# TESTING THE ANIMATE MONITORING HYPOTHESIS

**MITCHELL LAPOINTE**
**Bachelor of Arts (Hon.), St. Thomas University, 2008**

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

**MASTER OF SCIENCE**

Department of Psychology
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

This work is dedicated to Jesse Andrew LaPointe for his lessons of strength and determination.

# Abstract

The detection of human and non-human animals and their unique (and potentially dangerous) "animation" would have been important to our ancestors' survival. It would seem plausible that our ancestors would have required a vigilance above and beyond that dedicated to other, inanimate, objects. Considering the millions of years of expending extra energy to monitor these objects, it would also seem likely, at least as advocated by New, Cosmides, and Tooby (2007), that the human visual system would have developed mechanisms to allocate attention automatically and quickly to these objects. We tested the New et al. (2007) "animate monitoring" hypothesis by presenting viewers with a group of animate objects and a group of inanimate objects using the flicker task—a task that is assumed to be one that measures automatic visual attention. These objects were presented on a variety of backgrounds of natural scenes, including some backgrounds that were contextually inconsistent with the target objects. These objects were also presented in either a consistent location within each scene or a location that violated that consistency. Only people objects were consistently more readily detected, not animated objects in general. Detection in this task was affected by more than just the information provided by the target object. Both results provide a serious challenge to the "animate monitoring" hypothesis. Furthermore, the results were shown not to be due to peculiarities of our stimulus set, or by how interesting the members of each object category were.

# Acknowledgments

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 Guiding Visual Attention

Being able to find and accurately identify specific objects in our visual world is an important skill. There are times when this perceptual behaviour seems to be driven by internal goals, such as trying to find an empty table at a busy restaurant, and times when it seems as though something in the environment is the driving force in grabbing our attention, such as a growling dog. In either case, this perceptual feat is a daily occurrence and is one that is intrinsically tied to our proficiency as human beings, not only from a social standpoint, but for our personal and genetic survival. Our ability to detect and identify a relevant object requires that we disambiguate that object from surrounding objects and from the background in which it is embedded. How is it that we are able to recognize and identify particular objects over others in a complex visual scene? Although it seems obvious that a growling dog would visually jump out at a viewer, if that growling dog was standing amongst a group of growling dogs its salience as an individual object may not hold the same attention grabbing power. It is unclear how to define the degree of salience, as an object's salience does not stay consistent across scenes and situations. An object is not only defined by its surroundings, it can also jump out at us depending on its relevance to us as viewers (see, e.g., Pashler, 1988; Werner & Thies, 2000).

### *1.1.1  Stimulus-driven Attention*

The literature on visual attention and visual search has attempted to explain object detection as occurring in one of two ways (Zoest, Donk, & Theeuwes, 2004): goal-driven (Bacon

& Egeth, 1994) and stimulus-driven (Theeuwes, 1991; Yantis & Jonides, 1990). Stimulus-driven attention is described as being motivated by the properties of the stimulus, such as the object's movement, its colour, or luminance (Yantis & Jonides, 1990). In the case of the growling dog, the viewer's attention would be drawn to this object regardless of internal goals or biases, as the inherent nature of the object demands attention. It is important to recognize that objects never appear by themselves. Objects are encountered as pieces of larger scenes, in fact they are embedded within scenes. From a perceptual perspective, this relationship is unbreakable and cannot be ignored. An object's relevance, just as its salience, changes depending on the surrounding environment (Gibson, 1979; Oliva & Torralba, 2007). Very little work has been done on the relationship between an object and its background. Although the influence the background has on object detection is recognized (e.g., Hollingworth & Henderson, 2004), the strength of this interaction is largely unknown.

## *1.1.2   Goal-driven Attention*

The goal-driven explanation of object detection places importance on the internal state of the viewer rather than the intrinsic qualities of the object or its environment. From this perspective, the intentions of the viewer override the information being given by the environment, at least in large part. Taking the example of the patron at a busy restaurant, *where* the viewer looks and *how long* the viewer fixates on particular aspects of the scene are being dictated by the desire to find an empty table. It is clear in these cases that top-down processes are taking place: the viewer is actively looking rather than passively consuming.

However, the consensus that stimulus-driven and goal-driven processes are binary and mutually exclusive is not universal (e.g., Itti & Koch, 2000). The restaurant patron actively searches for a particular goal, in this case an empty table. However, where the viewer looks is being directed by the environment and as the viewer fixates at a particular point, the object

of focus gives the viewer information on whether the object is relevant or irrelevant to the ultimate goal of inquiry, leaving the viewer either satisfied or needing to continue with the search (Nothdurft, 2002; Theeuwes, 1992; Theeuwes, Atchley, & Kramer, 2000). From this perspective, it is clear that both goal-driven and stimulus-driven processes are shaping attention. Goal-driven object detection has been classified as a mainly top-down process as the viewer searches a visual scene with a particular object or location in mind. As a viewer searches, certain objects are disregarded until the object of desire is found. At this point, it has been argued, stimulus-driven perception takes over as the object grabs the attention of the viewer.

There has also been speculation that an individual's experience might shape what objects grab attention (Pashler, 1988; Smith, Hopkins, & Squire, 2006; Werner & Thies, 2000). People who have a great deal of experience with particular objects or scenes may be more adept at attending and detecting those objects in either similar or novel scenes in comparison with those people without that experience. Werner and Thies (2000), for example, found that football players were less likely to miss changes to football related objects and better at detecting changes in football related scenes compared with non-football players. Experience-based object detection is similar to goal-driven object detection in the sense that they are both being guided by top-down processes. In either case the viewer's internal state is overriding information being given by the environment. The experience of the viewer is bringing a level of expertise to viewing the scene. It is unclear how this advantage works exactly; for example, whether the experts know where to look or where not to look.

### 1.1.3  The Flicker Paradigm

Testing the way in which particular objects penetrate our focus of attention has been difficult. Much of the difficulty has revolved around creating experiments that are ecologically valid.

Most of the research in this area has used visual arrays where one object is quickly removed (e.g., Mitroff, Simons, & Levin, 2004), or line drawings where the participant must identify which object is different across two images (De Graef, Christiaens, & d'Ydewalle, 1990). For these types of experiments, reaction times are measured to assess which types of objects and features induce the quickest change detection (e.g., Reinecke, Becker, & Rinck, 2010). This line of research has identified a host of features that can affect change detection and object recognition (e.g., Boyce, Pollatsek, & Rayner, 1989; Gilbert, Ito, Kapadia, & Westheimer, 2000; Hoffman, Nelson, & Houck, 1983; Ro, Russell, & Lavie, 2001); however, due to the nature of the stimuli and paradigms used, most of these features are low-level. Few experiments have been able to tease apart the influence of the high-level features of a scene or its objects and the interplay between low and high-level features on change detection (e.g., Hollingworth & Henderson, 2004; Kelley, Chun, & Chua, 2003; Oliva & Torralba, 2006; Stirk & Underwood, 2007; Torralba, Oliva, Castelhano, & Henderson, 2006).

Unlocking the influence of high-level features, along with the influence of a scene's background, has led to the development of the flicker paradigm as a way of testing change-blindness, change detection, and object detection and recognition. The paradigm was first developed by Rensink, O'Regan, and Clark (1997), but has since been used by many investigators across several areas of research (Fletcher-Watson, Collis, Findlay, & Leekam, 2009; Kelley et al., 2003; McGlynn, Wheeler, Wilamowska, & Katz, 2008; Stirk & Underwood, 2007; Werner & Thies, 2000). The flicker paradigm has stayed relatively unchanged since it was first introduced by Rensink et al. (1997), although some slight variations have been used (e.g., New et al., 2007). The paradigm was developed based on visual memory tasks (Pashler, 1988; Phillips, 1974) and eye-movement experiments (Bridgeman, Hendry, & Stark, 1975; Grimes, 1996; McConkie & Zola, 1979), in which blank inter-stimulus intervals (ISI) were used between two visual stimulus presentations. In these experiments, an original image (*A*) is presented, followed by a brief ISI, followed by a slightly modified version of the
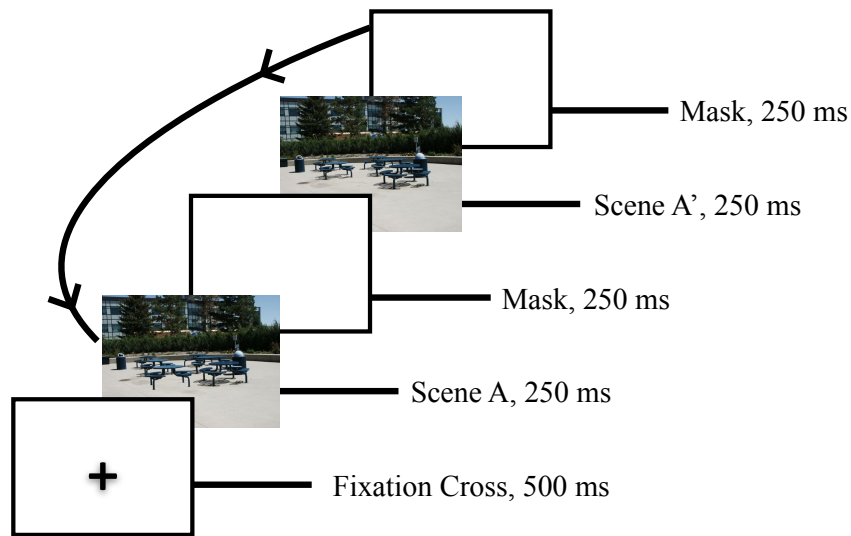
Figure 1.1: Flicker Task procedure used in Experiments 1–4,6

original image (*A'*). These experiments, as well as the flicker paradigm, take advantage of the phenomenon known as change-blindness. Change-blindness is used to describe viewers' inability to detect large and often important changes to visual scenes (Simons & Rensink, 2005). For the visual memory tasks and the eye-movement experiments, the two images (*A* and *A'*) are shown only once, whereas with the flicker paradigm the images continue to alternate with the ISIs. This slight change in presentation allows the viewer to partake in free viewing, does not require any particular eye-movements, and is argued to be more ecologically valid as the ISI models saccades more prevalent in a 3-dimensional space (Rensink et al., 1997). Figure 1.1 depicts the basic flicker task procedure.

Rensink et al. (1997) developed the flicker task primarily to examine change-blindness in its own right (see Simons & Rensink, 2005, for a review), but also as a method to investigate visual attention and perception (Rensink, 2002, 2004; Rensink et al., 1997; Rensink, O'Regan, & Clark, 2000). Most experiments that have used this paradigm have followed the same path toward similar goals (Fletcher-Watson et al., 2009; Hollingworth & Henderson, 2000, 2004; Kelley et al., 2003; Stirk & Underwood, 2007)

In a clever move, New et al. (2007) reversed the logic. Rather than use the flicker task to investigate change-blindness, why not use the phenomenon to investigate properties of the visual stimuli themselves?

## 1.1.4   The Animate Monitoring Hypothesis

This novel approach to using the flicker paradigm introduced by New et al. (2007) was intended to capture any innate biases humans may have toward particular groups of stimuli. Although this approach to the investigation of attention is certainly stimulus-driven, it is contingent upon innate biases on the part of the viewer, and is therefore thought of as a type of top-down processing. Much like the processes involved in goal-driven attention, here, it is assumed that the viewer's attention and ability to detect change is being guided by innate perceptual sensitivities to particular objects. From this perspective, and through the use of several experiments using the flicker paradigm, New et al. (2007) introduced what they called *The Animate Monitoring Hypothesis*. New et al. (2007) argued that there must be evolved biases within our visual attention system designed to capture and monitor objects that have played vital roles in our species' phylogeny. These objects must have required enough vigilance to have drawn the guiding hand of natural selection and, therefore, their vigilance must have been essential to our ancestors' survival. According to this logic, the group of objects that would have required the greatest amount of attention and whose vigilance would have been rewarded with one's survival and enhanced fitness would have been fellow humans, and other, non-human, animals.

Fellow humans played a vital role in our ancestors' survival as they braved the rough terrain and environment. Being able to identify correctly kin from non-kin, potential mates, and allies from enemies would have been essential to the individual's survival, if not the survival of the entire social group. Paying close attention to these objects would have

been important as they have the distinct ability of moving quickly, changing trajectory, position and orientation. These movements are driven by the objects themselves and are, therefore, unpredictable at times, requiring added vigilance. Non-human animals share these characteristics and would have served our ancestors equally as sources of food, but also as dangerous predators. According to New et al. (2007), this dynamic of importance and unpredictability is what has created the need for extra vigilance of animate objects over inanimate things in our world. They argue that a bias toward detecting animate objects would arise from a co-evolution with goal-driven visual attention, perhaps as natural selection favoured those who correctly detected an advancing predator, or recognized an enemy or potential mate earlier. The animate monitoring hypothesis predicts that the attention to animate objects would override other, more proximate perceptual biases such as individual experience, current goals, or internal state. This animate vigilance is argued to be innate and automatic and does not require any instructions, explicit or otherwise.

## 1.1.5   Testing the Animate Monitoring Hypothesis

In order to test the prediction that we are pre-wired to attend to human and non-human animals, New et al. (2007) placed images of natural scenes containing these objects into the flicker paradigm, to be compared against images containing normally fixed objects (e.g., a large rock), moveable, but inanimate objects (e.g., automobiles), and plant objects (e.g., trees). If we carry a disposition to detect animate objects, then changes to these objects should be noticed faster and more accurately than changes to inanimate objects. In order to avoid possible goal-driven confounds, New et al. (2007) presented these images within the flicker paradigm without any explicit instruction aside from detecting change. No instructions were given to look for any particular category, for example, or any particular object, leaving the innate bias presumably free to drive detection as different categories of

7

objects were encountered in the flicker task.

### 1.1.6 Do we have a bias to detect changes to animate objects?

New et al. (2007) presented the results of five experiments—all designed to test whether humans have a visual bias to attend to animate objects, and whether this bias is innately driven rather than based on individual experience. The first and second experiments were designed to distinguish any difference in change detection between animate and inanimate objects, namely people and animals on the one hand, and fixed, moveable and plant objects on the other. With 70 pairs of images (14 pairs from each object category) used for comparison, New et al. (2007) found that changes to animate objects were detected on average 1–2 seconds faster than changes to inanimate objects for both Experiment 1 and its replication, Experiment 2. Not only were the changes to animate objects detected faster, but also more accurately: for both Experiments 1 and 2, hit-rates for animate objects were 21–25% higher on average than hit-rates to inanimate objects. From these results, New et al. (2007) concluded that there is a bias to detect changes to animate objects over inanimate objects.

### 1.1.7 Is object identification the source of the bias?

If humans are biased to attend to animate objects, and if this mechanism was shaped through natural selection, it would seem logical that that skill would be driven by the ability to identify objects as members of their categories (e.g., animate vs. inanimate). To test this aspect of the animate monitoring hypothesis New et al. (2007) presented the same 5 groups of categories, embedded in natural scenes, within the flicker paradigm; however, in an

attempt to reduce object identification, they inverted the images. Inverting stimuli has been shown to be an effective way to reduce the identification and recognition of faces (Ro et al., 2001; Yin, 1969), as well as targets in natural scenes presented in a change detection task (Hollingworth & Henderson, 2004; Kelley et al., 2003). If object identification is the source of the effect, then disrupting object identification in the flicker task should reduce or eliminate the advantage seen with animate objects. This result is exactly what New et al. (2007) found: changes to animate objects were not detected any better than their inanimate counterparts when inverted.

Perhaps a better way to disrupt object identification is through the use of filtering (Bar, 2004; Oliva & Torralba, 2007). In their fourth experiment, New et al. (2007) blurred the images in an attempt to disrupt or reduce object identification. The same 70 image pairs were used, containing the same 14 objects from each of the 5 categories; however, the images were blurred using a gaussian filter. Again, if the animate monitoring advantage is due to the identification of the objects, as predicted by the animate monitoring hypothesis, than this bias should disappear when the identification is disrupted or reduced. As with their third experiment, New et al. (2007) found no advantage in change detection for animate objects. There was no difference in reaction time or accuracy between the animate and inanimate objects. For both experiments, then, disrupting viewers' ability to categorize the objects removed any bias towards the animate group of objects.

## 1.1.8   Is this bias a function of experience with animate objects as moving entities?

Is this perceptual bias to monitor animate objects based on our experience with these objects as moving entities? Although Rensink (2002) made a distinction between detecting change and detecting motion, where motion across an otherwise static background automatically

9

commands attention, the question here is not detecting motion, but detecting changes to objects with which we have experience as normally moving things. Our ontogeny revolves around people, from the day we are born until the day that we die, we are social creatures. According to New et al. (2007), it is quite possible that any change detection advantage for people could simply be due to our experience with these objects and their mobile behaviour. Although our personal experience with animals may not be quite as extensive, animals still play an important part of our lives. Many people spend countless hours looking for animals, documenting the species they see. Animals have also been invited into our homes as pets and family members. It is possible that our experience with these objects and their mobility, along with how interesting they are, could produce an animate advantage over inanimate objects. It is also quite possible that our attention to this group of objects is being driven by the fact that they have the ability to move, regardless of their animacy. New et al. (2007) point out that movement alone would require more perceptual vigilance than would objects that do not have the ability to change their location.

In order to test this issue of experience with motion, New et al. (2007) presented a fifth experiment. Using the flicker paradigm, New et al. (2007) presented a group of animate objects to be compared against a group of inanimate objects, each embedded within natural scenes. The inanimate group contained both objects that could move and appeared to be moving (i.e., vehicles) and objects that are fixed and cannot move (i.e., lamp post). To be clear, none of the target objects actually moved, but rather carry the potential to move as part of their category requirements. The animate group contained images of both animals and people, but half the images were taken of animate objects in motion, whereas the other half had animate objects that appeared to be stationary. New et al. (2007) argued that if the perceptual advantage to detect animate objects is due simply to experience with motion then the vehicle group of objects should also show a bias advantage, whereas the static group of animate images should no longer show the advantage. New et al. (2007) found that

10

changes to vehicles were detected faster than changes to static inanimate objects; however, they were not detected as fast or as accurately as changes to animate objects. There was no difference in reaction times or accuracy scores between stationary and dynamic animate objects, suggesting that it is not the mere possibility of motion per se that is guiding our visual attendance.

### 1.1.9   Is the bias towards animate objects based on how interesting the object is?

New et al. (2007) acknowledge that the bias could be driven by how interesting the categories (or the particular images used to depict them) were to the participants. It is possible animals are more interesting to human viewers compared with lamp posts. In fact, change detection experiments have shown a decided preference for interesting stimuli (Rensink et al., 1997; Kelley et al., 2003; Turatto & Galfana, 2000) and for interesting areas of a scene (Hollingworth & Henderson, 2000; Stirk & Underwood, 2007). To test whether differences in interest were the source of the bias, New et al. (2007) presented the target objects to a separate group of participants and had them rate each target on how interesting they thought it was. Although interest ratings of the targets did correlate with animacy, they did not predict individual reaction times in the flicker paradigm. New et al. (2007) argued that these findings suggest that although animate objects as a group are seen as more interesting, their individual differences in interest levels do not drive change detection. Similar results were found for target luminance, size, and eccentricity: none were found to predict individual reaction times.

It is unclear whether New et al. (2007) presented the targets as isolated objects or as pieces of larger scenes while gathering interest ratings. It is clear, however, that New et al. (2007) do place some importance on the surrounding scene. In each of their previous

experiments New et al. (2007) chose to use the flicker paradigm to present the target objects embedded within natural scenes rather than using a simpler task such as a visual array of objects. Yet given the importance New et al. (2007) place on object identification as being the driving force behind the animate advantage, it would be worthwhile investigating the strength of this bias when the target objects are either presented as isolated or in discordance with the supporting background.

Regardless, the results of New et al. (2007) suggest a distinct bias to detect changes to animate objects over inanimate objects. It would also appear as though this bias is based on differences in object identification, and not such factors as the luminance of the objects or their contrast with the background scene. Although animate objects were rated as being more interesting to viewers than inanimate objects, their level of interest does not seem to account for the differences in change detection. Based on these findings, it also seems likely that the change detection advantage is not produced by the simple motion aspect of animacy, as changes to vehicles did not induce the same reactions.

# Chapter 2

# Replicating New, Cosmides, & Tooby (2007)

The target objects used by New et al. (2007) were all presented in their natural contexts, presumably because that is how the objects would have been encountered throughout our evolutionary history. However, what would happen if we encountered an animate object that was out of place, such as a deer walking down the main drag of a city? Would the animate bias continue to dominate in this situation? There is evidence that detecting a changing object embedded in an inappropriate context becomes an easier task as the object seems to pop-out at the viewer (Hollingworth & Henderson, 2004). Although natural selection clearly did not work under the constraints of inappropriate environments, Hollingworth and Henderson (2004) have shown that detection for all objects is enhanced; therefore, the detection of animate objects might be expected to continue to dominate other objects even when encountered in inappropriate contexts or locations. In contrast, many have argued that context helps in identifying objects or their change (Chu & Jiang, 1998; Friedman, 1979; Kelley et al., 2003; Oliva & Torralba, 2007). The appropriate context, then, may provide information that guides one where to look in the scene (e.g., horizon line for terrestrial animals) and which objects are most likely to be encountered in particular scenes (e.g., a notebook on a desk). Thus, if the fact that natural selection did not operate under these unique constraints disrupts the animate bias, one would expect no change-detection advantage for animal and people objects under these circumstances. These possibilities are the focus of Experiments 1a and 1b.

