To Have Seen or Not to Have Seen: A Look at Rensink, O'Regan, and Clark (1997)

Ronald A. Rensink
Departments of Psychology and Computer Science, University of British Columbia

Abstract

Rensink, O'Regan, and Clark drew attention to the phenomenon of change blindness, in which even large changes can be difficult to notice if made during the appearance of motion transients elsewhere in the image. This article provides a sketch of the events that inspired that article as well as its subsequent impact on psychological science and on society at large.

Keywords
change blindness, visual attention, scene perception, consciousness

In *Perspectives on Psychological Science*, 2018, Vol. 13(2) 230–235. Penultimate draft.

To Have Seen or Not to Have Seen: A Look at Rensink, O'Regan, and Clark (1997)

Vancouver, Canada, spring of 1992. I had just finished my PhD and was preparing to move to the Vision Sciences Lab at Harvard to start my postdoctoral work with Patrick Cavanagh. A month before I left, a notice appeared announcing a conference—Vancouver Studies in Cognitive Science—which focused on interdisciplinary perspectives on visual perception.[1] Because I was interested in vision and was trying to develop such perspectives (going from a PhD in computer science to a postdoc in psychology), I thought such a conference might be interesting. "Interesting" turned out to be an underestimate.

Amid the prominent scientists and philosophers presenting at the conference was an unknown: John Grimes, a student of George McConkie at the University of Illinois at Urbana-Champaign. The McConkie group had been investigating the mechanisms underlying reading and, as part of their studies, had previously shown that a change in text case (e.g., from upper to lower case) during a saccade went largely unnoticed (e.g., McConkie & Zola, 1979). Although interesting, these results were thought to pertain only to reading. However, Grimes showed a similar failure to notice large changes in images of real-world scenes (see Grimes, 1996). The audience—myself included—was stunned: How could such a failure be reconciled with the belief that our visual system lets us see everything that happens in front of our eyes?

I moved this question to the back of my mind and began my postdoctoral work. However, the failure to see changes made during saccades continued to haunt me. Eventually, an opportunity to investigate this phenomenon appeared. As I was finishing my postdoctoral research, Ken Nakayama and Whitman Richards approached me about a position at Cambridge Basic Research (CBR), a new laboratory in Cambridge, Massachusetts. CBR was funded by the Nissan Motor Company, with the goal of carrying out scientific research into issues relevant to automobile driving. One of the biggest causes of driving accidents was "driver looked but failed to see," a poorly understood phenomenon in which drivers would collide with pedestrians and cars—and sometimes even trains!—directly in front of them, even when visibility was good. It seemed to me that this phenomenon might have some connection with the effect found by Grimes and

McConkie. So, when Ken and Whitman asked what problem I wanted to work on, I had an immediate answer.

The question now was how to investigate this phenomenon. Our lab could not afford the type of equipment used by Grimes and McConkie, which involved cameras running at 1000 Hz and computer systems that could change the contents at a particular location in the image the moment the eyes fixated on it.[2] Instead, I tried to develop a technique that would capture at least some of the effect using much simpler equipment. As part of this, I considered the situation from the point of view of the retina: an original image, followed by an image with considerable motion blur (because of the saccade), followed by the altered image, followed by another blurred image, and so on. It seemed to me that this situation might be approximated by showing an original image for a few hundred milliseconds, followed by a brief blank (essentially, an extreme case of blur), followed by the altered image, and then another blank. As to the dependent measure, I had done many studies in visual search, which typically used response time. It seemed like a good idea here as well—this quantity could be measured simply by cycling the image sequence until the observer reported the change. This approach was the origin of what would later be called the *flicker paradigm*.

I had hoped that the flicker paradigm would replicate part of the effect found by Grimes and McConkie; I recall hoping for a lag of a second or so until the change was seen. Because a change is easy to spot if you already know where it is, I could not try the technique out on myself. So, after creating a prototype, I asked one of my CBR colleagues to look at the flickering sequence and tell me when he saw the change. He looked. And looked. And looked. At that point, the hairs on the back of my neck began to rise. The paradigm seemed to have captured most—if not all—of the effect found by Grimes and McConkie: changes could be difficult to notice, even when fairly large (Figure 1).[3] The effect was therefore not limited to changes made during eye movements but involved mechanisms involving visual perception more generally.

