



ELSEVIER

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Event completion: Event based inferences distort memory in a matter of seconds

Brent Strickland*, Frank Keil

Department of Psychology, Yale University, 2 Hillhouse Avenue, New Haven, CT 06520, United States

ARTICLE INFO

Article history:

Received 21 May 2010
 Revised 20 April 2011
 Accepted 30 April 2011
 Available online xxxx

Keywords:

Events
 Event perception
 Causal reasoning
 Perception
 Memory

ABSTRACT

We present novel evidence that **implicit causal inferences distort memory for events only seconds after viewing**. Adults watched videos of someone launching (or throwing) an object. However, the videos omitted the moment of contact (or release). **Subjects falsely reported seeing the moment of contact when it was implied by subsequent footage but did not do so when the contact was not implied**. Causal implications were disrupted either by replacing the resulting flight of the ball with irrelevant video or by scrambling event segments. Subjects in the different causal implication conditions did not differ on false alarms for other moments of the event, nor did they differ in general recognition accuracy. **These results suggest that as people perceive events, they generate rapid conceptual interpretations that can have a powerful effect on how events are remembered.**

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

We tend to think and talk about our experiences in terms of discrete events even though they occur over a continuous time line. We impose boundaries on streams of activity that reflect conceptual schemes for interpreting and representing event-related information. Imagine, for example, observing someone setting down a coffee mug, releasing it and pulling one's hand back. Even though the time line during which this process unfolds is necessarily continuous, we tend to mentally represent this continuity as three discrete events with clear boundaries. Here, we present novel evidence that causal inferences related to these "event files" can distort perceptual memory in a matter of seconds.

Different factors have been proposed as cues for determining when an event boundary will be created: degree of physical change (Newtonson & Engquist, 1976), intentionality cues (Baldwin, Baird, Saylor, & Clark, 2001) and prediction error (Avrahami & Kareev, 1994; Swallow, Zacks, & Abrams, 2009). More recent literature has focused on the

downstream effects of segmenting events in these ways. For example, visual attention and memory have been shown to improve at event boundaries (Newtonson & Engquist, 1976), and recall for items from on-going events has been shown to be superior to memory from items in previous events, even after controlling for duration between exposure and test (Swallow et al., 2009).

However, much less theoretical attention has been paid to the internal structure of token event representations. Given that the mind is constantly setting up new event representations on the fly, there should also be sophisticated compression routines in place for efficiently packaging previous events as they are being sent to memory. Rapid conceptual inferences may help parse previous events into causally coherent packages in ways that could systematically distort memory. Demonstrations of such an effect could also have implications for false memory effects at much longer time scales (e.g., Loftus & Palmer, 1974).

One example of how disparate information can be made to cohere into a single representation comes from the literature on "causal bridging inferences" (Haviland & Clark, 1974). Readers are faster to verify the sentence "water extinguishes fire" when they read the passage: *Dorothy poured water on the bonfire. The bonfire went out* compared

* Corresponding author.

E-mail address: brent.strickland@yale.edu (B. Strickland).

to when they read the passage: *Dorothy poured water next to the bonfire. The bonfire went out.* This is because in the “on” case, but not the “next to” case, a reader must infer that the water caused the fire to go out in order to make the text “cohere.”

Here, we ask whether similar coherence based inferences might influence an observer’s memory of a recently perceived event only seconds after viewing.

2. Experiment 1

Observers watched videos depicting causal launching (e.g., kicking a ball; Michotte, 1946) and throwing events (e.g., throwing a card) that were missing the actual moment of contact (henceforth just “contact”). Participants also saw complete control videos containing the moment of contact.

In a between-subject manipulation, subjects appeared in one of three conditions. In the “with causal implication” condition, subjects saw all the moments of the event (either missing or containing the moment of contact depending on the video) and then saw the resulting flight of the ball. In the “without causal implication” condition, subjects saw something irrelevant from the same scene, like a person walking, instead of seeing the resulting flight of the ball. And in the scrambled condition, subjects saw identical video footage as those in the “with causal implication” condition except that the video segments were scrambled so as to disrupt causal cohesion (see Figs. 2–4 below).

