The Impact of Digital Change on Memory and Cognition

by

Sophie Jane Nightingale

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Psychology

University of Warwick, Department of Psychology

September 2017

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I would like to acknowledge my supervisors. Kimberley Wade, thank you for your endless support and advice, for encouraging me to take on new challenges, and for all the time and effort you have invested in my academic development. Derrick Watson, thank you for your continued and dedicated guidance, and for always offering an enthusiastic, thoughtful, and patient approach. Kim and Derrick, I have learnt so much from the both of you and I am forever grateful for all you have done to help me with my research and my future career. Thank you for your inspirational supervision, it has been an honour working with you both.

I would also like to thank a number of academic collaborators and mentors. Maryanne Garry, thank you for welcoming me into the Garry lab and taking the time to support my academic development—you are an inspiration. Thank you to Hany Farid for your invaluable advice and guidance; I am extremely grateful to you for investing your time to help me to further develop my research program. Thanks also to Linda Henkel, Rachel Zajac, and Andrew Mills—it has been a pleasure collaborating with you.

Many thanks to the Economic and Social Research Council and the Psychology Department at The University of Warwick for the financial support that has allowed me to conduct my research. Thank you to all of the academics, research students, and technical and support staff in the Psychology Department at Warwick University for providing a supportive and encouraging environment to work in. A special thanks to my lab mates, Melissa, Divya, Harriet, and Danni for your friendship and incredible willingness to help.

Finally, a massive thank you to my friends and family. Tam, Helen, Liz—thanks for always being there. To my parents: thank you for all of the opportunities you have given me and for your unwavering support. Thanks to my siblings for always giving excellent advice and wanting the best for me. Karen, thanks for being an incredible role model and for teaching me to never give up. My final thank you goes to Sarah, my partner and best friend: I am deeply grateful for your endless love and encouragement; above all, thank you for believing in me and helping me to believe in myself. I couldn’t have done this without you.

Declaration

This thesis is submitted to The University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work presented was carried out entirely by the author with one exception:

> The data from Experiment 1 in Chapter 7 were collected at the University of Otago by Andrew Mills and Dr Rachel Zajac. With their permission, I have re-analysed the data for use in this thesis.

Inclusion of Published Work

Parts of this thesis have been published by the author.

Chapter 3 includes the following publication:

> Nightingale, S. J., Wade, K. A., Watson, D. G. (2017). Can people identify original and manipulated photos of real-world scenes? *Cognitive Research: Principles and Implications*, *2*, 1–21. doi:10.1186/s41235-017-0067-2

Dr Kimberley Wade and Dr Derrick Watson contributed to the planning of this research and provided feedback on drafts of the manuscript.

In addition, Appendix A, "Saliency Analyses in Chapter 3", appears as a supplementary file in Nightingale et al. (2017).

Abstract

In the digital age, there has been a phenomenal rise in the number of photos people capture, share, and manipulate—a trend that shows no sign of slowing. Furthermore, research shows that photos—authentic and manipulated—are powerful; they can change people's memories for distant and recent experiences, beliefs about past actions, intentions for future actions, and judgements. Yet there is currently limited research exploring the effects of digital photography on memory, cognition, and behaviour.

Part One of this thesis comprises of a program of research that examines people's ability to discriminate between authentic and manipulated images. Advances in digital technology mean that the creation of visually compelling photographic fakes is growing at an incredible speed. Despite the prevalence of manipulated photos in our everyday lives, there is a lack of research directly investigating the applied question of people's ability to detect photo forgeries. The research in Chapter 3 addresses this question. Across two experiments, people showed an extremely limited ability to detect and locate manipulations of real-world scenes. Chapters 4 and 5 explore ways that might help people to detect image forgeries. Specifically, the research investigates the extent to which people can identify inconsistencies in shadows and reflections. The results suggest that people are reasonably insensitive to shadow and reflection information and indicate that such image properties might not help people to distinguish between authentic images and manipulated ones.

Part Two of this thesis examines how the act of taking photos can affect people's memory. Digital technology has revolutionised the ease with which people capture photos and accordingly there has been a remarkable rise in the number of photos that people take. The results of five experiments and a mini meta-analysis suggest that taking photos has only a small, or plausibly no, effect on people's memories.

Chapter 1 :
Introduction

"If I could tell the story in words, I wouldn't need to lug a camera."

Lewis Hine (1973)

People are able to rapidly and effortlessly process visual information. Just a glance at an image is often enough to glean its basic meaning (Potter, 1975; Thorpe, Fize, & Marlot, 1996). The ease with which people interpret visual information can help to explain why photography has had such a phenomenal impact on the world—as the cliché goes, a picture is worth a thousand words. Accordingly, photos are used across almost all domains, including law enforcement, photojournalism, politics, and advertising. More recently, even laypeople can easily upload and share their personal photos with the world. Research clearly demonstrates the powerful influence of photos: news stories with photos get more attention than text alone (Moses, 2002), and on social media, posts with photos get more likes and shares than those without (Bakhshi, Shamma, Kennedy, & Gilbert, 2015). Photography has now become so interwoven with modern society that it is almost impossible to imagine life without it. Yet cameras and photography are a relatively new technology. Less than 200 years have passed since the invention of photography, and now, it is ubiquitous.

A brief history of photography

Drawing with light

Centuries before photography was invented, people were aware of the basic principle of the camera: that light entering through a small hole in the wall of a dark room forms an inverted and horizontally reversed image of the scene outside (Galassi, 1981; Newhall, 1982). In fact, some of the earliest art—in the form of cave drawings—might have been created by tracing the light projections caused by sunlight entering through small holes in the cave walls (Gatton, 2009). The use of light projections to create images of the world was formalised in the 16th century with the invention of the *camera obscura* (Gernsheim, 1986; Newhall, 1982). The early camera obscura created blurry and undetailed projections but by the 18th century a vastly improved version was available and artists commonly used these devices as an aid for painting and drawing (Newhall, 1982; Steadman, 2001). Projecting the light, however, proved easier than preserving it: even when using a camera obscura, the resulting picture was still largely dependent on the individual's artistic ability (Newhall, 1982; Wright, 2016). Thus there remained a desire to make the image projections permanent, giving scientists the impetus to continue their search for a method to both produce and preserve images.

The chemical effects of light

Advances in the understanding of the chemical effects of light proved important for the development of photography (Gernsheim, 1986). Such scientific advances revealed a way to permanently fix images and led to the world's first successful permanent photograph from nature: Niépce's View from the Window at Le Gras (Gernsheim, 1977, 1986). Over the next two decades, attempts were made to refine and simplify the photographic process, the most successful of these attempts being the daguerreotype and calotype methods (Gernsheim, 1986; Graham, 1984; Larsen, 2008). Yet both of these processes were far from perfect—the main strength of the daguerreotype was its quality while the calotype marked the first successful attempt at a negative-positive photographic process. The appeal of having a negative from which numerous prints could be made was quickly recognised and scientists focused their interest on developing the negative-positive photographic process (Gernsheim, 1986; Hand, 2012; Mullins, 2013; Slater, 1991). A significant development came in the 1880's when Kodak's founder, George Eastman, introduced flexible photographic film and the first portable film camera (Larsen, 2008). The camera's accompanying advertising slogan "You press the button, we do the rest." conveyed that, for the first time in history, photographers were not required to process and develop their own photos—literally anyone could point the camera and shoot. Accordingly, a new era in photography began—an era that prevailed for the next century.

From chemicals to bits—the digitalisation of light

The invention of the charge-coupled device (CCD) marked one of the most important steps in the development of digital photography (Ghosh, 2017). A CCD image sensor is made up of a grid of light sensitive silicon cells, called pixels (Felber, 2002). When these pixels are exposed to light, they generate electrical charge that is proportional to the intensity of the light hitting them. Using electronic technology, the charge can be measured precisely and then converted into a digital copy of the light patterns hitting the sensor. This invention paved the way for digital photography. The first commercially available digital camera was released in 1990 and although this camera produced low-resolution, black and white photos, it nonetheless came with a hefty USD1000 price tag (Trenholm, 2007). Unsurprisingly, the combination of low quality and high price proved an unsuccessful formula. Over the next few years, companies worked to address the quality issue but digital cameras remained unaffordable to the masses. The key turning point came when the digital photography industry shifted to use cheaper image sensor technology—specifically, the complementary metal-oxide semiconductor (CMOS) image sensor (Felber, 2002; Ghosh, 2017). As a result, by 2002, a 2.1 megapixel camera could be purchased for USD100. And in 2003, for the first time, digital cameras outsold film cameras, a trend that has continued ever since (Camera & Imaging Products Association, 2017).

The success of the stand-alone digital camera industry has been relatively short-lived; the industry reached its peak in 2010, producing 121 million cameras, but this number fell drastically to just 23 million in 2016 (Camera & Imaging Products Association, 2017; Djudjic, 2017; Hillen, 2017). Does this mean that people are taking fewer photos now than in the past? Far from it. Estimates suggest that 1 trillion photos were taken in 2015—nearly triple the number taken in 2010 (Lee, 2016). Furthermore, projections indicate that the number of photos captured each year will continue to grow exponentially (Heyman, 2015; Lee, 2016). Yet the stand-alone digital camera is no longer the preferred device for capturing photos—most people now take photos with a camera phone (Hillen, 2017; Lee & Stewart, 2016). What is more, modern camera phones are not only capable of capturing high-quality photos but also allow users to instantly share the photos they take (Larsen & Sandbye, 2014).

A picture is worth a thousand words, or is it?

Clearly, digital technology has revolutionised photography in the sense that people, arguably, have more photos than they know what to do with. Digitalisation has also revolutionised photography in another sense; photos have never been easier to manipulate. Image manipulation is not a recent phenomenon; in fact, it is almost as old as photography itself (Guilshan, 1992). The famous portrait of Abraham Lincoln, shown in Figure 1.1, is a fake; it was created circa 1860 by joining together a lithograph print taken from a daguerreotype of Lincoln with a photo of Southern politician John C. Calhoun (Farid, 2006). Stalin was also a fan of image manipulation, requesting the use of cumbersome darkroom techniques to airbrush his enemies out of photos, and Hitler did the same. It is even possible to find early examples of airbrushing for aesthetic purposes: in a trade manual for photographers dating back to 1875, two portraits of a woman are shown, one printed from an untouched negative and the other from a retouched negative (Sheehan, 2014). The retouching successfully gives the impression of a smoother, lighter complexion, fuller cheeks and well-rested eyes. Thus, image manipulation did exist before the advent of digital photography, however, it is important to bear in mind that making alterations to analogue photos was—and still is—a complicated and costly process that was the exception rather than the rule (Parry, 2009). Therefore it was fairly reasonable to believe in the integrity of analogue photos because, more often than not, these photos provided a truthful depiction of reality (Farid, 2009a).

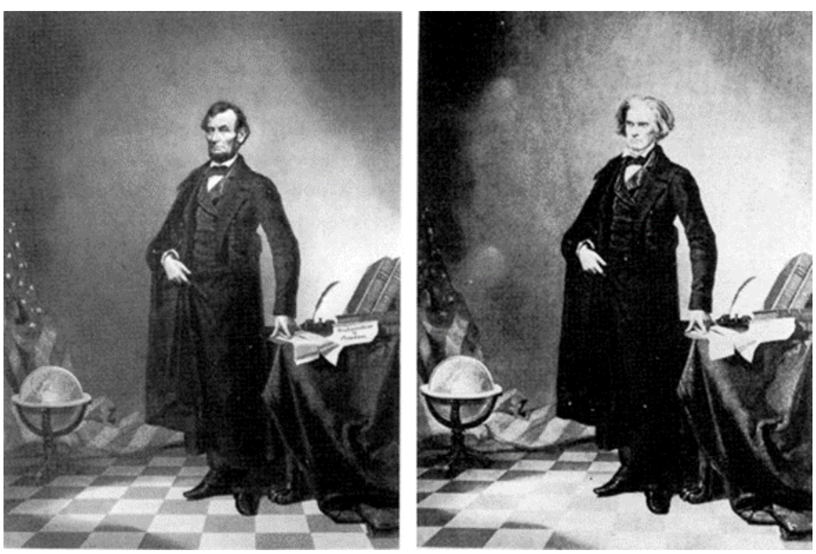


Figure 1.1. The portrait of Abraham Lincoln on the left is a fake, created by compositing Lincoln's head onto the photo of John C. Calhoun shown on the right. Photo source: Library of Congress.

Contrast this situation with that of today where digital technology allows photos to be manipulated in ways that were entirely impossible a few decades ago (Farid, 2009b). As image manipulation technology continues to develop, the creation of sophisticated and convincing forgeries will only become easier. Not only is this technology powerful, it is also more affordable than ever before. Nowadays, image manipulation software is available at little or no cost and allows users to create sophisticated alterations with ease—digital photos can be altered by almost anyone, it no longer necessarily takes skill or expertise (Parry, 2009). And even if these dedicated image manipulation software packages prove too complicated, most cameras, including camera phones, now come bundled with powerful manipulation tools that allow users to apply image enhancements at the click of a button. One of the most popular photo-editing phone apps, Facetune, allows users to reshape noses, whiten teeth, remove blemishes, perfect skin, and even add a smile (King, 2015). Furthermore, a newly purchased Samsung Galaxy S7 comes with a default beauty setting enabled, meaning that if the user does not change it, all photos are automatically altered to some extent (Cherrington, 2016).

As a result of the development of image manipulation technology, altering digital photos is now ubiquitous. News stories frequently reveal fraudulent photos or question the integrity of photos in the public domain (Tooth, 2011). What is more, these news stories reveal that image manipulation happens across a wide range of areas, including photojournalism, political campaigns, scientific publishing, sport, fashion, and music (e.g., Marszal, 2015; Pearson, 2006; Peters, 2010; Turner, 2016). These days, even casual photographers use image manipulation to their advantage. In one such case, shortly after Dinesh and Tarakeshwari Rathod were hailed as the first Indian couple to reach the summit of Mount Everest, the Nepalese authorities launched an official investigation to determine whether the pair had in fact manipulated their verification photos (Safi, 2016). The investigation revealed that the couple had superimposed their faces and the Indian flag onto another climber's photos. Likewise, there were many instances of fake photos circulating during the 2016 United States presidential campaign, including a manipulated image showing an immigration officer arresting an undocumented immigrant at a polling station (Garcia, 2016). Even after the conclusion of the presidential campaign, in March 2017, a crudely manipulated image depicting former President Barack Obama in handcuffs having been arrested for wiretapping

President Donald Trump circulated the web (Gillin, 2017). Despite the unsophisticated manipulation in the Obama example, the image was still shared thousands of times on social media.

In the digital age, people are inundated with photos, both authentic and manipulated. Digital photography is the result of a number of phenomenal technological achievements and is generally viewed as a positive development. But should we also be concerned by the changes brought about by digital photography? What impact, for example, does this constant exposure to photos have on human memory, cognition, and behaviour?

The effects of photos on memory and cognition

The constructive nature of memory

Remembering allows us to do something seemingly magical. We can relive our experiences by bringing the past into the present. In reality though, memory is fragile—decades of research has shown that memories are not an exact, unchanged record of the past (for a review, see Loftus, 2017). Rather, remembering involves a constructive process that is prone to a variety of errors and distortions (Schacter, 2013). As a result of this constructive process, memories can be drastically changed or created (Loftus, 2003). In a recent TED talk, Elizabeth Loftus (2013) proposed a new analogy for human memory, suggesting that, far from the misconception that memory acts as a recording device, it is instead more like an entry in Wikipedia—something that you and other people can change.

One of the earliest demonstrations of memory's constructive nature dates back to 1932 when Sir Frederic Bartlett showed that what people remember is not necessarily an accurate record of events, but rather is influenced by internal factors, specifically, people's expectations and beliefs. Bartlett's methods and findings were controversial at the time and largely dismissed (Johnston, 2001). Yet, the underlying idea that people's memories do not always perfectly match what actually happened was picked up decades later, and this time researchers were also interested in how external factors can influence memory.

A highly influential finding in psychology is the *misinformation effect* (Loftus, 1975; Loftus & Hoffman, 1989; Loftus, Miller, & Burns, 1978). This effect refers to the finding that exposure to misleading postevent information can influence a person's memory for an event. In the early misinformation effect studies, subjects witnessed an event, such as a traffic accident, and then later read descriptions of what they saw. Crucially, half of the subjects received misleading descriptions—for instance a stop sign in the original event was referred to as a yield sign in the description—while the other half received consistent descriptions. In a later test of their memory for the originally witnessed event, subjects given the misinformation tended to adopt that information as their memory—they were more likely to report that they originally saw a yield sign than subjects who were not given misinformation (Loftus et al., 1978). This finding demonstrates the alarming ease with which people can be misled about what they have seen and thus raises questions about the reliability of eyewitness testimony. Since this seminal study, many replications and extensions of the misinformation effect have been conducted (for a review, see Loftus, 2005).

Taking these misinformation effect findings a step further, in the 1990s memory scientists showed that people could construct false memories of entire events (e.g., Hyman, Husband, & Billings, 1995; Loftus & Pickrell, 1995; Porter, Yuille, & Lehman, 1999). One of the earliest attempts involved a relatively simple procedure in which adult subjects read narrative descriptions of four childhood events, three of which were genuine and one that was false. The false event, for instance, might describe how the subject got lost in a shopping mall for an extended period when they were about 5 years-old (Loftus & Pickrell, 1995). The researchers checked with the subjects' families to ensure that the subject had not experienced the false event as a child. After attending a series of suggestive interviews where subjects worked to remember details of all four events, 25% came to believe, wholly or partially, that as a child they got lost in a mall and were eventually helped by an elderly lady. This paradigm for introducing false memories of entire events is known as the familial-informant false-narrative procedure (Lindsay, Hagen, Read, Wade, & Garry, 2004) and has been used to explore the power of suggestive techniques and repeated retrieval. Thus, memory researchers have known for a long time that narrative descriptions of counterfactual childhood events can lead people to generate false memories for entire events. It was only recently, however, that scientists started to examine the power of photos to induce wholly false memories.

Photos and childhood memories

In the early 2000s, researchers adapted the familial-informant false-narrative procedure to determine whether doctored photos of childhood events might also lead people to develop false memories (Wade, Garry, Read, & Lindsay, 2002). Over a 1-2 week period, adult subjects worked at remembering moderately significant childhood events depicted in photos. Three of the photos were real, depicting one-off events like a special childhood birthday party, a school trip, or a cultural event, and one photo was a fake—a doctored image that depicted the subject taking a childhood hot air balloon ride. By the end of the study period, 50% of subjects claimed to have at least some memory of the hot air balloon ride even though their family confirmed this event never happened. Extending the lost in the mall research (Loftus & Pickrell, 1995), this finding shows that people can be led to remember wholly false childhood experiences using postevent information in the form of photos.

Critics have suggested, however, that subjects might not be constructing false memories, per se, but uncovering true memories of actual experiences. In response to such criticism, researchers have examined whether photos encourage people to develop false memories for implausible or impossible events (Braun, Ellis, & Loftus, 2002; Grinley, 2002; Strange, Sutherland, & Garry, 2006). In one study, subjects evaluated advertisements for a Disney resort (Braun et al., 2002). Half viewed a generic advert that did not show any cartoon characters, the other half viewed a fake advert that featured Bugs Bunny (an impossible scenario because Bugs Bunny is a Warner Brothers character, not Disney). After a delay, subjects were asked about their own experiences of visiting Disneyland as a child, and 16% of those who viewed the fake advert claimed that they remembered meeting Bugs Bunny on their trip.

Other research has shown that photos do not necessarily have to be manipulated in order to aid the construction of false memories: Authentic photos can facilitate the creation of false memories too (Garry, Strange, Bernstein, & Kinzett, 2005; Lindsay et al., 2004; Strange, Garry, Bernstein, & Lindsay, 2011). In one study, subjects attempted to “remember” a childhood event that never actually happened—putting a bright green gooey Slime toy in a teacher's desk in Grade 1 or 2 (aged 5 or 6 years-old) (Lindsay et al., 2004). Subjects worked to recall the event in one of two conditions: half of the subjects read a verbal description of the false event, the other half read the description and also viewed their real school class photo from the time. Just 23% of the description-only subjects stated that they remembered the false event, while 65% of subjects who also saw the class photo "remembered" putting Slime in a teacher's desk. Importantly, although the photo was related to the event it did not provide any evidence that the event really occurred, that is, the photo was non-probative. Yet it still inflated the likelihood that subjects "remembered" the event.

Can photos affect recent memories?

In addition, several other studies have shown that doctored images and videos are a powerful form of suggestion that can lead people to "remember" wholly false experiences, including false memories of recent events (Frenda, Knowles, Saletan, & Loftus, 2013; Nash & Wade, 2009; Nash, Wade, & Lindsay, 2009; Sacchi, Agnoli, & Loftus, 2007). For example, one study showed that photos can change people's memories of news and political events (Sacchi et al., 2007). Sacchi et al. manipulated two photos, one of the 1989 Tiananmen Square protest in Beijing, and another from the 2003 Iraq war protest in Rome. For the Beijing protest, the original photo showing a lone Chinese student facing the tank was manipulated to show crowds of spectators. For the Rome event the manipulations, for instance adding angry-looking protestors and police officers wearing riot gear, made the peaceful protest look more aggressive than it was in reality. After viewing a photo of each of the protests—one being the original version and the other being the manipulated version—subjects answered some questions about the two events. The results revealed that subjects who viewed the manipulated Beijing image estimated that more people had been involved in the event than those who viewed the original version. Similarly, subjects who viewed the manipulated Rome image described the protest as being more violent and negative than subjects who saw the original photo. Furthermore, subjects who saw the manipulated Rome image said they were less likely to participate in future protests than those who saw the original photo. These results indicate that visual information that contradicts the original event can change people's memories of the events, and also influence how they intend to act in the future.