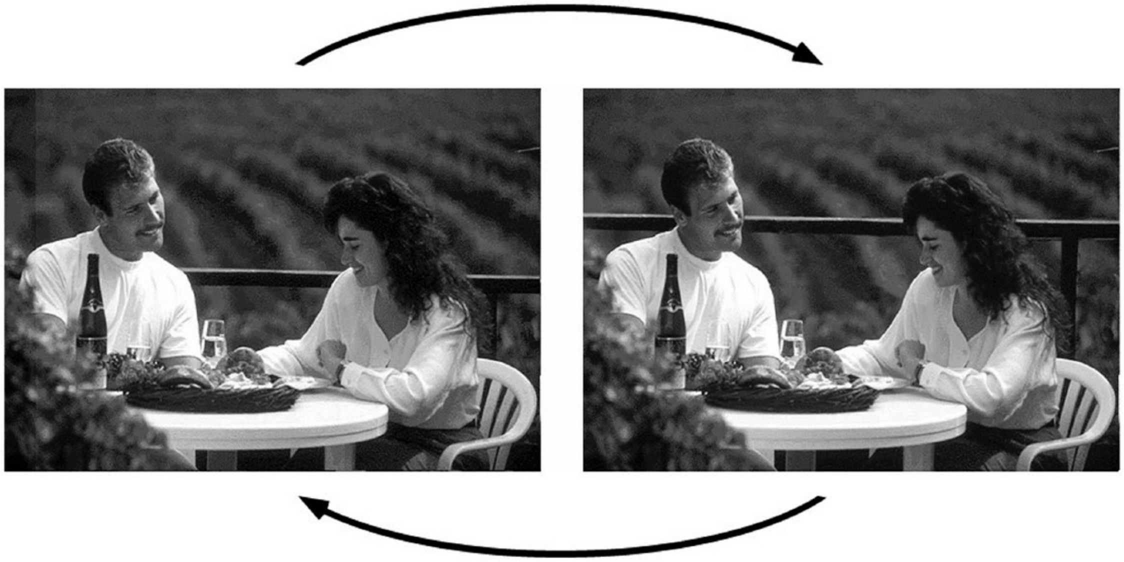

Figure 1. Example of a change in a scene. Here, the position of the background railing changes between images, which alternate every 640 ms. Despite this change being easy to see (once pointed out), observers take, on average, more than 10 s (16 alternations of the image) to notice it. Adapted from Rensink, O’Regan, and Clark (1997).

As to what these mechanisms might be, three observations were key: (a) When the brief blanks were omitted, changes became easy to see, presumably because the mechanism was drawn to the motion transients caused by the “blinking” of the changing items; (b) a considerable speedup could also occur via high-level control (e.g., based on interest); and (c) the long times often needed to see change suggested that the mechanism was severely limited in capacity. Together, these observations strongly suggested that the mechanism was visual attention. Once that had been worked out, the only remaining thing was to settle on a name for the effect. After trying out several candidate names on local researchers and students, a clear winner emerged: *induced change blindness*, later shortened to just *change blindness*.

Within a few weeks, I had an initial draft of a manuscript ready, with the assistance of two visiting researchers at CBR, Kevin O’Regan and Jim Clark. We thought the reviewers would love it. They did not.[4] However, the editor of *Psychological Science* at the time, John Kihlstrom, thought that change blindness was a valuable contribution, and his opinion prevailed. The article then made its way out into the world.

Impact

The finding that changes in scenes become difficult to see when made during an image flicker had a considerable impact on two sets of issues: (a) how visual attention relates to conscious visual perception, and (b) the nature of scene perception. With regard to the first issue, the article proposed that focused attention is needed to see change; more precisely, to consciously perceive an item as changing, focused attention must be allocated to its representation (a visual object) at the moment the change is made. With regard to the second issue, the limited capacity of visual attention indicated that—contrary to our impressions as observers—not all aspects of a scene in front of a viewer are perceived simultaneously even under normal viewing conditions and that the construction of a scene representation was therefore more about the coordination of various processes than it was about the simple accumulation of information.