## 2.1 Experiment 1a: The Role of Context Appropriateness and Location

The first experiment was designed to test the first two predictions and supporting results proposed by New et al. (2007). That is, that animate objects (both human and non-human animals) will command faster and more accurate change detection than their inanimate counterparts when presented in the flicker paradigm. The experimental design was constructed to be an exact replication of Experiment 1 in New et al. (2007); however, both the images and procedure in creating the stimuli differed slightly. The same 5 categories of objects used by New et al. (2007) were used as comparison for this experiment: animals and people (animate), and plants, moveable and fixed objects (inanimate). In line with the original article, 14 target objects from each category were presented to participants. Each target object was presented as part of a larger natural scene. This collection of images was presented to participants using the flicker paradigm as used by New et al. (2007).

### 2.1.1 Stimulus Creation

When constructing the images to be placed in the flicker paradigm, New et al. (2007) used the same photograph to create both the *A* and *A'* images. That is, both the background-only image and the background-plus-target image were created from the same photograph. In order to do this, the authors digitally removed the target object from its background in order for the background to serve as the *A'* image. Once the target object had been removed from the background, the prevailing gap had to be digitally touched-up; matching colour, luminance and line structure, as well as other features. A concern is that the process of touching-up the background would leave artifacts that may be grabbing viewers' visual attention and thus driving the change detection rather than the target object or its category.

In order to avoid this confound, when creating the images for the present experiments the target objects and background images were taken from different photographs. The target objects were digitally removed from their original photograph and superimposed onto the background image. This coupling served as the *A* image, to be compared against the *A'* image, which was composed of the background photograph only. This slight difference in stimulus creation allowed not only for stimuli less subject to editing artifacts, but also the opportunity to manipulate both the contextual appropriateness in which we placed the target object as well as the location at which the target was placed within the scene.

## 2.1.2    Manipulating Context and Location

As a direct replication of the first two experiments described by New et al. (2007), each of our participants were exposed to images in which the target object and background scene were contextually consistent and the target object was placed in a natural location within that scene. We termed this condition the Appropriate Context and Location. For examples of images presented in the Appropriate Context and Location see Figure 2.1. As a way of testing the role of contextual appropriateness, one-half of the participants were also subjected to images where the background did not support the target object. However, the target object continued to be placed in a correct location within the contextually inappropriate scene, creating the Inappropriate Context condition. For examples of images presented in the Inappropriate Context condition see Figure 2.2. The second half of participants received images where the background and target object were contextually consistent, however the target object was placed in an unnatural location within the scene. This coupling served as the Inappropriate Location condition. For examples of images presented in the Inappropriate Location condition see Figure 2.3.

Figure 2.1: Example of *A* and *A'* images from the Appropriate Context and Location condition for Experiment 1a.



Figure 2.2: Example of *A* and *A'* images from the Inappropriate Context condition for Experiment 1a.



Figure 2.3: Example of *A* and *A'* images from the Inappropriate Location condition for Experiment 1a.

### 2.1.3 Methods

### Participants

Forty-two psychology students from the University of Lethbridge participated in exchange for course credit. Participants varied in both gender and age, and all were naive to the hypotheses of the experiment.

### Stimuli

Each of the 3 experimental conditions consisted of 70 pairs of images. Each pair included a background image (*A'*) and a background image with a target object superimposed onto it (*A*). Each image was constructed from photographs taken in Southern Alberta, Canada. The size of each image was standardized to 859 pixels by 573 pixels, appearing as 23 cm by 15.5 cm on the computer screen.

All the background images (*A*) were constructed from photographs of natural scenes that were taken at various times of day under a variety of lighting conditions. The background images differed in the level of complexity and were of varying landscapes, including both indoor and outdoor scenes. The target object images (*A'*) were created by digitally removing the target object from a photograph that was distinct from its intended background pair and then superimposed onto the background photograph.

In line with Experiment 1 of New et al. (2007), the target objects were chosen as belonging to one of five categories: animal, people, plant, moveable, or fixed. The animal objects included both predators and prey, differing in both size and pose. The people objects included both sexes, a variety of sizes and ages, and were shown in a variety of poses, both looking at the camera and away from it. The plant objects included trees, flowers and other plant species in a variety of sizes, some of which were shown in pots or planters.

The moveable object category included items that could be easily moved such as cars and garbage cans, all of which differed in size and orientation. The fixed objects included things such as street signs and lamp posts and were shown in a variety of sizes.

The background photograph on which the target object was placed was chosen according to one of three conditions: Appropriate Context and Location, Inappropriate Context, and Inappropriate Location. For the Appropriate Context and Location condition, the target object was placed into a naturally fitting scene and into a naturally fitting location within that scene (e.g., a bear sitting on a forest floor, bottom left corner of the photograph). What was considered to be an appropriate context and what was considered to be an appropriate location necessarily differed across object categories. The choices for contextually appropriate backgrounds and appropriate locations were done by considering where that particular object category is normally found. For the Inappropriate Context condition, the target object was placed into a scene where that particular object would not be normally found, but was placed into a physically plausible location within the scene (i.e., a bear sitting on a living room floor, bottom left corner of the photograph). For the Inappropriate Location condition, the target object was placed into a naturally fitting scene, but was placed into an unnatural location within that scene (i.e, a bear floating in mid-air in a forest scene, top left corner of the photograph).

## *Apparatus*

Stimuli were presented on a 32-inch LED backlit computer monitor using Metacard experiment software. Participants were given headphones to hear the experiment instructions and to block out distractor noise. Participants indicated their responses to the stimuli by using a standard computer keyboard and microphone.

## *Procedure*

After reading and agreeing to an experimental consent form, participants were seated approximately 20 inches from a computer monitor. Instructions were explained to each participant individually by the experiment facilitator before the participant was asked to put on headphones to hear and to read the instructions again. Following the second presentation of experimental instructions, the participants were asked whether they had any questions before pressing a begin button on the keyboard to start the experiment.

Each trial began with the presentation of a fixation cross which appeared in the middle of a white box (30 cm x 20 cm) on the computer screen. The fixation cross stayed on the screen for 500 milliseconds (ms). Following the fixation cross, participants were exposed to the first image (*A*). This image included both the background photograph and the target object superimposed onto it. This image remained on the screen for 250 ms. Next, an inter-stimulus interval (ISI), in this case a blank white screen, appeared for 250 ms, followed by the second image (*A'*). The second image included the background photograph only and remained on the screen for 250 ms. The second image was followed by another blank white screen, which lasted for another 250 ms., and so on. Upon detecting the change participants were asked to press the spacebar, which ended the trial, and to name what it was they saw change. Verbal responses were recorded using a microphone. Within 1 second, the next trial began. Participants were at leisure to use whatever label they saw fit when naming the object of change, but were asked to keep their responses to one word. If detection had not been achieved by the time the second ISI was presented the process began again from the first image onward for up to 19 cycles. If after 19 cycles no change was detected the next trial began automatically. The presentation times of both the images and ISIs were created to be an exact replication of those used by New et al. (2007). Rensink et al. (1997, 2000) found that using a white ISI compared to black or grey, and stimulus presentation times of roughly 250 ms to be ideal.

Participants were randomly assigned to either the Inappropriate Context condition or the Inappropriate Location condition. In the Inappropriate Context condition participants were presented with 70 image pairs (*A'* and *A*) from the inappropriate context stimuli and 70 image pairs from the appropriate context and location stimuli. These two stimulus categories were shown as blocks in randomized order. The images within each of these blocks were presented in a randomized order. In the Inappropriate Location condition participants were presented with 70 image pairs from the inappropriate location stimuli and 70 image pairs from the appropriate context and location stimuli. Again, these two stimulus categories were shown as blocks in randomized order. And again, the images within each of these blocks were presented in a randomized order. All 42 participants were exposed to the Appropriate Context and Location condition with 20 of those participants being further exposed to the Inappropriate Context condition and the remaining 22 responding to the Inappropriate Location Condition.

## 2.1.4    Results

### Appropriate Context and Location

All 42 participants provided data for the Appropriate Context and Location condition. Hit rates (naming the correct object) were used to test accuracy as well as to assess reaction times. In line with the results of New et al. (2007), false alarm rates were extremely low. For the Appropriate Context and Location condition, only 2.6% of responses were false alarms. As with the results presented by New et al. (2007), false alarms were omitted from the analysis along with misses. The mean number of flickers (transitions between the object absent and object present images) for hits (correctly identified objects) for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant

Table 2.1: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for Experiment 1a.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
| --- | --- | --- | --- | --- | --- | --- |
| Appropriate | 3.33 | 3.60 | 3.76 | 4.94 | 4.45 | 0.38 |
| Inapp Context | 4.32 | 2.69 | 2.59 | 4.64 | 4.13 | 0.37 |
| Inapp Location | 4.27 | 3.24 | 4.25 | 4.53 | 4.16 | 0.45 |

ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table 2.1. There was a significant effect of Object Category, $F(4, 164) = 23.66, MSE = 0.77, p < 0.0001$. Also shown in Table 2.1 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.38$) for the comparison of the means of each object condition to each other object condition. The animal object category required fewer mean flickers for correct object detection than all other categories of objects with the exception of the people object category. In turn, the people and moveable object categories were detected significantly faster than both the plant and fixed object categories. Plant objects were also detected significantly faster than the fixed object category.

## *Inappropriate Context*

Twenty of the 42 participants provided data for the Inappropriate Context condition. Again, hit rates alone were used to test accuracy as well as to assess reaction times. For the Inappropriate Context condition, only 4.7% of responses were false alarms. In line with the results presented by New et al. (2007), the false alarms were omitted from the analysis along with misses. The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 2.1. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant

effect of Object Category, $F(4,76) = 53.44, MSE = 0.35, p < 0.0001$. In this case, given the $LSD_{.05} = 0.37$, the people and moveable object categories were detected significantly faster than the remaining categories of objects. Plant objects were detected significantly faster than the fixed object category, but did not differ significantly from the animal category. There was no difference in detection times between the animal and fixed object categories.

## Inappropriate Location

The remaining 22 participants provided data for the Inappropriate Location condition. Although false alarms were most prevalent in this condition they were still extremely rare, representing only 5.4% of all responses. Therefore, again, only hits were used to assess accuracy and to compare reaction times. The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 2.1. The mean number of flickers to hits per participants were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4,84) = 9.60, MSE = 0.56, p < 0.0001$. Given the $LSD_{.05} = 0.45$, people objects were identified significantly more quickly than each of the other object categories. There was no significant difference among the remaining 4 object categories.

## Cross Condition Comparison

To assess differences across conditions, the mean flickers of hits as a function of object category for the Appropriate Context and Location condition were compared to each of the Inappropriate conditions as within-participant manipulation (for different participants), and then compared the Inappropriate conditions to each other (as a between-participant

manipulation). Identification of objects in the Inappropriate Context condition approached significance as being faster in comparison to the Appropriate Context and Location condition, $F(1,19) = 9.32, MSE = 1.19, p = 0.07$, although there was a significant interaction with object category, $F(4,76) = 9.40, MSE = 0.62, p < 0.0001$. There was no significant difference in the speed at which participants identified objects in the Inappropriate Location condition in comparison to the Appropriate Context and Location condition, $F(1,21) = 1.39, MSE = 1.45, p = 0.25$. However, there was a significant interaction with object category, $F(4,84) = 5.50, MSE = 0.68, p = 0.0006$. There was a difference in the speed at which participants identified objects in the Inappropriate Context condition in comparison to the Inappropriate Location condition with these objects being detected faster, $F(1,40) = 4.55, MSE = 1.96, p = 0.04$, and there was also a significant interaction with object category, $F(4,160) = 12.49, MSE = 0.46, p < 0.0001$.

New et al. (2007) also presented their data as cumulative hit rates as a function of number of flickers. Accordingly, overall hit rates for each object category for each of the three experimental conditions are shown in Figure 2.4, along with the cumulative percentage of correct detections at each transitional point between flickers.

## 2.1.5 Discussion

### Replicating the Animate Monitoring Advantage

The results of Experiment 1a show a partial replication of those presented by New et al. (2007). When presented in the appropriate context and location, the animal object category was detected significantly faster than the remaining 4 object categories. In turn, the people object category was detected significantly faster than the remaining object categories with the exception of the moveable objects. The surprising reaction times for moveable objects
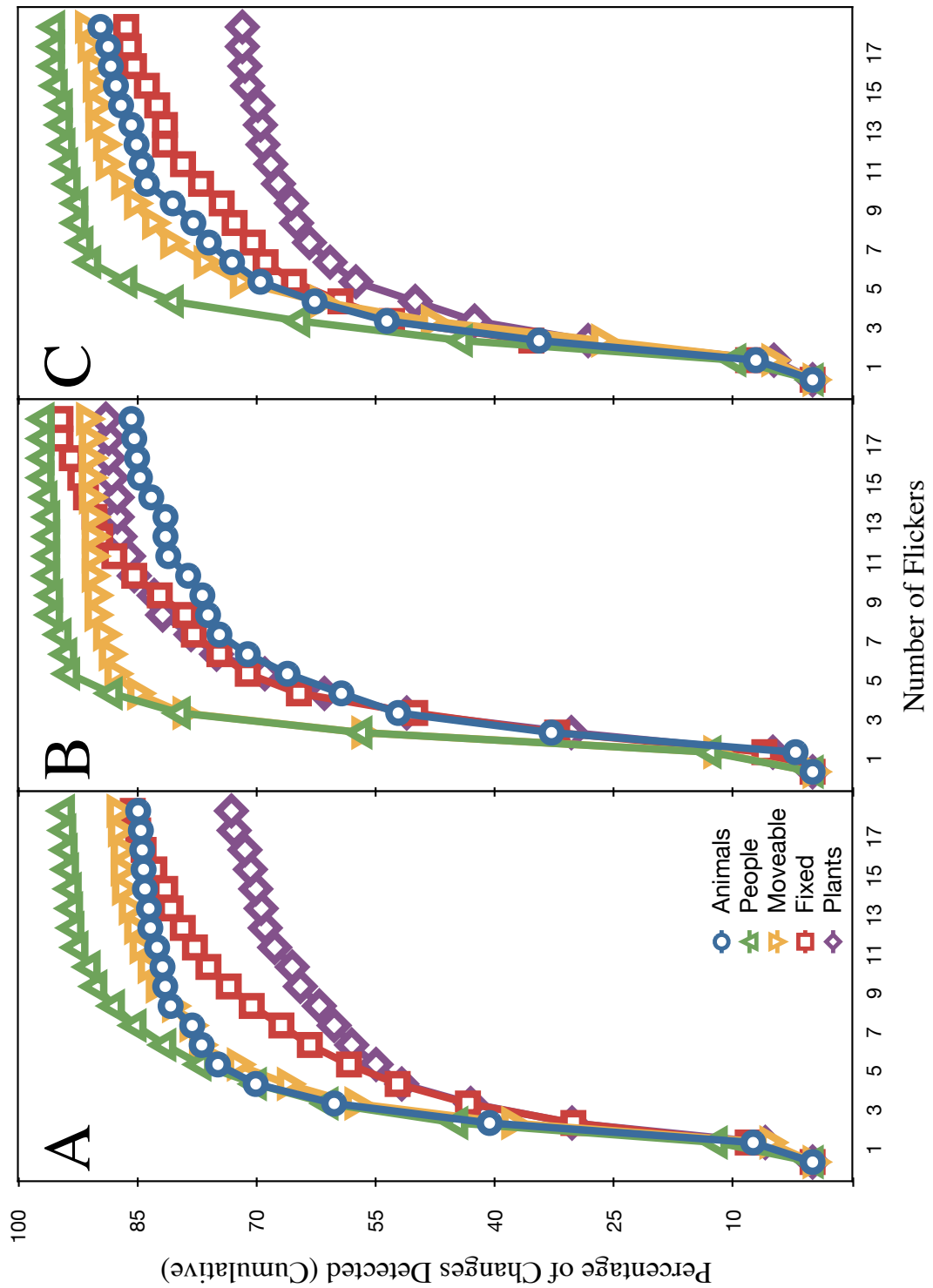
Figure 2.4: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A), the Inappropriate Context (B), and the Inappropriate Location (C) conditions for Experiment 1a.

may be due to the ambiguity of the moveable object category as it included a variety of items that could easily be classified in other ways. Nonetheless, the animal and people categories were detected with considerable ease compared to the inanimate object categories, moveable objects notwithstanding.

## *Disrupting the Animate Monitoring Advantage*

When the background on which the objects were placed no longer supported that particular object and when the location of the object violated the semantics of the scene, the animate advantage disappeared. Although the detection of people objects continued to dominate, along with the moveable object category in the Inappropriate Context condition, performance on animal objects suffered when the context of the background and location of the objects were manipulated. If the human visual attention system evolved a specialized mechanism to detect animate objects that is based on the identification of the objects and their category membership, manipulations to the surrounding background should have no differential bearing on the ease with which objects in different categories are detected. As detailed by Hollingworth and Henderson (2004), presenting objects in the inappropriate context typically speeds up reaction times. From this perspective, similar relative patterns to those found in the Appropriate Context and Location condition would be expected, however sped up, when the objects are presented in the inappropriate context.

Given that natural selection did not operate under these unique circumstances, it could be argued that the bias to attend to animate objects would not translate to scenes of inappropriate context or inappropriate locations. Under this assumption, no differences in reaction times among the 5 object categories would be expected. Even if differences in reaction times are to be expected, at the very least the animate group of objects should suffer or triumph equally. Yet, this was not the case for our group of participants. Viewers continued to detect

people objects quickly while reaction times to animal objects suffered under our contextual and locational manipulations.

In line with the arguments put forth by Hollingworth and Henderson (2004), reaction times for hits in the Inappropriate Context were faster than those found in the Inappropriate Location condition and approached significance when compared to the Appropriate Context and Location condition. There was no significant difference in reaction times for hits when comparing the Appropriate Context and Location condition to the Inappropriate Location condition. These findings suggest that objects or their change jump out faster when presented in inappropriate surroundings rather than when they are presented in an appropriate context (Hollingworth & Henderson, 2004). This difference across backgrounds also suggests that viewers rely on more than just the objects and their categories when detecting change, as these manipulations affected viewers' reaction times. Again, the differences across contextual and locational manipulations are curious given the claim of New et al. (2007) that the specialized visual attention mechanism relies on the identity of the target object and its category membership when preferentially attending to animate objects.

These findings, taken together, call into question some of the basic tenets proposed within the Animate Monitoring Hypothesis. The ease with which the animate bias was removed simply by manipulating background features highlights a lack of stability typically needed for an evolutionary explanation. The contrast in reaction times for hits across conditions questions some of the basic elements required under this hypothesis, specifically the focus on object identity and category-specific features. As in most cases, this experiment raises more questions than it answered. In particular, future research should tease apart the interplay between the object of focus and its background.

Given the contrasting results and interpretations we have found compared to those presented by New et al. (2007), these findings also question the efficacy of the flicker paradigm in measuring such differences among stimuli. The answers to some of these

questions are attempted in the experiments that follow. Although it is possible that our results are artifacts peculiar to our particular set of stimuli, this explanation seems unlikely given the changes in reaction times across conditions and object categories. We address this issue in Experiment 1b. An effort to pull apart the identifying features of the object and its background can be found in Experiments 2 and 3. The efficacy of the flicker paradigm design and its ability to measure perceptual biases is addressed in Experiment 4. An investigation into what other inherent qualities of the objects might be driving change detection can be found in Experiments 5a and 5b.

## 2.2 Experiment 1b: Testing the Animate Monitoring Hypothesis Across Stimulus Sets

Experiment 1a showed how easily the animate monitoring bias could dissolve. However, one major concern was that our findings were simply the product of our unique stimuli. It is possible that the reactions we recorded were simply a response to this particular stimulus set. Although we were able to replicate some of the major findings presented by New et al. (2007), alleviating some of this concern, we also found unique responses resulting from the introduction of our context and location manipulations. Because our manipulations and corresponding results from Experiment 1a were unique, it is not possible to determine whether these findings generalise beyond our stimulus set. It is possible that our Lethbridge participants were responding to the familiar stimuli in a unique way (e.g., the grey building being interpreted as the library building). In an attempt to alleviate these concerns, we replicated Experiment 1a using a separate set of stimuli created from photographs taken in Brisbane, Australia.

The current experiment was an exact replication of Experiment 1a; however, a different set of stimuli was presented to participants. We presented the same 5 categories with 14

objects presented from each. We presented the target objects within the same 3 conditions as those used in Experiment 1a: Appropriate Context and Location, Inappropriate Context, and Inappropriate Location. All images used in Experiment 1b were constructed in the same manner as those used in Experiment 1a, but they were created by a separate individual from photographs taken in Brisbane, Australia.

