Ward, R., Duncan, J., and Shapiro, K. (1996). The slow time-course of visual attention. *Cog. Psych.*, 30:79-109.

Wilken, P., Mattingley, J. B., Korb, K. B., Webster, W. R. and Conway, D. (1999). Capacity limits for detection versus reportability of change in visual scenes. Paper presented at the 26th Annual Australian Experimental Psychology Conference, April, 1999.

Wolfe, J. M. (1999). Inattentional amnesia. In V. Coltheart (Ed.), *Fleeting Memories*. pp. 71-94. Cambridge, MA: MIT Press.

Woodhouse, J. M., and Barlow, H. B. (1982). Spatial and temporal resolution and analysis. In H. B. Barlow and J. D. Mollon (Eds.), *The Senses*. pp. 133-164. Cambridge, England: Cambridge University Press.

10

The Role of Expectations in Change Detection and Attentional Capture

Daniel J. Simons
Stephen R. Mitroff

Imagine that you are driving your car and you come to a red light. While waiting, you play with the radio dial to find a more appealing station. You succeed, and return to waiting for the light to change. A few seconds later, the car behind you blares its horn. You look up only to realize that the light had changed to green while you were fiddling with the radio, and you did not notice when you looked up. Now imagine that you are trying to drive across a busy intersection. You carefully check for oncoming traffic in the near lane and then turn to check the far lane. Thinking your path is clear, you accelerate into traffic, only to realize that you had not seen the person in the crosswalk right in front of you. Take yet another case: you are a pilot trying to land an airplane under relatively foggy conditions. As you descend, you pay close attention to the elevation, speed, and pitch information projected onto your windshield's head-up display. You land the plane, never noticing there was another plane on the runway as you approached. Just this situation was studied in a flight simulator (Haines, 1991), and several professional pilots looked but did not see the other plane until they were too late to avoid it.

These examples have several things in common. First, all involve an attentionally demanding primary task (e.g., landing a plane, watching for traffic, etc.). Second, the unexpected object (e.g., the color of the stoplight, the pedestrian, or the plane) is not the currently attended aspect of the scene. Third, and most importantly for this chapter, the object change or appearance does not automatically capture attention when it is incidental to the primary focus of attention. Although we would like to believe that important, unanticipated events would automatically capture our attention, at least under these natural conditions, they sometimes do not.

Despite the importance of unexpected events for our understanding of attentional capture, research on attention generally has focused on the ability to detect (or ignore) briefly flashed objects when observers are intentionally looking for those objects. For example, the study of visual search has focused on how rapidly observers can spot an odd object among a set of different distractor objects (e.g., Treisman and Gelade, 1980). Similarly, the study of change detection has explored the ability to actively search for changes (e.g., Hollingworth and Henderson, 2000; O'Regan, Rensink, and Clark, 1999; Rensink, 2000; Rensink, O'Regan, and Clark, 1997; Simons, 1996; see Simons and Levin, 1997 for a review). Both lines of research have adopted what we will refer to as "the intentional approach." These studies explore the psychophysics of attention and visual memory–under precisely controlled conditions, what are the limits on our capacity, speed, and precision of detection? What are the thresholds of the attentional system? They determine what we can remember from one view to the next or how much we can attend to when we try. These traditional studies of attention and perception have focused exclusively on intentional search or change detection. The primary task is to find the target or the change. Yet, in the naturalistic examples discussed earlier, the "target" is secondary to the primary task, and detection is incidental to the observer's primary goal.

The intentional approach has provided a number of important insights into the limits of our abilities and the functioning of the visual system. However, such tasks do not account for the entirety of visual experience. To describe human performance fully, researchers need to adopt what we will refer to as "the incidental stance." We need to know not just what the visual system can do, but what it actually does. Although intentional visual search is central to our daily activities (e.g., finding your keys on a cluttered desk), the intentional approach may be suboptimal for exploring the capture of attention by unexpected or unattended objects. By adopting an incidental approach to the experimental study of attentional capture, we may better understand how perception works in conditions such as landing planes or driving through intersections.

Although the incidental approach has been relatively neglected in current empirical work on perception, recent research on change blindness from our lab (Simons and Levin, 1997) and more recent literature on inattentional blindness (Mack and Rock, 1998) have both adopted incidental approaches. Here we will review the findings from these two bodies of literature in the context of intentional approaches to change detection and attentional capture, noting where the intentional and incidental approaches produce similar results and where they differ. By combining both approaches, we may gain a better understanding of the functioning and the function of perception.

10.1 Change Blindness

Over the past few years, the study of change blindness has become a topic of intense study in the field of visual cognition (see *Visual Cognition* 7(1/2/3) for a special issue devoted to change blindness). In part, this focus on short-term visual memory has resulted from a series of striking demonstrations of our inability to notice large changes to complex displays (Grimes, 1996). The findings suggest that we consciously perceive and remember far less of our visual world than we might otherwise believe. The effects are particularly striking given our metacognitive beliefs about perception and vision. We feel that we retain a rich representation of our visual world and that large changes to our environment will draw our attention.

In fact, undergraduate students substantially overestimate their ability to detect change, predicting that they would notice changes that in fact few observers do (Levin, Momen, Drivdahl, and Simons, 2000). In this section, we briefly review evidence from the intentional approach to change detection, and then we discuss findings from the incidental approach.

10.1.1 The intentional approach

A comprehensive overview of the history of intentional approaches to change-detection is not central to the theme of this chapter (see Simons and Levin, 1997, for an overview of the literature on change blindness). Here we provide only a brief overview of more recent findings.