The implications were wide-ranging. For example, the proposal that attention is needed to see change meant that change blindness could occur whenever attention was prevented from being drawn to the location of the change. This proposal explained why changes could be difficult to see when made during a movie cut or when briefly occluded, as was found in work by Dan Simons and Dan Levin around that time (Simons, 1996; Simons & Levin, 1998). It also explained how change blindness could be induced in other ways too, such as making the change during a sudden translation of the image or during an eyeblink (see Rensink, 2002). Using simplified displays, researchers found that only about four items—or groups of items—could be seen to change at a time (Luck & Vogel, 1997; Rensink, 2000b), an estimate consistent with earlier work on change detection (e.g., French, 1953; Phillips, 1974). In addition, variants of these techniques eventually yielded considerable insight into the nature of iconic and visual short-term (or working) memory (e.g., Brady, Konkle, & Alvarez, 2011; Rensink, 2014; Sligte, Scholte, & Lamme, 2008). The proposal that attention is needed to consciously see change also had clear connections with the topic of consciousness. Among other things, it connected with work being done around that time by Arien Mack and Irv Rock (1998) on inattentional blindness, the failure to see an item when attention is deflected away from it. Later developments included the discovery of the ability to "sense" a change, even when the observer had no conscious picture of it (Rensink, 2004),[5] and the proposal of different kinds—or, at least, grades—of conscious visual experience, each involving a different kind of attentional process (Rensink, 2015). More generally, the phenomenon of change blindness (along with inattentional blindness) helped inspire a wave of studies on the relation between attention and conscious awareness, a topic that remains an active area of work in both psychology and philosophy (e.g., Cohen, Dennett, & Kanwisher, 2016; van Boxtel, Tsuchiya, & Koch, 2010).

Meanwhile, the finding that severe limits exist on what can be perceived about a scene at any moment connected four areas of research—visual attention, visual memory, eye movements, and scene perception—that had been investigated separately. One outcome was an ability to determine what an observer found interesting in a scene: If an item had no unusual visual structure (no unique colors, say), the time needed to see it change would indicate how quickly it drew attention on the basis of high-level factors, such as how interesting the observer considered it to be. This ability proved useful for the study of things such as expertise (Werner & Thies, 2000) and addiction (Jones, Jones, Smith, & Copley, 2003). A related development was the proposal that visual perception involved not one but two pathways: In addition to a selective, attentional stream was a nonselective, nonattentional stream that computed scene layout and gist, with these estimates guiding attention so that it arrived at the right item at the right time (Rensink, 2000a). This "just-in-time" system provided early support for the work on scene gist that was just emerging (e.g., Oliva & Torralba, 2001); indeed, it may have helped spark a renewed interest in the perception of scenes more generally,[6] setting the stage for further studies that showed scene perception to be a dynamic process that begins with a rapid global analysis, followed by the recognition of individual objects (e.g., Malcolm, Groen, & Baker, 2016; Oliva, 2013). This new view of scene perception in turn had practical implications. Under normal conditions, a just-in-time system could represent virtually everything in one's field of view. However, if attention were diverted—say, by cell phone use while driving—an observer could miss important objects and events. Our improved knowledge of these matters lets us understand not only how situations of this type might occur but also how to reduce them, via more effective interfaces for such things as automobiles, airplanes, and computers (e.g., Rensink, 2011; Ware, 2013).

Missteps

Nothing in life is perfect, and Rensink, O'Regan, and Clark (1997) is no exception: As with anything new, it takes a while to get completely clear on all the issues involved. For the most part, I think that the suggestions we put forward in the article have largely stood the test of time. However, there are a few things I would have done differently. The first involves the notion of a dense[7] spatiotopic buffer to accumulate the contents of individual eye fixations. Such a buffer was widely believed at the time to form the basis of the stable "picture" of the world we experience. However, some eye-movement researchers had begun to argue against the existence of such a buffer, and we suggested that the finding of change blindness supported their position. However, as pointed out by researchers such as Scott-Brown, Baker, and Orbach (2000), although the existence of change blindness does show that our ability to represent the dynamic aspects of a scene is limited (at least for the processes underlying conscious perception), it does not say much about limits on our ability to represent static structure. We should have simply stated that change blindness could be explained without requiring such a buffer and left it at that.[8]