## *2.2.1   Methods*

### *Participants*

Thirty-eight psychology students from the University of Lethbridge participated in exchange for course credit. Participants varied in both gender and age, but all were naive to the hypotheses of the experiment, and none had participated in Experiment 1a.

### *Stimuli and Apparatus*

Again, each of the 3 experimental conditions, Appropriate Context and Location, Inappropriate Context and Inappropriate Location, consisted of 70 pairs of images. And again, each pair included a background image (*A'*) and a background image with a target object superimposed onto it (*A*). The background images were chosen according to the same considerations detailed in Experiment 1a. The target objects were superimposed onto the backgrounds using the same digital procedure as that used in Experiment 1a and the objects were chosen as belonging to one of the same five categories used in the first experiment. The same computer and computer software used in Experiment 1a were employed for use in Experiment 1b. As was the case in the first experiment, each participant wore headphones to hear the experimental instructions and block out distractor noise.

## Procedure

The identical procedure to that used in Experiment 1a was used for Experiment 1b.

## 2.2.2 Results

### Appropriate Context and Location

All 38 participants provided data for the Appropriate Context and Location condition. Hit rates alone were used to test accuracy as well as to assess reaction times. The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table 2.2. There was a significant effect of Object Category, $F(4, 148) = 52.29, MSE = 0.79, p < 0.0001$. Also shown in Table 2.2 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.40$) for the comparison of the means of each object condition to each other object condition. The fixed object category required fewer mean flickers for correct object detection than all other categories of objects. In turn, the people and animal object categories were detected significantly faster than both the moveable and plant object categories. Moveable objects were also detected significantly faster than the plant object category.

### Inappropriate Context

Eighteen of the 38 participants provided data for the Inappropriate Context condition. Again, hit rates alone were used to test accuracy as well as to assess reaction times. The results from this condition in terms of mean flickers for hits as a function of object category are

Table 2.2: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for Experiment 1b.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 3.70 | 3.35 | 4.91 | 2.88 | 5.32 | 0.40 |
| Inapp Context | 3.56 | 1.93 | 3.34 | 2.78 | 3.69 | 0.35 |
| Inapp Location | 3.95 | 2.72 | 3.95 | 2.87 | 4.07 | 0.41 |

shown in Table 2.2. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4, 68) = 34.60, MSE = 0.27, p < 0.0001$. In this case, given the $LSD_{.05} = 0.35$, the people object category was detected significantly faster than the remaining 4 categories of objects. Fixed objects were detected significantly faster than the remaining 3 object categories. There was no difference in reaction times between the moveable and animal object categories; however, moveable objects were detected significantly faster than plants, whereas animals were not.

## *Inappropriate Location*

The remaining 20 participants provided data for the Inappropriate Location condition. The results from this condition in terms of mean flickers for hits as a function of object category are shown in Table 2.2. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4, 76) = 20.91, MSE = 0.42, p < 0.0001$. Given the $LSD_{.05} = 0.41$, people and fixed objects were identified significantly more quickly than each of the other object categories. There was no significant difference between the remaining 3 object categories. Overall hit
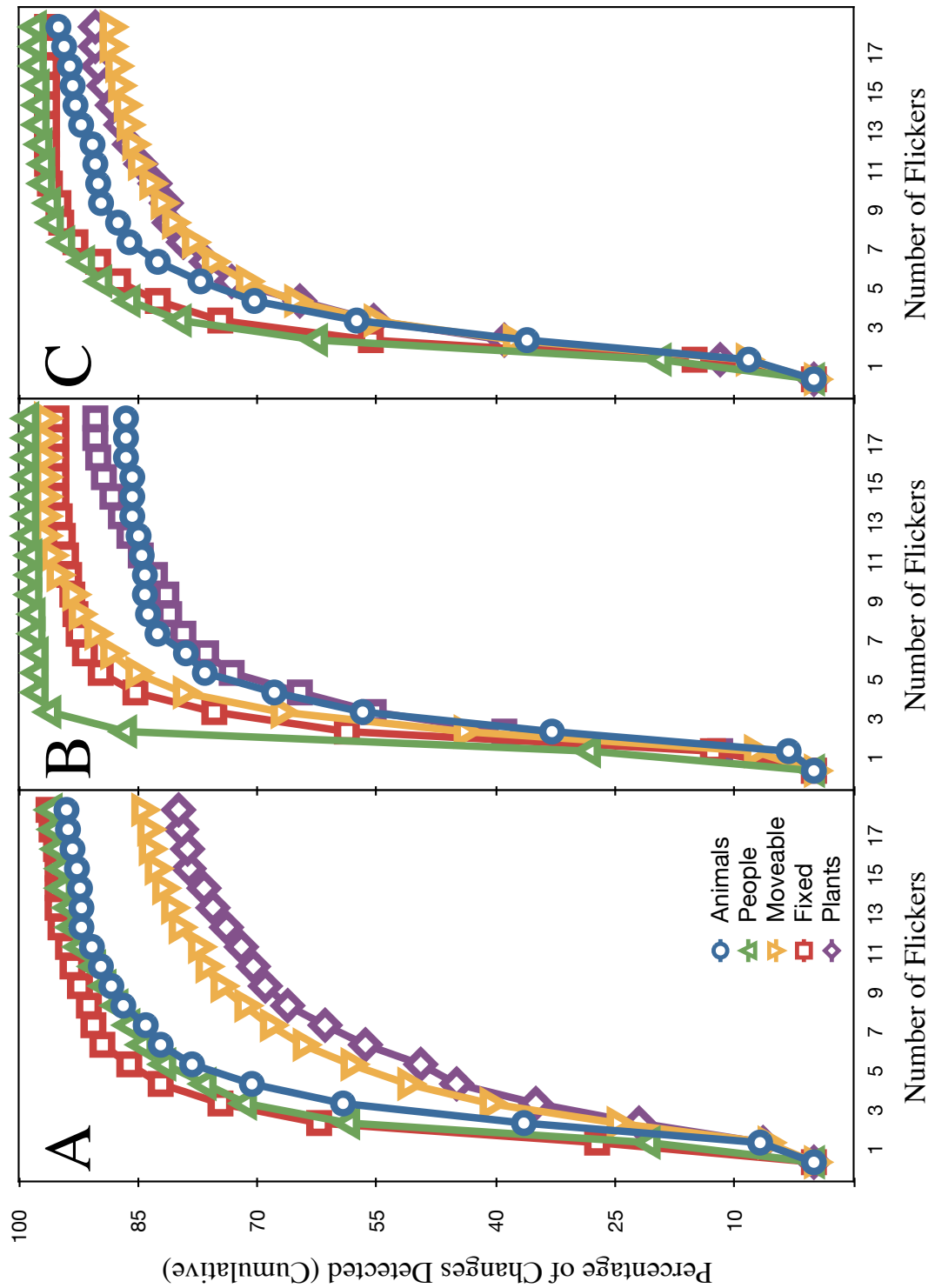
Figure 2.5: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A), the Inappropriate Context (B), and the Inappropriate Location (C) conditions for Experiment 1b.

rates for each object category for each of the three experimental conditions are shown in Figure 2.5, along with the cumulative percentage of correct detections at each transitional point between flickers.

## *Cross Condition Comparison*

Identification of objects in the Inappropriate Context condition were detected significantly faster than those in the Appropriate Context and Location condition, $F(1,17) = 25.08, MSE = 1.56, p = 0.0001$; there was also a significant interaction with object category, $F(4,68) = 8.64, MSE = 0.55, p < 0.0001$. Objects presented in the Inappropriate Location condition were also detected significantly faster than the objects presented in the Appropriate Context and Location condition, $F(1,19) = 17.36, MSE = 0.89, p = 0.0005$. There was also a significant interaction with object category, $F(4,76) = 6.92, MSE = 0.66, p < 0.0001$. There was a difference in the speed at which participants identified objects in the Inappropriate Context condition in comparison to the Inappropriate Location condition with these objects being detected significantly faster, $F(1,36) = 6.07, MSE = 1.62, p = 0.02$; however, no significant interaction with object category was found, $F(4,144) = 2.03, MSE = 0.35, p = 0.09$.

## 2.2.3 *Discussion*

## *Replicating the Animate Monitoring Advantage*

The results from Experiment 1b show another partial replication of the results presented by New et al. (2007). That is, when the target objects were presented in the appropriate context and location, reaction times for hits favoured the animate group of objects (animals

32

and people), with the curious exception of the fixed object category, which was detected faster than all other object categories. These results not only follow the same pattern as those described by New et al. (2007), but also the results found in Experiment 1a, using a separate set of stimuli. The odd reaction to moveable objects in Experiment 1a and fixed objects in Experiment 1b suggest that these are poorly constructed object categories that carry a degree of ambiguity.

## *Disrupting the Animate Monitoring Advantage*

According to the animate monitoring hypothesis, the human visual attention system has evolved a mechanism to monitor categories of objects that have been especially important throughout our evolutionary history, namely animate objects. The hypothesis argues that this perceptual bias is driven by the category itself, which is carried by the identity of the object. According to this view, even under contextual manipulations the preferential attention to animate objects should hold. However, in both Experiments 1a and 1b the bias for animals did not translate when the objects were presented on a contextually inappropriate background. Furthermore, if this mechanism had been evolved to attend to animate objects as an inclusive category then the reaction times to people and animal objects should be roughly the same. When the objects were presented in the inappropriate context, the people objects were detected relatively quickly, but the animal objects were not.

Similar patterns were found when the objects were placed in inappropriate locations within the appropriate context. As with the first experiment, the results for Experiment 1b showed that detecting changes to people objects, and fixed objects, was greater than the other object categories. This result raises the same issues as those found when the objects were placed in an inappropriate context. The Animate Monitoring Hypothesis suggests that change detection for animals and people should be better than all other categories

regardless of background or the location at which the object is placed. At the very least, attentional responses should be similar for animal and people objects, whether they are better, comparable or worse compared to inanimate objects. Both Experiments 1a and 1b showed that this is not the case.

As with the first experiment, objects were detected significantly faster when they were presented on a contextually inappropriate background in comparison to when the same objects were presented on an appropriate background. Similarly, an advantage to detect objects in the inappropriate context over objects in the inappropriate location emerged from both experiments. Unlike the first experiment, reaction times for objects presented in the inappropriate location also showed this advantage in comparison to the Appropriate Context and Location condition for Experiment 1b. Interaction effects for object category were found consistently across the two experiments with the exception of the Inappropriate Context compared to the Inappropriate Location condition in Experiment 1b.

## 2.3    General Discussion

*Generalizing Across Stimuli*

These results show a strong consistency across the first two experiments where two different sets of stimuli were used, created by two separate individuals. There was a concern that the familiar scenes used in Experiment 1a may have caused a shift in responses by native viewers. However, this concern was alleviated by using a separate set of images created from photographs taken in Brisbane, Australia for Experiment 1b. There is also a concern that the style of the photographer would emerge in the stimuli set used in Experiment 1a, where any unintended consistencies in the stimuli would be picked up by participants (i.e., always placing people objects in the same relative area within the larger image).

Experiments conducted by Wu, Tangen, and Vokey (2010) have shown this can be a real concern because photographs of people were consistently rated as 'people' photographs even when the people were removed. The same effect was subsequently found when comparing 'object' photographs. This potential confound does not seem to be making any difference as images created by a separate experimenter produced similar results. The consistency across Experiments 1a and 1b shows that the effects reported are not due simply to peculiarities in the stimulus set, from either the stimuli or their creation.

Both Experiments 1a and 1b were able to replicate some of the basic patterns reported by New et al. (2007), even though a new method of stimulus creation was introduced. However, when the contextual appropriateness of the background or the location of the objects were manipulated the animate bias did not appear. The animate monitoring hypothesis states that any bias for animate objects is based upon the category of these objects, contained within the identity of the object. According to this view, one would expect that the animate bias would continue to dominate change detection regardless of the background scene and regardless of where the object is placed within a scene. Although this manipulation and its effects was never explicitly addressed by New et al. (2007), it is possible that because of the uniqueness of finding an object in an inappropriate context, natural selection never had the opportunity to place pressure on our visual attention system in these situations. However, because the monitoring system is said to be vigilant of animate objects as a collective, one would expect that the reaction to animals and people would be similar regardless of background, whether they were detected faster than inanimate objects or not. This was not the case in our contextual and locational manipulations. The reaction to people objects and animal objects differed significantly when these objects were presented in an inappropriate context or inappropriate location.

# Chapter 3

# Disrupting Object Identification Through Inversion

Experiments 1a and 1b attempted, first, to replicate the basic findings put forth by New et al. (2007) and, second, to introduce new independent variable manipulations to test those assertions. The previous chapter detailed the successful replication in support of the animate monitoring hypothesis. Experiments 1a and 1b demonstrated that detected changes to animate objects were made more quickly than the changes to inanimate objects, with some exceptions. Through the use of contextual and locational manipulations the animate advantage was removed. The experiment introduced here was designed to test whether the bias towards animate objects is due to object identification. Although we were able to address this question in part in the first two experiments by manipulating the background and the location at which we placed the object, Experiment 2 takes a closer look at the relationship between an object and its background and what information that relationship may be giving us as we navigate our visual world.

## 3.1   Experiment 2: Inverting Aspects of a Scene

In the third experiment presented by New et al. (2007) identification of the target objects was disrupted by inverting the images. New et al. (2007) argued that if object identification is disrupted, any preferential monitoring of animate categories would disappear as the categorical information has been removed. This result is exactly what they found. When the images were inverted and placed in the flicker paradigm any bias to detect changes to animate objects disappeared. Reaction times and accuracy scores showed no difference among the 5 categories of objects.

### 3.1.1 Inverting the Images

However, when attempting to disrupt identification of the objects and their category membership New et al. (2007) inverted the entire scene, both target and background. Although the identification of the objects seem to be successfully disrupted by this procedure, so, inadvertently, is the identification of the backgrounds. Although the animate bias was removed through the use of this procedure, it does not seem to answer the question the authors set out to answer in the first place. That is, is the animate monitoring bias driven by the identification of the objects as members of the animate category, driven by the objects alone?

### 3.1.2 Inverting the Target Object

In order to address these issues, we first attempted to replicate the exact procedure and results detailed in the inversion experiment of New et al. (2007). We then attempted to disrupt only the identification of the objects, leaving the backgrounds completely intact. However, the second issue here offers another confusing element. When New et al. (2007) inverted the full images, the objects would have not only been inverted but they also would have changed their location on the computer screen. An object normally located in the bottom right-hand corner of the screen would have been moved to the top left-hand corner of the screen when the entire image is flipped (i.e., rotated by 180 degrees). In order to address this issue, we presented the inverted target objects in either their correct location in regard to the background image or in the location they would be found if the entire image had been inverted.

### 3.1.3   Methods

### Participants

Sixty-seven undergraduate psychology students from the University of Lethbridge volunteered to take part in Experiment 2, and each was given course credit for their service. Participants varied in both gender and age, but all were naive to the hypotheses of the experiment and none of which had participated in either of the previous experiments.

### Stimuli

The images were created from the same photographs used in Experiment 1b, with each condition containing 70 pairs of images. Also similar to Experiment 1b, each pair included a background image (*A'*) and a background image with a target object superimposed onto it (*A*).

Experiment 2 involved 3 experimental conditions: Full Inversion, Target Inversion-Correct Location, and Target Inversion-Incorrect Location. In each condition, the target object was presented in both the appropriate and inappropriate context. The Full Inversion condition was an exact replication of the design used by New et al. (2007), in which both the target object and its background image were turned upside down. For the two Target Inversion conditions only the target object was inverted, leaving the background intact. For the Target Inversion-Correct Location condition, the target object was inverted but remained in the same location within the background scene (e.g., a bear sitting on the forest floor in the bottom left-hand corner of the image was turned upside down, but remained in the bottom left-hand corner of the scene). In the Inverted Target-Incorrect Location condition, again, only the target object was inverted, but it was moved as if the background had also been inverted (e.g., the bear sitting on the forest floor in the bottom left-hand corner of the

image was turned upside down and was moved to the top-right hand corner of the scene).

## *Apparatus*

The same apparatus used in the first two experiments was employed for Experiment 2.

## *Procedure*

A similar procedure to those used in Experiments 1a and 1b were used in Experiment 2, where participants were given a standard consent form and instructions were detailed several times before starting the experiment.

Participants were randomly assigned to either the Full Inversion, Target Inversion-Correct Location, or Target Inversion-Incorrect Location condition. In each condition, participants were presented with targets embedded in both an appropriate and inappropriate context, with 70 image pairs (*A* and *A'*) in each. As with the previous experiments, these blocks were randomized and the images contained in each block were presented in a randomized order. Twenty-two participants took part in the Full Inversion condition, another 23 participated in the Target Inversion-Correct Location, and an additional 22 participants took part in the Target Inversion-Incorrect Location.

## 3.1.4   Results

## *Full Inversion*

Twenty-two participants provided data for the Inverted Background and Target condition. Each of these participants were exposed to 70 pairs of images from the appropriate context

Table 3.1: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for the fully inverted stimuli in Experiment 2.

| Condition | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 5.14 | 3.66 | 5.59 | 3.47 | 5.15 | 0.52 |
| Inapp Context | 4.60 | 2.89 | 4.07 | 3.76 | 4.33 | 0.40 |

stimuli and 70 pairs of images from the inappropriate context stimuli.

## *Appropriate Context*

The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table 3.1. There was a significant effect of Object Category, $F(4, 84) = 27.26, MSE = 0.75, p < 0.0001$. Also shown in Table 3.1 is the Fisher least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.52$) for the comparison of the means of each object condition to each other object condition. The categories of objects requiring the fewest mean flickers for correct object detection were the fixed and people object categories. Fixed and people objects were detected significantly faster than the remaining 3 categories of objects, with no significant differences between the plant, animal and moveable object categories.

## *Inappropriate Context*

The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 3.1. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants

crossing Object Category as the random error. There was a significant effect of Object Category, $F(4, 84) = 21.56, MSE = 0.45, p < 0.0001$. In this case, given the $LSD_{.05} = 0.40$, the people object category was detected significantly faster than the remaining 4 categories of objects. Fixed objects were detected significantly faster than both plant and animal objects, however there was no significant difference when compared to the moveable object category. The moveable object category, in turn, was detected significantly faster than the animal object category. There was no significant difference between the plant and animal object categories. Overall hit rates for each object category for both the Appropriate and Inappropriate Context conditions are shown in Figure 3.1, along with the percentage of correct detections (cumulative) at each transitional point between flickers.

## Cross Condition Comparison

In accordance with the two previous experiments, we compared the mean flickers to hits as a function of object category for the Appropriate Context condition to the Inappropriate context condition as a within-participant manipulation. Identification of objects in the Inappropriate Context condition was significantly faster than in the Appropriate Context condition when both the background and target object are inverted, $F(1, 21) = 10.57, MSE = 2.36, p = 0.004$. There was also a significant interaction with object category, $F(4, 84) = 8.95, MSE = 0.52, p < 0.0001$.

## Target Inversion-Correct Location

Twenty-three participants provided data for the Inverted Target Correct Location condition. Again, in this condition target objects were inverted and left in the same location on the untouched background. Each of these participants were exposed to 70 pairs of images from

Figure 3.1: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A) and the Inappropriate Context (B) conditions when the entire image is inverted for Experiment 2.

Table 3.2: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for the target-only inverted stimuli where the targets where placed in their correct location in Experiment 2.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 4.46 | 3.05 | 4.74 | 3.23 | 4.59 | 0.48 |
| Inapp Context | 4.21 | 2.68 | 3.50 | 3.07 | 3.48 | 0.37 |

the appropriate context stimuli and 70 pairs of images from the inappropriate context stimuli.

## *Appropriate Context*

The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table 3.2. There was a significant effect of Object Category, $F(4, 88) = 22.10, MSE = 0.68, p < 0.0001$. Also shown in Table 3.2 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.48$) for the comparison of the means of each object condition to each other object condition. The people and fixed object categories were detected significantly faster than the remaining 3 categories, with no significant difference between the animal, plant and moveable categories.

## *Inappropriate Context*

The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 3.2. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants

crossing Object Category as the random error. There was a significant effect of Object Category, $F(4,88) = 18.86, MSE = 0.40, p < 0.0001$. In this case, given the $LSD_{.05} = 0.37$, the people object category was detected significantly faster than the remaining 4 categories of objects. In turn, the fixed object category was detected significantly faster than the remaining 3 categories of objects. Next, it was the moveable and plant categories, both being detected significantly faster than the animal object category. Overall hit rates for each object category for both the Appropriate and Inappropriate Context conditions are shown in Figure 3.2, along with the percentage of correct detections (cumulative) at each transitional point between flickers.