After watching a video, subjects saw a series of still images. One such still image displayed the crucial contact picture like the one shown in Fig. 1.

If bridging inferences influence event memory, then subjects should be more likely to falsely report seeing the moment of contact after watching an incomplete video that implied the moment of contact compared to one that did not. However, false alarm rates on other plausible pictures for which the correct answer is “no” should not differ between conditions. In short, we predicted that people

would fill in missing elements in event perception in ways that plug gaps in specific causal conceptual structures, not merely filling in other likely elements suggested by the general context (e.g., Biederman, 1981).

2.1. Methods

2.1.1. Participants

Fifty-eight subjects over the age of 18 from around the New Haven, CT area participated in the experiment. Subjects were randomly assigned to condition. In each condition, one outlier was removed due to response times that were at least two standard deviations away from the mean.

2.1.2. Stimuli

Test videos were created and displayed on a computer monitor using a program written in Psychtoolbox for MATLAB (Pelli, 1997; Brainard, 1997). We employed 6 videos: throwing a ball, kicking a ball, slingshot, throwing a card, putting a golf ball, and badminton. Each video lasted around 30 s.

All videos had time-matched pairs (to within .56 s) consisting of complete and incomplete versions. The complete videos contained the moment of contact while the incomplete videos did not. A series of cuts made it possible to remove the moment of contact in a way that still fit in with the natural flow of the video. Videos were displayed at a frame rate of 30 frames/s. On average, 11.33 frames were removed from the contact part of the incomplete videos.

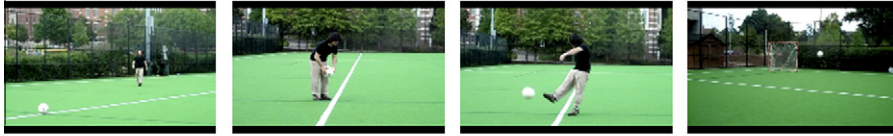
All videos were made either for the “with causal implication,” “without causal implication,” or “scrambled” condition. Video durations were time matched across conditions to within a second. The “with causal implication” videos contained footage of the resulting trajectory of the object being launched or thrown. The “without causal implication videos” contained irrelevant footage after the moment of contact (or non-contact) instead of the object’s resulting trajectory. The “scrambled” videos were created by segmenting each “with causal impli-



Fig. 1. The critical “contact” picture from the “kicking” video.

With Causal Implication

Complete



Incomplete

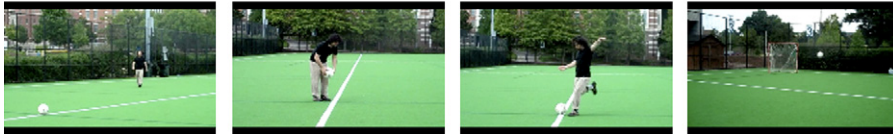
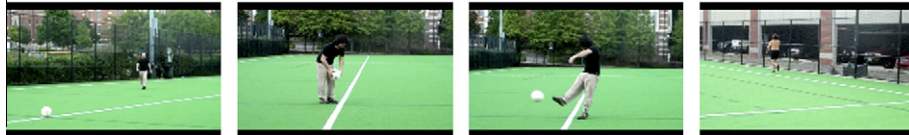


Fig. 2. Schematized version of the videos from the “with causal implication” condition.

Without Causal Implication

Complete



Incomplete

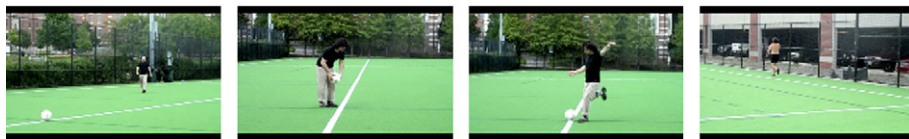


Fig. 3. Schematized version of the videos from the “without causal implication” condition.

Scrambled

Complete



Incomplete

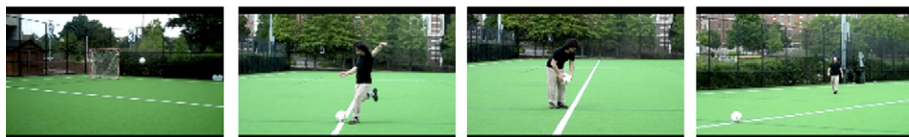


Fig. 4. Schematized version of the videos from the “scrambled” condition.

tion” video into 4 or 5 discrete segments and then playing the segments in reverse order.