Intentional change detection has adopted two different paradigms. In *discrete* change detection tasks, observers view an initial display, followed by some disruption (e.g., an eye movement: Grimes, 1996; Henderson and Hollingworth, 1999; a blink: O'Regan, Deubel, Clark, and Rensink, 2000; or a blank screen: Simons, 1996; Rensink et al., 1997) and then by a modified version of the display. Subjects typically are required to report the location of the change or to determine whether or not a change occurred. This paradigm is similar to a traditional recognition-memory paradigm in that it allows for the use of signal-detection analyses and has a study/test format. It differs in that recognition memory tasks typically present a complete study set prior to the test but in these discrete change-detection tasks, the test display immediately follows the study display on each trial. Experiments using discrete change detection tasks have demonstrated that the ability to detect large changes to photographs or objects is surprisingly poor when the original and modified scenes are separated by a brief disruption. In all cases, observers are looking for changes but cannot easily find them if the change occurs during the interruption. Of course, when there is no disruption and the modified scene immediately follows the original image, observers readily detect the change. Without the disruption, the motion or luminance transient captures attention.

The second type of intentional task used to study change detection is the "flicker" paradigm (Rensink et al., 1997). An original and modified photograph alternate repeatedly, separated by a brief blank interval. The dependent measure is the number of cycles observers take to spot the change. Even with relatively large changes and after practice with the task, observers often take many cycles to notice the difference. In this paradigm, observers tend to notice changes to the central objects in a scene more than changes to peripheral objects even when the changes are physically equal in magnitude (Rensink et al., 1997). This paradigm has recently been adapted to study the capacity of visual memory and the speed of search (see Rensink's chapter in this volume and also Rensink, 2000). In many ways, the flicker paradigm is better than traditional search tasks in that observers know only the dimension of the change, but not the particular feature. For example, observers may search for a color change, rather than for a particular color. To the extent that researchers are interested in how efficiently and effectively observers can selectively attend to a feature dimension, searching for the dimension of a change is preferable to searching for a particular instantiation of that dimension.

Together, these intentional approaches have produced the following important results: (1) observers often miss large changes to simple displays and to photographs; (2) changes to central objects are noticed more quickly and more often than changes to peripheral objects; and (3) provided that the change coincides with some disruption (e.g., an eye movement, blank screen, eye blink, or blot), the transients that would otherwise capture attention are effectively masked. The fact that observers show change blindness when actively trying to spot changes suggests that they should be equally if not more blind to changes when change detection is not their primary task. However, intentional change detection requires subjects to encode scenes differently than they might under more natural encoding conditions. Given subjects' inaccurate metacognition about change detection (Levin et al., 2000), they might actually adopt less effective coding strategies when intentionally searching for change than when performing some other primary task.

10.1.2 The incidental approach

In the incidental approaches to change detection, observers are not forewarned that a change will occur and are simply asked to perform some other primary task. The displays in these tasks are often dynamic motion pictures or even real world events. A further difference from the intentional change detection tasks is that these incidental paradigms typically include only one critical test trial. Thus, the primary statistic is the proportion of people noticing a given change. This approach is not well-suited to the systematic study of the limits or thresholds of our change-detection abilities; systematic manipulation of the nature of the change in a one-trial study would not provide sufficient experimental power without a prohibitively large number of subjects. The primary purpose of these studies is to see whether the findings of intentional tasks generalize to real-world behaviors and to explore the extent of change blindness when we are not actively searching for change. These findings often provide striking examples of change blindness that appear to contradict our intuitive belief that we represent our world in great detail and that large changes will capture our attention. Further, they allow an exploration of the detection of changes to "attended" objects (Levin and Simons, 1997; Simons and Levin, 1997, 1998). Experiments using intentional change detection tasks consistently reveal better detection of changes to central objects in a scene (Rensink et al., 1997). Of course, in intentional tasks, observers are actively searching for change and are likely to focus on the central objects first. When they are searching for change, they encode the features of an object in order to try to detect something that differs. However, in the real world, we do not necessarily encode objects (even central objects) in such a way that would allow us to notice changes to features.

Although relatively few researchers have explored the perception of motion pictures, much attention has been paid to the fact that people fail to detect continuity errors or editing mistakes (Hochberg, 1986). People are often fascinated by these mistakes, and the popular media draws attention to them as if they were an anomaly.

For example, in a recent episode of ABC television's show *Dateline* that focused on editing errors in Academy Award-winning movies, the reporter asks "what is it about filmmakers that they can shoot so carefully, so many takes, and still miss something so obvious, something the audience can see clearly?" The reporter's intuition is that these changes are often noticed by the audience and that it is hard to imagine how the filmmaker failed to see them. Yet, in most motion pictures, continuity errors occur regularly despite the best efforts of the continuity editors. The anomaly is not that people miss these changes, but that they sometimes actually do see them. Levin and Simons have explored this aspect of change blindness empirically (Levin and Simons, 1997). In one experiment, they created a brief (approximately 1 minute) motion picture of a conversation between two actors. Every time the camera instantaneously changed (or "cut") to a new position, they intentionally inserted at least one editing mistake. For example, in one shot, a woman was wearing a scarf, but in the next shot it had disappeared, only to reappear in the following shot. Other changes included shifts to the positions of the actors' arms, a change in the color of plates on the table from red to white and then back to red, and a shift in the position of food from one plate to the other. They explored the ability to notice these changes by using an incidental task. Subjects were told to "watch this brief video and pay close attention." When the video ended, the subjects were asked if they had noticed any changes to the objects, clothing, or body positions in the video. Of the then subjects, only one reported noticing any of the nine intentional changes, and even that change was described only vaguely. Even on a second viewing, after being warned by the questions following the first viewing about what features to focus attention on, subjects still saw fewer than two of the nine changes on average. Consistent with findings from the intentional search tasks, subjects were generally unable to detect changes to a natural scene across a disruption (in this case, a film cut).