## Cross Condition Comparison

Identification of objects in the Inappropriate Context condition was significant faster in comparison to the Appropriate Context condition, $F(1,22) = 18.26, MSE = 1.24, p = 0.0003$. There was also a significant interaction with object category, $F(4,88) = 5.46, MSE = 0.54, p = 0.0006$.

## Target Inversion-Incorrect Location

Twenty-two participants provided data for the Inverted Target-Incorrect Location condition. Here, the target objects were inverted and moved as if the background had been inverted, however the background was left unchanged. Each of these participants were exposed to 70 pairs of images form the appropriate context stimuli and 70 pairs of images from the inappropriate context stimuli.

Figure 3.2: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A) and the Inappropriate Context (B) conditions when only the target object is inverted and placed in the correct location for Experiment 2.

## Appropriate Context

The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table 3.3. There was a significant effect of Object Category, $F(4, 84) = 6.71, MSE = 0.50, p < 0.0001$. Also shown in Table 3.3 is the Fisher least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.43$) for the comparison of the means of each object condition to each other object condition. People objects were detected significantly faster than the remaining 4 categories of objects. There was no significant difference in reaction times between plants, moveable, animal and fixed objects.

## Inappropriate Context

The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 3.3. The mean number of flickers to hits per participants were subject to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4, 84) = 30.61, MSE = 0.37, p < 0.0001$. In this case, given the $LSD_{.05} = 0.37$, the people object category was again detected significantly faster than the remaining 4 categories of objects. There was no significant difference in detection rates between fixed and plant objects, however fixed objects were detected significantly faster than both moveable and animal objects, and plants were detected significantly faster than the animal object category. The animal object category was detected significantly slower than the other 4 categories of objects. Overall hit rates for each object category for both the Appropriate and Inappropriate Context conditions are shown in Figure 3.3, along with the percentage of correct detections (cumulative) at each transitional point between flickers.

Table 3.3: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for the target-only inverted stimuli where the targets where placed in a location on the screen as if the entire scene had been inverted in Experiment 2.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 4.02 | 3.12 | 3.79 | 4.13 | 3.72 | 0.43 |
| Inapp Context | 4.25 | 2.32 | 3.78 | 3.37 | 3.60 | 0.37 |

## *Cross Condition Comparison*

Identification of objects in the Inappropriate Context condition did not differ significantly in comparison to the Appropriate Context condition, $F(1,21) = 2.60, MSE = 1.79, p = 0.12$, although there was a significant interaction with object category, $F(4,84) = 5.59, MSE = 0.43, p = 0.0005$.

# *3.1.5 Discussion*

When inverting the images, New et al. (2007) found that reaction times and accuracy scores were worse overall compared to when the images were presented fully intact, although they offered no analysis on how the individual object categories compared across manipulations. When New et al. (2007) inverted their images, the animate change detection advantage was eliminated. People, animal, fixed and plant objects all elicited comparable reaction times and accuracy scores, with the change detection for moveable objects being significant worse. When we inverted our images we found that overall performance equaled that found in Experiment 1a. Unlike New et al. (2007), when the full images were inverted, we continued to find category differences and the results were comparable across contextual manipulations with the people object category being seen faster with the animal category typically eliciting the worst performance. This trend continued across the two target inversion conditions:
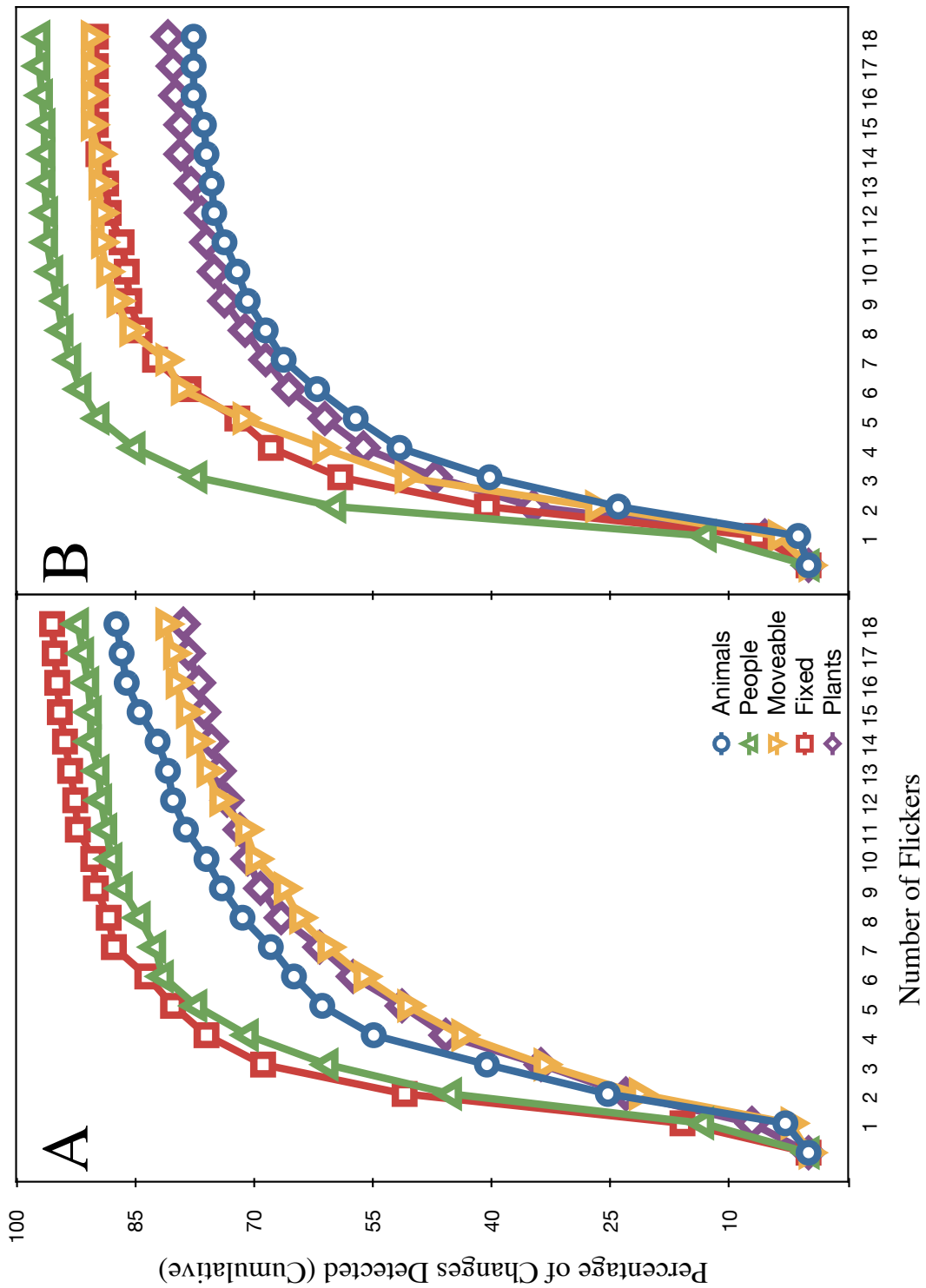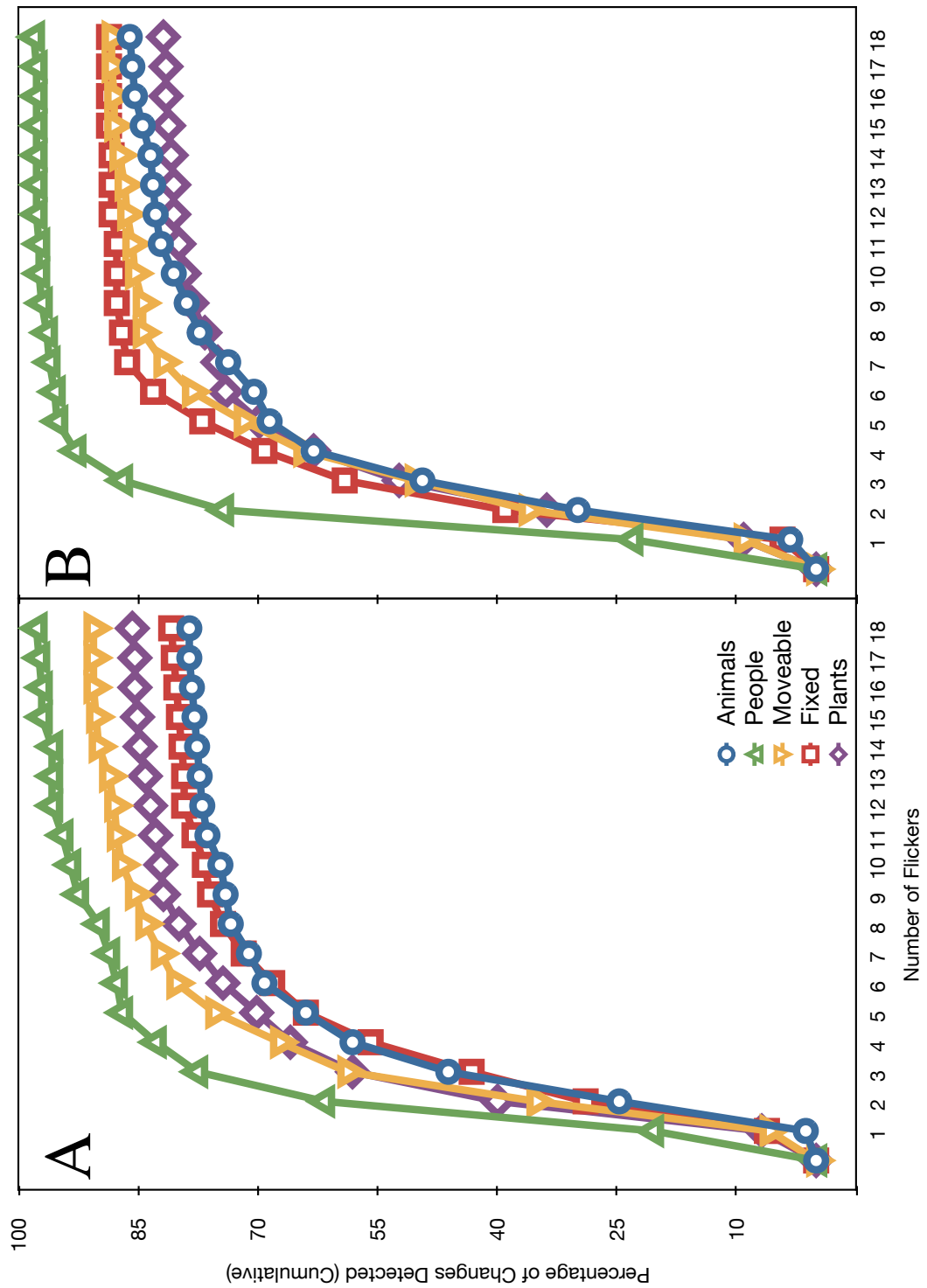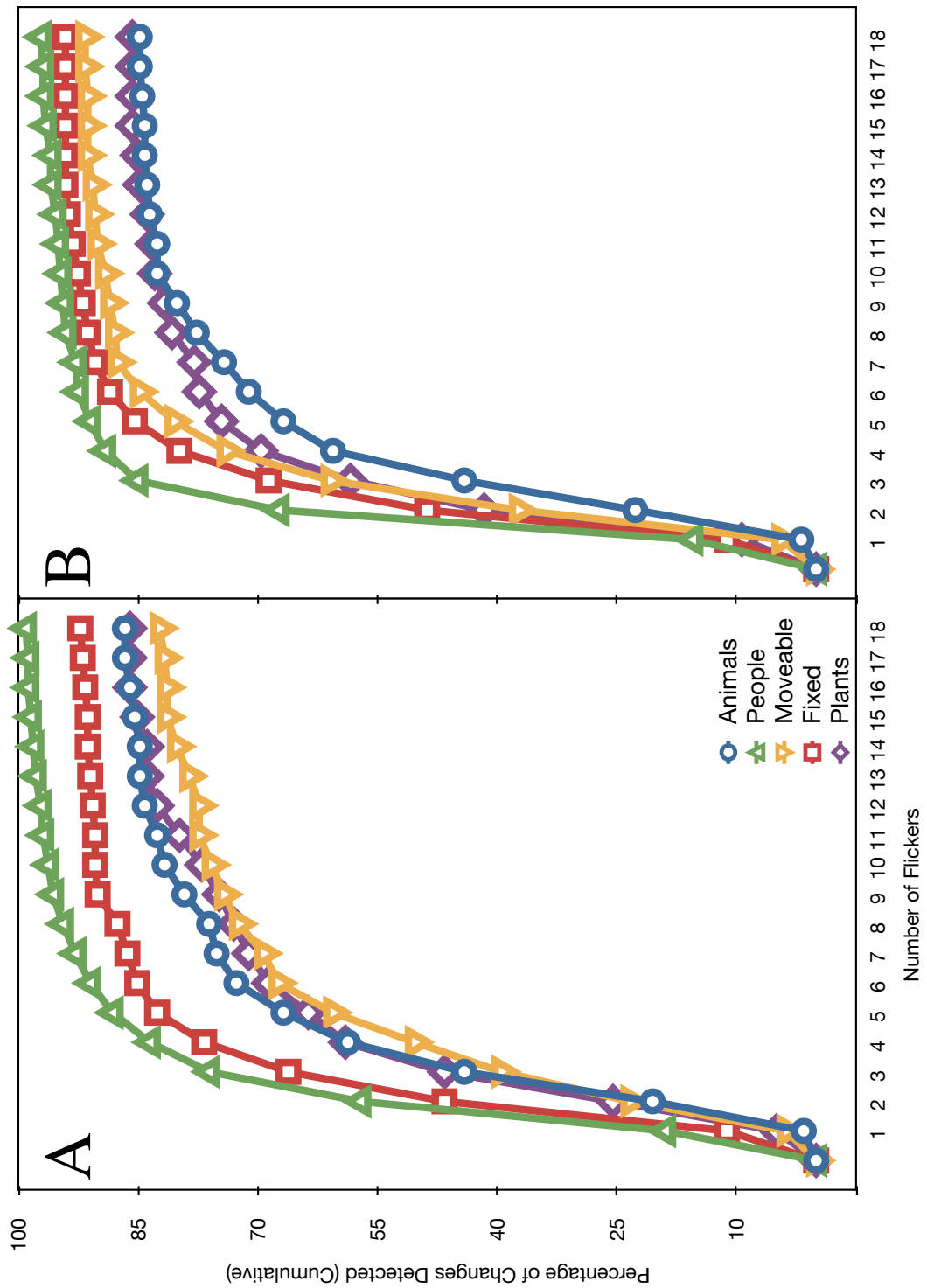
Figure 3.3: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A) and the Inappropriate Context (B) conditions when only the target object is inverted, but placed as if the entire image had been inverted, for Experiment 2.

changes to people objects were consistently detected faster across inversion conditions and contextual manipulations, with performance on the animal category of objects being consistently worse than the other object categories.

The problems raised by these findings are similar to the issues faced in light of the results found in the Inappropriate conditions of Experiments 1a and 1b. Proponents of the animate monitoring hypothesis base their arguments for a specialized visual monitoring system around the successful identification of the individual objects we encounter in our visual world. By disrupting identification of the objects, argue New et al. (2007), any animate bias should equally be removed; in fact, the performance on animate objects should be no different from performance on inanimate objects. Our findings do not support these assertions. Despite replicating the exact procedure of inverting the images in full and then inverting only the target objects, we continue to see differences across object categories. What is more troublesome for the animate monitoring hypothesis is that responses to the animate group of objects did not even move together.

It is quite possible that the process of inverting is not disrupting the identification of either the background or the target as intended. This possibility would explain the lack of deterioration in performance scores in comparison to Experiments 1a and 1b. Although there is a host of literature detailing the effects of inverting faces on recognition and memory (e.g., Ro et al., 2001; Yin, 1969), it is possible that the more complex natural scenes used in these experiments are processed in a much different way from faces. In order to account for this possibility, Chapter 4 details our attempt to disrupt object identification through the use of a gaussian blur filter.

# Chapter 4

# Disrupting Object Identification Through Blurring

In Chapter 3, an experiment was introduced where we attempted to disrupt the identification of the target objects by replicating a procedure used by New et al. (2007), that is, by inverting the images. The findings recovered from Experiment 2 did not support those found by New et al. (2007) and there is some concern that, despite replicating their procedure entirely, object identification was not disrupted adequately. Although inverting faces has been shown to interrupt recognition and memory, it is possible that these same interruptions do not translate to the complex set of natural images used here. This chapter will introduce an experiment where we tried a slightly different approach to disrupting the identification of the objects, a procedure also used by New et al. (2007), in this case, using a gaussian filter to blur the object.

## *Blurring the Image versus Blurring the Target Object*

In their fourth experiment, New et al. (2007) attempted to disrupt identification of the objects by placing a gaussian blur filter over each image. However, as with the inversion technique described in the previous chapter, the authors blurred the entire image rather than just the target object. This procedure makes it difficult to address the question they set out to answer, that is, is the animate monitoring bias driven by each object category, through information contained in the identity of the objects? Another problem with blurring and removing the category-specific features of the objects is that it necessarily makes it difficult to then name the objects when attempting to detect their change. New et al. (2007) did not explain how their participants were instructed to react to the stimuli when the information to categorize them was reduced.

New et al. (2007) found that when the images were blurred performance dropped for all object categories. Unfortunately, this result could be due simply to participants' inability to describe the change. If viewers were having trouble identifying the object of change, their performance could have plummeted as a consequence, regardless of when they actually saw the object change in the alternating images. New et al. (2007) also found no significant difference between animate and inanimate objects when the images were blurred. They interpreted these results as being evidence that the animate monitoring advantage was shaped by natural selection placing pressure on the visual system to attend to these categories, driven by object identification. Without the characteristics to identify the object readily at hand, the relevant information is lacking and the animate bias is lost.

## 4.1   Experiment 3: Blurring Aspects of a Scene

In Experiment 3, we presented participants with an exact replication of the fourth experiment introduced by New et al. (2007). That is, the same 5 categories of objects were introduced through the use of the flicker paradigm, with the images completely blurred using a gaussian filter. The target objects and backgrounds were the same as those used in Experiment 2. In an attempt to aid in describing the change, we gave participants a 4-box grid before presenting the blurred stimuli. We asked participants to try their best to name the object; however, if, due to the blurring, they were unable to identify the object of change, to rely then on the grid to identify in which part of the screen the change was happening. We also presented the images to participants where only the target object was blurred, leaving the background entirely intact. For both the Full Blur and Target-only Blur conditions the target objects were presented on both a contextually appropriate and contextually inappropriate background.

### 4.1.1   Methods

*Participants*

Forty-four undergraduate psychology students from the University of Lethbridge participated in exchange for course credit. Participants varied in both gender and age, but all were naive to the hypotheses of the experiment. Twenty-two of the 44 who participated took part in the Full Blur condition, whereas the remaining 22 took part in the Target-Only Blur condition.

*Stimuli*

As in the first two experiments, each of the experimental conditions (Full Blur-Appropriate Context, Full Blur-Inappropriate Context, Target Blur-Appropriate Context, and Target Blur-Inappropriate Context) consisted of 70 pairs of images. The images used were the same as those used in Experiments 1b and 2, created from photographs taken in Brisbane, Australia. The two Full Blur conditions were comprised of the images from the Appropriate and Inappropriate conditions used in Experiment 1b; however, here the images underwent a gaussian blur filter set at a PhotoShop 4.0 intensity of 1.0, similar to the filter used by New et al. (2007). The two Target Blur conditions used the same group of images; however, the gaussian blur application was applied to the target objects only, leaving the background image completely intact. For the Target Blur conditions, placing these images in the flicker paradigm gave the impression of a blurred mass changing somewhere within a fully intact background image.

## *Apparatus*

The same materials used in Experiments 1b and 2 were used for Experiment 3; however participants were also given a 4-box grid to help them explain where the object change was happening on the screen in the event that the blurring filter made it difficult to describe the object of change.

## *Procedure*

The procedure used for Experiment 3 was similar to the experiments previously described. Participants were randomly assigned to either the Full Blur or the Target-only Blur condition. However, unlike previous experiments and because of the blurring technique used here, participants were asked to try their best to name the object of change, but if they were unable to do so to then rely on the 4-box grid provided to describe the location of change.

## *4.1.2   Results*

## *Full Blur*

Twenty-two participants provided data for the Full Blur condition, each being exposed to both the appropriate and inappropriate context stimuli.

## *Appropriate Context*

The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits

Table 4.1: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for the fully blurred stimuli in Experiment 3.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 5.95 | 4.30 | 5.34 | 3.50 | 5.71 | 0.65 |
| Inapp Context | 6.00 | 3.67 | 5.55 | 5.34 | 4.27 | 0.52 |

are shown in Table 4.1. There was a significant effect of Object Category, $F(4, 84) = 20.17, MSE = 1.16, p < 0.0001$. Also shown in Table 4.1 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.65$) for the comparison of the means of each object condition to each other object condition. Fixed objects were detected significantly faster than the remaining 4 categories of objects. In turn, detection of people objects was significantly faster than moveable, plants and animal objects, with no significant differences among the remaining 3 categories of objects.