In the largest false memory study to date, over 5,000 subjects viewed descriptions of three true and one (of five) fabricated political events (Frenda et al., 2013). The true event descriptions were shown alongside an authentic photo, while the fabricated event description was shown with an image that had been manipulated to portray the false event. One false event, for example, depicted President Obama shaking the hand of Iranian President Mahmoud Ahmadinejad. Overall, 50% of subjects reported that they remembered the fabricated event happening. Furthermore, of those subjects who said they remembered the false event, 27% stated that they even remembered seeing the false event happen on the news. Importantly, this study used a large, diverse sample and real-world political content, and the results converge with previous findings showing that photos can influence people's memories.

So far, the research discussed illustrates the incredible power of photos, both authentic and manipulated, to influence people's memories, beliefs, and behavioural intentions. An important question, however, concerns whether photos can also influence how people think they have acted in the past. Could photos make people believe they have behaved in a way they have not? In one study exploring this question, subjects performed some actions and imagined performing others (e.g., open the envelope) (Henkel, 2011). A week later subjects viewed photos of the actions in their completed state (e.g., an opened envelope). The photos included a mixture of the actions the subject performed and imagined performing, as well as new actions that were neither performed nor imagined. Another week later, subjects completed a surprise source memory test in which they had to determine which actions they had originally performed and which they had imagined. The results revealed that seeing the photos of the completed actions made subjects more likely to falsely claim that they had performed, rather than simply imagined performing, the action. Further, seeing photos of the new actions made subjects more likely to claim that they had originally performed or imagined performing the new actions. To rule out the possibility that the effect was caused by repeated exposure to the actions and thus increased familiarity, Henkel ran another study using a similar procedure. The results indicated that viewing a photo of an action once had a similar effect on memory errors to reading a textual description of the action four times. Therefore, photos can lead people to believe they have acted in ways that they have not.

Of course, remembering opening an envelope when in fact you only imagined doing so might seem mundane and inconsequential. Consider, though, that law enforcement officers sometimes use photos to cue suspects' and witnesses' memories in criminal investigations (Kassin et al., 2007). In the laboratory, research has demonstrated that witnesses are more likely to falsely identify an innocent person in a line-up if they viewed that person beforehand in a mug shot (Deffenbacher, Bornstein, & Penrod, 2006). Therefore, this research suggests that the use of such techniques could potentially lead someone to falsely remember performing or witnessing an action; an error that could have extremely serious consequences.

Photos and judgements

Not only do photos influence people's memories of past and recent events, they can also affect people's judgements about the truth of a claim (Newman, Garry, Bernstein, Kantner, & Lindsay, 2012). More specifically, subjects judged whether a series of trivia claims, for instance "Macadamia nuts are in the same evolutionary family as peaches", were true or false. Trivia statements accompanied by a non-probative photo (i.e., a photo that was related to the claim but did not reveal whether the claim was true or false) were more frequently judged to be true than statements presented without a photo, an effect called "*truthiness*" (Newman et al., 2012). Therefore, when making rapid judgments about the truth of a claim, the mere presence of a non-probative photo can lead people to believe that claim. This truthiness effect also persists over time; subjects returning to the lab 48-hours after making judgements about trivia statements were still more likely to judge the statements that had previously been presented with photos to be true (Fenn, Newman, Pezdek, & Garry, 2013). Perhaps what seems most surprising about these findings is that the photos are related to the event but provide no evidence that the event actually happened—photos can nudge people towards believing the claims are true, regardless of whether those claims are true or not. Non-probative photos can also nudge people to "remember" fabricated childhood events—recall the Slime prank (Lindsay et al., 2004). Taken together, this body of research shows that even non-probative photos can act as powerful cues that people will draw on to inform their judgements and memories.

Summary

In the digital age, there has been a phenomenal rise in the number of photos people capture, share, and manipulate—a trend that shows no sign of slowing. Research clearly demonstrates that photos—authentic and manipulated—are powerful; they can change people's memories for distant and recent experiences, beliefs about past actions, intentions for future actions, and judgements. Therefore, it is critically important to gain further understanding of the effects of digital photography on memory, cognition, and behaviour. Chapter 2 identifies some of the real-world implications of photography in the digital age and outlines the current program of research.

Chapter 2 :

Real-World Implications of Photo Manipulation in the Digital Age

"Photography allows us to uncritically think. We imagine that photographs provide a magic path to truth."

Errol Morris (2011)

Consequences of digitally manipulating images

The rise of photo manipulation has consequences across almost all domains, from law enforcement and national security, through to scientific publishing, politics, media, and advertising. As outlined in Chapter 1, photos are extremely persuasive. People have a misplaced trust in the integrity of photos—often placing too much faith in the reliability and accuracy of photos (Kelly & Nace, 1994; Parry, 2009; Sturken & Cartwright, 2009). Worryingly, however, it has never been easier for images to misrepresent the truth.

Photo evidence on trial

One reason to be concerned about the rise of photo manipulation is that photos are routinely used in legal proceedings. Photos can be used to explain or illustrate the testimony of a witness (Mnookin, 1998; Parry, 2009; Peterson, 2010); for example, a photo of the crime scene can help a witness to more precisely explain where and how the crime occurred. In addition, photos can be used as substantive evidence; that is, to prove the truth of a fact at issue. For example, a defendant's claim to have never met the victim could be disproved by photo evidence showing these two people together. To give another example, if a defendant testifies that they were at home on the evening of the crime, photo evidence showing them anywhere other than at home that evening disproves that aspect of their testimony. In the United States, the admissibility of photographic evidence has long been governed by the Federal Rules of Evidence (Robinson, 2013). When these rules were enacted in 1975, photos were presumed to be reliable sources of information for two main reasons. First, the complex and costly task of altering analogue photos meant that it was reasonable to assume that photos were, more often than not, truthful. Second, even if analogue photos were altered, these manipulations were relatively easy to detect (Parry, 2009). In the 1990s, however, the criminal justice system started to use digital photography and now digital has become the norm (Johnson, 2012). Digital photos are far easier to manipulate, and the resulting fakes can be extremely difficult to detect (Farid, 2006). Yet despite these important differences an advisory group to the Federal Rules of Evidence has concluded that changes to the rules are not necessary (Hannon, 2014; Johnson, 2012; Robinson, 2013). Therefore, digital photos are admissible as evidence on the same grounds as analogue photos (Facciola & Barrett, 2016; Federal Rules of Evidence, 1975; Galves, 2000; J. L. Moore, 2010; Parry, 2009; Paul, 2008).

Typically, then, admitting digital photos simply requires a witness to testify that the photo portrays the subject matter fairly and accurately (Federal Rules of Evidence, 1975; Peterson, 2010). When photos are used as substantive evidence, as opposed to illustrate witness testimony, the rules further stipulate that the photo should be the original—this is known as the best evidence rule (Federal Rules of Evidence, 1975; Parry, 2009). Of course, with analogue photography it is possible to use the negative to determine whether the photo is the original, but identifying the original version of a digital photo can be extremely difficult. As such, the best evidence rule is outdated and, even for substantive photo evidence, witness testimony often provides the basis for determining whether a photo is admitted into evidence or not (Parry, 2009).

Furthermore, courts will admit digitally enhanced photos into evidence providing that there is adequate expert testimony concerning the digital enhancement process (Guilsham, 1992; Parry, 2009). Essentially, someone who is considered to hold expert knowledge about the program used to enhance the photo is required to testify about the method of enhancement and to confirm that the enhancement has not, intentionally or through negligence, resulted in the manipulation of data (Guthrie & Mitchell, 2007). Therefore, digital photos are usually admissible as evidence providing that either a lay or expert witness testifies that the photo offers a fair and accurate representation of the subject matter. These reasonably general and relaxed criteria concerning the admission of photographic evidence might have been sufficient in the analogue age when it was reasonable to believe in the integrity of images (Mnookin, 1998; Parry, 2009). In the digital age, however, the integrity of images has been eroded due to the ease of photo manipulation and these guidelines may no longer be viable (e.g., Facciola & Barrett, 2016; Farid, 2009b; Parry, 2009; Porter, 2014). Indeed, the attempt to simply assimilate digital within the existing rules for analogue photo evidence has inevitably put the courts at a heightened risk of fraud (Mnookin, 1998; Paul, 2004; Porter, 2014).

Worryingly, the availability of powerful and affordable digital-editing tools means that nearly anyone can create a manipulation that is difficult to detect. For instance, with only a small amount of practice using manipulation software, a layperson can remove a bruise or scar, change a license plate, or even add a person to a scene (Farid, 2008). Furthermore, with this ease of creating sophisticated digital fakes, even honest witnesses might inadvertently testify to the authenticity of a photo, unaware that it has been changed (Parry, 2009). This scenario is not that farfetched given that witnesses often testify about the information captured in photos months, or even years, after a crime. Therefore, a number of legal scholars have suggested that the standards set for authentication are not sufficient to prevent manipulated photos becoming evidence that a jury will trust (e.g., Mnookin, 1998; Parry, 2009; Porter, 2014). Despite the danger of admitting fraudulent digital photos into evidence, lawyers and courts very rarely challenge them. One possible explanation is a lack of awareness in the legal community regarding the features of digital photos that make them less reliable than analogue photos; regardless of the format, photos are still considered inherently trustworthy and objective (Mnookin, 1998; Sturken & Cartwright, 2009). Other than running the risk of photo forgeries being used unfairly as evidence in legal cases, this lack of challenge to the way that digital photos are admitted as evidence means that the issues perpetuate. It is possible, then, that the inappropriate use of digital photos could contribute to wrongful convictions.

The legal case of Alfred Swinton illustrates the potential dangers of having lax criteria for admitting digital photos in court and having judges and jurors who are too willing to trust photos. In 2001, Swinton was sentenced to 60 years in prison for the murder of Carla Terry (Baron, 2008; *Connecticut v. Swinton,* 2004). Terry was murdered a decade earlier in 1991, and although Swinton was arrested at the time, the charges were dismissed based on insufficient evidence (*Connecticut v. Swinton,* 2004). By the end of the 1990s, technology had advanced and a new image-enhancing software, Lucis Pro, became available to Connecticut's State Police Forensic Science Lab (Baron, 2008).

The prosecution used this software to enhance photos of bite marks on the victim's body (*Connecticut v. Swinton*, 2004). With these photos now clearly showing the victim's bite marks, the prosecution were able to go a step further and overlay Swinton's bite pattern using Adobe Photoshop® software. Dr Constantine Karazulas, a forensic odontologist, presented this new photo evidence to corroborate his testimony that the defendant's bite pattern was a match for the bite marks on the victim's body. With this new evidence, Swinton was convicted.

Swinton maintained his innocence and in 2004 appealed to the Connecticut Supreme Court contesting, primarily, the use of the new digitally enhanced photo evidence (*Connecticut v. Swinton*, 2004; Guthrie & Mitchell, 2007). To adhere to the Federal Rules of Evidence, enhanced photos must be corroborated through expert testimony to assure the court about the adequacy and reliability of the digital enhancement process (Guilsham, 1992; Parry, 2009). The defence stated that such testimony was not provided and thus the digitally enhanced photo evidence had been improperly admitted as evidence. The Court ruled that there was sufficient expert testimony given in support of the Lucis Pro enhanced photos of the victim's bite marks but not for the overlay photos created using Adobe Photoshop (*Connecticut v. Swinton*, 2004; Guthrie & Mitchell, 2007). Therefore the Court held that the trial court had acted improperly when admitting these overlay photos. That said, the Court also determined that the defendant had been able to effectively discredit the use of these photos through the testimony of his own expert witness. In conclusion, the Court considered the improper admission of these photos had been harmless and rejected the defendant's appeal (*Connecticut v. Swinton*, 2004; Guthrie & Mitchell, 2007).

At the time of writing the Swinton case has again made the news (Altimari & Owens, 2017). New DNA testing indicates that Swinton is highly unlikely to be the source of the bites. Furthermore, the prosecution's forensic odontologist, Karazulas, has recanted his original testimony, admitting in a signed affidavit that many thousands of individuals could be responsible for the bite marks (Altimari & Owens, 2017). The Innocence Project has presented a petition for a new trial and are seeking to overturn Swinton's murder conviction.

Although it remains possible that Swinton is guilty of murdering Terry, new DNA evidence suggests it is near impossible that Swinton was the source of the bite marks.

And yet it was the digitally altered photos of these bite marks that provided the necessary evidence to take the Swinton case to trial. As outlined in Chapter 1, research shows that photos do not need to offer evidence of an event or fact to affect people's judgements (Lindsay et al., 2004; Newman et al., 2012). In this case, it is possible that the digitally altered photos biased the jury to believe the claims made by Karazulas. As the Supreme Court pointed out, the defence did have an opportunity to discredit Karazulas's testimony yet when jurors are given conflicting claims about the truth, photos often help them to judge where the truth lies (Newman & Feigenson, 2013; Parry, 2009). Even when photos are used as an illustration, to clarify something rather than to corroborate it, they can reinforce beliefs, or even persuade a jury to believe in the truth of the statements they accompany (Parry, 2009).

The Swinton case also demonstrates the ease with which courts will admit digitally altered photos. What does this mean for photos that are manipulated intentionally to deceive the court? Would those involved in the case be able to determine the authenticity of the photo? Assuming people are unable to detect that these photos have been altered, it suggests that they too will act as powerful, albeit fraudulent, cues for the jury to make judgements of truth in a court of law. Thus it is important to understand whether people are able to identify photo forgeries.

Digital photo manipulation and well-being

Another reason to be concerned about photo manipulation is that it has become standard for advertisers and magazine editors to use digital photo-editing software to alter a person's appearance (Kee & Farid, 2011). On the one hand, these alterations can include relatively minor retouching, for instance whitening teeth, and removing a few wrinkles and blemishes. Or the alterations can be more extreme and incorporate liberal changes to body shape, for instance lengthening legs, and trimming the waist and stomach (Sheehan, 2014). Continual exposure to photos in the media of impossibly thin, tall, wrinkle- and blemish-free models and celebrities contributes to the formation of unrealistic beauty goals and aspirations such as the creation of a thin-body ideal (Fallon, 1990; Heinberg, 1996; Morry & Staska, 2001; Owen & Spencer, 2013; Stice, Schupak-Neuberg, Shaw, & Stein, 1994; Stice & Shaw, 1994). A body of research now suggests a link exists between the internalisation of unattainable physical beauty and body dissatisfaction, negative affect, or even eating disorders (e.g., Grabe, Ward, & Hyde, 2008; Groesz, Levine, & Murnen, 2002; Thompson & Heinberg, 1999; Thompson & Stice, 2001). Some researchers have proposed that the media's portrayal of the ideal appearance sets difficult-to-achieve, if not impossible, standards of physical beauty and thinness, thereby creating a discrepancy between expectation and reality (Thompson, Heinberg, Altabe, & Tantleff-Dunn, 1999; Tiggemann & McGill, 2004). It is this discrepancy that can lead to psychological problems as well as put people at risk of engaging in dangerous eating and exercise behaviour.

Unfortunately, the problem of setting impossible standards for beauty has been exacerbated by the fact that the magic of digital retouching is no longer reserved for those appearing in the fashion, entertainment, and advertising industries—literally anyone can airbrush a photo (Sheehan, 2014), and they regularly do. In fact, it has become so normal to airbrush photos that in 2012 Jesse Rosten, a TV-commercial director, created a parody commercial for a beauty product dubbed *Fotoshop by Adobé.* The commercial mocks the ease with which digital software can be used to implement a variety of techniques to drastically transform appearance (Sheehan, 2014). The parody clearly struck a chord with the general public, reaching 5 million views in just a few days. Although Rosten's parody commercial takes a sarcastic and humorous view of extreme digital retouching, we know that such practices can have extremely serious consequences. People are increasingly bombarded with airbrushed photos, not only of a specific group of celebrities and models, but also of their friends and family (Sheehan, 2014). The barrage of photos depicting an idealised standard of beauty that is unattainable, unhealthy, and of course unreal, is ever-growing and reinforces the perception of a discrepancy between expectation and reality. Ultimately, then, people are judging themselves against something that is not real, against a digitally created ideal.

Evidence suggests that these fake photos carry meaning and influence people both psychologically and physically. This issue of continually subjecting people to idealised standards of beauty is widely recognised. A handful of celebrities are rejecting digital retouching in an attempt to promote positive body image messages to young girls and boys around the world (Baron, 2008). For example, when GQ featured an excessively airbrushed photo of Kate Winslet on its cover the actor famously responded to affirm that she does not look like the photo of her, and nor does she desire to (Baron, 2008;

Sheehan, 2014). More recently, Winslet signed a new contract with global beauty brand L'Oréal but stipulated a clause to prevent airbrushing being used on any of her photos (Saul, 2015). Taking a similar stance, Dame Helen Mirren refused to allow airbrushed photos of her to be used in her first campaign for L'Oréal (Young & Silverman, 2014).

The pervasiveness of digital retouching to alter a person's appearance has also led to the implementation of more formalised interventions (Gladstone, 2016). In 2012, Israeli lawmakers adopted the Weight Restriction Law in an attempt to diminish the unrealistic standards of physical beauty in two ways; the body mass index (BMI)-based hiring requirement and the advertisement-labelling requirement (Levush, 2012). Of particular relevance here is the advertisement-labelling requirement which focuses on digitally manipulated images and stipulates that any image showing a model whose appearance has been altered must feature a clear and prominent disclaimer label to indicate this is the case. Similar laws have been passed in Italy, Spain, and most recently France ("France Bans Extremely Thin Models," 2017; Wallwork, 2015). The United Kingdom are also considering such legislation (Clark, 2016; Hooton, 2016).

Many applaud these legislative movements to challenge the permeation of unrealistic standards of physical beauty, however, the changes are not without criticism (Gladstone, 2016). Advertisers and publishers continue to resist any such legislation and criticise the limitations it places on free expression and artistic freedom (Kee & Farid, 2011; Turley, 2012). Yet, criticism aside, an important outstanding question is whether the strategy of using disclaimer labels to indicate to viewers that the photos have been retouched actually works. Research does suggest that viewing healthy weight models—as in the Dove Campaign for Real Beauty—results in a more healthy body ideal and more positive affect than viewing very thin models (Owen & Spencer, 2013). Furthermore, there is some evidence to suggest that women at risk of developing eating disorders are less likely to act on diet urges after viewing images of healthy weight models as opposed to underweight models (Fister & Smith, 2004). Yet simply adding a label to a photo to say it has been digitally altered does not prevent people from viewing a model who is impossibly thin or attractive. In addition, the labels do not differentiate between minor alterations and modifications that dramatically alter a person's appearance (Kee & Farid, 2011). Therefore, is a disclaimer label enough? Recent findings suggest not. Several studies have shown that adding disclaimer labels to indicate an image as altered failed to offer protective effects across various measures including body dissatisfaction, negative affect, or intent to diet (Ata, Thompson, & Small, 2013; Bissell, 2006; Harrison & Hefner, 2014; Tiggemann, Slater, Bury, Hawkins, & Firth, 2013). Moreover, in a recent study, researchers found that exposure to a thin-body ideal image with a disclaimer label that used the specific wording suggested by the French law—"This image has been altered to modify a person's bodily appearance."—resulted in increased accessibility to negative thoughts (Selimbegović & Chatard, 2015). Some researchers have hypothesised that disclaimer labels might actually encourage viewers to direct more, rather than less, visual attention to the model's body than they normally would (Selimbegović & Chatard, 2015; Tiggemann et al., 2013). These results suggest that using disclaimer labels on airbrushed photos might have counter-productive effects and accentuate the precise problems they were intended to address.

It is encouraging that steps are being taken to address the negative impact of the ubiquity of idealised and unrealistic representations of physical beauty. It is unfortunate, however, that the implemented "solutions" have not been formulated on the basis of sound empirical research. With several countries choosing to enact disclaimer-related legislation, it is important to note that doing so might have undesirable consequences and act to reinforce the very problems the legislation was intended to mitigate. As it stands, there is a real risk of inadvertently accentuating the problem rather than improving it. The lack of empirical research showing positive effects of disclaimer labels highlights an urgent need for a more comprehensive research program. One area of research that has not been explored is whether people can in fact identify when photos have been airbrushed. A better understanding of how people perceive digitally altered photos will be helpful to inform and guide policy makers and legislators towards the most effective forms of intervention.

Consequences of taking so many photos

Although the frequency with which photos are altered has grown phenomenally in the digital age, the actual practice of manipulating images is nearly as old as photography itself. A relatively recent development, however, is people's tendency to take so many photos. Digital technology has revolutionised the ease with which people capture photos (Whittaker, Bergman, & Clough, 2010). Given the remarkable rise in the number of photos that people take, scientific research has recently begun to explore the ways that taking so many photos might be affecting people.

Of particular interest is the question of how the act of taking photos might affect people's memory. One popular suggestion is that people take photos to help them to remember—in the future the photos can act as a memory trigger to help bring to mind past experiences (e.g., Chalfen, 1987; Harrison, 2002; Whittaker et al., 2010). Others have suggested, however, that taking photos might in fact do the opposite; that is, taking photos might impair people's ability to remember the past because their attention is focused on capturing their experiences on camera rather than on the experience per se (Mols, Broekhuijsen, van den Hoven, Markopoulos, & Eggen, 2015). In support of the latter suggestion, a recent study revealed that the act of taking photos can impair people's memory for the photographed content (Henkel, 2014). The new effect, termed the *photo-taking impairment effect* is an interesting finding and has great importance given that the number of photos taken continues to increase exponentially each year (Heyman, 2015). Of course, the question of how photography can influence human memory has only been an important consideration since people began to take so many photos. Therefore, there are many outstanding empirical questions to explore. Further research is required to help gain a better understanding of why and how taking photos can impair memory and also to more generally advance theoretical knowledge about the workings of human memory.

Thesis aims and outline

This thesis comprises of two parts and aims to extend current understanding of the effect of digital photography on memory and cognition. Although it is likely that people regularly make judgements about image authenticity, there is surprisingly little research looking at people's ability to perceive image manipulation. Further, it is evident that in some situations incorrectly accepting a manipulated image as authentic can have extremely serious consequences. Therefore, Part One of this thesis comprises of a program of research that examines people's ability to discriminate between authentic and manipulated images. Part Two of this thesis examines how the act of taking photos can affect people's memory.