All the changes in this study were to peripheral objects, and we know from studies using the flicker task that peripheral changes are generally noticed more slowly than central changes. This incidental task can also be used to explore the detection of changes to central objects in a scene. In another version of the conversation film (Simons, 1996), the only change involved a central object. The scene opened with a shot of one actor pouring cola from a bottle into her cup. As she set down the bottle, the camera panned to show the approach of the other actor. When the camera panned back to the table, the bottle was gone and in its place was a cardboard box of roughly the same dimensions. When ten observers were asked to view this film and then to describe what they had seen, none noticed the change. Interestingly, several did mention the bottle in their descriptions of the movie. They just did not notice its disappearance. This finding suggests that even changes to the central object in a scene can go unnoticed if subjects are not explicitly looking for changes. The fact that observers described the bottle suggested that it was central to the scene and that they had focused attention on it during the film. However, the failure to describe the box suggested that it might not have been a central object when the camera returned. Perhaps subjects only notice changes when the target object is attended both before and after the change.

To investigate this possibility, another set of motion pictures was created in which the changed object was central throughout the film (Levin and Simons, 1997). A single character performed a simple action such as getting up to answer the phone or walking through a classroom and sitting in a chair. Given that the character was the only person in the film and was the only moving object, subjects are likely to attend to this character throughout the film (Dmytryk, 1984). During the action sequence, the camera cut to a new view of the character, and after the cut, the original actor was replaced with a different person who then completed the action. By cutting on the motion of the actor, the film appeared to depict a single, continuous event. Subjects were instructed to watch the brief (10sec) movie and to pay close attention. Immediately after the movie ended, subjects wrote a description of what they had seen. Any indication of change detection in these written descriptions was taken to indicate successful detection. In other words, subjects did not need to report that the person had changed identity, but only to note that something about the actor had changed (there were also clothing differences between the two actors). Across eight different change films (with four pairs of actors), on average only 35% of subjects noticed a change. The written responses often included detailed descriptions of the actors and always mentioned the person both before and after the cut, suggesting the actor was the central object throughout the film. Yet, most subjects failed to see the change. Under incidental encoding conditions, observers fail to represent or compare those features that change from one view to the next. Another set of observers viewed all of the change films and an equal number of no-change films under intentional conditions – they successfully detected which films included a change.

Under incidental conditions, focusing attention on the central object does not guarantee that all of the features of that object are encoded and retained (see Simons and Levin, 1997). This result appears discrepant from a recent finding that used an intentional task with simple arrays of objects (Luck and Vogel, 1997). In this short-term memory task, observers tried to remember the properties of up to four objects. Although they generally could only remember four individual features in isolation, they could remember many more if the features were tied to individual objects. In other words, subjects could remember all of the properties of each of four objects; performance was limited by the number of objects and not the number of features in the display, and attention to an object allowed subjects to remember all of its feature dimensions. This finding raises an interesting possibility: perhaps observers only retain those properties that are the focus of their intentional efforts to remember. When subjects know they could be tested on any of a number of feature dimensions, they encode the display so that they will remember those features more accurately. In contrast, when they do not know that they may be tested, they encode only those features necessary to comprehend the meaning. They do not automatically encode those features that would discriminate one actor from another.

Although motion pictures are a good medium for studying incidental change detection and they are closer to actual experience than simple arrays or photographs, they still may not adequately test how we encode objects in the real world. When encountering another person, we may well encode and retain more information than we would from a brief glimpse of an actor in a passively observed motion picture. Simons and Levin tested this possibility by using a real-world change detection task (Simons and Levin, 1998). In these studies, an experimenter approached a pedestrian (the subject) to ask for directions. During their conversation, two people carrying a door passed between the experimenter and the pedestrian, and during that interruption, the first experimenter was replaced by a second experimenter who was wearing different clothing. Even though subjects engaged in an interaction with both the first and the second experimenter, 50% of the subjects did not notice the person change. The two experimenters in this study were both approximately 30-year-old white men with short dark hair, but in other respects they were quite different. For example, they were 6 cm different in height, and the shorter experimenter had a much deeper voice.

Interestingly, the 50% who missed the change (typically faculty or staff) were all older than the experimenters and the 50% who noticed (typically students) were all the same age or slightly younger than the experimenters (Simons and Levin, 1998). Although this effect could be attributed to aging, most of the "older" subjects were not elderly. It seemed more likely that the effect resulted from a difference in how people encode members of their own social group as opposed to members of a different social group. When viewing members of their own social group, people tend to encode features that differentiate individuals. In contrast, when viewing members of a different social group, people tend to encode group membership information and to ignore differentiating features (Rothbart and John, 1985). The older subjects simply coded the experimenter(s) as a student asking directions. Hence, they had not encoded any features that would change as a result of the switch in experimenters. In contrast, the younger subjects were more likely to encode features that would differentiate individuals, so they noticed the different features when the experimenters changed. To explore this possibility further, the experimenters were made members of an outgroup to the younger subjects by dressing them as construction workers. In this condition, only 35% of a new group of younger subjects noticed the change, down from 100% in the earlier version (Simons and Levin, 1998). This ingroup/outgroup difference likely would not have been noticed using an intentional change detection task. If observers knew that the person might change, they would likely code individuating features regardless of their group membership. When observers are not actively searching for a change, they tend to focus on the meaning of a scene – what is important for their immediate actions and goals. Such encoding is unlikely to lead subjects to focus on specific visual details. As a result, they are even less likely to detect changes, even if the changes are to central objects.