## *Inappropriate Context*

The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 4.1. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4, 84) = 26.83, MSE = 0.76, p < 0.0001$. In this case, given the $LSD_{.05} = 0.52$, the people object category was detected significantly faster than the remaining 4 categories of objects. The plant object category was detected significantly faster than the remaining 3 categories of objects. No significant difference was found between fixed and moveable objects; however, fixed objects were detected significantly faster than animals, whereas moveable objects were not. Overall hit rates for each object category for both the Appropriate

Table 4.2: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for the target-only blurred stimuli in Experiment 3.

| Condition | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
| --- | --- | --- | --- | --- | --- | --- |
| Appropriate | 4.55 | 3.08 | 4.75 | 3.74 | 3.56 | 0.48 |
| Inapp Context | 5.26 | 3.49 | 4.82 | 4.77 | 3.83 | 0.59 |

and Inappropriate Context conditions are shown in Figure 4.1, along with the percentage of correct detections (cumulative) at each transitional point between flickers.

## Cross Condition Comparison

Along with comparing object categories within condition we also compared responses across contextual manipulations. We compared the mean flickers to hits as a function of object category for the Appropriate Context and Location condition to the Inappropriate Context condition as a within-participant manipulation. Identification of objects in the Inappropriate Context condition was not significantly different from objects the Appropriate Context and Location condition, $F(1,21) = 0.0009, MSE = 2.77, p = 0.98$, although there was a significant interaction with object category, $F(4,84) = 18.55, MSE = 0.87, p < 0.0001$.

## Target Blur

Twenty-two participants provided data for the Target Blur condition. As in the Full Blur condition, each of these participants were exposed to 70 pairs of images form the appropriate context stimuli and 70 pars of images form the inappropriate context stimuli.

Figure 4.1: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A) and the Inappropriate Context (B) conditions when the entire image is blurred for Experiment 3.

## Appropriate Context

The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table 4.2. There was a significant effect of Object Category, $F(4, 84) = 16.72, MSE = 0.64, p < 0.0001$. Also shown in Table 4.2 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = .48$) for the comparison of the means of each object condition to each other object condition. The category of objects requiring the fewest mean flickers for correct object detection was the people object category. People objects were detected significantly faster than the remaining 4 categories of objects. In turn, detection of the plant and fixed object categories required fewer mean flickers for correct object detection than both the animal and moveable object category, with no significant difference between the two remaining categories of objects.

## Inappropriate Context

The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 4.2. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4, 84) = 12.41, MSE = 0.98, p < 0.0001$. In this case, given the $LSD_{.05} = 0.59$, the people and plant object categories were detected significantly faster than the remaining 3 categories of objects, with no significant difference between the fixed, moveable and animal object categories. Overall hit rates for each object category for both the Appropriate and Inappropriate Context conditions are shown in Figure 4.2, along with the percentage of correct detections (cumulative) at each transitional point between flickers.

## Cross Condition Comparison

Along with comparing object categories within conditions, we also compared responses across contextual manipulations. Again, we compared the mean flickers to hits as a function of object category for the Appropriate Context and Location condition to the Inappropriate Context condition as a within-participant manipulation. Identification of objects in the Appropriate Context condition was significantly faster in comparison to the Inappropriate Context and Location condition, $F(1,21) = 4.11, MSE = 3.28, p = 0.055$. There was, however, no significant interaction with object category, $F(4,84) = 2.08, MSE = 0.76, p = 0.09$.

## 4.1.3 Discussion

## Disrupting the Animate Advantage

As predicted by the animate monitoring hypothesis, no bias was given to animate objects as a collective category when the identification of the targets or images was disrupted. Not only should any advantage towards animate objects disappear, but people and animal objects should suffer equally. This was not the case when we blurred either the entire image or just the target object. Changes to people objects continue to elicit some of the quicker change detection, while the animal objects typically elicited the slowest detection. If our visual attention system has been shaped through natural selection to attend to these objects with added vigilance, any disruption should occur equally to both animals and people. This has not been the case with our contextual and spatial manipulations or in our attempt to disrupt object identification.
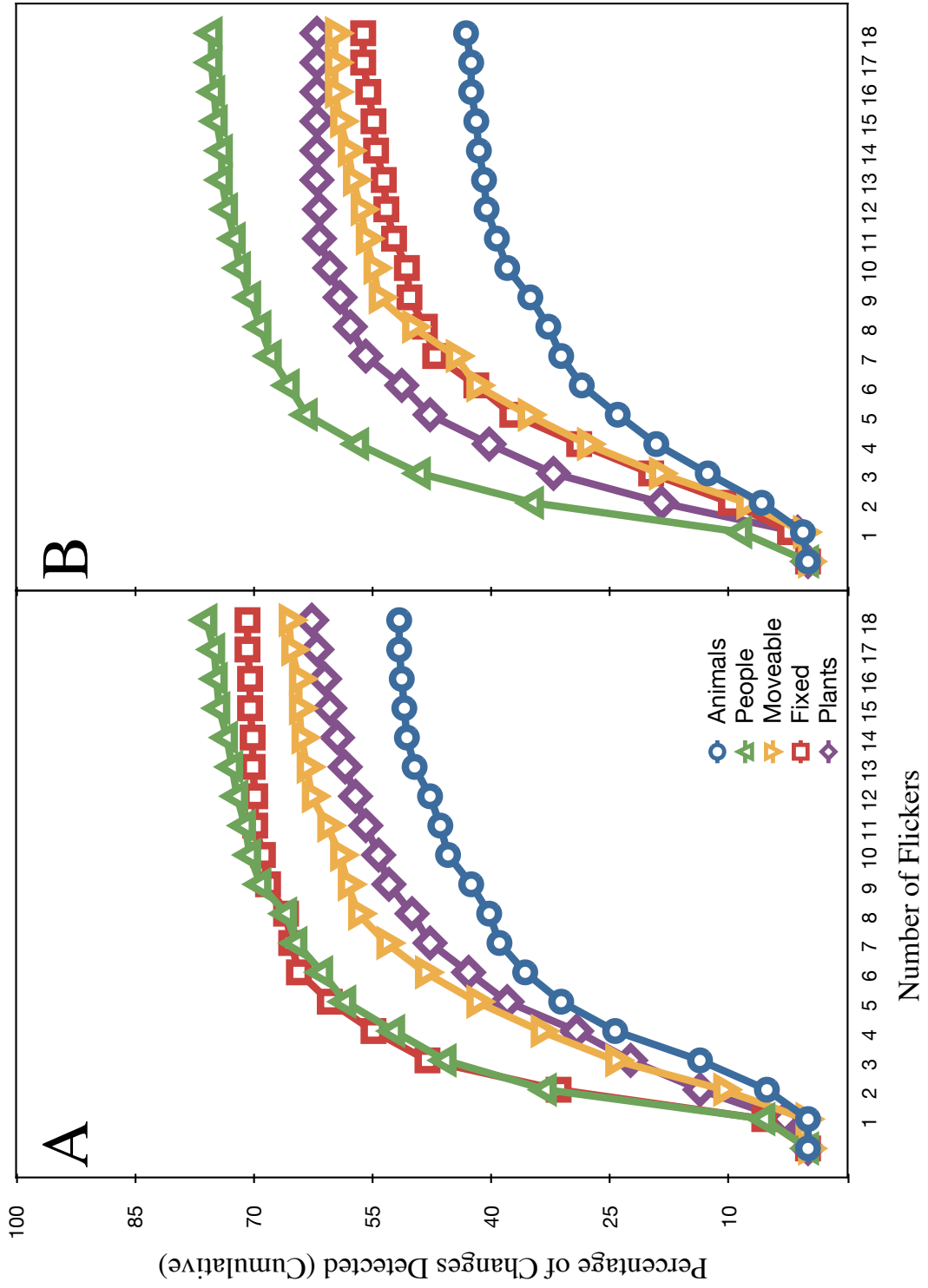
58
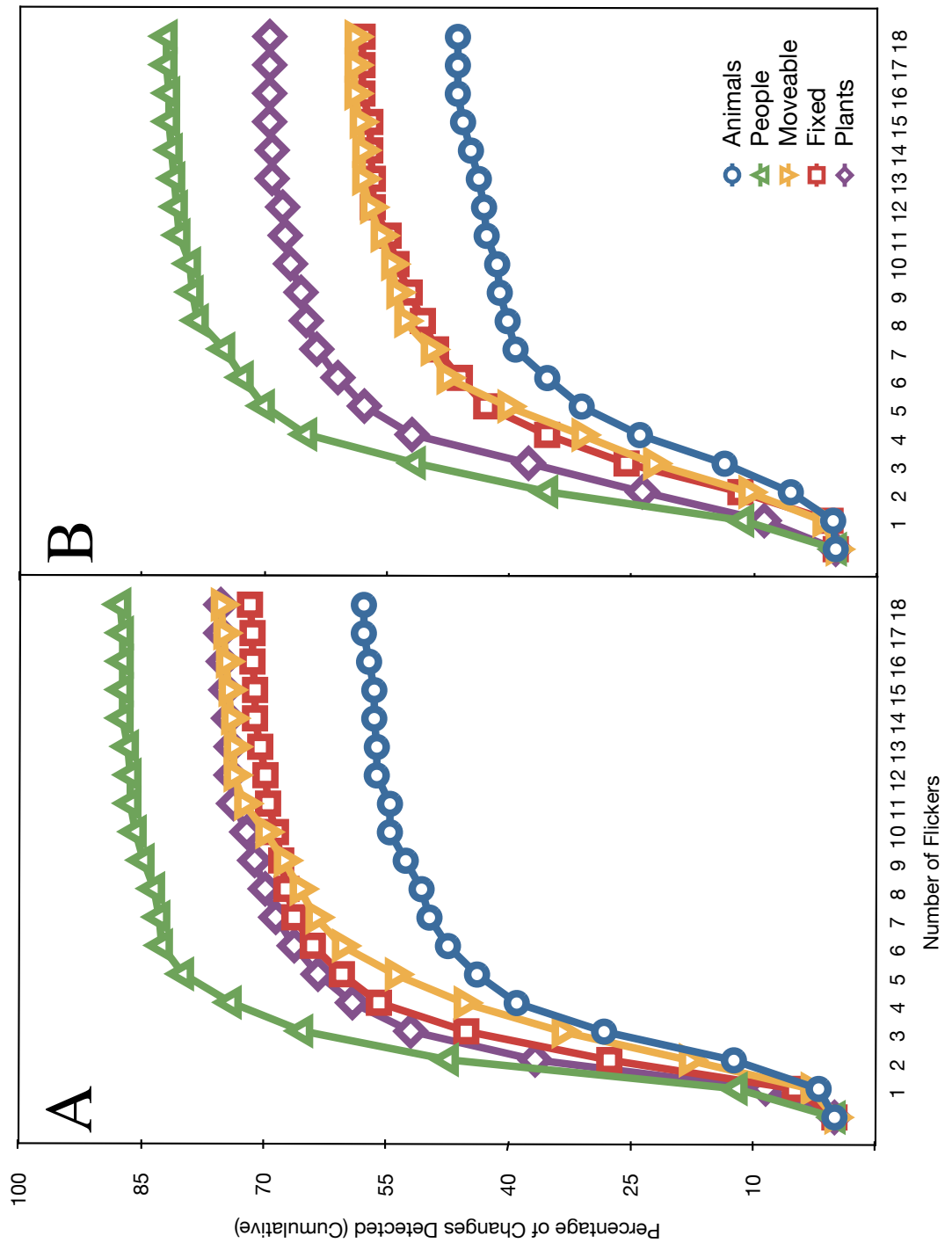
Figure 4.2: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A) and the Inappropriate Context (B) conditions when only the target object is blurred for Experiment 3.

## Contextual Effects

In each of our contextual manipulations prior to this experiment we have found a significant, or almost significant, change detection advantage for objects presented in the inappropriate context in comparison to change detection for objects presented on the appropriate background. Similarly, Hollingworth and Henderson (2000) found that the time it takes to detect change was shorter for changes to objects that were semantically inconsistent with the background scene and accuracy rates increased when presented in the flicker paradigm. Theoretically, this assertion has been questioned by several researchers where contextually consistent backgrounds are considered important sources of information when detecting change that should aid in both reaction times and accuracy scores when identifying an object (Chu & Jiang, 1998; Friedman, 1979; Kelley et al., 2003; Oliva & Torralba, 2007). By using contextual cueing, Chu and Jiang (1998) showed that change detection times were better for repeated visual arrays during an image comparison task. They argued that their viewers learned the contextual layout of the arrays and this helped them recognize and identify sources of change. However, these opposing results may be a product of the tasks used to assess the role of scene context. De Graef et al. (1990) argued that during free view or search related tasks, similar to the one used by Chu and Jiang (1998), the background context guides eye-gaze to likely locations for particular objects. In contrast, tasks that give the viewer only a short glance or use repeated interruptions may induce a more holistic way of gathering scene information and thus making violations more readily available.

It is possible that the nature of the flicker paradigm, with the short and interrupted image presentations, is inducing a holistic gathering of information compared to a free view or visual search approach. This perspective may explain the results found in our first 3 experiments where objects seemed to jump out at viewers when the objects violated the general style of the background scene. This perspective may also help to explain the opposite findings found in the fourth experiment resulting from blurring the target objects,

where these objects were detected faster when they were presented in the appropriate context. Blurring the objects may be inducing a guessing strategy based on the holistic representation of the background scene. Intuitively, this guessing strategy would seem to be more helpful when the blurred object is presented on a contextually appropriate background. This guessing strategy would be misguided when the blurred objects are presented on a contextually inappropriate background.

## *Detecting Change versus Identifying an Object of Change*

We introduced a slight modification of the procedure used by New et al. (2007) where we asked participants to rely on a 4-box grid in the event they were not able to identify the object of change. It is not clear from the original article presented by New et al. (2007) or the associated on-line appendices how their participants responded to the blurred stimuli. When the image is blurred identification becomes extremely difficult, but any difficulties with identification does not mean no change was detected. In order to ensure participants were able to indicate they had seen a change we removed the need to identify the object from simple change detection. Asking participants to indicate the location of change also helps to discount any questions of whether the viewers were simply guessing what the target object might be based on the surrounding scene. However, the question now becomes: Does naming the location of the change rather than identifying the object affect reaction times for hits? Experiment 4 was developed to answer this question.

## 4.2   Experiment 4: Detecting Change

The results collected from Experiment 3 raise some interesting issues. The animate monitoring hypothesis claims that an innate visual monitoring system produces a bias to see animate

objects and that this visual monitoring system works from the information contained through the identification of the target objects. It is clear from results found in Experiment 3 that change-blindness can still be avoided without the need for object identification. But were our viewers simply taking educated guesses based on the background scene? If the viewers were guessing their way through the flicker task, is this a strategy that can be translated to the real world or was this strategy induced by disrupting the identification of the objects?

Experiment 4 attempted to address some of these issues by presenting full backgrounds and target objects in the flicker paradigm and asking viewers to both name the object that is changing and to identify the location of the change relative to the rest of the screen. The same images used in the Experiment 1a were used for Experiment 4. Participants were given a 9-box grid rather than the 4-box grid used in Experiment 3 to allow for more confident accuracy scoring. We presented the objects in either the appropriate context and location, inappropriate context, or inappropriate location. We analyzed the responses in three different ways: if both the identification of the object and its location was correct it was counted as a hit, if the identification of the object was correct regardless of the location it was counted as a hit, and if the location described was correct, regardless of its identification it was counted as a hit.

## 4.2.1   Methods

### Participants

Twenty-four undergraduate psychology students from the University of Lethbridge participated in exchange for course credit. Participants varied in both gender and age, but all were naive to the hypotheses of the experiment.

## Stimuli

Experiment 4 used the exact same stimuli and conditions as those used in Experiment 1a.

## Apparatus

The same materials used in the first experiment were used here, with the exception of a 9-box grid given to participants in order to assess whether they could identify the location of the changing object on the screen.

## Procedure

The procedure used here was exactly the same as the procedure used in the first experiment; however, when giving their responses, participants were asked not only to name the object of change, but to also give its location relative to the screen with the aid of a 9-box grid provided for reference.

## 4.2.2 Results

All twenty-four participants provided data for the Appropriate Context and Location condition, with 11 of those participants being exposed to the Inappropriate Context condition and the other 13 participants being exposed to the Inappropriate Location condition.

## Category and Location

Here the responses to change-detection were counted as a hit if both the category of the object of change and its location in the scene were correct.

Table 4.3: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition when both the location of the change and the identity of the object of change are taken into account for Experiment 4.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 3.21 | 3.76 | 3.97 | 4.63 | 4.41 | 0.54 |
| Inapp Context | 3.35 | 2.43 | 2.53 | 4.24 | 3.43 | 0.72 |
| Inapp Location | 4.00 | 3.65 | 4.11 | 3.88 | 4.45 | 0.59 |

**Appropriate Context and Location**    The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table 4.3. There was a significant effect of Object Category, $F(4, 92) = 8.39, MSE = 0.89, p < 0.0001$. Also shown in Table 4.3 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.54$) for the comparison of the means of each object condition to each other object condition. The category of objects requiring the fewest mean flickers for correct object detection was animals. The animal category was detected significantly faster than the remaining 4 categories of objects. In turn, the people object category was detected significantly faster than the plant and fixed object categories, but did not differ significantly from the moveable object category. The moveable objects were detected faster than the fixed object category, but did not differ from the plant object category. There was no significant difference between the plant and fixed object categories.

**Inappropriate Context**    The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 4.3. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a

significant effect of Object Category, $F(4,40) = 8.71, MSE = 0.69, p < 0.0001$. In this case, given the $LSD_{.05} = 0.72$, the people and moveable object categories were detected significantly faster than the remaining 3 object categories. In turn, the animal and plant object categories were detected significantly faster than the fixed object category.

**Inappropriate Location**    The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 4.3. The mean number of flickers to hits per participant were, again, subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was no significant effect of Object Category found, $F(4,48) = 2.07, MSE = 0.55, p = 0.1$. Given the $LSD_{.05} = 0.59$, the people object category was detected significantly faster than the plant object category, but did not differ from the remaining 3 object categories. The plant object category was detected significantly slower than the people object category, but did not differ from the remaining 3 object categories. Overall hit rates for each object category for each of the 3 experimental conditions when both the location of change and the identity of the object of change are taken into account are shown in Figure 4.3, along with the percentage of correct detections (cumulative) at each transitional point between flickers.

**Cross Condition Comparison**    Identification of the objects in the Inappropriate Context condition were detected significantly faster than those objects in the Appropriate Context and Location condition, $F(1,10) = 20.77, MSE = 0.82, p = 0.001$; there was also a significant interaction with object category, $F(4,40) = 3.68, MSE = 0.80, p = 0.012$. There was no significant difference in the speed at which participants identified objects in the Inappropriate Location condition in comparison to the Appropriate Context and Location condition, $F(1,12) = 0.002, MSE = 2.90, p = 0.97$; there was also no significant interaction with object category, $F(4,48) = 1.61, MSE = 0.85, p = 0.19$. There was, however, a significant difference in the speed at which participants identified objects in the Inappropriate

Context condition in comparison to the Inappropriate Location condition with these objects being detected significantly faster, $F(1,22) = 6.14, MSE = 3.30, p = 0.02$; there was also a significant interaction with object category, $F(4,88) = 5.30, MSE = 0.62, p = 0.0007$.

## *Category*

Here, the responses to change detection were counted as a hit if the category given matched the object of change, regardless of where on the screen the viewer thought the change was occurring.

**Appropriate Context and Location**    The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table 4.4. There was a significant effect of Object Category, $F(4,92) = 10.85, MSE = 0.75, p < 0.0001$. Also shown in Table 4.4 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.50$) for the comparison of the means of each object condition to each other object condition. The category of objects requiring the fewest mean flickers for correct object detection was animals. The animal category was detected significantly faster than all of the remaining categories of objects except for people. In turn, the people object category was detected significantly faster than all other categories except the moveable object category. The moveable objects were detected faster than the fixed object category, but did not differ from the plant object category. There was no significant difference between the plant and fixed object categories.

**Inappropriate Context**    The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 4.4. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant

Figure 4.3: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A), Inappropriate Context (B), and Inappropriate Location (C) conditions when both the location of change and the identity of the object of change are taken in to consideration for Experiment 4.

Table 4.4: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition when only the identity of the object of change is taken into account for Experiment 4.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 3.36 | 3.71 | 4.02 | 4.83 | 4.42 | 0.50 |
| Inapp Context | 3.82 | 2.46 | 2.49 | 4.38 | 3.36 | 0.62 |
| Inapp Location | 4.14 | 3.61 | 4.56 | 4.19 | 4.38 | 0.54 |

ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4,40) = 14.87, MSE = 0.52, p < 0.0001$. In this case, given the $LSD_{.05} = 0.62$, the people and moveable object categories were detected significantly faster than the remaining 3 object categories. In turn, there was no difference between the plant and animal object categories; however, the plant category was detected significantly faster than the fixed object category, whereas the animal category was not.