The main aims of this thesis are to:

1. Examine the extent to which people are able to discriminate between authentic and manipulated photos of real-world scenes.
2. Explore people's ability to make use of two image properties—shadows and reflections—to help distinguish between authentic images and manipulated ones.
3. Investigate how taking photos affects people's memory of the photographed content.

Part One comprises of Chapters 3 to 5. Chapter 3 examines an important, yet previously unexplored, question: can people detect and locate manipulations in photos of real-world scenes? Chapters 4 and 5 then examine whether people can make use of information from two image properties—shadows and reflections—to help to distinguish between authentic and manipulated images. In Part Two of this thesis, Chapters 6 and 7 further explore people's recent tendency to take so many photos and the potential effect this has on memory. Finally, Chapter 8 presents a general discussion of these seven chapters, framing the findings within the context of the wider literature, identifying possible limitations, and outlining possible areas for further research.

Part One

Chapter 3 :

Can people identify original and manipulated photos of real-world scenes

"... faking has enjoyed a quantum leap with the advent of computerized manipulation. Now, with digital cameras, there is no "original" to compare... Fraudulent practice is easy and detection difficult, and photography will never be the same again."

Philip Jones Griffiths (n.d.)

Introduction

In 2015, one of the world's most prestigious photojournalism events—The World Press Photo Contest—was shrouded in controversy following the disqualification of 22 entrants, including an overall prize winner, for manipulating their photo entries. News of the disqualifications led to a heated public debate about the role of photo manipulation in photojournalism. World Press Photo responded by issuing a new code of ethics for the forthcoming contest that stipulated entrants "must ensure their pictures provide an accurate and fair representation of the scene they witnessed so the audience is not misled" (World Press Photo, n.d.). They also introduced new safeguards for detecting manipulated images, including a computerised photo-verification test for entries reaching the penultimate round of the competition. The need for such a verification process highlights the difficulties competition organisers face in trying to authenticate images. If photography experts cannot spot manipulated images, what hope is there for amateur photographers or other consumers of photographic images? This is the question we aimed to answer. That is, to what extent can lay people distinguish authentic photos from fakes?

Digital image and manipulation technology has surged in the previous decades. People are taking more photos than ever before. Estimates suggested that one trillion photos would be taken in 2015 alone (Worthington, 2014), and that, on average, more than 350 million photos per day are uploaded to Facebook—that is over 14 million photos per hour or 4000 photos per second (Smith, 2013). Coinciding with this increased popularity of photos is the increasing frequency with which they are being manipulated. Although it is difficult to estimate the prevalence of photo manipulation, a recent global survey of photojournalists found that 76% regard photo manipulation as a serious problem, 51% claim to always or often enhance in-camera or RAW (i.e., unprocessed) files, and 25% admit that they, at least sometimes, alter the content of photos (Hadland, Campbell, & Lambert, 2015). Together these findings suggest that we are regularly exposed to a mix of real and fake images.

The prevalence and popularity of manipulated images raises two important questions. First, to what extent do manipulated images alter our thinking about the past? We know that images can have a powerful influence on our memories, beliefs, and behaviour (e.g., Newman et al., 2012; Wade et al., 2002; Wade, Green, & Nash, 2010). Merely viewing a doctored photo and attempting to recall the event it depicts can lead people to remember wholly false experiences, such as taking a childhood hot air balloon ride or meeting the Warner Brothers character Bugs Bunny at Disneyland (Braun et al., 2002; Sacchi et al., 2007; Strange et al., 2006). Thus, if people cannot differentiate between real and fake details in photos, manipulations could frequently alter what we believe and remember.

Second, to what extent should photos be admissible as evidence in court? Laws governing the use of photographic evidence in legal cases, such as the Federal Rules of Evidence (1975), have not kept up with digital change (Parry, 2009). Photos were once difficult to manipulate; the process was complex, laborious, and required expertise. Yet in the digital age, even novices can use sophisticated image-editing software to create detailed and compelling fake images. The Federal Rules of Evidence state that the content of a photo can be proven if a witness confirms it is fair and accurate. Put another way, the person who took the photo, any person who subsequently handles it, or any person present when the photo was taken, is not required to testify about the authenticity of the photo. If people cannot distinguish between original and fake photos, then litigants might use manipulated images to intentionally deceive the court, or even testify about images, unaware they have been changed.

Unfortunately, there is no simple solution to prevent people from being fooled by manipulated photos in everyday life or in the criminal arena (Parry, 2009). But the newly emerging field of image forensics is making it possible to better protect against photo fraud (e.g., Farid, 2006). Image forensics uses digital technology to determine image authenticity, and is based on the premise that digital manipulation alters the values of the pixels that make up an image. Put simply, the act of manipulating a photo leaves behind a trace, even if only subtle and not visible to the naked eye (Farid, 2009a). Given that different types of manipulations—for instance, cloning, retouching, splicing—affect the underlying pixels in unique and systematic ways, image forensic experts can develop computer methods to reveal image forgeries. Such technological developments are being implemented in several domains, including law, photojournalism, and scientific publishing (Oosterhoff, 2015). The vast majority of image authenticity judgments, however, are still made by eye, and to our knowledge only one published study has explored the extent to which people can detect inconsistencies in images.

Farid and Bravo (2010) investigated how well people can make use of three cues—shadows, reflections, and perspective distortion—that are often indicative of photo tampering. The researchers created a series of computer-generated scenes consisting of basic geometrical shapes. Some scenes, for instance, were consistent with a single light source whereas others were inconsistent with a single light source. When the inconsistencies were obvious, that is, when shadows ran in opposite directions, observers were able to identify tampering with nearly 100% accuracy. Yet when the inconsistencies were subtle, for instance, where the shadows were a combination of results from two different light positions on the same side of the room, observers performed only slightly better than chance. These preliminary findings, based on computer-generated scenes of geometric objects, suggest that the human visual system is poor at identifying inconsistencies in such images.

In the current study we examined whether people are similarly poor at detecting inconsistencies within images of real-world scenes. On the one hand, we might expect people to perform even worse if trying to detect manipulations in real-world photos. Research shows that real-world photos typically contain many multi-element objects that can obscure distortions (Bex, 2010; Hulleman & Olivers, 2015). For example, people with the visual impairment metamorphopsia often do not notice any problems with their vision in their everyday experiences, yet the impairment is quite apparent when they view simple stimuli, such as a grid of evenly spaced horizontal and vertical lines (Amsler, 1953; Bouwens & Meurs, 2003). We also know that people find it more difficult to detect certain types of distortions, such as changes to image contrast, in complex real-world scenes than in more simplistic stimuli (Bex, 2010; Bex, Solomon, & Dakin, 2009). In sum, if people find it particularly difficult to detect manipulations in complex real-world scenes, then we might expect our subjects to perform worse than Farid and Bravo's (2010) subjects.

On the other hand, there is good reason to predict that people might do well at detecting manipulations in real-world scenes. Visual cognition research suggests that people might detect image manipulations using their knowledge of the typical appearance of real-world scenes. Real-world scenes share common properties, such as the way the luminance values of the pixels are organised and structured (Barlow, 1961; Gardner-Medwin & Barlow, 2001; Olshausen & Field, 2000). Over time, the human visual system has become attuned to such statistical regularities and has expectations about how scenes should look. When an image is manipulated, the structure of the image properties change, which can create a mismatch between what people see and what they expect to see (Craik, 1943; Friston, 2005; Rao & Ballard, 1999; Tolman, 1948). Thus, based on this real-world scene statistics account, we might predict that people should be able to use this "mismatch" as a cue to detecting a manipulation. If so, our subjects should perform better than chance at detecting manipulations in real-world scenes.

Although there is a lack of research directly investigating the applied question of people's ability to detect photo forgeries, people's ability to detect change in a scene is well-studied in the field of visual cognition. Notably, change blindness is the striking finding that, in some situations, people are surprisingly slow, or entirely unable, to detect changes made to, or find differences between, two scenes (e.g., Pashler, 1988; Simons, 1996; Simons & Levin, 1997). In some of the early studies, researchers demonstrated observers' inability to detect changes made to a scene during an eye movement (saccade) using very simple stimuli (e.g., Wallach & Lewis, 1966), and later, in complex real-world scenes (e.g., Grimes, 1996). Researchers have also shown that change blindness occurs even when the eyes are fixated on the scene: the flicker paradigm, for instance, simulates the effects of a saccade or eye blink by inserting a blank screen between the continuous and sequential presentation of an original and changed image

(Rensink, O'Regan, & Clark, 1997). It often requires a large number of alternations between the two images before the change can be identified. Furthermore, change blindness persists when the original and changed images are shown side by side (Scott-Brown, Baker, & Orbach, 2000), when change is masked by a camera cut in motion pictures (Levin & Simons, 1997), and even when change occurs in real-world situations (Simons & Levin, 1998).

Such striking failures of perception suggest that people do not automatically form a complete and detailed visual representation of a scene in memory. Therefore, to detect change, it might be necessary to draw effortful, focused attention to the changed aspect (Simons & Levin, 1998). So which aspects of a scene are most likely to gain focused attention? One suggestion is that attention is guided by salience; the more salient aspects of a scene attract attention and are represented more precisely than the less salient ones. In support of this idea, research has shown that changes to more important objects are more readily detected than changes made to less important objects (Rensink et al., 1997). Other findings, however, indicate that observers sometimes miss even large changes to central aspects of a scene (Simons & Levin, 1998). Therefore, the question of what determines scene saliency continues to be explored. Specifically, researchers disagree about whether the low-level visual salience of objects in a scene, such as brightness (e.g., Lansdale, Underwood, & Davies, 2010; Pringle, Irwin, Kramer, & Atchley, 2001; Spotorno & Faure, 2011) or the high-level semantic meaning of the scene (Stirk & Underwood, 2007) has the most influence on attentional allocation.

What other factors affect people's susceptibility to change blindness? One robust finding in the signal detection literature is that the ability to make accurate perceptual decisions is related to the strength of the signal and the amount of noise (Green & Swets, 1966). Signal detection theory has been applied to change detection. In one study, observers judged whether two sequentially presented arrays of coloured dots remained identical or if there was a change (Wilken & Ma, 2004). Crucially, the researchers manipulated the strength of the signal in the change trials by varying the number of coloured dots in the display that changed, while noise (total set size) remained constant. Performance improved as a function of the number of dots in the display that changed colour—put simply, greater signal resulted in greater change detection.

Given the lack of research investigating people's ability to detect photo forgeries, change blindness offers a highly relevant area of research. A key difference between the change blindness research and our current experiments, however, is that our change detection task does not involve a comparison of two images; therefore, representing the scene in memory is not a factor in our research. That is, subjects do not compare the original and manipulated versions of an image. Instead, they make their judgment based on viewing only a single image. This image is either the original, unaltered image or an image that has been manipulated in some way.

In the current study, we explored people's ability to identify common types of image manipulations that are frequently applied to real-world photos. We distinguished between physically implausible versus plausible manipulations. For example, a physically implausible image might depict an outdoor scene lit only by the sun with a person's shadow running one way and a car's shadow running the other way. Such shadows imply the impossible: two suns. Alternatively, when an unfamiliar face is retouched in an image it is quite plausible; eliminating spots and wrinkles or whitening teeth do not contradict physical constraints in the world that govern how faces ought to look. In our study, geometrical and shadow manipulations made up our implausible manipulation category, while airbrushing and addition or subtraction manipulations made up our plausible manipulation category. Our fifth manipulation type, super-additive, presented all four manipulation types in a single image and thus included both categories of manipulation.

We had a number of predictions about people's ability to detect and locate manipulations in real-world photos. We expected the type of manipulation—implausible versus plausible—to affect people's ability to detect and locate manipulations. In particular, people should correctly identify more of the physically implausible manipulations than the physically plausible manipulations given the availability of evidence within the photo. We also expected people to be better at correctly detecting and locating manipulations that caused more change to the pixels in the photo than manipulations that caused less change.

Experiment 1

Method

Subjects and design

A total of 707 subjects ($M = 25.8$ years, $SD = 8.8$, range = 14–82; 460 male, 226 female, 21 declined to respond) completed the task online. A further 17 subjects were excluded from the analyses because they had missing response time data for at least one response on the detection or location task. There were no geographical restrictions and subjects did not receive payment for taking part, but they did receive feedback on their performance at the end of the task. Subject recruitment stopped when we reached at least 100 responses per photo. We used a within-subjects design in which each person viewed a series of ten photos, half of which had one of five manipulation types applied, and half of which were original, non-manipulated photos. We measured people's accuracy in determining whether a photo had been manipulated or not and their ability to locate manipulations.

Stimuli

We obtained ten coloured images (JPEG format), 1600 × 1200 pixels, that depicted people in real-world scenes from Google Image search (permitted for non-commercial re-use with modification). We used GNU Image Manipulation Program® (GIMP, Version 2.8) to apply five different, commonly used manipulation techniques: (a) airbrushing, (b) addition or subtraction, (c) geometrical inconsistency, (d) shadow inconsistency, and (e) super-additive (manipulations a to d included within a single image). For the airbrushing technique, we changed the person's appearance by whitening their teeth, removing spots, wrinkles, or sweat, or brightening their eye colour. For the addition or subtraction technique, we added or removed objects, or parts of objects. For example, we removed links between tower columns on a suspension bridge and inserted a boat into a river scene. For geometrical inconsistencies, we created physically implausible changes, such as distorting angles of buildings or sheering trees in different directions to others to indicate inconsistent wind direction. For shadow inconsistencies, we removed or changed the direction of a shadow to make it incompatible with the remaining shadows in the scene. For instance, flipping a person's face around the vertical axis causes the shadow to appear on the wrong side compared with the rest of the body and scene. For the super-additive technique we presented all four previously described manipulation types in one photo. Figure 3.1 shows examples of the five manipulation types.

In total, we had ten photos of different real-world scenes. The non-manipulated version of each of these ten photos was used to create our original photo set. To generate the manipulated photos, we applied each of the five manipulation types to six of the ten photos, creating six versions of each manipulation for a total of 30 manipulated photos. This gave us an overall set of 40 photos. Subjects saw each of the five manipulation types and five original images but always on a different photo.

Image-based saliency cues can determine where subjects direct their attention; thus, we checked whether our manipulations had changed the salience of the manipulated area within the image. To examine this, we ran the images through two independent saliency models: the classic Itti-Koch model (Itti & Koch, 2000; Itti, Koch, & Niebur, 1998) and the Graph-Based Visual Saliency (GBVS) model (Harel, Koch, & Perona, 2006). To summarise, we found that our manipulations did not inadvertently change the salience of the manipulated regions. See Appendix A for details of these analyses.

Figure 3.1. Samples of manipulated photos. (a) Original photo; (b) airbrushing—removal of sweat on the nose, cheeks, and chin, and removal of wrinkles around the eyes; (c) addition or subtraction—two links between the columns of the tower of the suspension bridge removed; (d) geometrical inconsistency—top of the bridge is sheered at an angle inconsistent with the rest of the bridge; (e) shadow inconsistency—face is flipped around the vertical axis so that the light is on the wrong side of the face compared with lighting in the rest of the scene; (f) super-additive—combination of all previously described manipulations. Original photo credit: Vin Cox, CC BY-SA 3.0 license. Photos b–f are derivatives of the original and licensed under CC BY-SA 4.0.

Procedure

Subjects answered questions about their demographics, attitudes towards image manipulation, and experiences of taking and manipulating photos. Subjects were then shown a practice photo and instructed to adjust their browser zoom level so that the full image was visible. Next, subjects were presented with ten photos in a random order and they had an unlimited amount of time to view and respond to each photo. We first measured subjects' ability to detect whether each photo had been manipulated by asking "Do you think this photograph has been digitally altered?" Subjects were given three response options: (a) "Yes, and I can see exactly where the digital alteration has been made"; (b) "Yes, but I cannot see specifically what has been digitally altered"; or (c) "No." For the manipulated photos, we considered either of the "yes" responses as correct; for original photos we considered "no" as correct. Following a "yes" response, we immediately measured subjects' ability to locate the manipulation by presenting the same photo again with a 3×3 grid overlaid[1] (see Figure 3.2 for an example). Subjects were asked to: "Please select the box that you believe contains the digitally altered area of the photograph (if you believe that more than one region contains digital alteration, please select the one you feel contains the majority of the change)." On average, manipulations spanned two regions in the grid. For the analyses we considered a response to be correct if the subject clicked on a region that contained any of the manipulated area or a nearby area that could be used as evidence that a manipulation had taken place—a relatively liberal criterion. Subjects received feedback on their performance at the end of the study.

[1] In Experiment 2, subjects attempted to localise the manipulation regardless of their response in the detection task.

Figure 3.2. Example of a photo with the location grid overlaid. Photo credit: Vin Cox, CC BY-SA 3.0 license.

Results and discussion

An analysis of the response time data suggested that subjects were engaged with the task and spent a reasonable amount of time determining which photos were authentic. In the detection task, the mean response time per photo was 43.8 s ($SD = 73.3$ s) and the median response time 30.4 s (interquartile range 21.4, 47.7 s). In the location task, the mean response time was 10.5 s ($SD = 5.7$ s) and the median response time 9.1 s (interquartile range 6.5, 13.1 s). Following Cumming's (2012) recommendations, we present our findings in line with the estimation approach by calculating a precise estimate of the actual size of the effects.

Overall accuracy on the detection task and the location task

We now turn to our primary research question: To what extent can people detect and locate manipulations of real-world photos? For the detection task, we collapsed across the two "yes" response options such that if subjects responded either "Yes, and I can see exactly where the digital alteration has been made" or "Yes, but I cannot see specifically what has been digitally altered", then we considered this to be a "yes" response. Thus, chance performance was 50%. Overall performance on the detection task was better than chance; a mean 66% of the photos were correctly classified as original or manipulated, 95% confidence interval (CI)[2] [65%, 67%]. Subjects' ability to distinguish between original (72% correct) and manipulated (60% correct) photos of real-world scenes was reliably greater than zero, discrimination (d') = 0.80, 95% CI [0.74, 0.85]. Moreover, subjects showed a bias towards saying that photos were real; response bias (c) = 0.16, 95% CI [0.12, 0.19]. Although subjects' ability to detect manipulated images was above chance, it was still far from perfect. Furthermore, even when subjects correctly indicated that a photo had been manipulated, they could not necessarily locate the manipulation. Collapsing over all manipulation types, a mean 45% of the manipulations were accurately located, 95% CI [43%, 46%]. To determine chance performance in the location task, we need to take into account that subjects were asked to select one of nine regions of the image. Therefore, subjects had less chance of being correct by guessing in the location task than the detection task. On average, the manipulations were contained within two of the nine regions. But because the chance of being correct by guessing varied for each image and each manipulation type, we ran a Monte Carlo simulation to determine the chance rate of selecting the correct region. Table 3.1 shows the results from 1 million simulated responses. Overall, chance performance was 24%; therefore, collectively, subjects performed better than chance on the location task. Overall, the results show that people have some (above chance) ability to detect and locate manipulations, although performance is far from perfect.

[2] We report 95% confidence intervals to provide an estimate of the size of the effect—in 95% of cases, the population mean will fall within this range of values (Cumming, 2012).

Table 3.1

Mean number of regions (out of a possible nine) containing manipulation and results of Monte Carlo simulation to determine chance performance in the location task by manipulation type and overall

	Number of regions	% correct by chance	
Manipulation Type	*M*	*M*	95% CI
Airbrushing	1.83	20	[20, 21]
Add/Sub	1.33	17	[17, 17]
Geometry	1.50	19	[18, 19]
Shadow	1.67	15	[15, 15]
Super-Additive	4.33	48	[48, 48]
Overall	2.13	24	[24, 24]

Note. CI = confidence interval. For each manipulation type, we show the mean number of regions that contained the manipulation across all six images. The manipulation type "Overall" is the mean number of manipulated regions across all six images and all five manipulation types. To determine chance performance in the location task, we ran a Monte Carlo simulation of 1 million responses based on the number of regions manipulated for each image and manipulation type.

Ability to detect and locate by manipulation type

We predicted that people's ability to detect and locate manipulations might vary according to the manipulation type. Figure 3.3 shows subjects' accuracy on both the detection and the location task by manipulation type. In line with our prediction, subjects were better at detecting manipulations that included physically implausible changes (geometrical inconsistencies, shadow inconsistencies, and super-additive manipulations) than images that included physically plausible changes (airbrushing alterations and addition or subtraction of objects).

It was not the case, however, that subjects were necessarily better at locating the manipulation within the photo when the change was physically implausible. Figure 3.4

shows the proportion of manipulated photo trials in which subjects correctly detected a manipulation and also went on to correctly locate that manipulation, by manipulation type. Across both physically implausible and physically plausible manipulation types, subjects often correctly indicated that photos were manipulated but failed to then accurately locate the manipulation. Furthermore, although the physically implausible geometrical inconsistencies were more often correctly located, the shadow inconsistencies were only located equally as often as the physically plausible manipulation types—airbrushing and addition or subtraction. These findings suggest that people may find it easier to detect physically implausible, rather than plausible, manipulations, but this is not the case when it comes to locating the manipulation.