Although these findings could be taken to suggest that relatively little if any visual information is retained in daily experience, this conclusion would be premature and based on faulty logic. Successful change detection across an interruption does require a representation of both the initial and changed objects (or at least of the difference between them). However, change blindness does not imply the absence of a representation (see Simons, 2000a for a discussion). Subjects could fail to detect a change because they lack a representation of what had been present before, or they may be change-blind because they accurately represented the initial scene but not the modified scene. They may even accurately represent the features of both the initial and changed scene but still not detect the change if they fail to compare the scenes in the appropriate way (Scott-Brown, Baker, and Orbach, 2000; Simons, 2000a). In the real world case, subjects may not have made such a comparison because the change did not produce any inconsistency with the meaning of the scene and it did not capture attention. Further, the visual system likely assumes stability, so unless a transient signals a change, there would be no reason to try to compare the initial and modified representations. Several suggestions from research with motion pictures are consistent with the possibility that we actually do represent the details of the changed object when change detection fails. For example, in the bottle-to-box change (Simons, 1996), observers often described the bottle after the film ended. If change blindness results from the overwriting of the initial object by the changed object, subjects should not have been able to describe the bottle. Similarly, a number of subjects in the person-change films described the properties of the first actor rather than the second actor (Levin and Simons, 1997).

A more recent real-world experiment also examined the nature of our representations when change detection fails (Simons, Chabris, and Levin, 1999). In this study, an experimenter approached a pedestrian (the subject) to ask directions to a gymnasium. The experimenter was holding a red and white striped basketball and was wearing gym clothes. As the pedestrian was providing directions, a group of confederates walking on the sidewalk passed between the experimenter and pedestrian, and one of them surreptitiously took the basketball. If the initial representation is replaced following the disruption, subjects who fail to notice the change should not remember the basketball. Based on a series of probing questions, only 20% of subjects spontaneously reported the change. Surprisingly, when asked directly whether they thought the experimenter originally had a basketball, more than 50% said yes. Furthermore, most were able to describe the unusual appearance of the ball and none described a canonical basketball. Results of a no-change control condition suggest that this finding does not result from the leading questions. In essence, people may have accurate representations of both the original and modified features and still fail to notice changes.

10.1.3 Summary

Both incidental and intentional approaches have produced striking evidence for change blindness. Observers often fail to detect large changes to photographs or scenes from one view to the next, even when they are actively looking for changes. Although both intentional and incidental approaches provide evidence for change blindness, neither approach alone completely captures the phenomenon. Evidence from the intentional approach suggests that changes to central objects are detected more readily than changes to peripheral objects. From this finding, we might be tempted to conclude that change detection is driven solely by how central an object is to the scene. We would be tempted to infer that attention is drawn

to central objects and that attention is sufficient for the accurate representation and detection of changes. Yet, evidence from the incidental approach shows that even changes to attended objects can go undetected, suggesting that attention to an object is not sufficient for change detection. If we only considered evidence from the incidental approach, we would be similarly misled. Although we might intuitively expect some benefit of focused attention, the incidental approach is ill-suited for systematic exploration of the effects of centrality in a scene. If we relied only on findings from the incidental approach, we might incorrectly conclude that changes to central objects are no more likely to be detected than other changes. By adopting the intentional approach, we can gain a better understanding of our change-detection mechanisms, and by adopting the incidental approach, we can gain a better appreciation for how those mechanisms operate in the real world. By combining both approaches, we can avoid the pitfalls engendered by adopting either approach in isolation.

10.2 Attentional Capture

Perhaps the only domain to be studied exclusively using incidental tasks is that of "inattentional blindness." Inattentional blindness refers to the tendency not to see unattended objects (Mack and Rock, 1998). The study of inattentional blindness has its roots in research on selective attention and dichotic listening (Moray, 1959; Treisman, 1960, 1964; Neisser and Becklen, 1975; Neisser, 1979; Becklen and Cervone, 1983; Holender, 1986; Stoffregen and Becklen, 1989). Such studies focus on the ability to report what was presumably unattended information, often under conditions in which subjects do not expect any information to be present. Similarly, most studies of inattentional blindness explore the perception of unattended and unexpected information (e.g., Mack and Rock, 1998; Newby and Rock, 1998; Rock, Linnett, Grant, and Mack, 1992). In essence, these studies explore the tendency for different forms of information to capture attention when subjects have no prior expectation that the information will be presented at all. Once subjects become aware of the unexpected object, event, or sound (either because they noticed it or because they were questioned about it), they now expect the unexpected and the nature of the task has changed fundamentally. Consequentially, most studies of inattentional blindness have relatively few critical test trials. Because intentional tasks are rarely useful when they have only one trial, inattentional blindness has been explored almost exclusively with incidental tasks. However, the empirical study of attentional capture has employed intentional tasks, perhaps the most prominent of which is visual search (see Simons, 2000b for a more complete review of the links between attentional capture and inattentional blindness).