**Inappropriate Location** The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 4.4. The mean number of flickers to hits per participant were, again, subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4,48) = 3.56, MSE = 0.47, p = 0.01$. Given the $LSD_{.05} = 0.54$, the people object category was detected significantly faster than fixed, plant and moveable objects, but did not differ from the animal object category. There was no significant difference between animals, fixed, plants, and moveable objects. Overall hit rates for each object category for each of the 3 experimental conditions when only the identity of the object of change is taken into account are shown in Figure 4.4, along with the percentage of correct detections (cumulative) at each transitional point between flickers.

**Cross Condition Comparison**    Identification of the objects in the Inappropriate Context condition were detected significantly faster than those objects in the Appropriate Context and Location condition, $F(1,10) = 24.41, MSE = 0.71, p = 0.0006$, there was also a significant interaction with object category, $F(4,40) = 7.61, MSE = 0.56, p = 0.0001$. There was no significant difference in the speed at which participants identified objects in the Inappropriate Location condition in comparison to the Appropriate Context and Location condition, $F(1,12) = 0.26, MSE = 2.15, p = 0.62$, and there was also no significant interaction with object category, $F(4,48) = 1.60, MSE = 0.75, p = 0.19$. There was, however, a significant difference in the speed at which participants identified objects in the Inappropriate Context condition in comparison to the Inappropriate Location condition, $F(1,22) = 8.83, MSE = 2.56, p = 0.007$, and there was also a significant interaction with object category, $F(4,88) = 9.05, MSE = 0.49, p < 0.0001$.

## *Location*

Here, responses were counted as a hit if the viewer correctly identified the location at which the change was taking place, regardless of whether they were able to successfully identify the changing object's category membership.

**Appropriate Context and Location**    The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table 4.5. There was a significant effect of Object Category, $F(4,92) = 11.44, MSE = 0.76, p < 0.0001$. Also shown in Table 4.5 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.50$) for the comparison of the means of each object condition to each other object condition. The category of objects
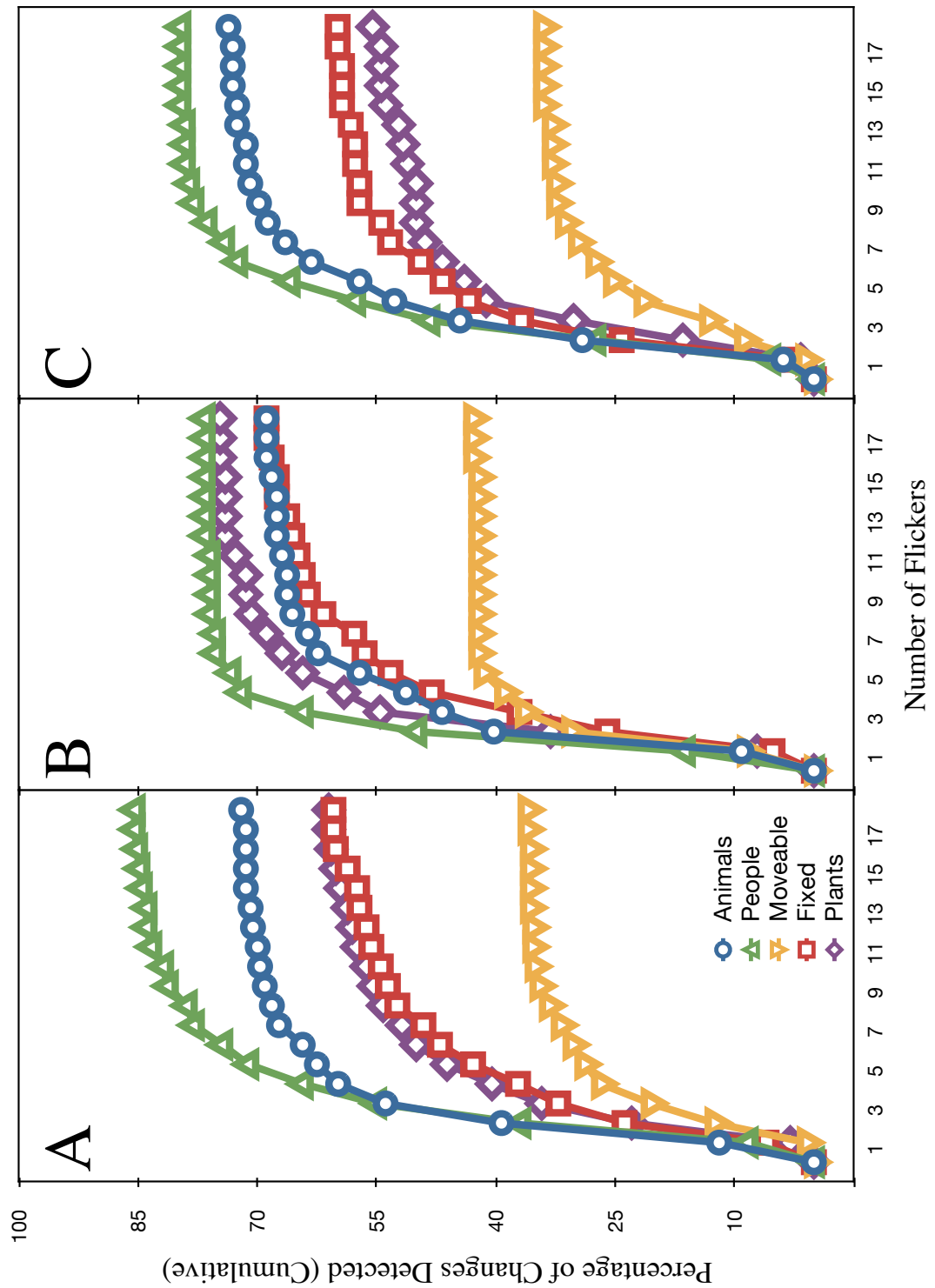
Figure 4.4: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A), Inappropriate Context (B), and Inappropriate Location (C) conditions when only the identity of the object of change is taken in to consideration for Experiment 4.

Table 4.5: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition when only the location of the changed object is taken into account for Experiment 4.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 3.24 | 3.81 | 3.78 | 4.68 | 4.56 | 0.50 |
| Inapp Context | 3.37 | 2.39 | 2.32 | 4.22 | 3.58 | 0.52 |
| Inapp Location | 4.09 | 3.63 | 4.49 | 4.04 | 4.41 | 0.53 |

requiring the fewest mean flickers for correct object detection was animals. The animal category was detected significantly faster than the remaining 4 categories of objects. In turn, the moveable and people object categories were detected significantly faster than the plant and fixed object categories. There was no significant difference between the plant and fixed object categories.

**Inappropriate Context**    The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 4.5. The mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4, 40) = 20.28, MSE = 0.36, p < 0.0001$. In this case, given the $LSD_{.05} = 0.52$, the moveable and people object categories were detected significantly faster than the remaining 3 object categories. In turn, the animal and plant object categories were detected significantly faster than the fixed object category.

**Inappropriate Location**    The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table 4.5. The mean number of flickers to hits per participant were, again, subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4, 48) = 3.31, MSE = 0.45, p = 0.018$. Given the

$LSD_{.05} = 0.53$, the people object category was detected significantly faster than the plant and moveable object categories, but did not differ from the fixed and animal categories. There was no significant difference between the fixed, animal, plant and moveable object categories. Overall hit rates for each object category for each of the 3 experimental conditions when only the location of change is taken into account are shown in Figure 4.5, along with the percentage of correct detections (cumulative) at each transitional point between flickers.

**Cross Condition Comparison**   Objects in the Inappropriate Context condition were detected significantly faster than those objects in the Appropriate Context and Location condition, $F(1,10) = 24.77, MSE = 0.71, p = 0.0006$; there was also a significant interaction with object category, $F(4,40) = 5.32, MSE = 0.59, p = 0.002$. There was no significant difference in the speed at which participants identified objects in the Inappropriate Location condition in comparison to the Appropriate Context and Location condition, $F(1,12) = 0.09, MSE = 2.92, p = 0.77$. However, there was a significant interaction with object category, $F(4,48) = 3.12, MSE = 0.66, p = 0.02$. There was also a significant difference in the speed at which participants identified objects in the Inappropriate Context condition in comparison to the Inappropriate Location condition, $F(1,22) = 8.81, MSE = 3.01, p = 0.007$, and there was also a significant interaction with object category, $F(4,88) = 10.54, MSE = 0.41, p < 0.0001$.

## 4.2.3   Discussion

The patterns found across our 3 criteria for hits were similar. As would be expected, the more conservative scoring technique uncovered fewer significant results, however the patterns for the speed for hits for each category stayed relatively consistent. There was some concern that guessing strategies might be one way to explain the results in Experiment 3 without
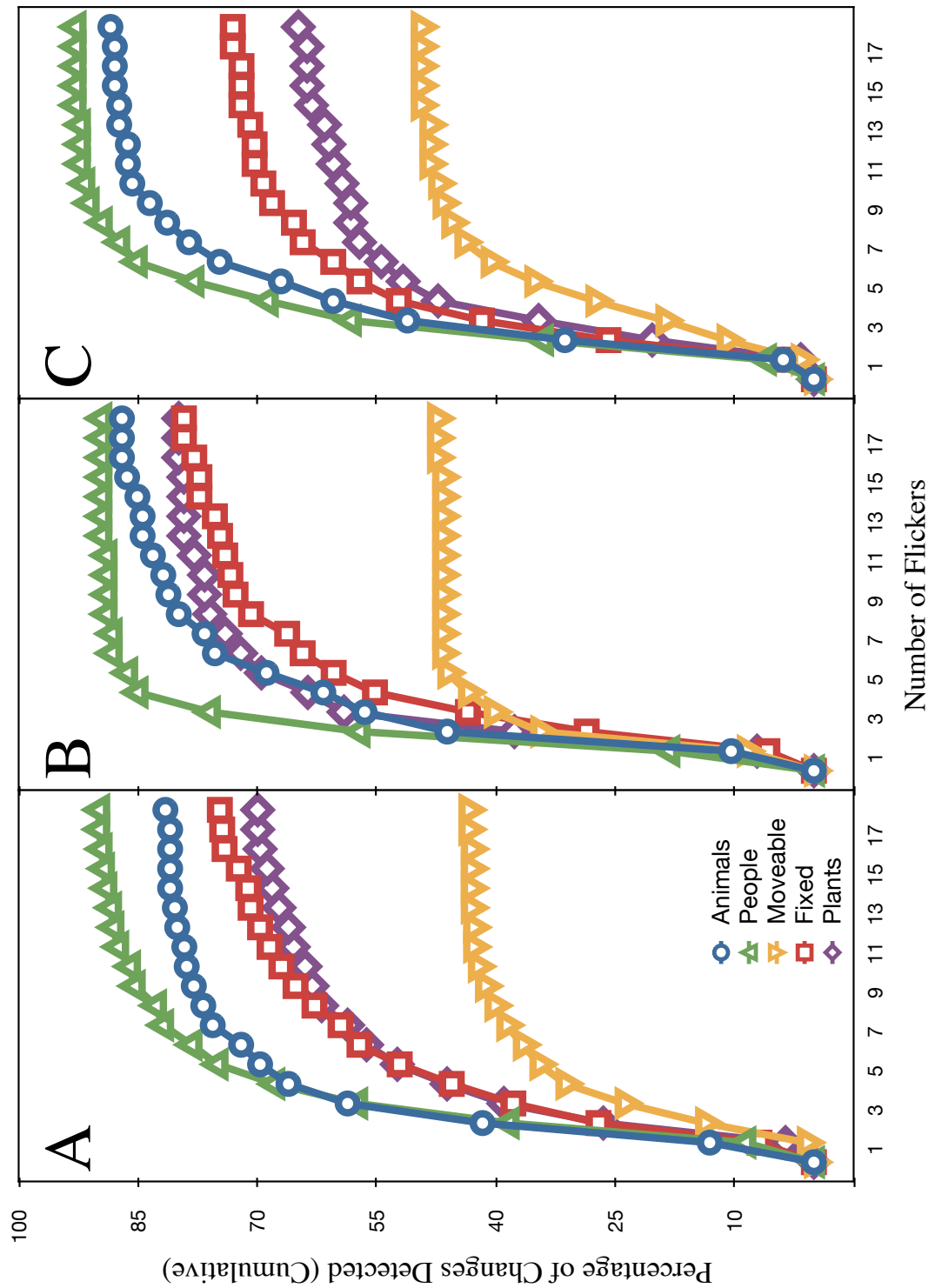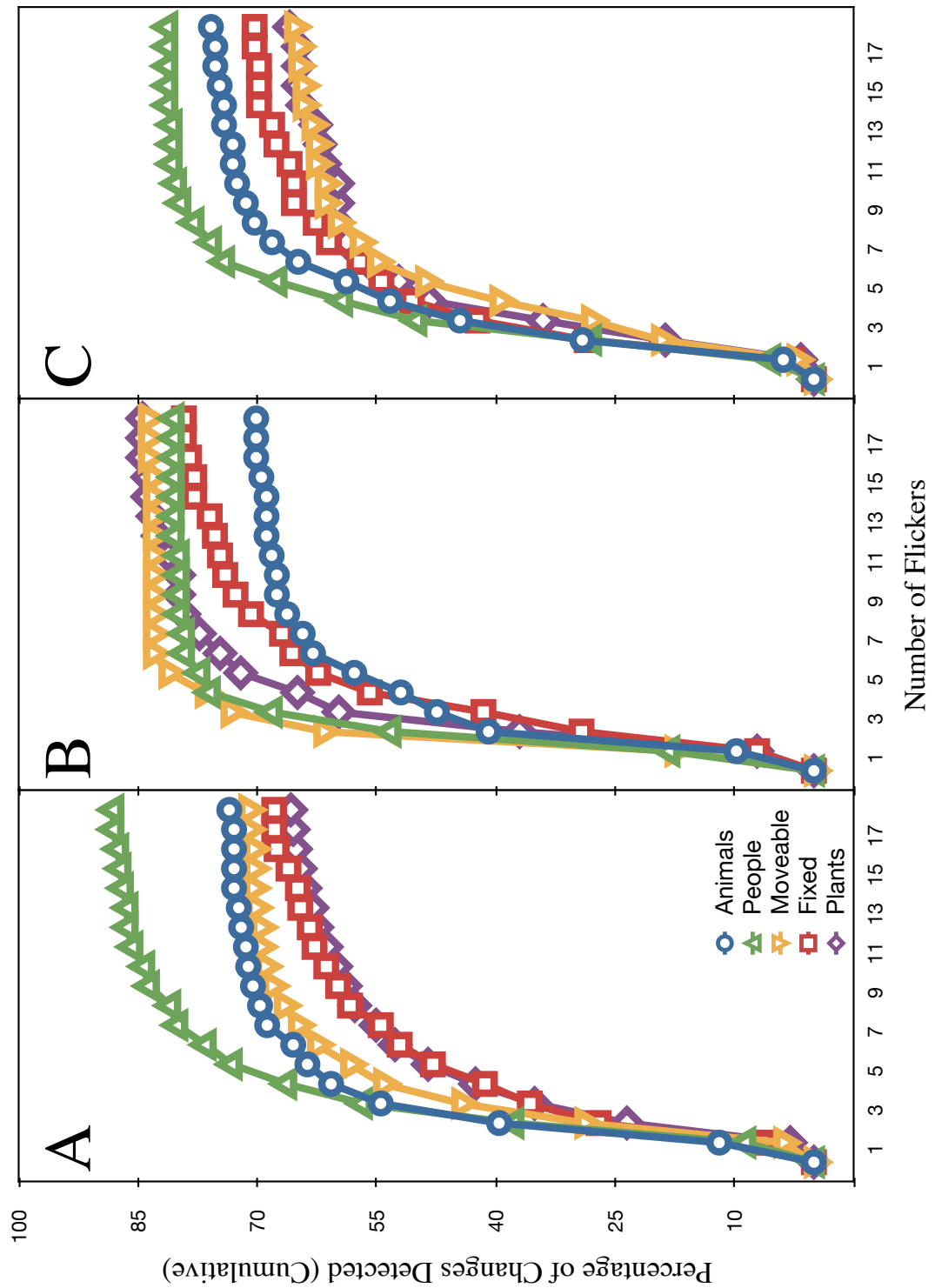
Figure 4.5: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A), Inappropriate Context (B), and Inappropriate Location (C) conditions when only the location of change is taken in to consideration for Experiment 4.

the need to detect any change. The consistency in results across our 3 scoring methods in Experiment 4 suggest that viewers are indeed seeing the change when they indicate such. It is important to acknowledge the power and limits of context and any subsequent guessing strategies. Although context may help guide a viewer to see relevant objects or their change, it does not seem to induce pure guessing. Although scoring for correct location only gave the most significant results compared to identifying the object, this result is not necessarily due to any guessing advantage. Taking away the difficult step of actually identifying the object of change may be enough to increase reaction times for hits.

## 4.3 General Discussion

### *4.3.1 Disrupting the Animate Advantage*

Experiment 3 failed to reproduce the results presented by New et al. (2007). Although our blurring techniques, for both the full image and target-only image, successfully removed any advantage for animal objects, the same was not true for people objects. This relative advantage for people objects has been a reoccurring theme for all of the experiments presented thus far. Any advantage for detecting animal objects is easily disrupted, whereas detecting changes to people objects continues to be done with relative ease. Through the lens of the animate monitoring hypothesis, New et al. (2007) predicted that by disrupting object identification any advantage for animal and people objects would be eliminated. It would seem that either our visual attention system is acting on different levels of information when dealing with animals versus people or perhaps, contrary to animate monitoring hypothesis, natural selection did not have a hand in shaping vigilance for these objects as a collective group.

### 4.3.2   Anomaly Detection

Although reaction to objects presented in the inappropriate context typically heightens recognition and change detection, this result was not the case for all categories of objects used here. Violations to a scene's style typically induces a holistic interpretation and gathering of information; as such, one would expect that any innate vigilance directed at particular categories would be exaggerated when presented on an odd background. This result has not been the case with animate objects as a collective category. Furthermore, if, as claimed by the animate monitoring hypothesis, change detection for animate objects is driven by the objects' categories, no differences should be found in detection when object identification is disrupted and presented on an appropriate background in contrast to an inappropriate background. The faster detection times for hits found in the Appropriate Context condition in Experiment 3 show that detection is based on more than just the target object. Information from the background is being gathered and used by viewers.

### 4.3.3   Using the Entire Visual Scene

The question remains whether this reliance on background information is something we use in our every day lives or whether it was induced by disrupting object identification. There is some strong evidence recently presented by Wu et al. (2010) that this is a strategy used even when the identifying features of the target object are readily available. Here, the authors presented photographs of either people or inanimate objects where each group of photographs was attached to one of two arbitrary categories. Participants' attention was not directed to this dichotomy, but were asked to remember the photographs and their category association. Participants not only classified new people and object photographs into the correct associative categories, but did so for photographs even where the person or object had been removed. Similar results have been found for artists' paintings. Tangen, Humphreys,

Vokey, and Cornwell (2010) presented reconstructed images of impressionist paintings by Monet and cubist paintings by Picasso in which all of the explicit identifying information was removed. In a series of experiments, Tangen et al. (2010) showed that participants were still able to classify the paintings despite the lack of semantic markers, presumably relying on the fundamental style of each artist. These results indicate that the background or style of a visual scene contains important information in addition to the target object.

# Chapter 5

## Assessing Effects of Interest on Change Detection

It is quite possible that our viewers' responses to our stimuli could be driven by something as simple as how interesting our particular set of images or target objects are. We attempted to address this question by introducing a second set of stimuli in Experiment 1b. The stimuli used in this experiment differed substantially from those used in Experiment 1a: they were created from photographs taken in Brisbane, Australia and were taken by a different photographer. The creation of the stimuli from photograph to flicker paradigm image was also orchestrated by two different individuals for each set. Using two sets of stimuli that differed in such ways may help to avoid common patterns such as the lighting of the photographs, locations of objects within the photographs or placement of the targets during image creation. Yet, given these differences in the images presented in Experiments 1a and 1b, consistent results were obtained across them, indicating that there was nothing peculiar with either set.

Aside from the image creation, it is possible that there is something inherent in the object categories themselves that makes one more interesting than another regardless of proximal experience or ultimate selection. New et al. (2007) acknowledge this point and indicated that given our phylogenetic history with animate objects they should, in fact, be considered more interesting. The animate monitoring hypothesis, however, states that any preferential attention bias is being driven by the importance these objects once played for our species, not by how interesting we now find them to be. In order to discount the possibility that any change detection advantage is being guided by the interest levels of the target objects or their categories, New et al. (2007) invited a separate group of participants to rate how interesting they found each target object to be. New et al. (2007) do not describe whether targets were presented as isolated objects for the judgement of interest, or embedded in the

background scenes. However, due to the constraints imposed by their process of stimulus creation, it seems most plausible that the target objects were presented embedded in the backgrounds. As expected, New et al. (2007) found a correlation between object category and interest ratings in the expected direction; however, the individual interest ratings did not predict individual reaction times. Therefore, New et al. (2007) concluded, the animate change detection advantage cannot be accounted merely by how interesting the objects or their categories are to the participants.