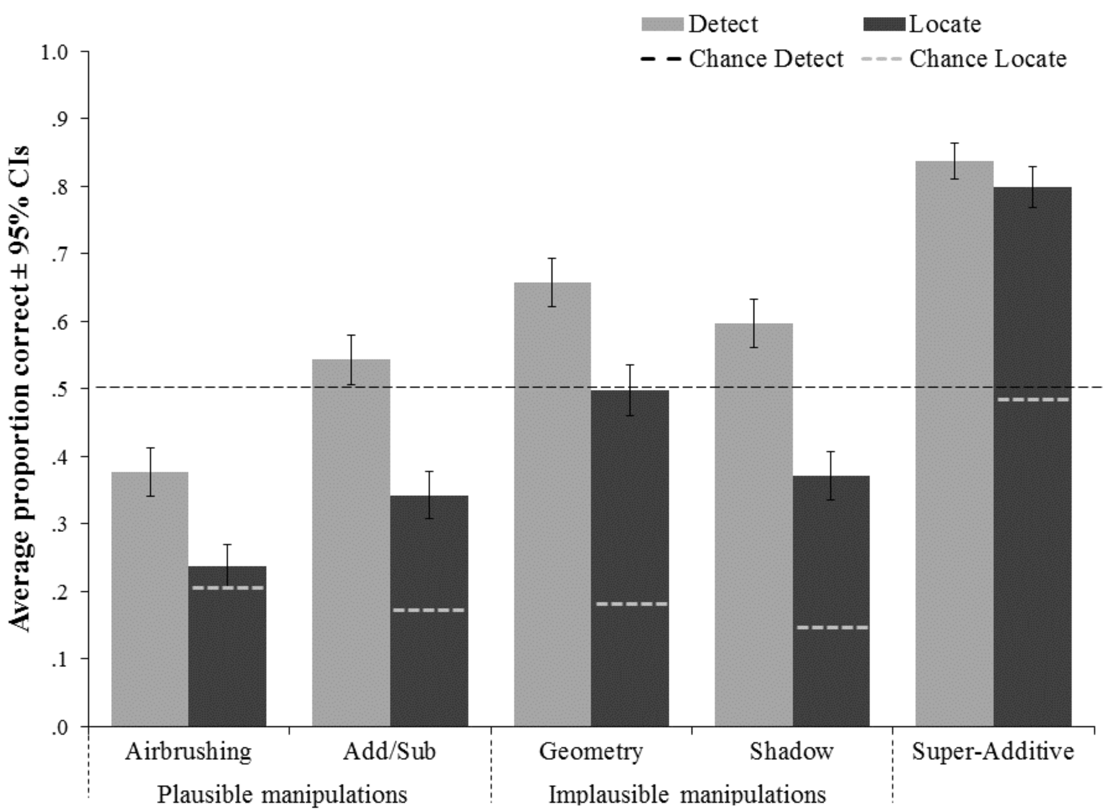


Figure 3.3. Mean proportion of correct “detect” and “locate” decisions by type of photo manipulation. The dotted line represents chance performance for detection. The grey dotted lines on the locate bars represent chance performance by manipulation type in the location task. Error bars represent 95% confidence intervals.

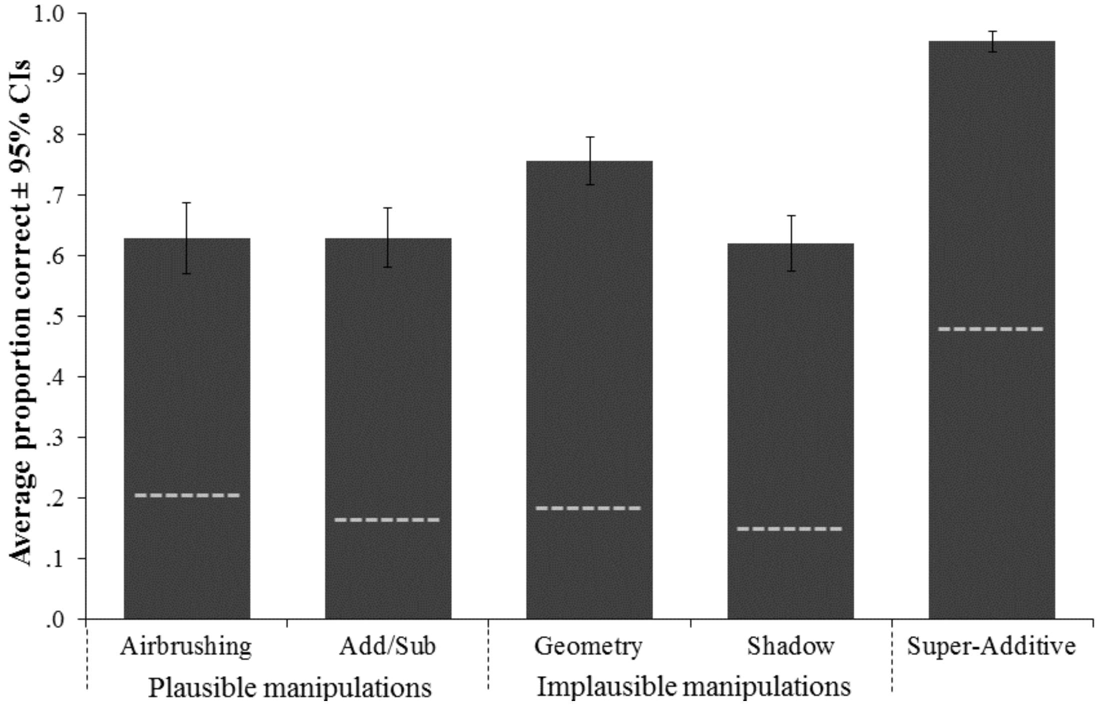


Figure 3.4. Mean proportion of correct "locate" decisions when subjects correctly detected that the photo was manipulated (i.e., correctly said "Yes" on the detection task). The grey dotted lines on the bars represent chance performance for each manipulation type. Error bars represent 95% confidence intervals.

Image metrics and accuracy

To understand more about people's ability to identify image manipulations, we examined how the amount of change in a photo affects people's accuracy in the detection and location tasks. When an image is digitally altered, the structure of the underlying elements—the pixels—are changed. This change can be quantified in numerous ways but we chose to use Delta-E76 because it is a measure based on both colour and luminance (Robertson, 1977). To calculate Delta-E, we first converted the images in Matlab® to L*a*b* colour space because it has a dimension for lightness as well as colour. Next we calculated the difference between corresponding pixels in the original and manipulated versions of each photo. Finally, these differences were averaged to give a single Delta-E score for each manipulated photo. A higher Delta-E value indicates a greater amount of difference between the original and the manipulated photo.[3] We calculated Delta-E for each of the 30 manipulated photos.

Figure 3.5 shows the log Delta-E values on the x-axis, where larger values indicate more change in the colour and luminance values of pixels in the manipulated photos compared with their original counterpart. The proportions of correct detection (Figure 3.5a) and location (Figure 3.5b) responses for each of the manipulated photos are presented on the y-axis. We found a positive relationship between the Delta-E measure and the proportion of photos that subjects correctly detected as manipulated, albeit not reaching significance: $r(28) = 0.34$, $p = .07$.[4] Furthermore, the Delta-E measure was positively correlated with the proportion of manipulations that were correctly located, $r(28) = 0.41$, $p = .03$. As predicted, these data suggest that people might be sensitive to the low-level properties of real-world scenes when making judgments about the authenticity of photos. This finding is especially remarkable given that our subjects never saw the same scene more than once and so never saw the original version of a manipulated image. This finding fits with the proposition that disrupting the underlying pixel structure might exacerbate the difference between the manipulated photos and people's expectations of how a scene should look. Presumably, these disruptions make it easier for people to accurately classify manipulated photos as being manipulated. We can also interpret these findings based on a signal detection account—adding greater signal (in our experiment, more change to an image, as measured by Delta-E) results in greater detection of that signal (Green & Swets, 1966; Wilken & Ma, 2004).

[3] One limitation of the Delta-E measure is that a global change to an image, for instance adjusting the brightness of the entire image, would result in a high Delta-E value, yet such a change is likely to be difficult to detect. That said, in our research we are only concerned with local image changes and therefore Delta-E provides a useful measure.

[4] This is based on a two-tailed test, given that we would predict that detection rates would increase with the amount of change, we might consider a one-tailed test to be appropriate. With a one-tailed test, the relationship between Delta-E and the proportion of photos correctly detected as manipulated would be significant at the .035 level.

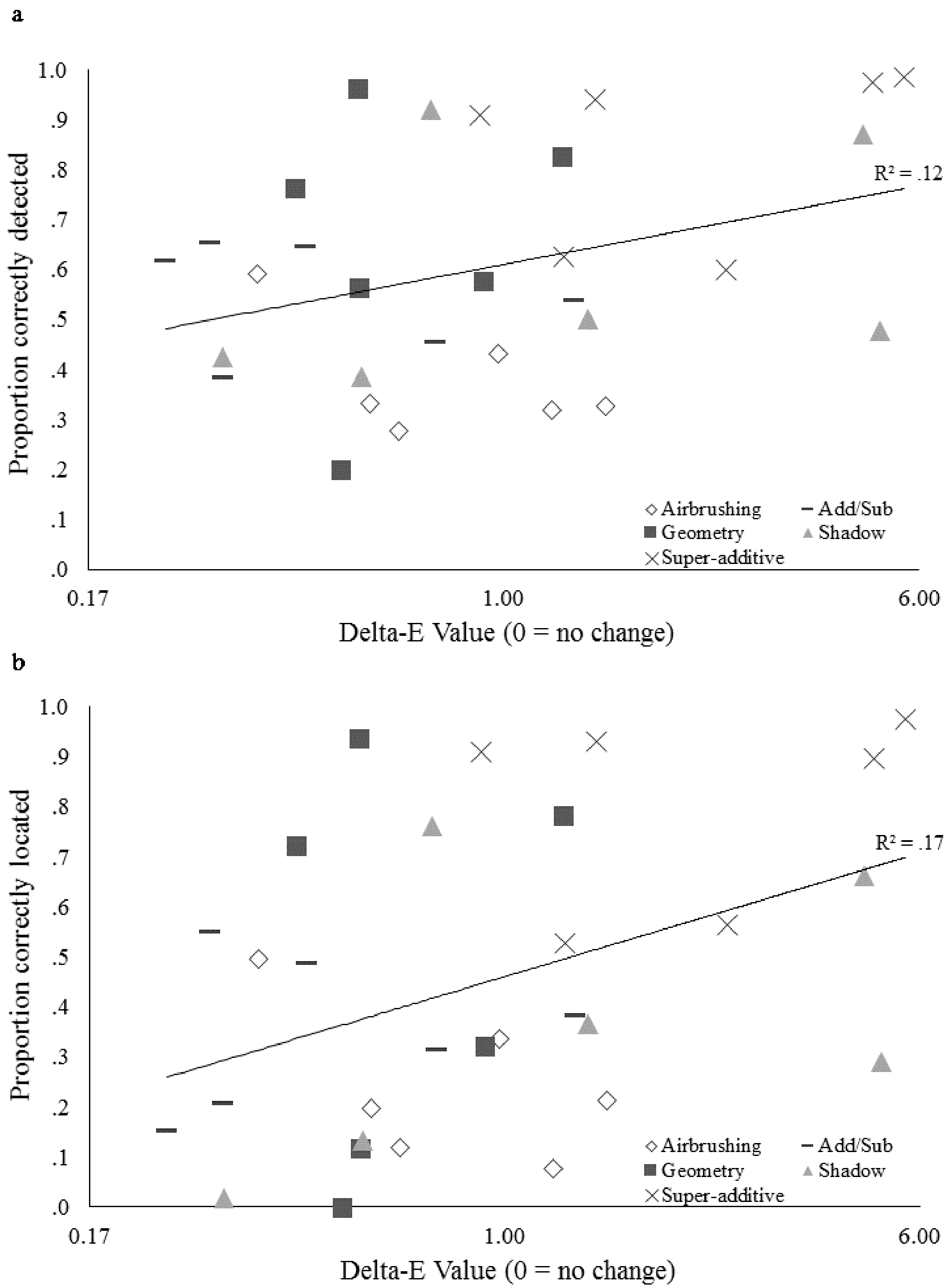


Figure 3.5. Mean proportion of correctly detected (a) and located (b) image manipulations by extent of pixel distortion as measured by Delta-E. The graphs show individual data points for each of the 30 manipulated images.

Next, we tested whether there was a relationship between the mean amount of change and the mean proportion of correct detection (Figure 3.6a) and location (Figure 3.6b) responses by the category of manipulation type. As Figure 3.6 shows, there was a numerical, but non-significant, trend for a positive relationship between amount of change and the proportion of photos that subjects correctly detected as manipulated: $r(3) = 0.68$, $p = .21$. There was also a numerical trend for a positive relationship between amount of change and the proportion of manipulations that were correctly located: $r(3) = 0.69$, $p = .19$.

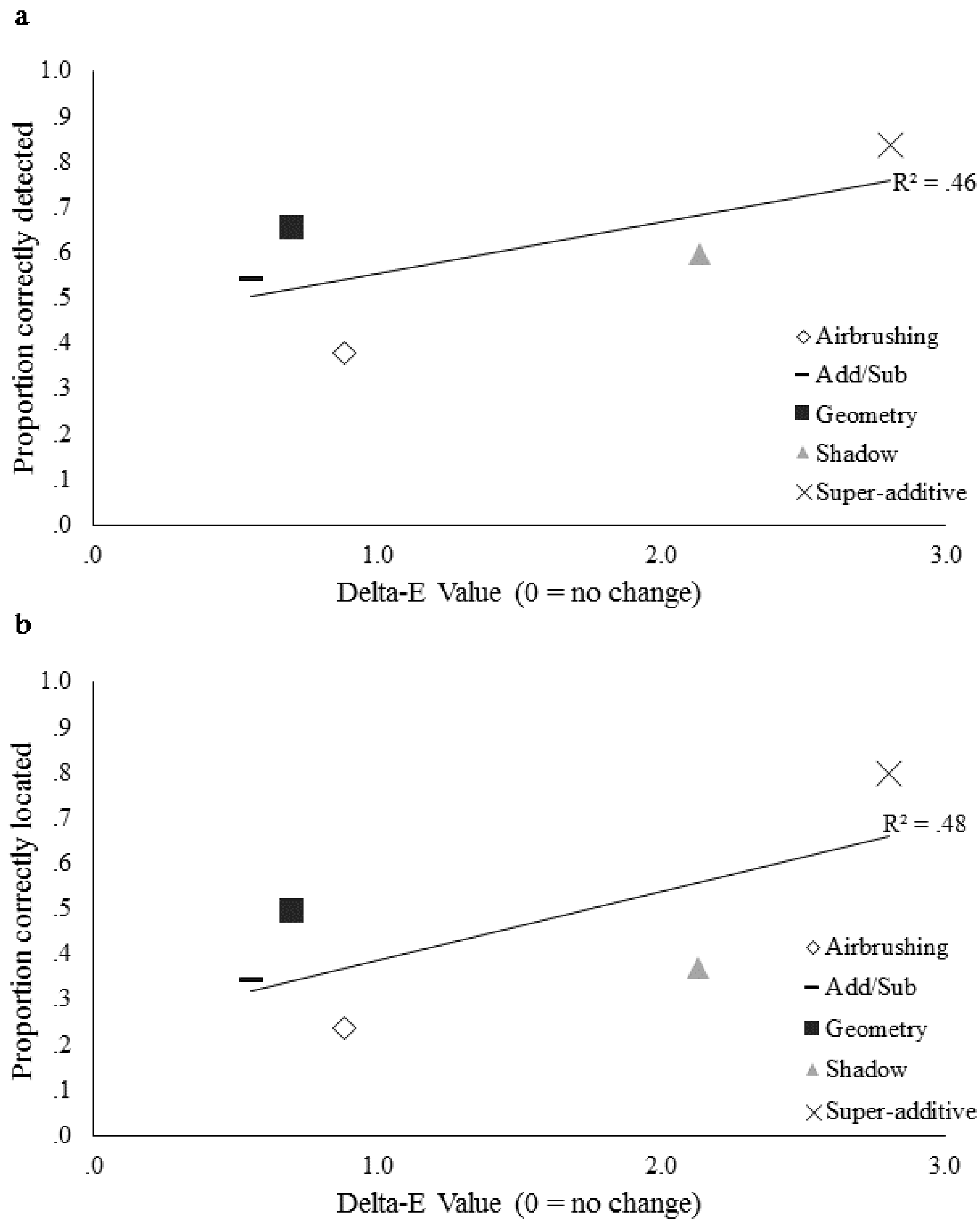


Figure 3.6. Mean proportion of correctly detected (a) and located (b) image manipulations by extent of pixel distortion as measured by Delta-E. The graphs show the mean values for each of the five categories of manipulation type.

Individual factors in detecting and locating manipulations

To determine whether individual factors play a role in detecting and locating manipulations, we gathered subjects' demographic data, attitudes towards image manipulation, and experiences of taking and manipulating photos. We also recorded subjects' response times on the detection and location tasks.

To determine how each factor influenced subjects' performance on the manipulated image trials, we conducted two generalized estimating equation (GEE) analyses—one for accuracy on the detection task and one for accuracy on the location task. Specifically, we conducted a repeated measures logistic regression with GEE because our dependent variables were binary with both random and fixed effects (Liang & Zeger, 1986). For the detection task, we ran two additional repeated measures linear regression GEE models to explore the effect of the predictor variables on signal detection estimates d' and c. The results of the GEE analyses are shown in Table 3.2. In the detection task, faster responses were more likely to be associated with accurate responses than slower responses. There was also a small effect of people's general belief about the prevalence of manipulated photos in their everyday lives on accuracy in the detection task. Those who believe a greater percentage of photos are digitally manipulated were more likely to correctly identify manipulated photos than those who believe a lower percentage of photos are digitally manipulated. Further, the results of the signal detection analysis suggest that this results from a difference in ability to discriminate between original and manipulated photos, rather than a shift in response bias—those who believe a greater percentage of photos are digitally manipulated accurately identified more of the manipulated photos without an increased false alarm rate. General beliefs about the prevalence of photo manipulation did not have an effect on people's ability to locate the manipulation. This pattern of results is somewhat surprising. It seems intuitive to think that a general belief that manipulated photos are prevalent simply makes people more likely to report that a photo is manipulated because they are generally sceptical about the veracity of photos rather than because they are better at spotting fakes. Although interesting, the small effect size and counterintuitive nature of the finding indicate that it is important to replicate the result prior to drawing any strong conclusions. The only variable that had an effect on accuracy in the location task was gender; males were slightly more likely than females to correctly locate the manipulation within the photo.

Table 3.2

Results of the GEE binary logistic and linear regression models to determine variables that predict accuracy on the detect and locate tasks

		Detect			Locate		
Predictor		*B*	*OR [95% CI]*	*p*	*B*	*OR [95% CI]*	*p*
Response time							
	Accuracy	0.11	1.11 [1.08, 1.15]	<.001	-	-	-
	d'	-0.01	0.99 [0.98, 1.01]	.31	-	-	-
	c	0.01	1.01 [1.00, 1.02]	.10	-	-	-
General beliefs about percentage of images manipulated = High (71-100%)							
	Accuracy	0.20	1.22 [1.06, 1.41]	.01	0.11	1.11 [0.98, 1.26]	.10
	d'	0.16	1.17 [1.05, 1.30]	.01	-	-	-
	c	-0.05	0.96 [0.90, 1.02]	.16	-	-	-
Gender = Female							
	Accuracy	0.05	1.05 [0.90, 1.23]	.50	-0.16	0.86 [0.75, 0.98]	.03
	d'	-0.06	0.95 [0.84, 1.06]	.35	-	-	-
	c	-0.05	0.95 [0.89, 1.02]	.15	-	-	-
Interest in photography = Interested							
	Accuracy	0.06	1.07 [0.92, 1.24]	.41	0.04	1.05 [0.92, 1.19]	.51
	d'	-0.02	0.98 [0.88, 1.10]	.73	-	-	-
	c	-0.05	0.96 [0.89, 1.03]	.20	-	-	-
Frequency of taking photos = daily/weekly							
	Accuracy	-0.15	0.86 [0.73, 1.01]	.07	-0.07	0.94 [0.81, 1.08]	.35
	d'	-0.08	0.92 [0.81, 1.04]	.18	-	-	-
	c	0.01	1.01 [0.94, 1.09]	.71	-	-	-

Note. CI = confidence interval. *B* and odds ratios (*OR*) estimate the degree of change in (a) accuracy on the task (based on the manipulated image trials), (b) *d'*, or (c) *c* associated with one unit change in the independent variable. An odds ratio of 1 indicates no effect of the independent variable on accuracy; values of 1.5, 2.5, and 4.0 are generally considered to reflect small, medium, and large effect sizes, respectively (Rosenthal, 1996). The category order for factors was set to descending to make the reference level 0. The reference groups are: General beliefs about percentage of images manipulated = Low (0–70%), Gender = Male, Interest in photography = Not Interested, Frequency of taking photos = Monthly/yearly/never. For response time (RT) we divided the data into eight equal groups (level 1 represents the slowest RTs (≥43.4 s) and level 8 the fastest RTs (≤8.4 s)). The 21 subjects who chose not to disclose their gender were excluded from these analyses leaving a total sample of $n = 686$. Given that subjects only responded on the location task if they said "yes", the photo had been manipulated, we did not have location response time data for all of the trials and therefore were unable to consider response time on the location task. Because we did not have a fixed number of choices per condition in the location task, we were unable to calculate the degree of change in *d'* or *c* associated with the predictor variables.

Together these findings show that individual factors have relatively little impact on the ability to detect and locate manipulations. Although shorter response times were associated with more correct detections of manipulated photos, we did not manipulate response time so we cannot know whether response time affects people's ability to discriminate between original and manipulated photos. In fact, our response time findings might be explained by a number of perceptual decision making models, for example, the drift diffusion model (Ratcliff, 1978). Yet determining the precise mechanism that accounts for the association between shorter response times and greater accuracy is beyond the scope of the current paper.

Experiment 1 indicates that people have some ability to distinguish between original and manipulated real-world photos. People's ability to correctly identify manipulated photos was better than chance, although not by much. Our data also suggest that locating photo manipulations is a difficult task, even when people correctly indicate that a photo is manipulated. We should note, however, that our study could have underestimated people's ability to locate manipulations in real-world photos. Recall that subjects were only asked to locate manipulations on photos that they thought were manipulated. It remains possible people might be able to locate manipulations even if they do not initially think that a photo has been manipulated. We were unable to check this possibility in Experiment 1, so we addressed this issue in Experiment 2 by asking subjects to complete the location task for all photos, regardless of their initial response in the detection task. If subjects did not think that the photo had been manipulated, we asked them to make a guess about which area of the image might have been changed.