A second misstep involved the notion of sparse representation. We had proposed that a scene could be represented by a relatively small set of structures entered by attention into visual short-term memory. We had also suggested that the (dense) contents of each eye fixation were briefly held in a simple form of visual memory that could be overwritten by the contents of subsequent fixations. However, we did not discuss this latter representation further. This decision was in line with a common convention at the time in which object representation, say, was concerned only with high-level structure (e.g., the general shape of the object) and not low-level "stuff" (e.g., the distribution of textures or colors within it).[9] However, not all readers followed this convention, and some took our failure to further mention dense representation as meaning that we did not believe that dense representations existed in visual perception. This was not so: We had simply not considered these to be part of scene perception proper. Our take on this issue should have been made clearer. A few years later, I put forward a more detailed proposal that stated explicitly that low-level representations were dense and—in the absence of attention—volatile, being overwritten or dissipating within a few hundred milliseconds (Rensink, 2000a). However, by then, the damage had been done. The mistaken notion that we had proposed that visual perception involved only sparse representations persisted far longer than it should have.

A related misunderstanding involved the erroneous belief that we had argued against the existence of long-term memory in visual perception. In the interests of simplicity, we had left connections to visual long-term memory an open issue, apart from an suggestion that it had to involve "relatively sparse" representations. Again, a few years later, I showed that the existence of a long-term memory in scene perception was compatible with the existence of change blindness (Rensink, 2000a). However, it took a while for all relevant parties to become aware of this compatibility.

Final Thoughts

Over the years, the phenomenon of change blindness has appeared in numerous science exhibits, shows on popular science, and textbooks (e.g., Wolfe, 2014). I had originally considered it simply as a tool—a powerful one, to be sure, but still just a tool—to investigate perception; its general appeal came as something of a surprise. In hindsight, I think three factors played a role in its popularity. First, it was simple and could be demonstrated using standard video setups. Second, there was an element of fun: After the viewer searched the flickering image for a short while, the changing item would suddenly jump into their conscious experience. Third, this sudden appearance was highly surprising for anyone who believed (as most did at the time) that their visual system always let them see everything that happened in front of them.

These factors may also help explain the fate of earlier studies on failures of conscious perception. Although such studies had received considerable attention at the time they appeared (e.g., Hochberg, 1968, 1978; Neisser,1976), they were not taken further. Why was this? Part of the reason may have been that they often used a "one-shot" paradigm, which measured the probability of an observer detecting a change in a single transition between two images (see Rensink, 2002). Although in many respects equivalent to the flicker paradigm (in which the images kept cycling), it did not include the sudden appearance of the change when the stimulus was finally attended, thereby losing much of the visceral impact. In addition, the special experimental setups typically used in those studies (e.g., eye tracking, superimposed images) may have suggested that their results were not relevant for visual perception under more normal conditions.

However, there may have been another factor that was even more important. The 1970s and 1980s saw the emergence and eventual dominance of a wave of vision research that focused on the rigorous treatment of the lowest levels of visual processing. Most researchers of the time may therefore have simply set aside issues of high-level structure for a later day. If so, the appearance of Rensink et al. (1997) and related works in the mid-1990s may have been the first signs of that day finally dawning.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Notes

1. The presentations at this conference appeared in Akins (1996).

2. This is likely why the original work of Grimes and McConkie was not followed up immediately.

3. Demonstrations of this effect can be found at various websites (e.g., http://www.gocognitive.net/demo/change-blindness).

4. Curiously, the reviewers disagreed as to why: Some thought that any failure based on attentional limits was obvious; others thought that the effect was impossible—an observer would always notice any change that was sufficiently large.

5. Skepticism was initially expressed about the possibility of a different mechanism for sensing (Simons, Navarez, & Boot, 2005). However, the results of later work (e.g., Ball & Busch, 2015; Galpin, Underwood, & Chapman, 2008) have tended to support the original proposal.

6. There had already been some interest in the nature of scene representation around that time (e.g., Boyce & Pollatsek, 1992;O'Regan, 1992). However, the scale of this interest increased considerably in the subsequent decade.

7. A dense representation has no significant gaps in its coverage; it contains at least some fraction of the information (in bits) in the incoming light. A sparse representation, in contrast, does have such gaps and contains just a small number of bits.

8. Indeed, most current theories of vision avoid any mention of such a buffer. Like the luminiferous ether in physics, it is difficult to find direct evidence against its existence. And like the ether, it simply is not needed.

9. This convention may have arisen in part from a fear that incorporating low levels into theories about objects or scenes would have been tantamount to a complete theory of visual perception. A similar separation existed in theories of low-level vision, where elements underlying texture or visual search, for example, were proposed with little regard as to how they might be incorporated into objects or scenes.