Video completeness was manipulated within subjects such that each subject saw 3 complete and 3 incomplete

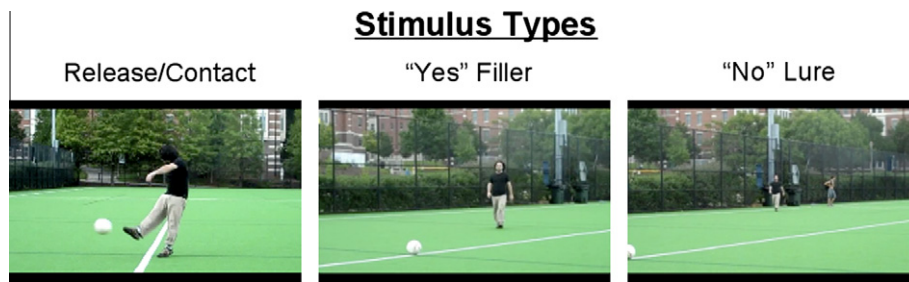


Fig. 5. The three stimulus types employed in Experiment 1.

videos. All video and completeness orders were randomized. The causal implication conditions were manipulated between subjects. The 3×2 design is schematized below in Figs. 2–4 below.

These are actual images from the “kicking” video. The third picture from the left for the “incomplete” versions in Figs. 2–4 depicts the moment at which the incomplete videos cut out before contact.

Each video was associated with 10–12 response pictures. Each picture set necessarily contained the moment of contact, “yes” fillers, and 3–4 “no” lures (see Fig. 5 below). The contact picture appeared in half of the videos (i.e. the complete videos). “No” lures depicted scenes that had not appeared in the video. They included minor changes to the background, changes to the clothing or hairstyle of the main actor in the video, or a change in the color of the object being launched. “Yes” fillers were pictures other than the contact picture that had appeared in the preceding video. All orders in each picture set were completely randomized.

2.2. Procedure

Participants were instructed to carefully watch each video on the computer screen. After the presentation of the video they were told that they would see a series of pictures and their task would be indicate whether the picture had appeared in the previous video by pressing the “y” or

“n” key. Participants were shown one practice video and picture set, and then moved onto the actual experiment.

2.3. Results

Subjects incorrectly responded “yes” to the contact picture significantly more often on incomplete videos with a causal implication ($M = .74$) than on incomplete videos without a causal implication ($M = .51$), $t(35) = 2.15$, $p < .05$ (see Fig. 6 below). Subjects in the scrambled condition also false alarmed to the contact picture significantly less often ($M = .47$) than in the “with causal implication” videos ($M = .74$), $t(35) = 2.64$, $p < .05$. False alarm rates on the contact picture in the “scrambled” and the “without implication” conditions did not differ significantly ($p = .74$).

There were no significant differences in false alarm rates on “no” lures across the three conditions $F(2, 53) = .39$, $p = .68$. There were also no significant differences on “yes” responses on “yes” filler items: $F(2, 53) = .65$, $p = .53$ (see Fig. 7).

Correct “yes” responses to the contact picture did not differ significantly between the different causal conditions, $F(2, 53) = .71$, $p = .50$.

Overall accuracy for the “with causal implication” ($M = .79$), “without causal implication” ($M = .79$), and the “scrambled” conditions ($M = .78$) did not differ significantly $F(2, 53) = .20$, $p = .82$. Overall response times were as follows (in seconds): “with” $M = 2.18$, “without”