We also created a new set of photographic stimuli for Experiment 2. Rather than sourcing photos online, we captured a unique set of photos on a Nikon D40 camera in RAW format, and prior to any digital editing, converted the files to PNGs. There are two crucial benefits to using original photos rather than downloading photos from the web. First, by using original photos we could be certain that our images had not been previously manipulated in any way. Second, when digital images are saved, the data are compressed to reduce the file size. JPEG compression is lossy in that some information is discarded to reduce file size. This information is not generally noticeable to the human eye (except at very high compression rates when compression artefacts can occur); however, the process of converting RAW files to PNGs (a lossless format) prevented any loss of data in either the original or manipulated images and, again, ensured that our photos were not manipulated in any way before we intentionally manipulated them.

Experiment 2

Method

Subjects and design

A total of 659 subjects ($M = 25.5$ years, $SD = 8.2$, range = 13–70; 362 male, 283 female, 14 declined to respond) completed the study online. A further 32 subjects were excluded from the analyses because they had missing response time data for at least one response on the detection or location task. As in Experiment 1, subjects did not receive payment for taking part but were given feedback on their performance at the end of the study. We stopped collecting data once we reached 100 responses per photo. The design was similar to that of Experiment 1.

Stimuli

We took our own photos in RAW format at a resolution of 3008 × 2000 pixels and converted them to PNGs with a resolution of 1600 × 1064 pixels prior to any digital editing. We checked the photos to ensure there were no spatial distortions caused by the lens, such as barrel or pincushion distortion. The photo manipulation process was the same as in Experiment 1. We applied the five manipulation techniques to six different photos to create a total of 30 manipulated photos. We used the non-manipulated version of these six photos and another four non-manipulated photos to give a total of ten original photos. Thus, the total number of photos was 40. As in Experiment 1, we ran two independent saliency models to check whether our manipulations had influenced the salience of the region where the manipulation had been made. See Appendix A for details of the saliency analyses. Similar to Experiment 1, our manipulations made little difference to the salience of the regions of the image.

Procedure

The procedure was similar to that used in Experiment 1, except for the following two changes. First, subjects were asked to locate the manipulation regardless of their response in the detection task. Second, subjects were asked to click on one of 12, rather than nine, regions on the photo to locate the manipulation. We increased the number of regions on the grid to ensure that the manipulations in the photos spanned two regions, on average, as per Experiment 1.

Results and discussion

As in Experiment 1, subjects spent a reasonable amount of time examining the photos. In the detection task, the mean response time per photo was 57.8 s ($SD = 271.5$ s) and the median 24.3 s (interquartile range = 17.3 to 37.4 s). In the location task, the mean response time was 10.9 s ($SD = 27.0$ s) and the median 8.2 s (interquartile range = 6.1 to 11.2 s).

Overall accuracy on the detection task and the location task

Overall accuracy in the detection task was slightly lower than that observed in Experiment 1, but still above chance: Subjects correctly classified 62% of the photos as being original or manipulated (cf. 66% in Experiment 1), 95% CI [60%, 63%]. Subjects had some ability to discriminate between original (58% correct) and manipulated (65% correct) photos, $d' = 0.56$, 95% CI [0.50, 0.62], replicating the results from Experiment 1. Again, this provides some support for the prediction that the match or mismatch between the information in the photo and people's expectation of what real-world scenes look like might help people to identify original and manipulated real-world photos. In contrast to Experiment 1, however, subjects did not show a bias towards saying that photos were authentic: $c = -0.07$, 95% CI [−0.10, −0.04]. It is possible that asking all subjects to search for evidence of a manipulation—the location task—regardless of their answer in the detection task, prompted a more careful consideration of the scene. In line with this account, subjects in Experiment 2 spent a mean of 14 s longer per photo on the detection task than those in Experiment 1.

Recall that the results from Experiment 1 suggested that subjects found the location task difficult, even when they correctly detected the photo as manipulated. Yet, we were unable to conclusively say that location was more difficult than detection because we did not have location data for the manipulated photo trials that subjects failed to detect. In Experiment 2 we gathered those data, but before we could directly compare subjects' ability to detect manipulated photos with their ability to locate the manipulations within, we had to correct for guessing. For the detection task, chance performance was the same as Experiment 1, 50%. For the location task, however, there were two differences to Experiment 1. First, subjects were asked to select one of 12, rather than one of nine, image regions. Second, we used a new image set; thus, the number of regions manipulated for each image and manipulation type changed. Accordingly, we ran a separate Monte Carlo simulation to determine the chance rate of selecting the correct region. Table 3.3 shows that overall chance performance in the location task was 17%.

Table 3.3
Mean number of regions (out of a possible 12) containing manipulation and results of Monte Carlo simulation to determine chance performance in the location task by manipulation type and overall

	Number of regions	% correct by chance	
Manipulation Type	*M*	*M*	95% CI
Airbrushing	1.50	12	[12, 13]
Add/Sub	1.33	11	[11, 11]
Geometry	1.33	11	[11, 11]
Shadow	1.33	11	[11, 11]
Super-Additive	4.67	39	[39, 39]
Overall	2.03	17	[17, 17]

Note. CI = confidence interval. For each manipulation type, we show the mean number of regions that contained the manipulation across all six images. The manipulation type "Overall" is the mean number of manipulated regions across all six images and all five manipulation types. To determine chance performance in the location task, we ran a Monte Carlo simulation of 1 million responses based on the number of regions manipulated for each image and manipulation type.

Subjects performed better than chance on the location task: a mean 56% of the manipulations were accurately located, 95% CI [55%, 58%]. Given that a mean 62% of the manipulated images were accurately detected and a mean 56% of the manipulations located, it seems that performance was very roughly similar on the two tasks. But this interpretation does not take into account how subjects would perform by chance alone. A fairer approach is to compare subjects' performance on the detection and location tasks with chance performance on those two tasks. For the detection task, subjects detected a mean 12% more manipulated images than would be expected by chance alone, 95% CI [10%, 13%]. Yet, somewhat surprisingly, subjects located a mean 39% more of the manipulations than would be expected by chance alone, 95% CI [38%, 41%]. This finding suggests that people are better at the more direct task of locating manipulations than the more generic one of detecting if a photo has been manipulated or not. Although this potential distinction between people's ability to detect and locate manipulations is an interesting finding, the reason for it is not immediately apparent. One possibility is that our assumption that each of the 12 image regions has an equal chance of being picked is too simplistic—perhaps certain image regions never get picked (e.g., a relatively featureless area of the sky). If so, including these never picked regions in our chance calculation might make subjects' performance on the location task seem artificially high. To check this possibility, we ran a second chance performance calculation.

In Experiment 2, even when subjects did not think that the image had been manipulated, they still attempted to guess the region that had been changed. Therefore, we can use these localisation decisions in the original (non-manipulated) versions of the six critical photos to determine chance performance in the task. This analysis allows us to calculate chance based on the regions (of non-manipulated images) that people actually selected when guessing rather than assuming each of the 12 regions has an equal chance of being picked. Using this approach, Table 3.4 shows that overall chance performance in the location task was 23%. Therefore, even based on this chance localisation level, subjects still located a mean 33% more of the locations than would be expected by chance alone, 95% CI [32%, 35%]. This finding supports the idea that subjects are better at the more direct task of locating manipulations than detecting whether a photo has been manipulated or not.

Table 3.4

Chance performance in location task by manipulation type and overall based on mean number of subjects choosing the manipulated region in the original version of the image

	% correct by chance						
	Image						
Manipulation Type	A	B	C	D	E	F	Overall
Airbrushing	19	31	28	28	23	20	25
Add/Sub	24	5	15	3	3	1	9
Geometry	11	12	17	2	26	12	13
Shadow	20	16	28	39	4	5	19
Super-Additive	74	63	44	72	33	26	53
Overall Image	30	25	27	29	18	13	23

Note. For each of the six critical images and each of the five manipulation types, we show the probability that the manipulated region of the image was selected by chance in the original version of the image. The "Overall" column denotes the mean probability of selecting the manipulated regions for that manipulation type across all six images A-F. The "Overall image" is the mean probability of selecting the manipulated regions for that image across all manipulation types. Each image had a minimum of 101 responses.

Ability to detect and locate by manipulation type

On the manipulated photo trials, asking subjects to locate the manipulation regardless of whether they correctly detected it allowed us to segment accuracy in the following ways: (i) accurately detected and accurately located (hereafter, DL), (ii) accurately detected but not accurately located (DnL), (iii) inaccurately detected but accurately located (nDL), or (iv) inaccurately detected and inaccurately located (nDnL). Intuitively, it seems most practical to consider the more conservative accuracy—DL—as correct, especially in certain contexts, such as the legal domain, where it is crucial to know not only that an image has been manipulated, but precisely what about it is fake. That said, it might be possible to learn from the DnL and nDL cases to try to better understand how people process manipulated images.

Figure 3.7 shows the proportion of DL, DnL, nDL, and nDnL responses for each of the manipulation types. The most common outcomes were for subjects to both accurately detect and accurately locate manipulations, or both inaccurately detect and inaccurately locate manipulations. It is interesting, however, that on almost a fifth (18%)

of the manipulated photo trials, subjects accurately detected the photo as manipulated yet failed to locate the alteration. For 10% of the manipulated trials, subjects failed to detect but went on to successfully locate the manipulation. Subjects infrequently managed to detect and locate airbrushing manipulations; in fact it was more likely that subjects made DnL or nDL responses. Although this fits with our prediction that plausible manipulations would be more difficult to identify than implausible ones, the pattern of results for geometrical inconsistency, shadow inconsistency, and addition or subtraction do not support our prediction. Subjects made more DL responses on the plausible addition or subtraction manipulation photos than on either of the implausible types, geometrical manipulations and shadow manipulations. Why, then, are subjects performing better than expected by either of the chance measures on the addition or subtraction manipulations and worse than expected on the airbrushing ones? One possibility is that people's ability to detect image manipulations is less to do with the plausibility of the change and more to do with the amount of physical change caused by the manipulation. We now look at this hypothesis in more detail by exploring the relationship between the image metrics and people's ability to identify manipulated photos.

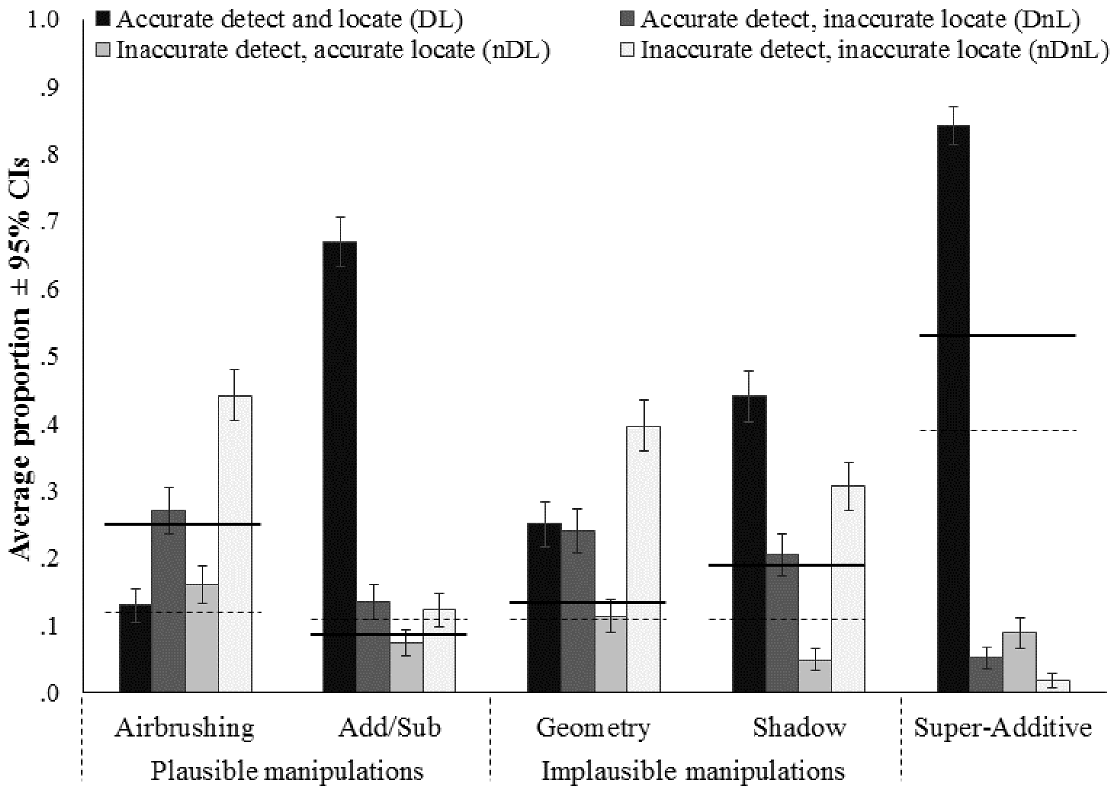


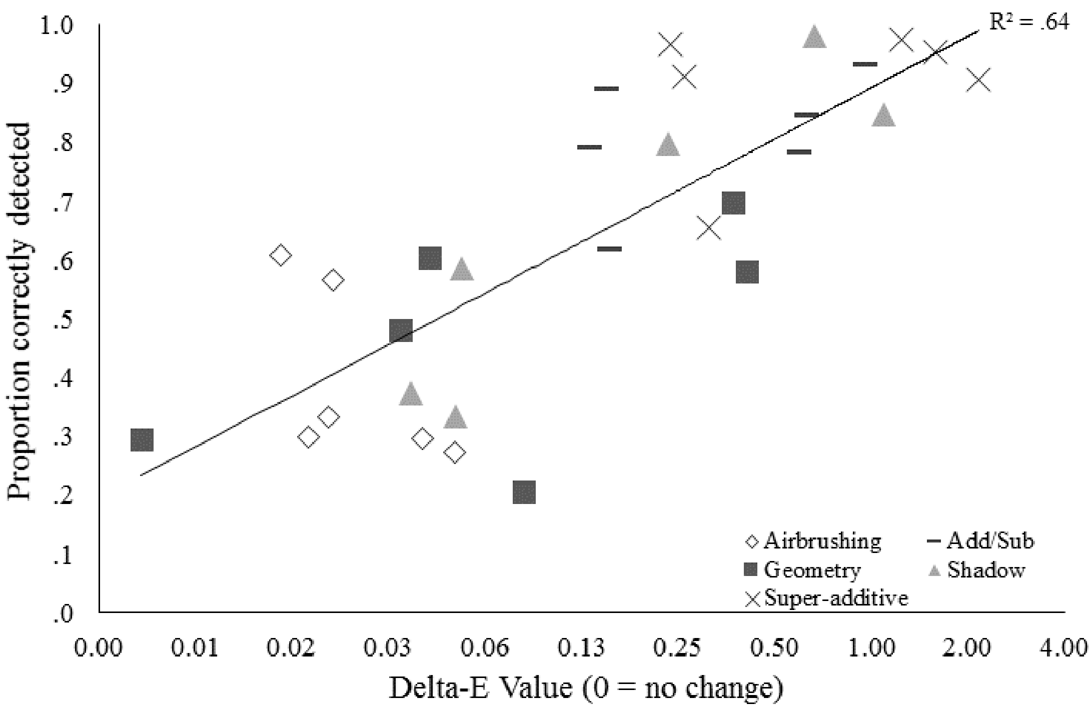
Figure 3.7. Mean proportion of manipulated photos accurately detected and accurately located (DL); accurately detected, inaccurately located (DnL); inaccurately detected, accurately located (nDL) and inaccurately detected, inaccurately located (nDnL) by manipulation type. The dotted horizontal lines on the bars represent chance performance for each manipulation type from the results of the Monte Carlo simulation. The full horizontal lines on the bars represent chance performance for each manipulation type based on subjects' responses on the original image trials. Error bars represent 95% confidence intervals.

Image metrics and accuracy

Recall that the results from Experiment 1 suggested a relationship between the correct detection and location of image manipulations and the amount of disruption the manipulations had caused to the underlying structure of the pixels. Yet, the JPEG format of the images used in Experiment 1 created some (re-compression) noise in the Delta-E measurements between different images; thus, we wanted to test whether the same finding held with the lossless image format used in Experiment 2. As shown in Figure 3.8, we found that the Delta-E measure was positively correlated with the proportion of photos that subjects correctly detected as manipulated ($r(28) = 0.80$, $p < .001$) and the proportion of manipulations that were correctly located ($r(28) = 0.73$, $p < .001$). These

Pearson correlation coefficients are larger than those in Experiment 1 (cf. detect $r = 0.34$ and locate $r = 0.41$ in Experiment 1). It is possible that the re-compression noise in the JPEG images in Experiment 1 obscured the relationship between Delta-E and detection and localisation performance. To check whether there was a stronger relationship between Delta-E and people's ability to detect and locate image manipulations in Experiment 2 than Experiment 1, we converted the correlation coefficients to z values using Fisher's transformation. There was a significantly stronger correlation between the Delta-E and detection in Experiment 2 than in Experiment 1: $z = -2.74$, $p = .01$. Yet because we had good reason to predict a stronger relationship in Experiment 2 than Experiment 1 (based on the JPEG re-compression noise), it might be fairer to consider the p value associated with a one-tailed test, $p = .003$. The correlation between Delta-E and accurate localisation was not significantly stronger in Experiment 2 than in Experiment 1 based on a two-tailed test ($z = -1.81$, $p = .07$), but was based on a one-tailed test ($p = .04$). Therefore, it is possible that the global (re-compression) noise in the Delta-E values in Experiment 1 weakened the association between the amount of change and people's ability to identify manipulations. This finding suggests that Delta-E is a more useful measure for local, discrete changes to an image than it is for global image changes, such as applying a filter.

a



b

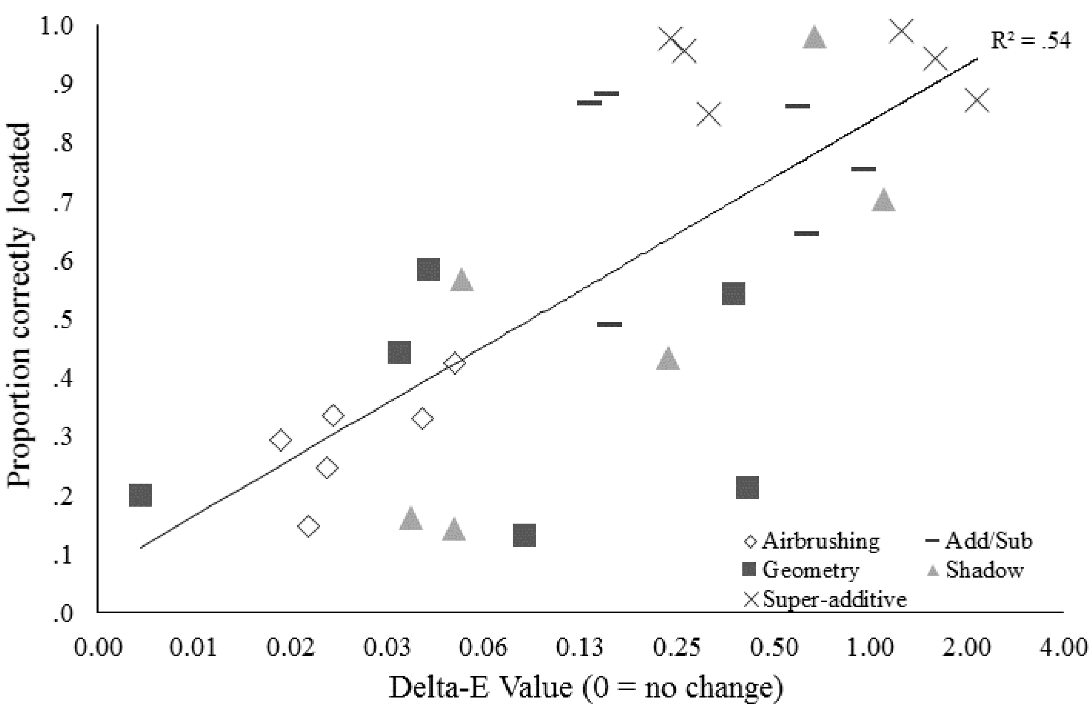


Figure 3.8. Mean proportion of correctly detected (a) and located (b) image manipulations by extent of pixel distortion as measured by Delta-E. The graphs show individual data points for each of the 30 manipulated images.

Of course, the whole point of manipulating images is to fool observers, to make them believe that something fake is in fact true. Therefore, it might not be particularly surprising to learn that people find it difficult to spot high quality image manipulations. Yet it is surprising to learn that, even though our subjects never saw the same image more than once, this ability might be dependent on the amount of disruption between the original and manipulated image. The positive relationship between the accurate detection and localisation of manipulations and Delta-E suggests that it might be possible to develop a metric that allows for a graded prediction about people's ability to detect and locate image manipulations. The possibility that a metric could be used to predict people's ability to identify image manipulations is an exciting prospect; however, further research is needed to check that this finding generalises across a wider variety of images and manipulation types. Our findings suggest that manipulation type and the technique used to create the manipulation, for instance, cloning or scaling, might be less important than the extent to which the change affects the underlying pixel structure of the image. To test this possibility, we next consider the relationship between the Delta-E values and the proportion of (a) correct detection and (b) location responses by the category of manipulation type.

Our findings in Experiments 1 and 2 show that subjects' ability to detect and locate image manipulations varied by manipulation type, yet, in Experiment 2 the differences were not adequately explained by the plausibility of the manipulation. That is, subjects accurately detected and located more of the addition or subtraction manipulations than the geometry, shadow, or airbrushing manipulations. One possibility is that the five categories of manipulation type introduced different amounts of change between the original and manipulated versions of the images. If so, we might expect these differences in amount of change to help explain the differences in subjects' detection and localisation rates across these categories.