10.2.1 The intentional approach

Intuitively, it would seem sensible for a visual system to detect unexpected or unusual events automatically, drawing attention to them, thereby allowing a rapid and appropriate response. For example, if a predator suddenly charged toward you, you would want to become aware of it without having to consciously and effortfully shift attention to it. Because we cannot introduce predators into a laboratory setting, researchers have adopted the methods of visual search as a proxy for this more natural situation. Visual-search studies of attentional capture look for indications that a target object was detected without any effort or that it drew attention away from other items (Yantis and Jonides, 1984; Folk, Remington, and Johnston, 1992; Theeuwes, 1994; Yantis, 1998). Effort in a search task is operationally defined by the effect of this target item on search latencies. If it captures attention, it should be processed before other items in the display. Thus, if the distinctive item is the target of the search task, then search latency will be unaffected by the number of other items in the display. Similarly, if the distinctive item is one of the distractor items, search latency will be relatively slowed because attention will be drawn to this incorrect item rather than to the target of the search.

Several different variants of this search task have been developed. In the "irrelevant-feature search task," observers perform a traditional search task (e.g., looking for an L among Ts), and one item in the display is different from all of the other items (e.g., it onsets later or is a different color). This distinctive item does not predict the location of the search target; it is no more likely to appear as the target than as one of the distractors. Observers know that the distinctive feature is irrelevant and that there is no reason to focus attention on it – it will not aid or impair their search performance. If this irrelevant feature captures attention when it happens to coincide with the search target, detection latency should be unaffected by the number of other items in the display. This methodology allows a systematic exploration of the sorts of features that might capture attention. Across a number of studies, the one feature that consistently appears to capture attention is a sudden and late onset of the target item (see Yantis and Hillstrom, 1994; Yantis and Jonides, 1996, although there is still debate about whether this result is based entirely on bottom-up attentional capture or whether the observer's attentional set influences search and contributes to speeded search; see, Folk et al., 1992; Yantis, 1993). Abrupt onsets reduce search latencies, suggesting that they can automatically draw attention.

In another intentional approach to attentional capture, observers view an irrelevant spatial cue prior to a search task. The cue does not predict the target location. Yet, when the properties of this cue match those of the target (e.g., a color precue for a color target), and if it happens to appear in the location of a target, observers show speeded search performance (Folk, Remington, and Wright, 1994). Similarly, when the spatial precue happens to appear in the location of a distractor item, performance is impaired. Presumably, attention is drawn to this cue automatically. When the cue is consistent with the nature of the search (e.g., subjects are looking for a color target and they receive a color cue), they cannot help but pay attention

to it even though it is known to be irrelevant to the primary task (Folk et al., 1994).

In both of the preceding cases, observers perform an intentional search task, fully expecting this irrelevant item to appear on every trial. They know that focusing attention on the distinctive feature will not improve their performance. Yet, they still shift attention to it. Such findings have been taken to suggest that some distinctive features (e.g., abrupt onsets) draw attention automatically. Even though the feature is irrelevant, observers cannot help but attend to it.

Unfortunately, an intentional task may not adequately reflect the sorts of attentional capture that we would need in order to avoid a charging predator. These intentional tasks effectively focus on the ability not to attend to something that we know to be irrelevant. The distinctive item is always present but does not help in the primary task performance. In the real world, however, the predator does not always appear, and when it does, it is rarely irrelevant or expected. Do such unexpected stimuli capture attention? Only one experiment using the irrelevant-feature search paradigm has explored this question (Gibson and Jiang, 1998). For the first 192 trials of this study, observers performed a traditional conjunction search task. On the 193rd trial, one of the items in the search display was a different color from the others. Yet, this distinctive and unexpected item failed to capture attention; search performance was no different from what was predicted by the preceding trials. The absence of attentional capture in this study raises a somewhat radical possibility: in the absence of expectations, unusual and distinctive objects may not capture attention. In other words, attention may not automatically be drawn to the sudden appearance of a predator. The intentional approach is not well-suited to explore whether unexpected objects capture attention. Hence, recent experiments on inattention have adopted incidental approaches.

10.2.2 The incidental approach

Studies of inattentional blindness have used two different paradigms. In an approach introduced by Mack and Rock (Mack, Tang, Tuma, and Kahn, 1992; Rock et al., 1992; Mack and Rock, 1998), subjects engaged in a primary task of determining which line of a cross (the horizontal or vertical) was longer. On each trial, the cross was flashed for 200 msec. On the critical trial (typically the third or fourth trial), another object appeared along with the cross. Subjects were then asked to report whether they had seen the unexpected object. Attentional capture is indicated by higher rates of detection, and inattentional blindness is indicated by failed detection. Once subjects were asked about the unexpected object, they knew to look for it on subsequent trials. A subsequent trial with the object was therefore a divided-attention trial (divided between searching for anything other than the cross and performing the line-judgment task). Subjects were consistently better able to detect the object on these divided attention trials, suggesting that the absence of expectations plays a role in inattentional blindness.

Experiments using this paradigm reveal a surprising degree of blindness for the unexpected object (Mack and Rock, 1998). Even when the object is a different color it is no more likely to be noticed than a black object (both are noticed by 25 – 75% of subjects, depending on the conditions). Of course, there are differences in which features are noticed and which are not. First, objects that are relatively close to the focus of attention (i.e., the cross) are more likely to be noticed (Mack and Rock, 1998). Furthermore, stimuli that are meaningful to observers are often noticed (Newby and Rock, 1998). For example, subjects are more likely to notice their own name than they are their own name with one letter changed (e.g., Jake vs. Jeke). They also tend to notice some schematic objects such as smiley faces.