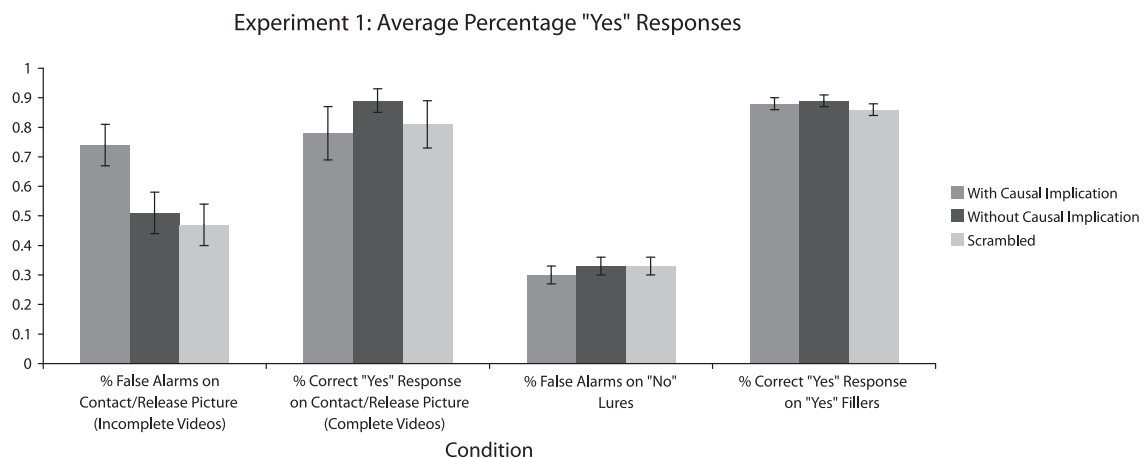


Fig. 6. Average percentage of “yes” responses on all picture types and across all conditions. Bars represent standard error.

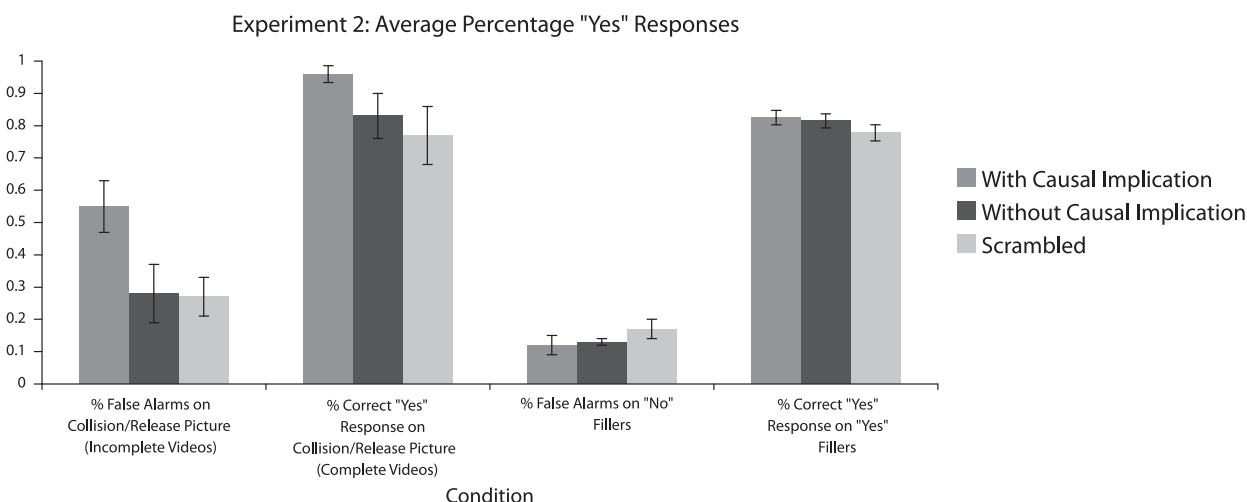


Fig. 7. Average percentage of "yes" responses on all picture types and across all conditions. Bars represent standard error.

$M = 2.13$, "scrambled" $M = 2.23$. These did not differ significantly across the three conditions $F(2, 51) = .12, p = .89$.

On average the contact picture appeared 11.41 s after the offset of the video: "with" $M = 10.40$; "without" $M = 12.13$; "scrambled" $M = 11.65$. These values did not differ significantly: $F(2, 53) = .92, p = .40$. Average response times on the contact picture also did not differ significantly (in seconds): "with" $M = 2.46$, "without" $M = 2.64$, "scrambled" $M = 2.81$. $F(2, 53) = .59, p = .56$.