To check this, we calculated the mean proportion of correct detections, localisations, and Delta-E values for each of the five categories of manipulation type. As Figure 3.9 shows, there was a positive correlation between the amount of change and the proportion of correct detections ($r(3) = 0.92$, $p = .03$) and the proportion of correct localisations ($r(3) = 0.95$, $p = .01$). These results suggest that the differences in detection and localisation rates across the five manipulation types are better accounted for by the extent of the physical change to the image caused by the manipulation, rather than the plausibility of that manipulation. Yet, given that subjects did not have the opportunity to compare the manipulated and original version of the scene, it is not entirely obvious why amount of change predicts accuracy.

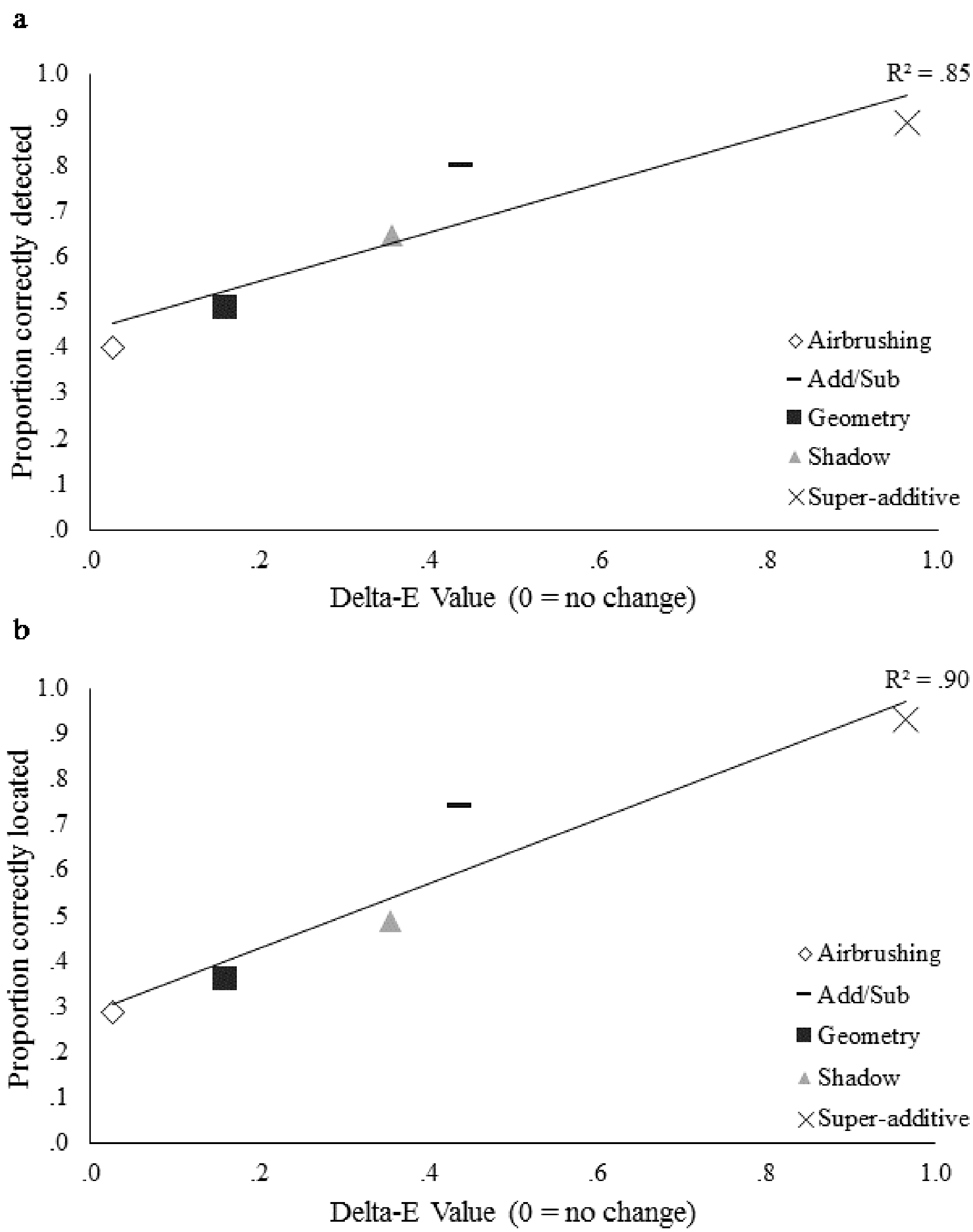


Figure 3.9. Mean proportion of correctly detected (a) and located (b) image manipulations by extent of pixel distortion as measured by Delta-E. The graphs show the mean values for each of the five categories of manipulation type.

Our results suggest that the amount of change between the original and manipulated versions of an image is an important factor in explaining the detectability and localisation of manipulations. Next we considered whether any individual factors are associated with improved ability to detect or locate manipulations.

Individual factors in detecting and locating manipulations

Using GEE analyses, we again explored various factors that might affect people's ability to detect and locate manipulations. As discussed, we were able to use liberal or stringent criteria for our classification of detection and location accuracy on the manipulated image trials. Accordingly, we ran three models: the first two used the liberal classification for accuracy (and replicated the models we ran in Experiment 1), and the other examined the more stringent classification, DL. As in Experiment 1, for the detection task, we also ran two repeated measures linear regression GEE models to explore the effect of the predictor variables on signal-detection estimates d' and c. We included the same factors used in the GEE models in Experiment 1. The results of the GEE analyses are shown in Table 3.5.

Table 3.5

Results of the GEE binary logistic and linear regression models to determine variables that predict accuracy in the detect and locate tasks

Predictor		*B*	*OR [95% CI]*	*p*
			Detect (DL and DnL)	
Response time				
	Accuracy	0.13	1.14 [1.10, 1.18]	<.001
	d'	-0.01	0.99 [0.98, 1.01]	.40
	c	0.004	1.00 [0.99, 1.01]	.42
General belief about percentage of images manipulated = High (71-100%)				
	Accuracy	0.16	1.18 [1.02, 1.36]	.03
	d'	0.09	1.09 [0.97, 1.23]	.14
	c	-0.04	0.96 [0.90, 1.03]	.25
Gender = Female				
	Accuracy	-0.01	0.99 [0.86, 1.15]	.92
	d'	-0.03	0.97 [0.86, 1.09]	.60
	c	-0.01	0.99 [0.93, 1.06]	.82
Interest in photography = Interested				
	Accuracy	0.17	1.19 [1.02, 1.39]	.03
	d'	0.04	1.04 [0.92, 1.18]	.56
	c	-0.05	0.95 [0.89, 1.02]	.18
Frequency of taking photos = daily/weekly				
	Accuracy	-0.01	0.99 [0.84, 1.17]	.91
	d'	-0.07	0.93 [0.82, 1.07]	.31
	c	-0.05	0.95 [0.88, 1.02]	.18
	Accuracy		Locate (DL and nDL)	
Response time		0.10	1.11 [1.08, 1.14]	<.001
General belief about percentage of images manipulated = High (71-100%)		-0.01	0.99 [0.87, 1.12]	.84
Gender = Female		-0.10	0.91 [0.80, 1.03]	.14
Interest in photography = Interested		0.16	1.17 [1.02, 1.34]	.02
Frequency of taking photos = daily/weekly		-0.08	0.92 [0.80, 1.06]	.27
	Accuracy		Detect and locate (DL)	
Response time: Detect		0.17	1.19 [1.15, 1.23]	<.001
Response time: Locate		0.13	1.14 [1.11, 1.18]	<.001
General belief about percentage of images manipulated = High (71-100%)		0.05	1.05 [0.91, 1.20]	.51
Gender = Female		-0.13	0.88 [0.77, 1.01]	.07
Interest in photography = Interested		0.20	1.22 [1.06, 1.41]	.01
Frequency of taking photos = daily/weekly		-0.09	0.92 [0.78, 1.07]	.28

Note. CI = confidence interval. *B* and odds ratios (*OR*) estimate the degree of change in (a) accuracy on the task (based on the manipulated image trials), (b) *d'*, or (c) *c* associated with one unit change in the independent variable. An odds ratio of 1 indicates no effect of the independent variable on accuracy; values of 1.5, 2.5, and 4.0 are generally considered to reflect small, medium, and large effect sizes, respectively (Rosenthal, 1996). The category order for factors was set to descending to make the reference level 0. The reference groups are: General beliefs about percentage of images manipulated = Low (0–70%), Gender = Male, Interest in photography = Not Interested, Frequency of taking photos = Monthly/yearly/never. For response time (RT) we divided the data into eight equal groups with level 1 representing the slowest RTs (detect ≥47.1 s; locate ≥18.9 s) and level 8 the fastest (detect ≤8.1 s; locate ≤2.7 s). The 14 subjects who chose not to disclose their gender were excluded from these analyses, leaving a total sample of *n* = 645. Because we did not have a fixed number of choices per condition in the location task, we were unable to calculate the degree of change in *d'* or *c* associated with the predictor variables.

Using the more liberal accuracy classification, that is, both DL and DnL responses for detection, we found that three factors had an effect on likelihood to respond correctly: response time, general beliefs about the prevalence of photo manipulation, and interest in photography. As in Experiment 1, faster responses were more likely to be correct than slower responses. Also replicating the finding in Experiment 1, those who believe a greater percentage of photos are digitally manipulated were slightly more likely to correctly identify manipulated photos than those who believe a lower percentage of photos are digitally manipulated. In addition, in Experiment 2, those interested in photography were slightly more likely to identify image manipulations correctly than those who are not interested in photography. For the location task, using the more liberal accuracy classification, that is, both DL and nDL responses, we found that two factors had an effect on likelihood to respond correctly. Again there was an effect of response time: In the location task, faster responses were more likely to be correct than slower responses. Also those with an interest in photography were slightly more likely to correctly locate the manipulation within the photo than those without an interest. Next we considered whether any factors affected our more stringent accuracy classification, that is, being correct on both the detection and location tasks (DL). The results revealed an effect for two factors on likelihood to respond correctly. Specifically, there was an effect of response time with shorter response times being associated with greater accuracy. There was also an effect of interest in photography, with those interested more likely to correctly make DL responses than those not interested.

Our GEE models in both Experiments 1 and 2 revealed that shorter response times were linked with more correct responses on both tasks. As in Experiment 1, this association might be explained by several models of perceptual decision making; however, determining which of these models best accounts for our data is beyond the scope of the current paper.

Conclusion

In two separate experiments we have shown, for the first time, that people's ability to detect manipulated photos of real-world scenes is extremely limited. Considering the prevalence of manipulated images in the media, on social networking sites, and in other domains, our findings warrant concern about the extent to which people may be frequently fooled in their daily lives. Furthermore, we did not find any strong evidence to suggest that individual factors, such as having an interest in photography or beliefs about the extent of image manipulation in society, are associated with improved ability to detect or locate manipulations.

Recall that we looked at two categories of manipulations—implausible and plausible—and we predicted that people would perform better on implausible manipulations because these scenes provide additional evidence that people can use to determine if a photo has been manipulated. Yet the story was not so simple. In Experiment 1, subjects correctly detected more of the implausible photo manipulations than the plausible photo manipulations, but in Experiment 2, the opposite was true. Further, even when subjects correctly identified the implausible photo manipulations, they did not necessarily go on to accurately locate the manipulation. It is clear that people find it difficult to detect and locate manipulations in real-world photos, regardless of whether those manipulations lead to physically plausible or implausible scenes.

Research in the vision science literature may help to account for these findings. We know that people might have a simplified understanding of the physics in our world (Cavanagh, 2005; Mamassian, 2008). Studies have shown, for instance, that the human visual system is relatively insensitive to the physically impossible cast shadows created by inconsistent lighting in a scene (Ostrovsky, Cavanagh, & Sinha, 2005; see also Chapter 4 of this thesis). It is not necessarily the case that people ignore shadows altogether, but rather that the visual system processes shadows rapidly and uses them only as a generic cue. Put simply, as long as the shadow is roughly correct then we accept it as being authentic (Bonfiglioli, Pavani, & Castiello, 2004; Ostrovsky et al., 2005; Rensink & Cavanagh, 2004). Similarly, people use shortcuts to interpret geometrical aspects of a scene; if the geometry is close enough to people's expectation, then it is accepted as accurate (Bex, 2010; Howe & Purves, 2005; Mamassian, 2008). Furthermore, the change blindness literature also highlights people's insensitivity to shadow information. Research has shown that people are slower to detect changes to cast shadows than changes to objects (Wright, 2005), even when the shadow changes affect the overall meaning of the scene (Ehinger, Allen, & Wolfe, 2016). It follows, then, that when trying to distinguish between real and manipulated images, our subjects do not seem to have capitalised on the evidence in the implausible manipulation photos to determine whether they were authentic or not. It remains to be seen whether it is possible to train people to make use of physically implausible inconsistencies; perhaps one possibility would entail "teaching" the visual system to make full use of physical properties of the world as opposed to automatically simplifying them.

Although the plausibility of a manipulation might not be so important when it comes to detecting manipulated images, we found that the extent to which the manipulation disrupts the underlying structure of the pixels might be important. Indeed, we found a positive correlation between the image metric (Delta-E) we used to measure the difference between our original and manipulated photos and the likelihood that the photo was correctly classified as manipulated. In other words, the manipulations that created the most change in the underlying pixel values of the photo were most likely to be correctly classified as manipulated. Of course, from the perspective of signal detection theory, it follows that adding greater signal results in greater detection of that signal (Green & Swets, 1966; Wilken & Ma, 2004).

Although this might seem intuitive, recall that our subjects never saw the same scene more than once. That is, they never saw the non-manipulated versions of any of the manipulated photos that they were shown; despite this, their ability to detect the manipulated photos was related to the extent of change in the pixels. It seems possible that our subjects might have been able to compare the manipulated photo with their expectations about what the scene "should" look like in terms of scene statistics. In doing this, subjects might have found the manipulated photos with less change, and thus smaller Delta-E values, were more similar to their prior expectations of what the world looks like—resulting in those photos being incorrectly accepted as authentic more often. At the same time, the manipulated photos with more change, and thus larger Delta-E values, might have been more difficult to match to a prior expectation—resulting in these photos more often being correctly identified as manipulated. It seems that this difference in ease of finding a match to prior knowledge and expectation for the manipulated photo helped subjects to make an accurate decision. If this is the case, then one might speculate that it could be possible to develop a metric that will predict people's ability to detect and locate manipulations of real-world scenes. A future investigation using a wider range of stimuli where subjects see more than one of each manipulation type might consider whether there is an interaction between Delta-E and manipulation type.

On a different note, our research highlights a potential opportunity to improve people's ability to spot manipulations. In Experiment 2, we were able to compare subjects' ability on the two tasks: detection and location. We were surprised to find that subjects performed better on the location task than on the detection task. Although this is an interesting finding, the reason for it is not immediately apparent. One possibility is that these two tasks might encourage subjects to adopt different strategies and that subjects are better at the more direct task of locating manipulations than the generic one of detecting whether a photo has been manipulated or not.

Our research provides a first look at people's ability to detect and locate manipulations of real-world images. A strength of the current method—applying each of the five different manipulation types to the same image—is that we know the differences in subjects' performance is owing to the manipulation itself rather than the specific image. A drawback, however, is that the difficulty of finding or generating a set of suitable images that allowed all of the manipulation types to be applied reduced the total number of photos that could be tested to some degree. Although, ideally, future work might extend the range of images tested, we nonetheless note the close consistency in results that we obtained across the two different and independent image sets used in Experiments 1 and 2.

Future research might also investigate potential ways to improve people's ability to spot manipulated photos. Our findings suggest, however, that this is not going to be a straightforward task. We did not find any strong evidence to suggest there are individual factors that improve people's ability to detect or locate manipulations. That said, our findings do highlight various possibilities that warrant further consideration, such as training people to make better use of the physical laws of the world, varying how long people have to judge the veracity of a photo, and encouraging a more careful and considered approach to detecting manipulations. What our findings have shown is that a more careful search of a scene, at the very least, might encourage people to be sceptical about the veracity of photos. Of course, increased scepticism is not perfect because it comes with an associated cost: a loss of faith in authentic photos. Yet, until we know more about how to improve people's ability to distinguish between real and fake photos, a sceptical approach might be wise, especially in contexts such as law, scientific publication, and photojournalism where even a small manipulation can have ethically significant consequences.

But what should we be sceptical about? Are some changes acceptable and others not? Should the context of the manipulation be taken into account? Though we are unable to answer these complex questions here, we can offer some points for thought. Although it is true that all image manipulations are to some extent deceptive, not all manipulations are intentionally deceptive. This distinction is an important one and raises the possibility that people do not set out to detect all image manipulations but instead are primarily concerned about forgeries that have been created with the intention to deceive the viewer. Of course, people might expect that all images provided as evidence, for instance news images, have been subjected to rigorous validation processes. It is unlikely, however, that people set themselves the same standard for detecting manipulation in everyday contexts. Perhaps more important than being able to identify all instances of manipulation, people are most concerned about the extent to which they can trust the message conveyed from the image. Although this poses an interesting question, our results suggest that people might struggle to detect image manipulations based on either of these definitions. In the current research, not only did subjects find it difficult to accurately locate the specific aspects of the image that had been altered, they also found it difficult to distinguish original, truthful photos from manipulated, untruthful ones.

In light of the findings presented in this paper, it is not surprising that World Press Photo have introduced a computerised photo-verification test to their annual photo contest. But ultimately, this is only a competition. What do our findings mean for other contexts in which an incorrect decision about the veracity of a photo can have devastating consequences? Essentially, our results suggest that guidelines and policies governing the acceptable standards for the use of photos, for example, in legal and media domains, should be updated to reflect the unique challenges of photography in the digital age. We recommend that this is done soon, and that psychological scientists work together with digital forensic experts and relevant end-users to ensure that such policies are built on sound empirical research.

Chapter 4 :

Can people identify geometric inconsistencies in cast shadows?

"For me, a landscape does not exist in its own right, since its appearance changes at every moment; but the surrounding atmosphere brings it to life - the air and the light, which vary continuously."

Claude Monet (1891)

Introduction

On May 23rd 2016, Dinesh and Tarakeshwari Rathod were hailed as the first Indian couple to conquer Mount Everest (Boone, 2016). Yet the couple's celebrations were short lived; three weeks after their incredible triumph, fellow mountaineers filed a complaint stating that the couple never made it to the summit and that the photos provided to evidence their success were forgeries. The complainants revealed several contradictions in the couple's summit photos. Of particular interest was the date and time stamp on the photos—6.25am on May 23rd 2016. Crucially, these camera setting details did not match the time indicated by the direction of the shadows in the scene. Instead, the direction of the shadows suggested the photo was taken around 5 hours later in the day, closer to noon (Boone, 2016). Following an investigation, the Nepalese government confirmed that the couple had indeed faked their summit photos and subsequently imposed a ban to prevent them from mountaineering in Nepal for the next 10 years (Safi, 2016). This example highlights the possibility that shadow information could offer a useful means to determine whether photos are authentic or manipulated. Thus the work in the current chapter explores whether people can identify when the shadows in a scene are consistent or inconsistent with a single light source, and considers the theoretical implications for detecting image forgeries and our understanding of how the visual system processes shadow information.

Illumination conditions, that is the position and intensity of the light source, can strongly influence the appearance of a scene. Shadows, for example, are not intrinsic or stable properties of a scene but instead are the result of the interaction between the world and its illumination (Baxandall, 1995; Rensink & Cavanagh, 2004). An object's shadow will appear differently, or disappear altogether, when lighting conditions change for instance. This chapter focuses on cast shadows that are formed when an opaque object obstructs the light and prevents it from illuminating an external surface, such as the ground. Since light travels in a straight line, it is known that a point in a shadowed region, its corresponding point on the shadow-casting object, and the light source must all lie on a single straight line (Farid, 2016; Farid & Bravo, 2010; Kee, O'Brien, & Farid, 2013). This relationship means that shadows provide information about the geometry of the scene and can be used to reason about the location of the illuminating light source (Casati, 2004; Farid, 2016; Farid & Bravo, 2010). The physical laws that constrain the behaviour of light in the 3-D world—either real or virtual—also apply to 2-D images (Farid, 2016; Kajiya, 1986). That is, when taking a photo (or rendering an image from a virtual environment) the interaction of light and the 3-D objects in the scene is captured in the geometry of the 2-D image.[5]

In fact, the constraint that connects the shadow, the shadow-casting object, and the light source permits a surprisingly simple image-based geometric technique for objectively verifying the authenticity of shadows. As Figure 4.1a shows, to use this technique, simply locate any point on a shadow and its corresponding point on the object, then draw a line through them. Repeat this process for as many corresponding shadow and object points as possible, and the point at which these lines intersect is the location of the light source. Now consider if a new object is added to the scene, for example a bus stop. As Figure 4.1b shows, the geometric analysis offers a powerful technique to objectively analyse the plausibility of this scene. Using the same principle, the line connecting the bus stop's shadow and the corresponding point on the object does not intersect the scene's light source. Therefore, this inconsistency indicates that the image has been tampered with—and demonstrates how shadows can be helpful in detecting forgeries (Farid, 2016; Kee et al., 2013). Research has shown that this shadow-based analysis provides a useful forensic tool (Kee et al., 2013), but the question of whether people can use shadow information to help identify image forgeries remains largely unexplored.

---

[5] In a 3-D scene a line connects the shadow point and object point, and intersects the light source. The transformation of the 3-D world coordinates to 2-D image coordinates means that in a 2-D image of the scene, the line connects the images of the shadow point and object point, and intersects the projected image of the light source (e.g., Kee et al., 2013).