One concern about drawing strong conclusions from Mack and Rock's line-judgment task is that the target object is flashed only briefly. Under more natural conditions, unexpected objects would likely be visible for more than 200 msec, and subjects would not be fixating a single point. For example, pilots in the flight-simulator study described at the beginning of the chapter had an extended opportunity to view the unexpected target object (Haines, 1991).

Studies using a different paradigm, "selective looking," have explored the capture of attention by unexpected objects and events in displays that last substantially longer than 200 msec (Neisser and Becklen, 1975; Neisser, 1979). In these studies, observers view a display with two ongoing events, but they are only required to monitor one of them. During their ongoing performance of this task, an unexpected event occurs. At the end of the trial, subjects are asked to report what they saw.

This method for studying attentional capture was developed by Ulric Neisser and colleagues during the 1970s and 1980s (Neisser and Becklen, 1975; Littman and Becklen, 1976; Neisser, 1976; Neisser, 1979; Becklen and Cervone, 1983; Stoffregen and Becklen, 1989). For example, Neisser and Becklen (1975) showed observers two ongoing events, each partially transparent and superimposed on top of the other. Thus, the events occupied the same spatial locations on the display and therefore on the retina. In this initial experiment, one event was a hand slapping game with two players and the other was a group of three people passing a basketball to each other while moving around. Subjects were required either to press a button whenever the basketball players made a pass or to press a button when a particular event happened in the hand game (in some cases, subjects tried to do both tasks simultaneously). After several trials of this task, something unexpected occurred in the unattended event. For example, the people playing the hand game would stop and shake hands or the people in the basketball game would temporarily have no basketball but would continue to fake passes. Surprisingly, many subjects failed to notice these unexpected events (see Neisser, 1979).

In a further variant of this experiment (Becklen and Cervone, 1983), the two superimposed events were identical basketball games with one group of players wearing white and the other wearing black. As subjects were selectively monitoring one team, a woman carrying an open umbrella walked slowly across the display and off the far side. All three events were partially transparent, and because the umbrella woman walked across the screen, she occupied the same spatial locations as the players. Yet, even though this unexpected event lasted nearly 5 sec, many subjects failed to notice it.

Neisser and colleagues conducted a number of additional variants on this task,

but unfortunately, many of them were not published and were described only in a summary book chapter (see Neisser, 1979). The original work on these issues was difficult to incorporate into the theory of the time (see Simons, 2000), but in light of more recent work on inattentional blindness it has taken on new significance (Simons, 1999; 2000b). Not only can we miss briefly flashed objects away from the focus of attention, but we also fail to see ongoing dynamic events, provided that attention is focused on another event or object. Furthermore, attention in this task need not be focused on a different spatial location to produce inattentional blindness, an important difference from work using the line-judgment task (Newby and Rock, 1998). In selective looking, attention seems to be directed to objects and events rather than to spatial locations.

Interestingly, in both the line-judgment and selective looking tasks, stimuli that might capture attention in an intentional task do not do so (Simons, 2000b). This fundamental difference has important implications for models of visual capture and visual representations. Color singletons may only capture attention if we know they may appear. Models of attention and vision are often based on the notion that some features will pre-attentively capture attention (Treisman and Gelade, 1980). Yet, if under more typical viewing conditions, features thought to be "primitive" fail to capture attention, these models will only account for performance when subjects are aware that a stimulus may appear. The idea behind intentional studies of attentional capture is that unexpected or unusual objects should capture our attention. Yet, the methods used to study capture may be sub-optimal for testing whether they do capture attention.

Both the line-judgment and selective looking tasks suggest that without attention, some objects and events will not be consciously perceived (see Mack and Rock, 1998). However, neither speaks particularly well to the perception of unattended objects in naturalistic displays. The line-judgment task involves stimuli flashed only briefly on a computer monitor and typically does not involve a dynamic unexpected event (with the exception of briefly presented stroboscopic motion sequences, see Mack and Rock, 1998). The superimposed selective looking task is also somewhat unnatural. Although subjects could see the umbrella woman if they looked for her (Becklen and Cervone, 1983), perhaps the limited detection is due to the unnatural transparent display. Perhaps inattentional blindness only occurs for briefly flashed displays or for events that are difficult to view.

Simons and Chabris in the late 1990's conducted a new series of studies in an effort to revive interest in the study of inattentional blindness for prolonged, dynamic events as a tool for understanding the perception of unattended objects (Simons and Chabris, 1999). These studies used a task quite similar to those used by Neisser and colleagues. Observers viewed a dynamic display of one team of three players in white shirts and one team in black shirts, each passing a basketball. Subjects counted how many passes one of the teams made. In the easy version of the task, they simply counted the total number of passes. In the difficult version, they kept two running totals, one of the number of aerial passes and one of the number of bounce passes. After approximately 45 seconds of performing this task, an unexpected event occurred. Two distinct unexpected events were used: a woman carrying an open umbrella (as in the original studies by Becklen and Cervone, 1983) and a woman wearing a black gorilla suit. Finally, to explore the possibility that the degree of inattentional blindness in the earlier dynamic displays resulted from the odd superimposition, two versions of each event were used. In one version, both teams and the unexpected event were superimposed and partially-transparent, much as they had been in the earlier studies. In the other version, both teams and the unexpected event were filmed using a single camera and choreographed action. In this ³opaque² version, all of the objects were fully visible and clear, and the players could occlude one another. Consistent with the earlier work using transparent dynamic displays, observers often did not see the partially-transparent umbrella woman or the gorilla when their attention was focused on another object or event. Averaging across both unexpected events, both colors of attended team, and both easy and difficult tasks, nearly 60% missed the unexpected event when the teams were partially transparent, a level roughly consistent with the findings of Neisser and colleagues (Neisser, 1979). Yet, even when the teams and unexpected object were fully visible and opaque, nearly 35% still did not see them! As expected, more subjects noticed the unexpected object when performing the easy counting task than the difficult counting task.

