending a visual angle of 25.7° x 19.4°). The video depicted a pair of hands (belonging to author PL) in close-up dealing cards onto a black table top (see Figure 1). The video used for all three experiments was identical except for the addition of a superimposed pink outline in Experiment 2 and 3. The video was accompanied by an audio narration by PL stating:

1. **PL unpacks cards**: "This is a task using some playing cards. I'm going to deal cards face up on to the table. Your job is to count exactly how many red cards are dealt. Let's see how you can do."
2. **PL deals cards face up on to the table one at a time.**
3. **PL has one card left in his hand:** "One card left. It has a blue back, but is it red?"
4. **Turns card over**: "So, how many red cards did you see. 10? 15?"
5. **Gathers all cards together, drops first few face-up on the table then turns rest over:** "In fact, there are more than that. All of these cards are red cards."

<u>Procedure</u>

Across three experiments, participants watched a video of a pair of hands viewed from above as they dealt cards on to a table (see video 1). Each participant was instructed to "*count exactly how many red cards are dealt*". The cards were turned over from a clearly visible deck of cards with blue backs. The rest of the pack is left in clear view on the side of the table along with the box. Both display the blue card backs. In the reveal, the backs of all of the cards were shown to have changed color from blue to red. After the video was finished participants were asked how many red cards they had counted, whether they had noticed the card backs change color, and if so, to describe what they had seen. They were then shown the video for a second time without having to count the cards. After the second presentation they were asked if they had seen the card backs change color and to describe what they had seen.

Analysis

Raw gaze data from the eyetracker was parsed for blinks (lost data) and saccades (eye velocity >30°/s and acceleration >9000°/s^2) then converted into frame-based gaze coordinates for each video and participant. Gaze of multiple participants was visualized on top of the original video as raw gaze spots (see Figure 1).

Supplementary Materials

Supplementary video 1: Card Trick (Original)

http://youtu.be/8rtos_8bg0

Supplementary video 2: Card Trick and experiment 1 eye movements of a fifteen participants during 1st viewing. Diameter of ellipse indicates fixation duration and heatmap indicates degree of gaze clustering (more clustered = hotter colour).

http://youtu.be/pgLbFuPa6fY

Supplementary video 3: Card Trick and experiment 1 eye movements of fifteen participants during 2nd viewing. Diameter of ellipse indicates fixation duration and heatmap indicates degree of gaze clustering (more clustered = hotter colour).

http://youtu.be/_SeKNe8xLgE

Supplementary video 4: Card Trick and experiment 2 eye movements of fifteen participants during 1st viewing. Diameter of ellipse indicates fixation duration and heatmap indicates degree of gaze clustering (more clustered = hotter colour).

http://youtu.be/FBpUpXNW844

Supplementary video 5: Card Trick and experiment 3 eye movements fifteen participants during 1st viewing. Diameter of ellipse indicates fixation duration and heatmap indicates degree of gaze clustering (more clustered = hotter colour).

http://youtu.be/rWSgjmo7vNg