In order to make sure change detection in our experiments was not due to how interesting the target objects might be, we collected interest ratings for each. However, we were unsure of how New et al. (2007) presented their target objects to viewers and what question(s) were asked when they collected interest ratings. In an attempt to get the best approximation possible, we conducted two separate experiments. In the first experiment, the target objects were explicitly noted before participants gave their interest rating. In the second experiment participants were asked to say which part of the scene they found most interesting without any mention of target objects.

## 5.1  Experiment 5a: How Interesting?

For this experiment we presented the full (i.e., *A*) images from each of the conditions used in the first two experiments (Appropriate Context and Location, Inappropriate Context, and Inappropriate Location). However the target objects were circled in each image. We asked participants how interesting they found each of the target objects on a 12-point sliding scale. It was left to the viewers whether they let the background influence their decisions or not. In this case a score of *1* indicates that the target object was found to be *very uninteresting*, with a score of *12* indicating that the viewer found the object to be *very interesting*.

### 5.1.1 Methods

## Participants

Twenty-eight undergraduate psychology students from the University of Lethbridge offered their participation in exchange for course credit. The participants varied in both age and gender and all were naive to the hypotheses of the experiment.

## Stimuli & Apparatus

The background-plus-target images (*A*) from each of the 3 experimental conditions used in Experiment 1a were used for the current experiment; however, a red circle was placed around each target object for each image. The same apparatus used in Experiment 1a was used for the present experiment.

## Procedure

Participants were asked to read a standard experimental consent form and were given experimental instructions by the experimental facilitator. Target objects were circled and presented in either the appropriate context and location, inappropriate context, or inappropriate location. The images in each of these conditions were presented in a random order, ensuring that no two participants saw the same order of images. Participants were then asked to indicate how interesting they found each target object. They were not given any instruction on what information could be used. Interest ratings were indicated using a 12-point sliding scale and each participant was given as much time as needed for each image.

Table 5.1: Mean interest ratings for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for Experiment 5a.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 6.46 | 4.27 | 4.26 | 5.13 | 4.06 | 0.66 |
| Inapp Context | 7.98 | 7.98 | 6.91 | 6.60 | 5.87 | 0.62 |
| Inapp Location | 7.53 | 8.06 | 6.88 | 5.82 | 5.74 | 0.95 |

## 5.1.2 Results

### Appropriate Context and Location

Ten of the 28 participants were presented with images from the Appropriate Context and location condition. The mean interest rating for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random variate. The mean interest ratings for each object category are shown in Table 5.1. There was a significant effect of Object Category, $F(4, 36) = 18.62, MSE = 0.54, p < 0.0001$. Also shown in Table 5.1 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.66$) for the comparison of means of each object category to each other object category. With a potential score of '12', indicating the object is considered 'very interesting', the animal group of objects was rated to be significantly more interesting than all other categories of objects. In turn, fixed objects were rated as being significantly more interesting than the remaining 3 categories, with no significant difference among moveable, people, and plant objects.

## Inappropriate Context

Eight participants were presented with images from the Inappropriate Context condition. Again, the mean interest rating for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random variate. The mean interest ratings for each object category are shown in Table 5.1. There was a significant effect of Object Category, $F(4,28) = 18.38, MSE = 0.37, p < 0.0001$. In this case, given the $LSD_{.05} = 0.62$, the people and animal object categories were rated as being significantly more interesting than the remaining 3 categories of objects. There was no difference in interest ratings for the moveable and fixed object categories; however, each were rated as more interesting than the plant object category.

## Inappropriate Location

The remaining 10 participants were presented with images from the Inappropriate Location condition. The mean interest rating for each participant to each of the five object categories were again subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random variate. The mean interest ratings are shown in Table 5.1. There was a significant effect of Object Category, $F(4,36) = 9.49, MSE = 1.10, p < 0.0001$. Given the $LSD_{.05} = 0.95$, the people object category was rated as being more significantly interesting than each of the other object categories aside from animals. The animal and moveable objects were rated as more interesting than the remaining two object categories. There was no significant difference in interest ratings between the fixed and plant object categories.

*Cross Condition Comparison*

To assess differences across conditions we compared the mean interest ratings for the Appropriate Context and Location, the Inappropriate Context, and the Inappropriate Location as a between participant manipulation. There was a significant difference in interest ratings among conditions, $F(2,25) = 5.73, MSE = 12.35, p = 0.009$. There was also a significant interaction with object category, $F(8,100) = 6.16, MSE = 0.69, p < 0.0001$. With a mean difference of 2.23, target objects presented on a contextually inappropriate background were rated as being significantly more interesting than objects presented on a contextually appropriate background. Images presented from the Inappropriate Location condition were also rated as being significantly more interesting than those presented from the Appropriate Context and Location condition with a mean difference of 1.97. However, there was no significant difference in the interest ratings for Inappropriate Context condition in comparison to the Inappropriate Location condition with a mean difference of 0.26.

### 5.1.3 Discussion

Interest ratings collected in Experiment 5a show a similar pattern across experimental conditions as reaction times for hits found in the first two experiments, where the Inappropriate Context condition elicited the highest interest ratings and the fastest reaction times. The Inappropriate Location condition followed, again, for both interest ratings and reactions times, with the worst ratings and reaction times being reserved for the Appropriate Context and Location condition. Although this pattern reflects those patterns for reaction times found by New et al. (2007) and from our first two experiments, these patterns do not hold once we look at interest ratings for individual object categories within each condition. This result suggests that something other than how interesting an object is found to be is driving change detection.

## 5.2 Experiment 5b: Where does the interest lie?

In Experiment 5b we attempt, again, to gather the level of interest of each of our target objects; however, because of the ambiguity in the procedure used by New et al. (2007), we are forced to try several approaches. In this experiment, rather than bringing the viewer's attention to the target object, we presented the images with a 12-box grid superimposed onto them. Participants were asked to indicate which part of the screen they found most interesting. This technique allows us to assess whether the target objects are the most interesting aspect of each scene. Again, according to New et al. (2007), we should find that viewers rate the animate group of objects as more interesting than the inanimate group of objects. New et al. (2007) found a correlation between interest ratings and reaction times to the objects in the flicker task; however, they also found that the interest ratings did not predict the individual object category reaction times. According to this view, we should find the highest correlation of area of interest with target object location for the animate group of objects.

### *5.2.1 Methods*

### *Participants*

Thirty undergraduate psychology students from the University of Lethbridge participated in this experiment in exchange for course credit. None of the participants had previously participated in Experiment 5a.

## Stimuli and Apparatus

The background-plus-target images (*A*) from the three conditions in Experiment 1a were used in the current experiment; however, a 12-box grid was superimposed onto each image. The same apparatus and materials used in Experiment 5a were used for Experiment 5b.

## Procedure

The procedure used here resembled the procedure used in the previous experiment; however, rather than rating the target object on its interest level, participants were asked to indicate which part of the screen they found most interesting by referring to the 12-box grid superimposed onto each image.

## 5.2.2 Results

## Appropriate Context and Location

Thirteen of the 30 participants were presented with images from the Appropriate Context and Location condition. In order to assess interest ratings, hits were scored when the box containing the target object was rated as being the most interesting part of the screen. The mean hit rate for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random variate. The mean hit rates for each object category are shown in Table 5.2. There was a significant effect of Object Category, $F(4, 48) = 3.23, MSE = 0.02, p = .02$. Also shown in Table 5.2 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.11$) for the comparison of means of each object category to each other object category. Animal target objects were rated as being in the most interesting part

Table 5.2: Mean hit rates for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for Experiment 5b.

| | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| Condition | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
|---|---|---|---|---|---|---|
| Appropriate | 0.40 | 0.34 | 0.25 | 0.25 | 0.25 | 0.11 |
| Inapp Context | 0.71 | 0.61 | 0.64 | 0.54 | 0.39 | 0.15 |
| Inapp Location | 0.71 | 0.90 | 0.77 | 0.69 | 0.64 | 0.09 |

of the scene significantly more often than each of the other object categories, except for people objects. There were no significant differences in mean hit rates among the remaining 4 categories of objects.

## *Inappropriate Context*

Nine participants were presented with images from the Inappropriate Context condition. Again, in order to assess interest ratings, hits were scored when the box containing the target object was rated as being the most interesting part of the screen. The mean hit rate for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random variate. The mean hit rate for each object category are shown in Table 5.2. There was a significant effect of Object Category, $F(4,32) = 5.48, MSE = 0.03, p = .002$. In this case, given the $LSD_{.05} = 0.15$, animal target objects were rated as being in the most interesting part of the scene significantly more often than both fixed and plant objects. Moveable and people objects were rated as being in the most interesting part of the scene significantly more often than plants, whereas no difference was found among fixed and plant objects.

## Inappropriate Location

The remaining 8 participants provided data for the Inappropriate Location condition. Again, the mean hit rates for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random variate. The mean hit rates for each object category are shown in Table 5.2. There was a significant effect of Object Category, $F(4, 28) = 9.82, MSE = 0.01, p < 0.0001$. Given the $LSD_{.05} = 0.09$, the people target objects were rated as being in the most interesting part of the scene significantly more often than each of the remaining 4 categories of objects. In turn, moveable targets were rated as being in the most interesting part of the scene significantly more often than plant targets, which, however, did not differ significantly from animal and fixed targets. There was no significant differences among animal, fixed, and plant target objects.

## Cross Condition Comparison

To assess differences across conditions we compared the mean hit rates for the Appropriate Context and Location, the Inappropriate Context, and the Inappropriate Location as a between participant manipulation. There was a significant difference in hit rates among conditions, $F(2, 27) = 23.66, MSE = 0.11, p < 0.0001$. There was also a significant interaction with object category, $F(8, 108) = 2.84, MSE = 0.02, p = 0.007$. With a mean difference of 0.28, target objects presented on a contextually inappropriate background were rated as being a part of the most interesting aspect of the scene significantly more often than objects presented on a contextually appropriate background. Images presented from the Inappropriate Location condition were also rated as being a part of the most interesting aspect of the scene significantly more often than those presented from the Appropriate Context and Location condition with a mean difference of 0.45. Objects presented in the

inappropriate context were rated as being part of the most interesting area of the screen significantly more often than objects presented in the inappropriate location with a mean difference of 0.16.

### *5.2.3   Discussion*

As with the previous experiment, we did not find similar patterns between the areas of interest for target object location and the reaction times for hits found in the first two experiments. Interest ratings did not produce any advantage for animate objects as a collective either. The results from both Experiments 5a and 5b indicate that something other than how interesting an object is found to be is driving change detection. If an object's level of interest was guiding change detection, we would have expected results from these two experiments to reflect the same or similar patterns found for reaction times for hits in the first two experiments.

## 5.3   General Discussion

The interest ratings collected in Experiments 5a and 5b seem intuitive in the sense that objects presented in an inappropriate context or placed in the inappropriate location seem odd and, therefore, interesting. The oddness may explain the interest ratings, but the interest ratings do not seem to explain the reactions to change. Although we found similar results to those presented by New et al. (2007), these ratings do not predict reaction times for change detection. This is an important finding and one that allows us to speak to any suggestions that our viewers' reactions to these stimuli are based on our particular set or how interesting each category seems to be. New et al. (2007) interpreted their interest ratings and their lack of predictive power as a signal that the change detection must be a product of our

innate visual monitoring system, tuned to animate objects. We are not prepared to make that assertion, especially in light of our previous findings that seriously question the animate monitoring hypothesis.

# Chapter 6

# Conclusion

## 6.1 Stimulus Creation

We began by introducing a slightly different approach to stimulus creation from that used by New et al. (2007). We were concerned that the process of digitally removing and filling in the prevailing gap could attract unwanted attention that might be confused with object change detection. We were especially concerned that if this confound was creating an artifactual effect it would surely be for the animal objects as their typical backgrounds are perhaps the most difficult to fill in and create. In order to avoid this confound, we created our stimuli from two separate sources, one to serve as the backgrounds and the other to provide the target objects. We think this approach strengthens our findings. We are confident that photo-editing artifacts are not drawing attention to our target objects.

Each image for each experiment was created from photographs taken from the same camera, taken by the same photographer. The settings on the camera stayed consistent for all images, including both the backgrounds and the target objects. Each pairing of background and target object was done by the same person, adding consistency to the stimuli. It is acknowledged that this consistency could actually serve as a different confound (Tangen et al., 2010) of photographic style. Although effort was taken to avoid this possibility, it is conceivable that certain idiosyncrasies left by the photographer could have translated into the images. For example, all the people objects could have a certain orientation or relative placement in each of the backgrounds. In order to avoid this confound, a second set of images was created. The second set was created by a different person from the first set. The images taken for the second set were taken by a separate photographer from a separate camera, set at different settings. Using these two sets of stimuli and finding consistent results

should negate any concerns of such confounding effects.

## 6.2   Replicating the Animate Monitoring Advantage

When testing a hypothesis it is important to first replicate the original findings. We were able to do this, with the exception of some outliers. By using the same procedure, object categories and backgrounds we were able successfully to replicate the findings of New et al. (2007) with two independent sets of stimuli. With both sets, a clear change detection advantage for people and animals was found. This advantage prevailed despite differences in lighting conditions, orientation, pose, or busyness of the background scenes. Our participants responded consistently to images taken from both Canadian and Australian photographs. These results would appear to support the animate monitoring hypothesis.

## 6.3   Disrupting the Animate Monitoring Advantage

However, by manipulating either the contextual appropriateness of the background or the location at which the target object was placed any bias toward animate objects as a whole dissolved. It should be noted that New et al. (2007) did not test or speculate as to how the animate monitoring system might react to target objects being encountered out of place. From the framework of the animate monitoring hypothesis, it seems quite plausible that because natural selection did not shape the visual system under these conditions, then the preferential attention system would simply not react when confronted with these types of situations.

It is curious, however, that in our experiments people objects were typically detected most rapidly among the five categories of objects, despite contextual or locational manipulations, while the performance on animal objects routinely deteriorated. If we take the position

that natural selection did not guide our visual attention system through contextually odd scenes, and therefore no bias toward animate objects should be expected when these objects are encountered out of the ordinary, then why is performance so high for people objects? Not only is the high performance for people objects curious, but also is the fact that the animate objects separated. If we take as fact that natural selection placed pressure on our species to be aware of animate objects and that this vigilance ensured individual and genetic survival, one would expect that change detection for people and animals would be roughly the same, regardless of contextual or locational manipulations. In fact, New et al. (2007) do not distinguish between the two categories. It is animate objects as a collective group for which our visual attention system has been shaped they argued, not people objects, not animal objects. Animacy is the criterion. Therefore, detection of change to these objects should be the same; they should elicit strong performance as a collective or suffer in the same manner.

## 6.4   Disrupting Object Identification

According to the animate monitoring hypothesis, the mechanism involved in the bias toward animate objects works by using the categorical information contained in the identity of the individual object. In order to test this prediction, New et al. (2007) disrupted object identification by inverting or blurring the images. As was addressed in chapters 3 and 4, by inverting and blurring the entire images the authors where inadvertently disrupting the identification information for the backgrounds as well. This procedure not only creates a confound and confuses the original question posed, but it also makes identifying the object of change excessively difficult. How participants in these experiments responded to the changing objects once all the identifying information was removed is unclear. New et al. (2007) cite the poor performance in these experiments as evidence that the animate monitor-

ing system relies on the categorical information carried in the identity of the target objects. When object identification is disrupted, argue the authors, so is the animate advantage. However, these findings could be the result of the difficulty participants undoubtedly had in naming the objects once identifying information was removed.

We replicated the procedure used in New et al. (2007); that is, we inverted or blurred the entire image. Performance in these experiments plummeted, but we still found a distinction in reaction times among our categories of objects. In fact, it was the people objects that were consistently identified among the fastest of all the object categories. This result is an issue for the animate monitoring hypothesis. First, according to this perspective, when object identification is disrupted there should be no difference among the object categories. Second, if a distinction is found it should not be among the animate objects—they should be moving in tandem. The same mechanism should be working on both the people and animal objects, under the same conditions. Responses to these categories should be similar.

In an attempt to disambiguate the information being given by the target object and the background we inverted or blurred only the target objects, leaving the backgrounds completely intact, disrupting only object identification. In this case, the animate monitoring hypothesis would predict that any bias towards animate objects should disappear. Not only did we continue to find distinctions among object categories, but also significant differences among the categories of animate objects in both the inverted and blurred experiments.

When the target objects were blurred, we also found a difference in response times across contextual manipulations, but they were the opposite to the contextual effects found when no blurring was used. When object identification was disrupted it seemed to force viewers to rely on the background scene as a guide to what the object might be. The blurred objects presented in the Appropriate Context produced faster response times than those same objects presented in the Inappropriate Context. In contrast, when the objects' categorical information is left intact, the target objects typically jump out when they are presented

on an inappropriate background. By disrupting object identification through blurring the interaction between object and background features that typically provides the pop-out effect was eliminated.

## 6.5   A Preference for People

The results of these experiments do not support a preferential bias to attend to animate objects as a collective. The reduction in performance on animal objects when the scene's context is manipulated, when the location of the object is manipulated, and when the identification of the object is disrupted does not support an innate bias towards these objects. The ease with which reaction times for animal and people objects could be separated questions the idea that our visual attention system has developed a preferential bias to attend to animate objects. Yet, despite the apparent ease with which reaction times for animals could be slowed, change detection for people objects proved exceedingly difficult to influence, at least relative to the other object categories.

The are a number of possibilities that could be the source of this ability to see changes to people objects, and the explanation does not need to have an innate underpinning. The people objects used in these experiments were not particularly large in comparison with other categories, in fact they were quite average. While there is a rich line of research indicating our abilities to see, recognize, and remember faces, most of our people objects had their faces partially or fully obscured. Although our people objects were ranked as being relatively interesting they were not always rated as being the most interesting among our object categories, and these scores did not predict reaction times. People objects were not embedded in a small set of images, in fact they enjoyed the most diverse set of backgrounds. In fact, when creating the images it was difficult to find inappropriate places to put people objects. But, this would not necessarily help identification, it would actually hinder the pop-

out effect for the Inappropriate Context condition if the context was not as inappropriate for these objects as it was for other objects. Could it be our nature as social animals, constantly surrounded by other people, that has created a particular attention for these objects? This possibility does not necessarily need to be an innate argument, it could very well be that we have more experience with people than any other category of objects. But does this explain our being drawn to people or to detecting their change?

## 6.6   Detecting Change

Whether viewers are actually seeing the object of change when giving their responses has been a nagging issue throughout each of these experiments. As Wu et al. (2010) have shown, people can pick up on the subtleties of a scene's context and make accurate guesses as to what is most likely to be present in a particular scene. This possibility could account for some of the effects gathered when the objects were presented in the appropriate context. If this were the case, our data would not be a measure of change detection, but rather how quickly our participants can classify different types of natural scenes. We attempted to address this issues in a number of different ways. We used two different sets of stimuli, created by two different individuals. We removed the identifying features of both the objects and their backgrounds. We asked participants to indicate not only the object that changed, but also the location at which the change occurred on the computer screen. Yet, perhaps the most obvious way to test this possibility would be to present the images to participants using the flicker paradigm while gathering eye-gaze information.

Tracking eye-gaze could provide essential information which could be used to assess whether viewers are actually seeing the object they say they are seeing or whether they are instead classifying the style of the background scene. Gaze location at the time of response could be analyzed to uncover any potential correlations with change detection hit

rates. Fixation times could be compared at the time of response, as well as throughout each trial. Patterns of objects commanding the viewer's focus of attention at onset, detection and throughout the trial could be compared across object categories and background scenes.

Being the natural next step, we actually attempted this experiment. However the nature of our experimental design did not translate smoothly into the eye-gaze tracking software. Unfortunately, we were not able to capture eye-gaze patterns, but fixation points and times are described in the Appendix for objects presented in the appropriate context and location, the inappropriate context, and inappropriate location.

Translating our findings to the real world is a difficult task. Our world is not a static two-dimensional space like the images used in these experiments. As Rensink (2002) noted, much of change detection is driven by motion in our visual world. Although our animate group of objects possess the ability to move as being an intrinsic part of the category, their motion is not translated on to static images. Motion does not negate change blindness (Simons, Chabris, Schnur, & Levin, 2002; Levin & Simons, 1997; Simons & Levin, 1998), but it does typically attract attention (Rensink, 2002). Change without motion is rare, but it is something we do face from time to time, such as a new traffic sign. However, changes to animate objects without motion is even rarer. The flicker paradigm is useful in the way that New et al. (2007) employed it—that is whether there are particular biases for certain categories of objects or stimuli or differences among them, but without the other real-world characteristics attached to these objects, it is difficult to assess how and whether these findings would appear in our every day lives.