One interesting and unexpected finding was that more people noticed the gorilla when counting the passes made by the team wearing black than when counting passes by the team wearing white. This difference may be due to the greater similarity of the gorilla to the black team than the white team (the umbrella woman was wearing brown clothes so she was different from both the attended and ignored teams). To explore this possibility, our lab recently conducted a new series of studies using more tightly controlled computer displays (Most et al., accepted). In these experiments, subjects viewed a display with four white shapes and four black shapes. The shapes moved pseudo-randomly, periodically bouncing off the sides of the display window. The subject's task was to count the total number of times that one set of shapes (either black or white) bounced off the sides. Each trial of this task lasted for 15 sec, and on the third trial, an unexpected event occurred: after 5 sec, a cross began to move linearly across the center of the display, passing the fixation point, and exiting the other side of the display 5 sec later. After this trial, subjects were asked whether they had seen the unexpected object. Unlike the video experiments, in this study the similarity of the unexpected object to the attended and ignored items could be controlled precisely by varying the luminance of the cross.

The results bore out the effect of similarity we found in the video studies: when attending to the white shapes, subjects generally noticed a white cross and almost never noticed a black cross. Similarly, when attending to black shapes, they almost always noticed a black cross and almost never noticed a white cross. Detection of gray unexpected crosses was intermediate and noticing varied with the similarity of the luminance to the attended items and the dissimilarity to ignored items.

These findings clearly show the importance of visual similarity in the detection of unexpected objects. However, based on this effect of luminance alone, we cannot determine whether detection is based on how similar the cross is to the attended items or how different it is from the ignored items – the two explanations were perfectly confounded. In a final experiment, subjects attended to a set of gray shapes against a colored background and ignored either a set of white shapes or a set of black shapes (only one ignored set was present in a given experiment). In this case, the unexpected cross was either white or black; it was either the same as the ignored items or it was equally different from the attended objects but in a direction different from the ignored items. Surprisingly, subjects were less likely to notice the cross when it was the same color as the ignored items than when it was different from the ignored items. In these cases, the unexpected object was equally different from the attended items. This finding suggests an important role for the ignored items in determining what captures attention. Subjects apparently are actively inhibiting conscious perception of items similar to the ignored items when performing a selective attention task (Watson and Humphreys, 1997, 1998). It is interesting to note that subjects were not required to ignore the irrelevant items. They were just told to selectively attend to one group of items. Our lab is currently exploring whether inattentional blindness for dynamic events requires that subjects selectively ignore some of the items or if they will still fail to notice unexpected objects, provided that attention is focused on other objects or events.

10.2.3 Implications

Subjects often fail to see unexpected objects in situations ranging from brief flashes on a computer display to ongoing, naturalistic, dynamic events. This failure seems to depend on two factors: (1) observers do not expect the object or event, and (2) observers are engaged in an attentionally demanding task. In the line-judgment task, observers do not selectively ignore any other items, but their attention is focused on the cross. Under these conditions, inattentional blindness is related to the proximity to the cross. In the dynamic tasks, subjects selectively ignore one set of items while attending to other items. Under those conditions, we find inattentional blindness even when the unattended and attended items occupy the same space. However, the similarity in appearance of the unexpected object to the attended and ignored items does seem to play a role.

Although these two paradigms produce somewhat different results, the overall pattern using these incidental tasks is consistent and, more importantly, different from the results of intentional tasks. When observers do not expect an object to appear, they often do not see it at all. Yet, in an intentional search task, observers often see the odd object – it captures attention. The factors that produce attentional capture in an intentional task do not necessarily produce the same results in an incidental task. The fundamental difference appears to be the expectations of the observer. In the real world, our attention is typically focused on some goal and we do not always expect the unexpected.

10.3 Conclusions

The *incidental stance* approach to perception asks what people do under typical perceptual conditions. Models based solely on intentional tasks may describe the capabilities and capacities of attention, but they may not adequately describe what people actually perceive and how they perceive it. For example, models of perception that require the capture of attention by primitive features may not generalize to naturalistic viewing conditions; attentional capture may depend on the observer's expectation that something might suddenly appear. Similarly, when observers do not intentionally search for changes, they often fail to notice when the central object in a scene is replaced. If we relied exclusively on intentional change-detection tasks, we might have concluded that changes to central objects are typically detected.

Of course, without intentional tasks, our understanding of attention would be quite limited. Incidental tasks are ill-suited to the exploration of the systematic variation in performance and cannot provide much information about sensory or attentional thresholds. With only one critical trial, the incidental approach is fundamentally limited to studies of the average performance of groups of subjects rather than the performance of individual subjects. By using both the intentional and incidental approaches in concert, we can gain a better understanding of the full range of change blindness and of visual perception in general.