False alarm rates on the contact picture for subjects in the "with causal implication" condition did not differ significantly for the first half of the pictures in each trial ($M = .61$) compared to the second half in each trial ($M = .79$), $t(55) = 1.54, p = .13$ (computed over individual trials). False alarm rates were particularly high for the first picture, which appeared 1.03 s after the offset of the video ($M = .78$).

2.4. Discussion

The results of Experiment 1 supported our original hypothesis. Participants were more likely to falsely alarm on the contact picture when this event was highly implied compared to when it was not. They did so about 11 s after viewing, and even as quickly as 1.03 s after the implied moment of contact, false alarm rates were very high. However, overall accuracy rates for the other "no" lures did not differ significantly between the different causal implication conditions. This rules out any possibility that the differences in false alarm rates on the contact picture between the different experimental conditions are driven by a general response bias.

3. Experiment 2

Experiment 2 replicated the effect found in Experiment 1 with an entirely novel set of stimuli.

3.1. Methods

Experiment 2 was identical to Experiment 1 with the following exceptions.

3.1.1. Participants

Fifty-eight subjects over the age of 18 from around the New Haven, CT area participated in the experiment. Two outlying subjects were removed on the basis of response time (two standard deviations away from the mean) from the "with causal implication" and "without causal implication" conditions.

3.1.2. Stimuli

Five new test videos were created: basketball, billiards, kicking, throwing and bowling. Video completeness was manipulated within subjects such that each subject saw either 2 complete and 3 incomplete videos or 3 complete and 2 incomplete videos. All videos had time-matched pairs (to within .93 s) consisting of complete and incomplete versions.

For each video, a set of either 10 or 11 still pictures was created. The videos had between 6 and 9 still pictures that had been directly extracted from the video. In addition, between 1 and 5 "no" lure stills were created that the subject did not actually view.

3.2. Results

Participants again false alarmed to the contact picture significantly more often in the incomplete videos with a causal implication ($M = .55$) than in the incomplete videos without a causal implication ($M = .28$), $t(34) = 2.27, p < .05$. Subjects in the scrambled condition also false alarmed to the contact picture significantly less often ($M = .27$) than in the "with causal implication" videos ($M = .55$), $t(34) = 2.38, p < .05$.

Despite the difference in false alarm rates on the contact picture, subjects did not differ across conditions

on overall accuracy, $F(2, 51) = 2.22, p = .12$ or on overall response times, $F(2, 51) = 1.03, p = .36$. Average response times (in seconds) on the contact picture also did not differ significantly: “with” $M = 1.93$, “without” $M = 2.31$, “scrambled” $M = 2.07, F(2, 51) = .816, p = .448$.

There were no significant differences in false alarm rates on “no” lures across the three conditions $F(2, 51) = 1.26, p = .26$. Nor were there any significant differences on “yes” responses for filler items which had actually appeared in the preceding video: $F(2, 51) = 1.237, p = .3$.

Correct “yes” responses to the contact picture did not differ significantly between the “with” ($M = .96$) and the “without” ($M = .83$) causal implication conditions $t(34) = 1.71, p = .10$. However they did differ significantly between the “with causal implication” and “scrambled” ($M = .77$) conditions, $t(34) = 3.25, p = .002$. This latter finding did not replicate in Experiment 1 and was likely a false positive.

False alarm rates were again were again particularly high when the contact picture had appeared as the first picture in the test set ($M = .8$). In these cases, the contact picture appeared only 1.03 s after the offset of the video.

3.3. Discussion

These results successfully replicated the findings from Experiment 1 on a completely novel set of stimuli. Participants were again significantly more likely to falsely alarm on a release or contact picture when this event was highly implied compared to when it was not. This pattern provides evidence for the robustness of the basic effect.

4. General discussion

When people observe real-world events they spontaneously and rapidly construct conceptually coherent interpretations that enable them to package continuous streams of visual information into discrete event units. These experiments suggest that coherence-based inferences induce false recognitions via “event extensions” on a relatively quick timescale of seconds. In the two studies reported, subjects falsely remembered seeing a moment of contact only in videos where such a moment was highly implied. When no evidence of the contact immediately followed a “non-contact,” subjects did not falsely remember this event.