# Appendix A

# Tracking Eye-Gaze

We have shown that the effects of context are powerful. In Experiment 3, objects that had their category specific information stripped were better recognized when presented on the appropriate contextual background compared to when they were embedded in a contextually inappropriate background. Are viewers using the background to navigate their attention to adequate areas within the image or are they simply using a guessing strategy based on the information being gathered through the background scene? We attempted to answer this question in part by asking participants not only to name the object of change, but to also detail the area on the screen in which the change was occurring. Naming the location of change resulted in the most accurate responses. This makes sense as we removed one of the steps of the process—that is naming the object of change. It is an issue, however, whether viewers are not actually seeing the change, but are simply guessing. The animate monitoring bias is contingent on visual attention being drawn to certain objects and their changes in a natural scene.

## A.1   Experiment 6: Tracking Eye-Gaze and Fixations

In order to address this issue further, we presented a replication of our first experiment to a new group of participants while also tracking their eye-gaze. The question this experiment set out to answer was to see whether viewers were fixating on the target object when they indicated they had noticed a change. If viewers are actually seeing the object of change, their eye-gaze should be fixated on the target object as they give their response. If, however, viewers' responses to the change are being induced by guessing strategies being guided by the background scene, then we would not necessarily expect eye-gaze to be fixated on the

target object when giving their response. Tracking eye-gaze also affords the opportunity to monitor any patterns that might appear as viewers approach detection. That is, whether there are certain objects that most viewers attend to just prior to change detection. It is possible that there are fundamental cues contained within the background scene that might help viewers determine the object of change.

## *A.1.1   Methods*

## *Participants*

Twenty undergraduate psychology students from the University of Lethbridge volunteered to participate in exchange for course credit. Participants varied in both gender and age, but all were naive to the hypotheses of the experiment. Unfortunately, due to problems with the gaze tracking software most of the data collected for this experiment was incomplete and had to be discarded. We gathered data from all 20 participants to use for the flicker task, however data from only 12 participants was complete enough to use in the eye-gaze tracker analysis.

## *Stimuli*

The same target objects, categories, and backgrounds used in Experiment 1a were used for the present experiment.

## Apparatus

Stimuli were presented on a standard desktop computer using Metacard experiment software. Facelab software was used to gather facial information from each of our participants and Gaze-Tracker software was used to track eye-gaze. Participants indicated their responses to stimuli by using a standard computer keyboard and microphone.

## A.1.2   Results

## Change Detection

## Appropriate Context and Location

All 20 participants provided data for the Appropriate Context and Location condition. Hit rates alone were used to test accuracy as well as to assess reaction times. The mean number of flickers for hits for each participant to each of the five object categories were subjected to a one-way (Object Category), within-participant ANOVA with participants crossing Object Category as the random error. The mean number of flickers for hits are shown in Table A.1. There was a significant effect of Object Category, $F(4, 19) = 11.14, MSE = 0.61, p < 0.0001$. Also shown in Table A.1 is the Fisher Least Significant Difference at the $p = .05$ level ($LSD_{.05} = 0.49$) for the comparison of the means of each object condition to each other object condition. The moveable, people, plant and animal object categories were all detected significantly faster than the fixed group of objects.

Table A.1: Mean number of flickers for hits for each object category and experimental condition, as well as the Fisher Least Significant Difference at the 0.05 level for each experimental condition for Experiment 6.

| Condition | Object Category | | | | | |
| | Animate | | Inanimate | | | |
| | Animals | People | Moveable | Fixed | Plants | $LSD_{.05}$ |
| --- | --- | --- | --- | --- | --- | --- |
| Appropriate | 3.79 | 3.46 | 3.42 | 4.85 | 3.76 | 0.49 |
| Inapp Context | 3.99 | 2.52 | 2.51 | 4.24 | 3.39 | 0.56 |
| Inapp Location | 4.54 | 3.90 | 4.33 | 4.17 | 4.47 | 0.75 |

## *Inappropriate Context*

Twelve of the 20 participants provided data for the Inappropriate Context condition. The results from this condition in terms of mean flickers for hits as a function of object category are also shown in Table A.1. Again, the mean number of flickers to hits per participant were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. There was a significant effect of Object Category, $F(4, 11) = 17.07, MSE = 0.46, p < 0.0001$. In this case, given the $LSD_{.05} = 0.56$, the moveable and people objects were detected significantly faster than the remaining three object categories. In turn, the plant objects were detected significantly faster than the animal and fixed groups of objects.

## *Inappropriate Location*

The remaining 8 participants provided data for the Inappropriate Location condition. The results from this condition in terms of mean flickers for hits as a function of object category are shown in Table A.1. The mean number of flickers to hits per participants were subjected to a one-way (Object Category) within-participant ANOVA with participants crossing Object Category as the random error. An effect of Object Category approached significance, $F(4, 7) = 2.576, MSE = 0.534, p = 0.059$. In this case, given the $LSD_{.05} = 0.75$, the plant

category of objects were detected significantly faster than animal objects, but did not differ from the remaining 3 categories of objects. There was no significant difference between people, fixed, moveable, and animal objects. Overall hit rates for each object category for each of the three experimental conditions are shown in Figure A.1, along with the percentage of correct detections (cumulative) at each transitional point between flickers.

## Cross Condition Comparison

Identification of objects in the Inappropriate Context condition was significantly faster in comparison to the Appropriate Context and Location condition, $F(1, 11) = 12.59, MSE = 0.81, p = 0.005$, and there was also significant interaction with object category, $F(4, 44) = 3.40, MSE = 0.46, p = 0.02$. There was, however, no significant difference in the speed at which participants identified objects in the Inappropriate Location condition in comparison to the Appropriate Context and Location condition, $F(1, 7) = 1.54, MSE = 1.30, p = 0.26$. However, there was a significant interaction with object category, $F(4, 28) = 4.43, MSE = 0.49, p = 0.007$. Objects presented in the Inappropriate Context condition were detected significantly faster than objects presented in the Inappropriate location condition, $F(1, 18) = 5.74, MSE = 2.36, p = 0.03$; there was also a significant interaction with object category, $F(4, 72) = 6.66, MSE = 0.49, p = 0.0001$.

## A.1.3   Eye-Gaze

Unfortunately, due to issues with the software package data were not adequately recorded for eye-gaze in any useful way. However, fixation times for each objects location on the screen were recorded and analyzed.
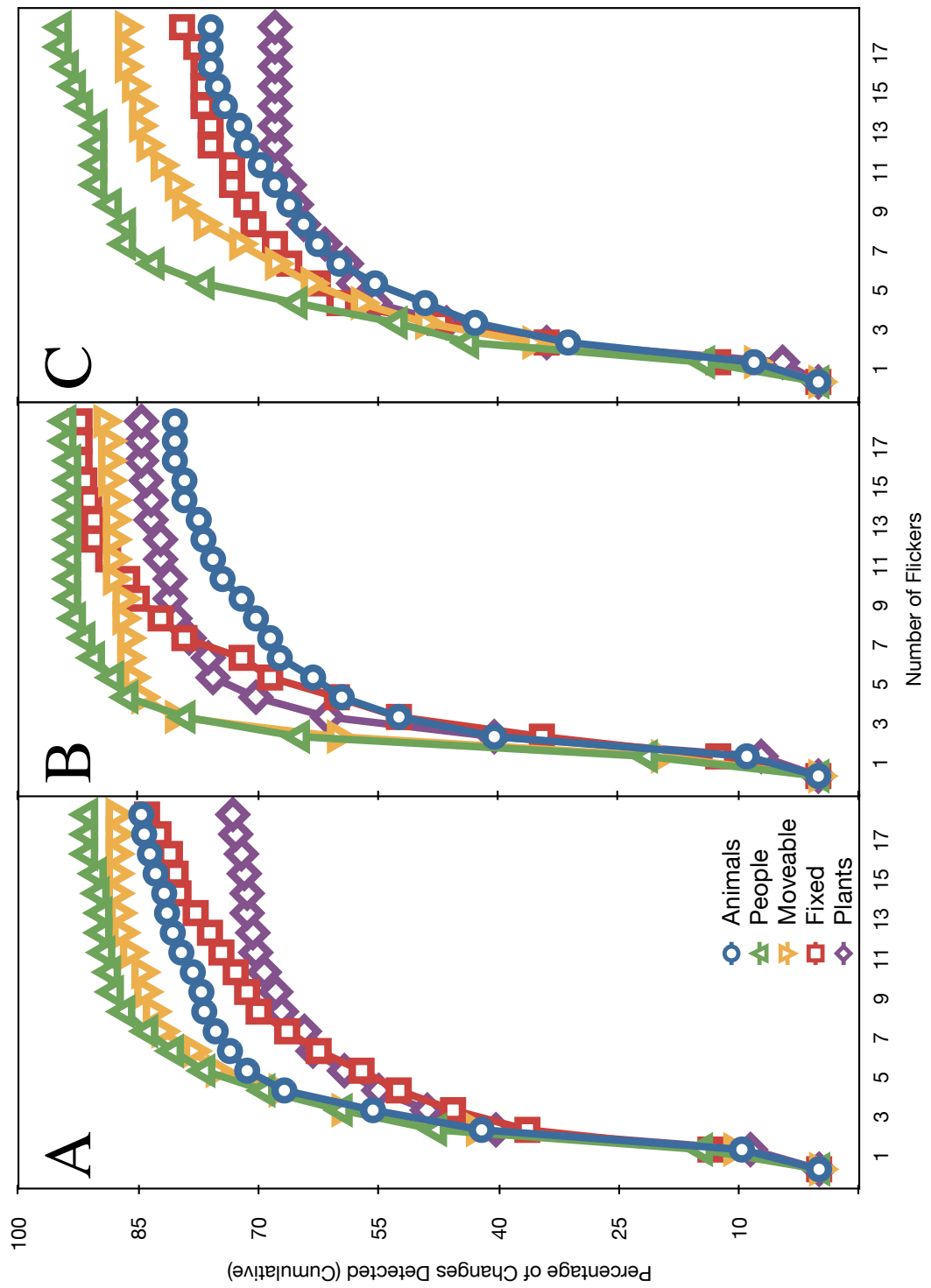
Figure A.1: Percentage of correct responses (cumulative) for each object category at each transition point between flickers for the Appropriate Context and Location (A), the Inappropriate Context (B), and the Inappropriate Location (C) conditions for Experiment 6.

*Fixations*

Data from only 12 of the 20 who participated were used for the fixation analysis. In order to access the amount of time viewers spent looking at the target object during image presentation we split the computer screen up into 24 look-zones. Whenever a viewers gaze stopped in a particular look zone for 250 ms or longer it was counted as a fixation. Once the eye-gaze had been determined to be a fixation we then calculated how long on average viewers fixated on the target object across object categories. It is important to note that some objects are harder to detect and therefore these trials last longer than trials for objects where the change is quickly detected. Therefore, some trials provide more time to fixate on the object even though that object's change has not yet been discovered.

For the Appropriate Context and Location condition, viewers spent the most time fixated on the fixed object category with a mean fixation time of 1.38 seconds, the category of objects eliciting the next longest mean fixation time was people (0.89), followed by animals (0.86), plants (0.81), and finally moveable objects (0.79). For the Inappropriate Context condition, the animal category of objects elicited the longest mean fixations with a mean fixation time of 1.29 seconds per trial, followed by fixed (1.17), people (0.96), plants (0.94), and moveable objects (0.80). In contrast, the moveable group of objects elicited the longest mean fixation times in the Inappropriate Location condition with a mean fixation of 1.59 seconds, followed by animals (1.54), fixed (1.49), plants (1.48), and people (1.35). Figure A.2 shows the mean fixation time for each object category for each of the three experimental conditions.

## A.1.4 Discussion

It is important to note that the results obtained from the flicker paradigm task across the three experimental conditions loosely reflect the results obtained from the first experiment which
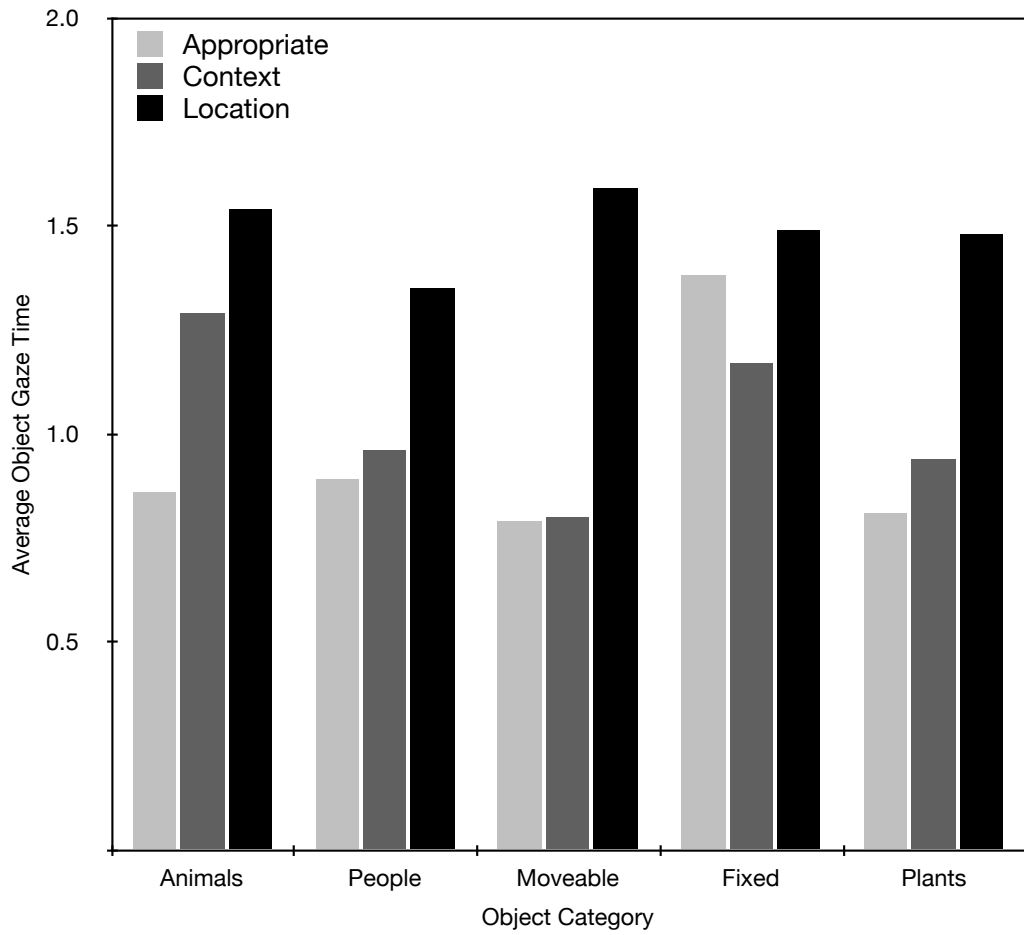
Figure A.2: Mean fixation time (s) for each category of object across each of the three experimental conditions presented in Experiment 6.

used the same stimuli and procedure. Although there were some differences across these experiments, these were most likely due to limited number of participants we were able to run using the FaceLab and Gaze Tracker software. Replicating the results from Experiment 1a is important to show that the procedure used to calibrate and maintain eye-gaze did not disrupt the attention of viewers when performing the flicker task.

## *Fixations*

The patterns uncovered for gaze fixations are curious in contrast to reaction times for change detection. Although reaction times have often been faster for objects presented in the Inappropriate Location condition in comparison to objects presented in the Appropriate Context and Location condition, fixation times were longer for objects that were presented in the inappropriate location. De Graef et al. (1990) found a similar result when presenting line drawings of complex scenes for free viewing. In this experiment the viewers spent more time fixating on objects that were located in an inappropriate location within the scene. Similar fixation times were found for objects that were presented without a background scene. De Graef et al. (1990) interpret these results as evidence of a bottom-up process for guiding visual attention and object perception. From this point of view, the background scene guides viewers to important aspects of that scene. It is important to note that the task used by De Graef et al. (1990) allowed for free viewing, whereas our short and interrupted presentation of images does not. Similar to the discusion in the Chapter 4, free viewing may be inducing a different way of gathering information compared to short, interrupted splashes of information.

It is interesting to note that although we found longer fixations for objects in the inappropriate location, the change detection for these objects was faster in comparison to changes to objects in the appropriate context and location. If viewers are using context to

direct their gaze and gather information, why are they not noticing changes faster when they are presented in the appropriate context compared to the inappropriate context? Again, it could be that when images are presented in short interrupted sequences the viewer must employ a holistic gathering of information. From this perspective, violations would immediately have been identified, in contrast to building a scene piece by piece, leading a viewer to the anomaly.

# References

Bacon, W. T., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*, 485–496.

Bar, M. (2004). Visual objects in context. *Nature Reviews: Neuroscience*, *5*, 617–629.

Boyce, S. J., Pollatsek, A., & Rayner, K. (1989). Effect of background information on object identification. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 556–566.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displaements of the visual world during saccadic eye movements. *Vision Research*, *15*, 719–722.

Chu, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.

De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research*, *52*, 317–329.

Fletcher-Watson, S., Collis, J. M., Findlay, J. M., & Leekam, S. R. (2009). The development of change blindness: Children's attentional priorities whilst viewing naturalistic scenes. *Develpmental Science*, *12*, 438–445.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*, 316–355.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin Company.

Gilbert, C., Ito, M., Kapadia, M., & Westheimer, G. (2000). Interactions between attention, context and learning in primary visual cortex. *Vision Research*, *40*, 1217–1226.

Grimes, J. (1996). *On the failure to detect changes in scenes across saccades* (k. Akins ed.). New York: Oxford University Press.

Hoffman, J. E., Nelson, B., & Houck, M. R. (1983). The role of attentional resources in automatic detection. *Cognitive Psychology*, *51*, 379–410.

Hollingworth, A., & Henderson, J. M. (2000). Semantic informativeness mediates the detection of changes in natural scenes. *Visual Cognition*, *7*, 213–235.

Hollingworth, A., & Henderson, J. M. (2004). Sustained change blindness to incremental scene rotation: A dissociation between explicit change detection and visual memory. *Perception & Psychophysics*, *66*, 800–807.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.

Kelley, T., Chun, M., & Chua, K. (2003). Effects of scene inversion on change detection of targets matched for visual salience. *Journal of Vision*, *3*, 1–5.

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

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

McGlynn, F. D., Wheeler, S. A., Wilamowska, Z. A., & Katz, J. S. (2008). Detection of change in threat-related and innocuous scenes among snake-fearful and snake-tolerant participants: Data from the flicker task. *Journal of Anxiety Disorders*, *22*, 515–523.

Mitroff, S. R., Simons, D. J., & Levin, D. T. (2004). Nothing compares 2 views: Change blindness can occur despite perserved access to the changed information. *Perception & Psychophysics*, *66*, 1268–1281.

New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences of the United States of America*, *66*, 16598–16603.

Nothdurft, H. C. (2002). Attention shifts to salient targets. *Vision Research*, *42*, 1287–1306.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, *155*, 23–36.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*, 520–527.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*, 369–378.

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

Reinecke, A., Becker, E. S., & Rinck, M. (2010). Visual working memory and threat monitoring: Spider fearfuls show disorder-specific change detection. *Behaviour Research and Therapy*, *48*, 770–778.

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

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (2000). On the failure to detect changes in scenes across saccades. *Visual Cognition*, *7*, 127–145.

Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: A detection advantagein the flicker paradigm. *Psychological Science*, *12*, 94–99.

Simons, D. J., Chabris, C. F., Schnur, T., & Levin, D. T. (2002). Evidence for preserved representations in change blindness. *Consciousness and Cognition*, *11*, 78–97.

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

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*, 16–20.

Smith, C. N., Hopkins, R. O., & Squire, L. R. (2006). Experience-dependent eye movements, awareness, and hippocampus-dependent memory. *The Journal of Neuroscience*, *26*, 1304–11312.

Stirk, J. A., & Underwood, G. (2007). Low-level visual saliency does not predict change detection in natural scenes. *Journal of Vision*, *7*, 1–10.

Tangen, M. J., Humphreys, M. S., Vokey, J. R., & Cornwell, T. B. (2010). Nonanalytic cognition: Generalizing without thinking. *Journal of Experimental Psychology: Human Perception and Performance*, *submitted*.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and ofsets. *Perception & Psychophysics*, *49*, 83–90.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*,

*51*, 599–606.

Theeuwes, J., Atchley, P., & Kramer, A. F. (2000). *On the time course of top-down and bottom-up control of visual attention*. Cambridge, MA: MIT Press.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*, 766–786.

Turatto, M., & Galfana, G. (2000). Color, form and luminance capture attention in visual search. *Vision Research*, *40*, 1639–1643.

Werner, S., & Thies, B. (2000). Is "change blindness" attentuated by domain-specific expertise? An expert-novices comparision of change detection in football images. *Visual Cognition*, *7*, 162–173.

Wu, W., Tangen, M. J., & Vokey, J. R. (2010). On discriminating people: Nonanalytic concept learning of image styles. *In preparation*.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 121–134.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.

Zoest, W. V., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 746–759.