Acknowledgments

See the enclosed CD rom for videos from some of the incidental tasks described in this chapter. Videos are also available on our laboratory web site http://www.mjh.harvard.edu/~ viscog/lab. The writing of this manuscript was supported in part by NSF grant #BCS-9905578 and by a Research Fellowship from the Alfred P. Sloan Foundation.

References

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

Dmytryk, E. (1984). *On Film Editing: An Introduction to the Art of Film Construction*. Boston, MA: Focal Press.

Folk, C. L., Remington, R. W., and Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *J. Exp. Psych. Hum. Percept. Perf.*, 18:1030-1044.

Folk, C. L., Remington, R. W., and Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *J. Exp. Psych. Hum. Percept. Perf.*, 20:317-329.

Gibson, B. S., and Jiang, Y. (1998). Surprise! An unexpected color singleton does not capture attention in visual search. *Psychol. Sci.*, 9:176-182.

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Perception,* Vancouver Studies in Cognitive Science Vol. 2, pp. 89-110. New York: Oxford University Press.

Haines, R. F. (1991). A breakdown in simultaneous information processing. In G. Obrecht and L. W. Stark (Eds.), *Presbyopia Research: From Molecular Biology to Visual Adaptation*, pp. 171-175. New York, NY: Plenum.

Henderson, J. M., and Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychol. Sci.*, 10:438-443.

Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance Vol. 1: Sensory Processes and Perception*, pp. 22.21-22.64. New York, NY: John Wiley and Sons.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behav. and Brain Sci.*, 9:1-66.

Hollingworth, A., and Henderson, J. M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Vis. Cog.*, 7:213-235.

Levin, D. T., Momen, N., Drivdahl, S. B., and Simons, D. J. (2000). Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Vis. Cog.*, 7:397-412.

Levin, D. T., and Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychon. Bull. Rev.*, 4:501-506.

Littman, D., and Becklen, R. (1976). Selective looking with minimal eye movements. *Percept. Psychophys.*, 20:77-79.

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390:279-281.

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

Mack, A., Tang, B., Tuma, R., and Kahn, S. (1992). Perceptual organization and attention. *Cog. Psych.*, 24:475-501.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Q. J. Exp. Psych.*, 11:56-60.

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

Neisser, U. (1976). *Cognition and reality: Principles and Implications of Cognitive Psychology*. San Francisco, CA: W. H. Freeman.

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

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

Newby, E. A., and Rock, I. (1998). Inattentional blindness as a function of proximity to the focus of attention. *Percept.*, 27:1025-1040.

O'Regan, J. K., Deubel, H., Clark, J. J., and Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Vis. Cog.* 7:191-211.

O'Regan, J. K., Rensink, R. A., and Clark, J. J. (1999). Change-blindness as a result of "mudsplashes". *Nature*, 398:34.

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

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

Rock, I., Linnett, C. M., Grant, P., and Mack, A. (1992). Perception without attention: Results of a new method. *Cog. Psych.*, 24:502-534.

Rothbart, M., and John, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact. *J. Soc. Iss.*, 41:81-104.

Scott-Brown, K. C., Baker, M. R., and Orbach, H. S. (2000). Comparison blindness. *Vis. Cog.* 7:253-267.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychol. Sci.*, 7:301-305.

Simons, D. J. (1999). To see but not to see: Review of Inattentional Blindness by A. Mack and I. Rock (1998). *J. Math. Psych.*, 43:165-171.

Simons, D. J. (2000a). Current approaches to change blindness. *Vis. Cog.* 7:1-15.

Simons, D. J. (2000b). Inattentional blindness and attentional capture. *Trends Cog. Sci.*.

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

Simons, D. J., Chabris, C. F., and Levin, D. T. (1999). Change blindness is not caused by later events overwriting earlier ones in visual short term memory. manuscript in preparation.

Simons, D. J., and Levin, D. T. (1997). Change blindness. *Trends Cog. Sci.*, 1:261-267.

Simons, D. J., and Levin, D. T. (1998). Failure to detect changes to people in a real-world interaction. *Psychon. Bull. Rev.*, 5:644-649.

Stoffregen, T. A., and Becklen, R. C. (1989). Dual attention to dynamically structured naturalistic events. *Percept. Mot. Skills*, 69:1187-1201.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *J. Exp. Psych. Hum. Percept. Perf.*, 20:799-806.

Treisman, A. M. (1964). Monitoring and storage of irrelevant messages in selective attention. *J. Verbal Learn. Verbal Behav.*, 3:449-459.

Treisman, A. M. (1960). Contextual cues in selective listening. *Q. J. Exp. Psych.*, 12:242-248.

Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cog. Psych.*, 12:97-136.

Watson, D. G., and Humphreys, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psych. Rev.*, 104:90-122.

Watson, D. G., and Humphreys, G. W. (1998). Visual marking of moving objects: A role for top-down feature-based inhibition in selection. *J. Exp. Psych.*, 24: 46-962.

Yantis, S. (1993). Stimulus-driven attentional capture and attentional control settings. *J. Exp. Psych. Hum. Percept. Perf.*, 19:676-681.

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention*, pp. 223-256. East Sussex, UK: Psychology Press.

Yantis, S., and Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *J. Exp. Psych. Hum. Percept. Perf.*, 20:95-107.

Yantis, S., and Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. *J. Exp. Psych. Hum. Percept. Perf.*, 10:601-621.

Yantis, S., and Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking? *J. Exp. Psych. Hum. Percept. Perf.*, 22:1505-1513.