We began by suggesting that event perception may involve a process of dividing the continuous stream of visual information into meaningful chunks. That process, however, seems to result in especially strong memory distortions in which illusory components of the event are inserted so as to link together the observed components into a more causally coherent memory. Although it has long been shown that verbally presented information can be distorted in ways that increase causal coherence (Bransford & Johnson, 1972), the studies described here are the first to show that visually presented information can be quickly distorted by high level conceptual and causal factors that are divorced from bottom up perceptual cues.

Our working hypothesis is that this causal filling in effect results from the particular way in which the compression algorithms for event files are set up. As a new token event representation is being set up in working memory, the outgoing event representation is sent to memory. However, saving all the information from that outgoing representation would be too costly in terms of speed and memory capacity. So it is likely that there are compression routines in place that efficiently package information. In many cases, this can lead to a loss perceptual detail (see Swallow et al., 2009). However, in some circumstances, conceptual packaging can induce the perceiver to insert unseen information in order to fulfill structural requirements. This was the case in the present study.

At first glance, our results may seem similar to those found in the representational momentum paradigm (Hubbard, 1995). However any superficial similarities are misleading. In representational momentum studies, the specific location of a given object is usually mis-remembered to be located slightly forward along an anticipated trajectory. The effects presented here differ in two important respects. Firstly, our effects are *postdictive* in that memory is strongly influenced by what occurs after the moment in question (i.e. the implied moment of contact). In representational momentum however, memory is influenced only by what comes before the moment in question (i.e. the moment at which the object disappears). Secondly, in representational momentum, what participants falsely remember seeing is qualitatively similar (or identical) to what they just saw. Upon seeing movement along a trajectory, subjects falsely remember seeing a little more movement along that trajectory. Here however, people falsely report having seen a qualitatively different type of occurrence than what they had actually seen (contact vs. simple motion).

The effects from this study could however been seen as the temporal analog to *amodal completion* (Rauschenberger & Yantis, 2001), which denotes the phenomenon whereby the mind automatically fills in spatially occluded parts of objects. In “event completion”, one might instead conceive of the mind as filling in temporally occluded parts of events. Our findings should probably not be interpreted as the temporal analog of “boundary extension” (Intraub & Richardson, 1989) since it is the middle of the event, as opposed to its boundaries, that is falsely being inserted in memory.

The results presented here are compatible with the idea that people are confusing on-line predictions (Zacks, Speer, Swallow, Braver, & Reynolds, 2007) with truly seen elements. However, it is also possible that the false memory in these tasks is due to schema- or principle-based post hoc inferences. These could potentially be related to encoding or recall mechanisms in memory. The precise underlying machinery responsible for this “causal filling in” awaits more thorough examination in follow-up experiments.

Acknowledgements

We would like to thank Ray Xiong, Sarah Hailey, Andrea Klestadt who helped run participants in this study. We

would also like to thank Brian Scholl, Amit Almor, Alex Shaw and Brandon Liverence for helpful comments.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.cognition.2011.04.007](https://doi.org/10.1016/j.cognition.2011.04.007).

References

- Avrahami, J., & Kareev, Y. (1994). The emergence of events. *Cognition*, *53*, 239–261.
- Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development*, *72*, 708–717.
- Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual Organization* (pp. 213–263). Hillsdale, New Jersey: Lawrence Erlbaum.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Bransford, J., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigators of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*, 717–726.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*, 512–521.
- Hubbard, T. L. (1995). Environmental invariants in the representation of motion: Implied dynamics and representational momentum, gravity, friction, and centripetal force. *Psychonomic Bulletin & Review*, *2*, 322–338.
- Intraub, H., & Richardson, M. (1989). Wide angle memories of close-up scenes. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *15*, 179–187.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585–589.
- Michotte, A. (1946/ English transl. 1963) *The Perception of Causality*, Basic Books.
- Newton, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, *12*, 436–450.
- Pelli, D. G. (1997). The videotoolbox software for visual psychophysics. *Spatial Vision*, *10*, 437–442.
- Rauschenberger, R., & Yantis, S. (2001). Masking unveils pre-amodal completion representation in visual search. *Nature*, *410*, 369–372.
- Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*, 236–257.
- Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind/brain perspective. *Psychological Bulletin*, *133*, 273–293.