References

Akins, K. A. (Ed.). (1996). *Vancouver Studies in Cognitive Science: Perception* (Vol. 5). New York, NY: Oxford University Press.

Ball, F., & Busch, N. A. (2015). Change detection on a hunch: Pre-attentive vision allows "sensing" of unique feature changes. *Attention, Perception, & Psychophysics*, 77, 2570–2588.

Boyce, S. J., & Pollatsek, A. (1992). Identification of objects in scenes: The role of scene background in object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 531–543.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, 11(5), Article 4. doi:10.1167/11.5.4

Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences*, 20, 324–335.

French, R. S. (1953). The discrimination of dot patterns as a function of number and average separation of dots. *Journal of Experimental Psychology*, 46, 1–9.

Galpin, A., Underwood, G., & Chapman, P. (2008). Sensing without seeing in comparative visual search. *Consciousness and Cognition*, 17, 672–687.

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Vancouver Studies in Cognitive Science*: *Perception* (Vol. 5), pp. 89–110. New York, NY: Oxford University Press.

Hochberg, J. (1968). In the mind's eye. In R. N. Haber (Ed.), *Contemporary Theory and Research in Visual Perceptio*n (pp. 309–331). New York, NY: Holt, Rinehart & Winston.

Hochberg, J. E. (1978). *Perception* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Jones, B. T., Jones, B. C., Smith, H., & Copley, N. (2003). A flicker paradigm for inducing change blindness reveals alcohol and cannabis information processing biases in social users. *Addiction*, 98, 235–244.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–280.

Mack, A., & Rock, I. (1998). *Inattentional Blindness*. Cambridge, MA: MIT Press.

Malcolm, G. L., Groen, I. I. A., & Baker, C. I. (2016). Making sense of real-world scenes. *Trends in Cognitive Sciences*, 20, 843–856.

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics*, 25, 221–224.

Neisser, U. (1976). *Cognition and Reality*. San Francisco, CA: W. H. Freeman.

Oliva, A. (2013). Scene perception. In J. S. Werner & L. M. Chalupa (Eds.), *The New Visual Neurosciences* (pp. 725–732). Cambridge, MA: MIT Press.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal in Computer Vision*, 42, 145–175.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, 46, 461–488.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception &Psychophysics*, 16, 283–290.

Rensink, R. A. (2000a). The dynamic representation of scenes. *Visual Cognition*, 7, 17–42.

Rensink, R. A. (2000b). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition*, 7, 345–376.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, 53, 245–277.

Rensink, R. A. (2004). Visual sensing without seeing. *Psychological Science*, 15, 27–32.

Rensink, R. A. (2011). The management of visual attention in graphic displays. In C. Roda (Ed.), *Human Attention in Digital Environments* (pp. 63–92). Cambridge, England: Cambridge University Press.

Rensink, R. A. (2014). Limits to the usability of iconic memory. *Frontiers in Psychology*, 5, Article 971. doi:10.3389/fpsyg.2014.00971

Rensink, R. A. (2015). A function-centered taxonomy of visual attention. In P. Coates & S. Coleman (Eds.), Phenomenal qualities: Sense, perception, and consciousness (pp. 347–375). Oxford, England: Oxford University Press.

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

Scott-Brown, K. C., Baker, M. R., & Orbach, H. S. (2000). Comparison blindness. *Visual Cognition*, 7, 253–267.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science*, 7, 301–305.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 5, 644–649.

Simons, D. J., Navarez, G., & Boot, W. R. (2005). Visual sensing IS seeing: Why "mindsight," in hindsight, is blind. *Psychological Science*, 16, 520–524.

Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PLOS ONE*, 3(2), Article e1699. doi:10.1371/journal.pone.0001699

van Boxtel, J. J., Tsuchiya, N., & Koch, C. (2010). Consciousness and attention: On sufficiency and necessity. *Frontiers in Psychology*, 1, Article 217. doi:10.3389/fpsyg.2010.00217

Ware, C. (2013). *Information Visualization: Perception for Design* (3rd ed.). New York, NY: Morgan Kaufmann.

Werner, S., & Thies, B. (2000). Is "change blindness" attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*, 7, 163–173.

Wolfe, J. M. (2014). *Sensation & Perception* (4th ed.). Sunderland, MA: Sinauer Associates.