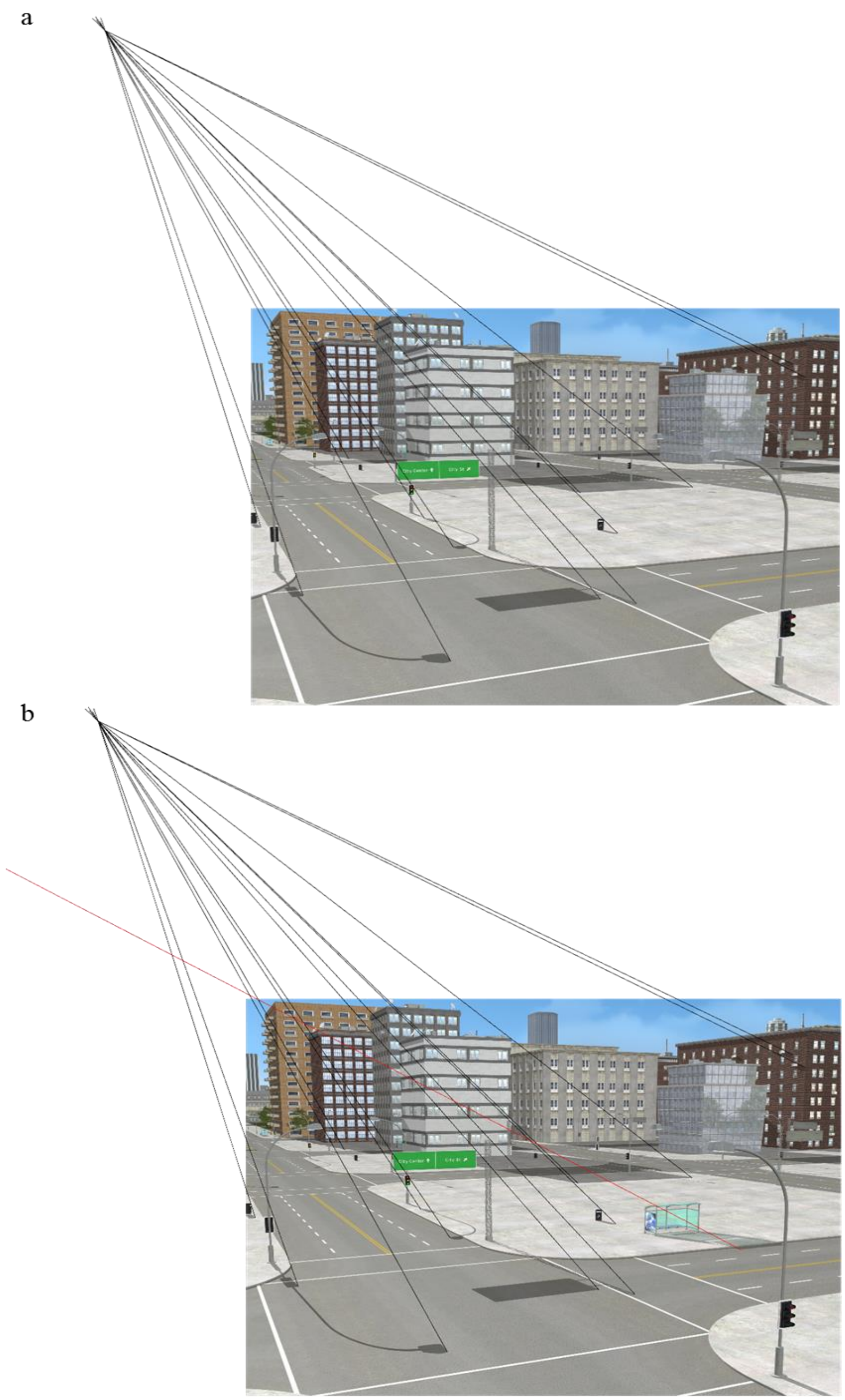


Figure 4.1. Example of using the shadow-based analysis technique: (a) the lines connecting the corresponding points of the shadows and objects intersect at a single point, indicating that the shadows are consistent with a single light source; (b) the same scene is shown with a bus stop added—the line connecting the bus stop's shadow and the corresponding point on the object does not intersect the scene light source, highlighting an inconsistency.

On the one hand, we might predict that people will be able to make use of shadow information to help identify image forgeries. Shadows convey important information about the arrangement and spatial position of objects in a scene and numerous studies have demonstrated that the human perceptual system makes use of such information to understand the world (Dee & Santos, 2011). In particular, studies have shown that shadow information supports the perception of depth and spatial position of objects in a scene (Allen, 1999; Kersten, Knill, Mamassian, & Bülthoff, 1996; Kersten, Mamassian, & Knill, 1994; Mamassian, Knill, & Kersten, 1998; Tarr, Kersten, & Bülthoff, 1998). And although the evidence is mixed, there is some suggestion that shadow information supports object recognition (Castiello, 2001; Cavanagh & Leclerc, 1989; Leek, Davitt, & Cristino, 2015; Tarr et al., 1998; but see Braje, Legge, & Kersten, 2000). Furthermore, a number of studies have shown that people are sensitive to changes in lighting direction in controlled settings using simplistic stimuli (e.g., Khang, Koenderink, & Kappers, 2006; Koenderink, Van Doorn, & Pont, 2004; O'Shea, Agrawala, & Banks, 2010). In one of the first studies to investigate the perception of inconsistent shadows, people completed a visual search task to determine whether a target cube that was illuminated from a different direction than the distractor cubes was present or absent in a series of displays (Enns & Rensink, 1990). Subjects rapidly identified the presence or absence of the target cube, suggesting that the human visual system can process complex visual properties, such as lighting direction, at a preattentive stage and in parallel across the image. This remarkable ability to perceive shadow information suggests that such information might also help in the detection of image forgeries.

On the other hand, there is also evidence to support the opposite prediction: that people will not be able to make effective use of shadow information. In Enns and Rensink's (1990) study, all of the cubes in the display were identical in shape, size, and orientation and it has been proposed that this object homogeneity might account for subjects' ability to automatically identify illumination inconsistencies. To check, other researchers ran an extension of the original study, this time using cubes that were randomly orientated in space to better represent real-world situations (Ostrovsky et al., 2005). In contrast to the results of the original study, people had difficulty recognising when the objects in the display were illuminated from a consistent or inconsistent direction. In a follow-up experiment, using real-world scenes that were presented for either 1000, 2000, or 5000 ms, people again had surprising difficulty identifying illumination inconsistencies. Interestingly, however, subjects' performance improved with longer presentation times suggesting that detecting lighting inconsistencies requires a relatively slow scan of the scene and involves processing objects serially rather than in parallel. Perhaps people are able to estimate the lighting direction for an individual object (Enns & Rensink, 1990), but difficultly arises when the task involves computing and accumulating the lighting direction for a number of different objects across the entire scene (Ostrovsky et al., 2005).

More recently, Farid and Bravo (2010) investigated whether people could make use of shadow cues in a scene to detect lighting inconsistencies. The researchers created a number of computer-generated scenes depicting simple geometric shapes. Half of the scenes were illuminated by a single light source that generated consistent shadows in the scene; the other half were illuminated by two different light sources that generated inconsistent shadows in the scene. Subjects were given an unlimited amount of time to judge whether the scenes portrayed shadows that were consistent or inconsistent with a single light source. When the inconsistencies were obvious—lights on opposite sides of the scene and shadows that ran in opposite directions—subjects showed an almost perfect ability (95.5%) to detect the inconsistent scenes. Yet when the inconsistencies were subtle—shadows that were a result of two light sources in slightly different locations on the same side of the scene—subjects detected just 52.8% of the inconsistent shadow scenes. Although subjects had difficulty determining whether the shadows in these scenes were consistent or inconsistent with a single light source, the near-perfect performance on trials where the lights were on opposite sides of the scene suggests that there might be a point at which lighting inconsistencies become noticeable.

Indeed, attempts have been made to quantify the point at which people notice lighting inconsistencies in images of outdoor scenes (Lopez-Moreno, Sundstedt, Sangorrin, & Gutierrez, 2010; Tan, Lalonde, Sharan, Rushmeier, & O'Sullivan, 2015). For example, Tan et al. created inconsistent scenes with varying degrees of error between the original and second light source. The researchers used real-world scenes in Experiment 1 but owing to difficulties controlling the lighting conditions they ran a second experiment using computer-generated scenes. Generally, subjects detected more of the inconsistent scenes when the angle difference between the two lighting positions was larger than when it was smaller. Yet the results were not conclusive; there were instances in which the smallest illumination changes were noticed and the largest changes were not. These results indicate that the extent of error between the two light sources is not the only factor that affects the sensitivity of the human perceptual system to lighting inconsistencies—other factors might also be important, such as scene content and layout (Xia, Pont, & Heynderickx, 2016). It is important to note, however, that in this experiment the inconsistent light source was moved in relatively coarse 30° increments from the original light source (Tan et al., 2015). Therefore, it remains to be seen whether more granular angle differences between the original and inconsistent light position affect the judgement of lighting inconsistencies in a similar way.

In sum, previous research has shown that people can make effective use of shadow information for certain perceptual tasks, including estimating the lighting direction in scenes consisting of simple geometric shapes. Yet a growing body of research suggests that the ability to estimate lighting direction does not extend to more complex stimuli. It remains unknown, however, whether people might be able to identify consistent and inconsistent shadows when there is enough information in the scene and sufficient time to make a judgment. The research presented in the current chapter examined this possibility. New image stimuli were created for Experiment 1a that contained a number of objects with well-defined edges and clearly visible shadows. These features made it possible for subjects to use the shadow-based analysis and determine the position of the light source by finding the intersection of the lines that connect the corresponding points of shadows and shadow-casting objects. Indeed, subjects were instructed to use the shadows in the scene to guide their judgement. Under these circumstances, people might be able to make effective use of the shadow information in the scenes to judge whether the shadows are consistent or inconsistent with a single light. In addition, we further explore the possibility of a perceptual threshold for detecting illumination inconsistencies in scenes.

Experiment 1a

Method

Subjects and Design

A total of 69 students (*M* = 18.9 years, *SD* = 2.2, range = 17-31; 62 women, 7 men) from Warwick University completed the experiment online in return for course credit. A further 11 subjects were excluded from the analyses, 8 who had missing response time data for at least one response on the task and 3 who failed to understand instructions. Subject recruitment continued until we reached a minimum of 20 responses per image (the stimulus set consisted of four consistent shadow images and eight inconsistent shadow images, further details of the images are provided in the following Stimuli section). The design was within-subjects: each subject viewed a series of four computer-generated images, half of which had consistent shadows, and half of which were manipulated to show inconsistent shadows. Using a two-alternative forced choice method (2AFC), we measured people's accuracy in determining whether an image had consistent or inconsistent shadows.

Stimuli

To create five different outdoor city scenes, we used a 3-D cityscape model from turbosquid.com and a 3-D animation software called Maya® (2016; Autodesk, Inc.). To represent a real-world outdoor environment lit by the sun, each scene was illuminated by a single light source. Each scene included a target object—a lamppost—and its corresponding shadow. In addition, to ensure subjects could use the shadow-based analysis technique outlined in the introduction, we also made a number of other non-target objects and their corresponding shadows visible in the scene. Recall that when a scene is illuminated by a single source all of the shadows in the scene must be consistent with that light; if any shadow is inconsistent with the light source, then the scene is physically impossible (Farid, 2016; Kee et al., 2013). We rendered[6] each of the five 3-D scenes from Maya® to generate TIF image files with a resolution of 960×720 pixels. To ensure that the shadows in the 2-D images were physically accurate and therefore representative of the shadows that people experience in the real world we rendered the

[6] Rendering is a computer process to automatically convert 3-D models into 2-D images

images with raytraced[7] shadows. These five scenes comprised our original, consistent image set—each illuminated by a single source and thus containing only consistent shadows.

To create the inconsistent shadow scenes we rendered each of the five scenes from Maya® (2016; Autodesk, Inc.) two more times: once with the light moved to the left of its original position (-800 m on the x-axis) and once with the light moved to the right of its original position (+800 m on the x-axis). The scene layout, for instance the position of the objects, remained identical across each version of the scene, yet the three different light positions—original, left and right—meant that each version had a different shadow configuration. For each of the five scenes, we selected a single lamppost and its corresponding shadow to manipulate. The manipulation process involved three stages, all carried out using GNU Image Manipulation Program® (GIMP, Version 2.8). First, we removed the target lamppost's shadow in the original version of the scene. Second, we cut the shadow of that same target lamppost from the version of the scene with the light moved left of the original position. Third, we overlaid this shadow onto the original version of the scene. We then repeated stages two and three for the version of the scene with the light moved right of the original position (see Figure 4.2 for an example of the editing process). We exported the images as PNGs which is a lossless format. We repeated this manipulation process for the other four scenes.

Overall, we had three versions of each of the five city scenes, a total of 15 images. The original version of each scene was used to create our consistent shadow image set. And the two manipulated versions of each scene were used to create our inconsistent shadow image set. Subjects saw two consistent and two inconsistent shadow images but always in a different city scene. The fifth city scene was used in the practice.

[7] Raytracing is a type of shadow rendering that calculates the path of individual light rays from the light source to the camera—it produces physically accurate shadows that are like shadows in the real world (Autodesk, 2016)

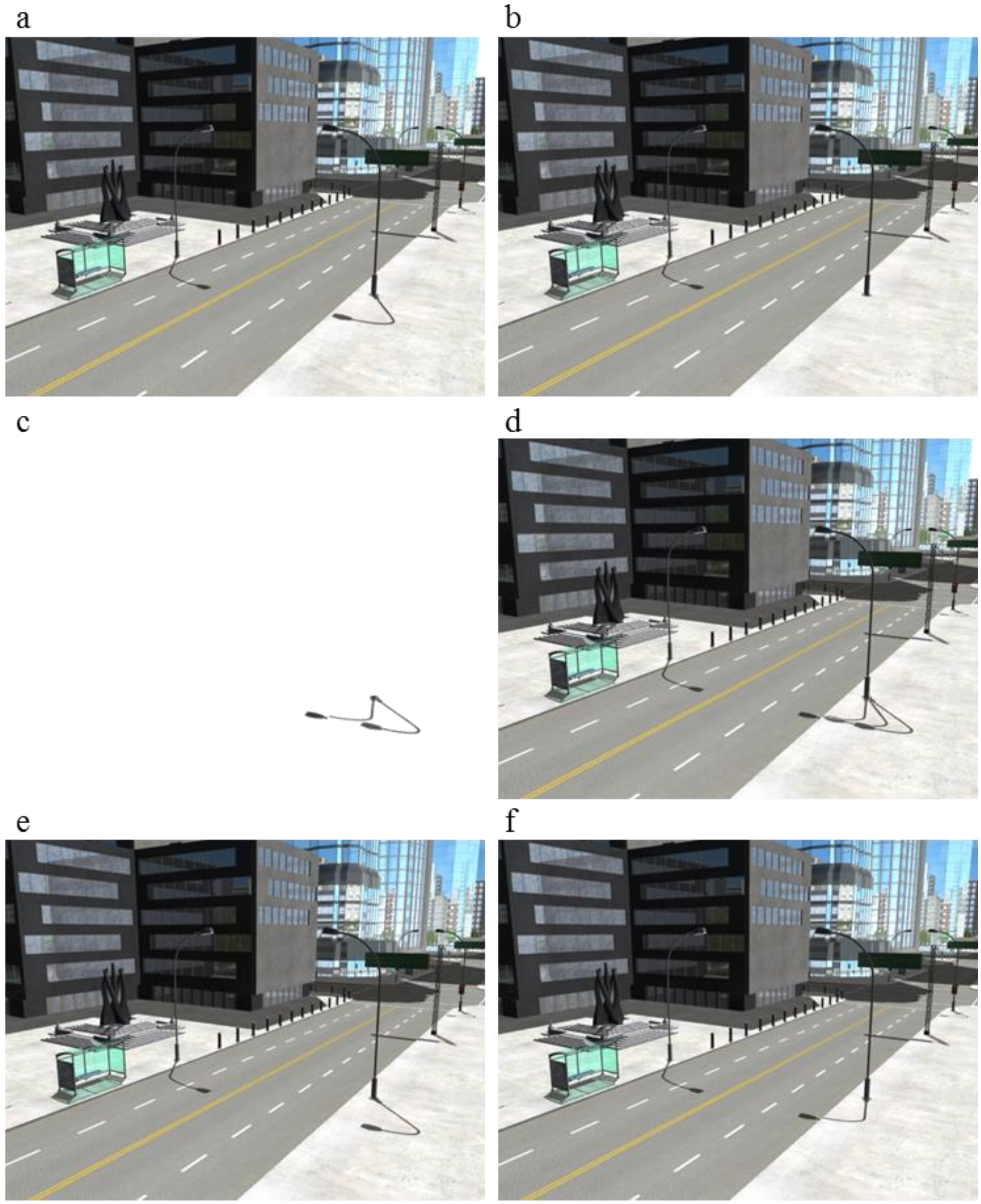


Figure 4.2. Example of the image manipulation process: (a) original scene with consistent shadows; (b) original scene with the target lamppost shadow removed; (c) isolated shadows cut from the scenes with the light position moved left and right of the original light position; (d) original, left, and right light position shadows all shown in the scene; (e) left light position shadow added to the original scene, the shadow of the target lamppost is inconsistent with all of the other shadows in the scene; (f) right light position shadow added to the original scene, the shadow of the target lamppost is inconsistent with all of the other shadows in the scene. Each subject saw this city scene just once, they were randomly shown a, e, or f.

Procedure

First, **s**ubjects received detailed instructions about the experiment. Importantly, we informed subjects they could assume that "…each of the scenes is illuminated by a single light source, such as the sun." Before the experiment proper began, subjects were given a practice trial. To cue subjects' attention to the target lamppost, they were first shown an almost entirely greyed out image with only the target lamppost fully visible and highlighted in a red ellipse. After four seconds, the full scene automatically became visible. We also added a small yellow dot on the base of the target lamppost to ensure subjects did not forget which lamppost to base their response on. Subjects were asked "Is the lamppost's shadow consistent or inconsistent with the shadows in the rest of the scene?" They were given unlimited time to select between two response options: (a) "Consistent", (b) "Inconsistent." They were then asked to rate their confidence in their decision using a 100-point Likert-type scale ranging from 0 (*Not at all confident*) to 100 (*Extremely confident*).

Then the experiment proper began. Subjects were presented with the four city scenes in a random order. Each subject saw two consistent shadow scenes and two inconsistent shadow scenes, however, they were unaware of this 50:50 ratio. For each scene, subjects completed exactly the same procedure as described for the practice trial. After completing the shadow task, subjects were asked a series of questions about their demographics, interest in photography, and video gaming experience. They were also asked whether they had experienced any technical difficulties while completing the experiment. Subjects received feedback on their performance at the end of the study.

Results and Discussion

An analysis of the response time data suggested that subjects were engaged with the task and spent a reasonable amount of time determining whether the shadows in the scenes were consistent or inconsistent with a single light source. The mean response time per image was 17.9 s (*SD* = 18.8 s) and the median response time 11.6 s (interquartile range: 7.3, 21.4 s).

Overall accuracy on the shadow task

We now turn to our primary research question: Can people identify whether scenes have consistent or inconsistent shadows? Overall, a mean 54% of the scenes were correctly classified, 95% CI [48%, 60%]. Subjects' ability to distinguish between consistent (64% correct, 95% CI [56%, 73%]) and inconsistent (44% correct, 95% CI [36%, 53%]) shadow scenes was not reliably greater than zero, d' = 0.17, 95% CI [-0.06, 0.39]. Furthermore, subjects showed a bias towards accepting the shadow scenes as consistent, c = 0.20, 95% CI [0.09, 0.30]. These results indicate that subjects' ability to identify whether the shadows in a scene are consistent or inconsistent with a single light source was extremely limited. Given this result, it appears that subjects did not make use of the information available within the scene to objectively work out the answer. Instead, the results suggest that subjects relied on a subjective visual analysis of the scene and, as indicated by their bias, were often willing to accept that the shadows in the scenes were consistent. We next consider whether any individual factors or image metrics were associated with an improved ability to identify consistent and inconsistent shadow scenes.

Individual factors and image metrics

To determine whether individual factors play a role in identifying consistent and inconsistent shadows, we gathered subjects' demographic data, as well as details about their interest in photography and video gaming experience. We also asked subjects to rate their confidence for each of their decisions and recorded their response time. In addition, we checked whether three properties of the image itself affected people's accuracy on the task. One image property was simply whether the light position had moved left or right of the original light position. The second image property was the location of the light source: For each of the four scenes, we measured the distance from the centre of the scene to the light source. The third image property was a measurement of the rotation, in degrees, from the consistent shadow position to the inconsistent shadow position.

To check how each factor influenced subjects' performance, we conducted two generalized estimating equation (GEE) analyses—one for the inconsistent shadow scenes and one for the consistent shadow scenes. Specifically, we conducted a repeated measures logistic regression with GEE because our dependant variables were binary with both random and fixed effects (Liang & Zeger, 1986). The results of the GEE analyses are shown in Table 4.1.

The GEE analyses revealed that none of the variables had an effect on subjects' ability to accurately identify inconsistent shadow scenes. Only one variable had an effect on subjects' ability to accurately identify consistent shadow scenes, and that was subjects' confidence in their decision. More confident responses were slightly more likely to be associated with accurate responses than less confident responses.

Table 4.1
Results of the GEE binary logistic regression models to determine variables that predict accuracy in the shadow task

	Inconsistent			Consistent		
Predictor	*B*	*OR [95% CI]*	*p*	*B*	*OR [95% CI]*	*p*
Confidence	0.00	1.00 [0.98, 1.01]	.56	0.03	1.03 [1.01, 1.05]	<.001
Video gaming = Frequent (at least once a month)	0.83	2.30 [0.64, 8.28]	.20	1.06	2.88 [0.72, 11.60]	.14
Response time	0.00	1.00 [0.98, 1.02]	.87	-0.01	0.99 [0.97, 1.01]	.23
Gender = Female	-0.12	0.89 [0.21, 3.66]	.87	0.66	1.93 [0.46, 8.13]	.37
Interest in photography = Interested	0.39	1.47 [0.76, 2.85]	.25	-0.29	0.74 [0.34, 1.65]	.47
Distance to light source	0.00	1.00 [0.98, 1.01]	.38	0.01	1.01 [1.00, 1.02]	.07
Light position = Left	0.41	1.51 [0.69, 3.28]	.30	-	-	-
Angle difference	-0.01	0.99 [0.97, 1.02]	.56	-	-	-

Note. CI = confidence interval. *B* and odds ratios (*OR*) estimate the degree of change in accuracy associated with one unit change in the independent variable. An odds ratio of 1 indicates no effect of the independent variable on accuracy, values of 1.5, 2.5, and 4.0 are generally considered to reflect small, medium, and large effect sizes, respectively (Rosenthal, 1996). The category order for factors was set to descending to make the reference level 0. The reference groups are: Video game playing = Infrequent (never/less than once a month), Gender = Male, Interest in photography = Not Interested, Light position = Right. Response time, confidence, distance of light source from the scene, and angle difference were added as continuous variables. All subjects were included in these analyses $N = 69$. The light position and angle difference predictor variables were not applicable in the consistent shadow scenes.

Our results align with the growing body of literature suggesting that people are quite insensitive to lighting inconsistencies (e.g., Farid & Bravo, 2010; Ostrovsky et al., 2005). Extending on previous research, we demonstrate that this insensitivity to lighting inconsistencies persists even when there is information available in the scene to objectively determine the correct answer. Further, our results do not provide any strong evidence to support an association between a range of individual factors and image properties with an improved ability to identify consistent and inconsistent shadow scenes. One outstanding question is whether these findings generalise to a wider group of people. The subjects in Experiment 1a were first year Psychology students who participated for course credit, it is possible that we might find different results amongst individuals who choose to complete the task without an extrinsic incentive. One possibility is that individuals who choose to take part are interested in the task and perhaps are more motivated to perform well. We conducted Experiment 1b to examine whether these findings are tied to a specific sample.

Experiment 1b

Method

Subjects and Design

A total of 102 subjects (*M* = 25.5 years, *SD* = 9.0, range = 14-57; 60 men, 39 women, and 3 chose not to disclose their gender) were recruited off campus and completed the task online. A further 4 subjects were excluded from the analyses, 3 who had missing response time data for at least one response on the task and 1 who experienced technical difficulties. There were no geographical restrictions and subjects did not receive payment for taking part, but they did receive feedback on their performance at the end of the task. Subject recruitment continued until we reached a minimum of 20 responses per image. The design was identical to that of Experiment 1a.

Stimuli and Procedure

The stimuli and procedure were unchanged from Experiment 1a.

Results and Discussion

An analysis of the response time data suggested that subjects were engaged with the task and spent a reasonable amount of time determining whether the shadows in the scenes were consistent or inconsistent with a single light source. The mean response time per image was 16.9 s (*SD* = 9.1 s) and the median response time was 14.1 s (interquartile range: 10.2, 23.1 s).

Overall accuracy on the shadow task

Overall, subjects correctly classified a mean 61% of the shadows scenes (cf. Expt 1a: 54%), 95% CI [56%, 65%]. Subjects had some ability to discriminate between consistent (75% correct, 95% CI [70%, 81%]) and inconsistent (46% correct, 95% CI [38%, 53%]) shadow scenes, d' = 0.41, 95% CI [0.22, 0.59]. These findings offer further empirical support for the idea that people are quite insensitive to lighting inconsistencies (e.g., Farid & Bravo, 2010; Ostrovsky et al., 2005). Although subjects' ability to tell the difference between consistent and inconsistent shadow scenes was slightly better than in Experiment 1a (cf. Expt 1a: d' = 0.17), the difference did not reach significance $t(163)$ = 1.64, p = .10, d = 0.26. As in Experiment 1a, subjects showed a bias towards accepting the shadow scenes as consistent, c = 0.29, 95% CI [0.20, 0.38]. This bias tells us that subjects had a relatively conservative criterion for judging that shadows were inconsistent with the scene light source and typically accepted them as consistent.

Individual factors and image metrics

As in Experiment 1a, we conducted two GEE analyses—one for the inconsistent shadow scenes and one for the consistent shadow scenes. We included the same factors used in the GEE models in Experiment 1a. The results of the GEE analyses are shown in Table 4.2. This time we found that two of the variables had an effect on the likelihood of responding correctly: video gaming and angle difference. Those who play video games frequently were more likely to correctly identify inconsistent shadow scenes than those who do not play video games frequently. There was also a small effect of angle difference—inconsistent shadows positioned further from the correct position were more likely to be associated with accurate responses than inconsistent shadows positioned closer to the correct position. It seems, then, that there might be a discernible point at which the inconsistent shadow becomes different enough from its consistent position to make the inconsistency noticeable—lending support to the notion of a perceptual threshold for detecting lighting inconsistencies (Lopez-Moreno et al., 2010; Tan et al., 2015). In other words, our subjects appeared to hold a basic understanding about where an object's shadow must cast to be consistent with the light source. Yet this understanding was not very precise—subjects were willing to accept a shadow as consistent when cast in a range of locations that were relatively near to its correct location.

For the consistent shadow scenes, the results revealed that the distance of the light source from the scene had a small effect on likelihood to respond correctly. Specifically, scenes in which the light was closer to the scene were more likely to be identified as consistent compared with scenes in which the light was further from the scene. One possible reason for this effect is that subjects were better able to determine the accuracy of shadows in a scene when the light source was more readily available to use as a guide. Perhaps, then, our subjects were able to make use of the shadow-based analysis technique, but only when it was relatively easy to calculate the location of the light source.

Table 4.2

Results of the GEE binary logistic regression models to determine variables that predict accuracy in the shadow task

	Inconsistent			Consistent		
Predictor	*B*	*OR [95% CI]*	*p*	*B*	*OR [95% CI]*	*p*
Confidence	-0.01	0.99 [0.98, 1.01]	.44	0.01	1.01 [0.99, 1.03]	.16
Video gaming = Frequent (at least once a month)	0.75	2.12 [1.05, 4.24]	.03	0.15	1.16 [0.49, 2.75]	.73
Response time	0.00	1.00 [0.98, 1.02]	.93	0.00	1.00 [0.98, 1.03]	.76
Gender = Female	-0.04	0.96 [0.48, 1.91]	.91	0.24	1.28 [0.54, 3.04]	.58
Interest in photography = Interested	0.42	1.52 [0.76, 3.05]	.24	0.33	1.39 [0.55, 3.55]	.49
Distance to light source	0.00	1.00 [0.99, 1.01]	.78	-0.02	0.98 [0.97, 0.99]	<.001
Light position = Left	0.11	1.12 [0.65, 1.92]	.68	-	-	-
Angle difference	0.02	1.02 [1.00, 1.05]	.03	-	-	-

Note. CI = confidence interval. *B* and odds ratios (*OR*) estimate the degree of change in accuracy associated with one unit change in the independent variable. An odds ratio of 1 indicates no effect of the independent variable on accuracy, values of 1.5, 2.5, and 4.0 are generally considered to reflect small, medium, and large effect sizes, respectively (Rosenthal, 1996). The category order for factors was set to descending to make the reference level 0. The reference groups are: Video game playing = Infrequent (never/less than once a month), Gender = Male, Interest in photography = Not Interested, Light position = Right. Response time, confidence, distance of light source from the scene, and angle difference were added as continuous variables. The three subjects who chose not to disclose their gender were excluded from these analyses leaving a total sample of n = 99. The light position and angle difference predictor variables were not applicable in the consistent shadow scenes.

Recall that the stimuli used in Experiments 1a and 1b were identical. Why, then, did we find that two of the image properties—angle difference and distance of the light—had an impact on subjects' performance in Experiment 1b but not 1a? One possibility is that the incentive to take part was an important factor: In Experiment 1a, subjects were incentivised with course credit; in Experiment 1b the only incentive was that subjects received their results at the end of the task. As such, we can speculate that the subjects in Experiment 1b might have been more interested in the task and more motivated to do well than the subjects in Experiment 1a. Offering some support for this suggestion, previous research has shown that using shadow information to determine the position of the light source in complex scenes is not an automatic process but instead requires effortful encoding (Langer & Zucker, 1997; Ostrovsky et al., 2005). It follows, then, that if people are more interested in the task, they might take a more considered and effortful approach and perhaps therefore give more attention to the image properties.

As well as a small effect of the two image properties, video gaming influenced the likelihood of responding correctly: those who play video games frequently were more likely to correctly identify inconsistent shadow scenes than those who do not play video games frequently. This finding makes sense when considering the substantial body of evidence showing that video gamers outperform non-video gamers across a range of perceptual measures (e.g., Boot, Kramer, Simons, Fabiani, & Gratton, 2008; Feng, Spence, & Pratt, 2007; Green & Bavelier, 2003, 2006, 2007; for a review, see Green & Bavelier, 2012). For example, research has demonstrated that video gaming improves people's selective attention, that is, their ability to choose which aspects of a stimulus are task-relevant and should receive additional processing, while filtering out task-irrelevant items (e.g., Green & Bavelier, 2003, 2006, 2007). Given that detecting lighting inconsistencies is not an automatic process but instead requires a relatively slow scan of the scene to carefully process objects and their shadows in a serial manner (Langer & Zucker, 1997; Ostrovsky et al., 2005), it is likely that selective attention is important in the shadow task. Accordingly, our finding that video gamers were more likely to correctly identify inconsistent shadow scenes than non-gamers is concordant with the broader literature (e.g., Green & Bavelier, 2003, 2006, 2007; Ostrovsky et al., 2005).

In Experiment 1b, subjects were slightly more likely to identify the inconsistent shadows when the angle difference from the correct shadow location was larger compared to when it was smaller. Yet the design of the experiment meant that there were only eight inconsistent shadow scenes and thus only eight angle differences to consider. In Experiment 2a, to more precisely estimate the perceptual threshold for identifying lighting inconsistencies, we asked subjects to move a target shadow to the exact position that they thought was consistent with the lighting of the scene.

Experiment 2a

Method

Subjects and Design

A total of 109 subjects (*M* = 25.1 years, *SD* = 9.7, range = 14-64; 41 women, 65 men, and 3 chose not to disclose their gender) completed the task online. Two additional subjects were excluded from the analyses because they experienced technical difficulties. There were no geographical restrictions and subjects did not receive payment for taking part. Subject recruitment continued until we reached a minimum of 100 responses per scene. In a within-subjects design, each subject viewed a series of four computer-generated images and, for each image, attempted to position a target shadow so that it was consistent with the scene light source. We measured subjects' accuracy in determining the correct position for the target shadow.

Stimuli

We used the same five original city scenes as in Experiments 1a and 1b. In GNU Image Manipulation Program® (GIMP, Version 2.8), we removed the target lamppost's shadow from each of the original scenes and saved the shadows as new PNG image files. We developed a program in HTML to display the original scene with the shadow image overlaid but rotated at a random angle between -45° and +45° of the correct shadow location. This time then, there was one version of each scene, but the initial position for the target shadow was randomised.

Procedure

Subjects were first given detailed instructions about the experiment and informed that they could assume that "…each of the scenes is illuminated by a single light source, such as the sun." Before the experiment proper began, subjects were given a practice trial. To cue subjects' attention to the target lamppost (i.e., the lamppost they would base their response on), the scene first appeared with a red ellipse around the target lamppost then the ellipse disappeared automatically after 2 s. Subjects were asked to "Please use the left and right arrow keys to change the shadow rotation and then press enter when you think it is consistent with the other shadows in the scene." Subjects were given unlimited time to rotate the target shadow and provide their answer. They were then asked to rate their confidence in their decision using a 100-point Likert-type scale ranging from 0 (*Not at all confident*) to 100 (*Extremely confident*).

Next the experiment proper began. Subjects were presented with the four city scenes in a random order. For each scene, subjects completed exactly the same procedure as described for the practice trial. After completing the shadow task, subjects were asked whether they had experienced any technical difficulties while completing the experiment.

Results and Discussion

An analysis of the response time data suggested that subjects were engaged with the task and spent a reasonable amount of time positioning the target shadow in the scenes. The mean response time per image was 57.3 s (*SD* = 419.3 s) and the median response time was 23.9 s (interquartile range: 15.8, 39.7 s).

Overall accuracy on the shadow task

In the shadow task, subjects rotated the target shadow about its base to the position they judged to be consistent with the scene's light source. The target shadow could be rotated 360°, with 0° representing the position in which the shadow was consistent with the scene's light source. Clockwise rotation gave positive values from 0.1° to 179.9°, indicating that the shadow was placed left of its consistent position. Counter-clockwise rotation gave negative values from -0.1° to -179.9°, indicating that the shadow was placed right of its consistent position. This range of shadow positions makes it possible to consider accuracy on the task in various ways. First, taking an extremely conservative approach and considering an accurate response to be between -1° and +1°, across the four scenes, subjects were accurate just 7% of the time, 95% CI [5%, 10%]. Taking a slightly more lenient approach, 51% of the shadows were positioned between -10° and +10°, 95% CI [46%, 56%]. And 95% of the shadows were positioned between -40° and +40° of the correct location, 95% CI [93%, 97%]. The remaining 5% of shadows were positioned as widely as -150° and +150° of the correct location. As in Experiment 1b, these results indicate that subjects had only a basic understanding about where an object's shadow must cast to be consistent with the light source. Our results also offer empirical support for the notion of a perceptual threshold for noticing lighting inconsistencies in images (Lopez-Moreno et al., 2010; Tan et al., 2015). To more precisely estimate the perceptual system's sensitivity to lighting inconsistencies than in previous research, we allowed subjects to rotate the target shadow through 360° at 0.1° increments. Despite this high level of control, subjects were still willing to rotate the target shadow to a relatively wide range of positions that were inconsistent with the scene lighting. Therefore, people's perceptual threshold for accepting shadows as consistent with a single light source appears to be surprisingly wide.

To further understand subjects' perception of illumination inconsistencies we calculated accuracy by scene. Figure 4.3 displays the angle difference from the correct shadow position (0°) on the x-axis where smaller angle differences indicate that the shadow was positioned closer to its correct position than larger angle differences. The cumulative proportion of responses that were made by each angle difference level are presented on the y-axis. The light grey line with circle markers shows the overall proportion of responses made by each angle difference level and therefore includes both negative and positive values, for example 10° on the graph represents responses between -10° and +10°. We also calculated the number of positive and negative responses made by each angle difference level and displayed these as a proportion of the total responses. The black line with triangle markers shows the proportion of negative angle difference responses (shadow to the right of the consistent position). The dark grey line with square markers shows the proportion of positive angle difference responses (shadow to the left of the consistent position). This analysis revealed that subjects' performance varied by scene. Specifically, subjects were most accurate in Scene 4 where the target shadow was positioned between -5° and +5° of the correct location by 50% of subjects, 95% CI

[41%, 60%]. On the other hand, subjects were least accurate in Scene 3, where the target shadow was positioned between -5° and +5° of the correct location by only 15% of subjects, 95% CI [8%, 21%]. In line with this finding, previous research has shown that the detection of illumination inconsistencies is affected by various factors, such as the scene content and layout (Tan et al., 2015; Xia et al., 2016). Specifically, if a scene contains a number of objects that are all orientated in the same direction, then lighting inconsistencies are typically easier to notice (Enns & Rensink, 1990). Therefore, our results offer support for the idea that the perception of shadow inconsistencies is influenced by the configuration of the scene (Ostrovsky et al., 2005; Tan et al., 2015; Xia et al., 2016). Next we consider whether subjects showed a preference to position the target shadow to the left or to the right of the consistent position.

Preference for shadows to the left or right

Across the four scenes, we classified each response according to whether subjects positioned the shadow to the left or right of the correct position. The results revealed that a mean 20% more of the shadows were positioned to the left of the correct location than to the right of the correct location, M_{diff} 95% CI [12%, 28%]. Yet, as Figure 4.3 shows, this trend was not found across all four scenes. Subjects were more likely to rotate the shadow to the left (positive angle difference) of the correct position in Scenes 1 and 3 (Scene 1: 84%, 95% CI [77%, 91%]; Scene 3: 75%, 95% CI [67%, 83%]). Conversely, in Scene 2, 73% of subjects rotated the shadow so that it appeared to the right (negative angle difference) of its correct position, 95% CI [65%, 82%]. And in Scene 4, a similar proportion of subjects positioned the shadow to the left of its correct location (54%) as to the right (46%). These findings contradict research suggesting a leftward bias for illumination position (Mamassian & Goutcher, 2001; Sun & Perona, 1998; Symons, Cuddy, & Humphrey, 2000). Yet the research showing that people hold a prior assumption that light comes from above-left has typically relied on simplistic stimuli, such as shaded disks. Thus it is possible that the leftward bias for light position might not apply to more complex stimuli. In fact, more recent studies have shown that prior assumptions about lighting direction are unimportant in everyday perception (Morgenstern, Geisler, & Murray, 2014; Morgenstern, Murray, & Harris, 2011).

Specifically, any prior assumptions people might hold are easily overridden by lighting direction cues such as shading and shadows—our findings support this account.

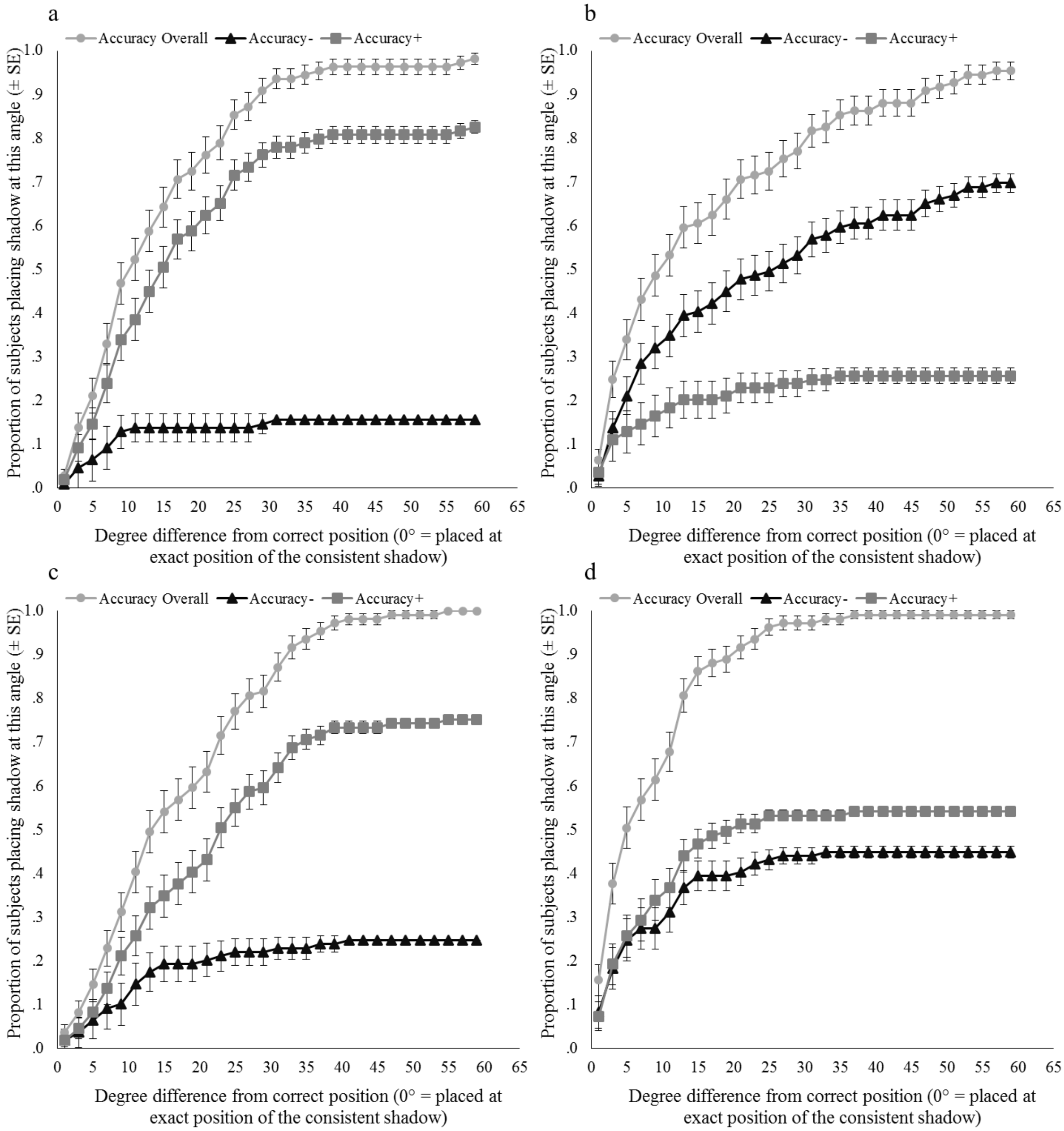


Figure 4.3. Cumulative proportion of responses made by each angle difference level for (a) Scene 1 (b) Scene 2 (c) Scene 3 and (d) Scene 4. The light grey line with circle markers shows the overall proportion of responses made by each angle difference level and therefore includes both negative and positive values—for example, 10° on the graph represents responses between -10° and +10°. The black line with triangle markers shows the proportion of negative angle difference responses (shadow to the right of the consistent position). The dark grey line with square markers shows the proportion of positive angle difference responses (shadow to the left of the consistent position). Error bars represent standard error of the mean.

In sum, our results in Experiment 2a further support the idea that subjects typically make imprecise judgements about where shadows must be positioned to be consistent with a single light source. This result is somewhat surprising because there was sufficient information available in the scene to determine the location of the light and thus to allow an objective judgement. Given that only 51% of the shadows in Experiment 2a were positioned within +/- 10° of their correct location it suggests that subjects might not easily perceive this lighting information, thereby making it difficult to localise the light source. That said, because the target shadow used in Experiment 2a was taken from the consistent version of the scene it is possible that the shape of the target shadow provided subjects with a cue to its correct position. Therefore these results might actually *overestimate* people's ability to position a target shadow so it is consistent with the scene light source. To check this possibility we ran Experiment 2b in which subjects were not only able to rotate the shadow but also change the scale of the shadow.

Experiment 2b

Method

Subjects and Design

A total of 102 subjects (*M* = 25.1 years, *SD* = 10.2, range = 14-59; 47 women, 50 men, and 5 chose not to disclose their gender) completed the task online. Four additional subjects were removed because they experienced technical difficulties. There were no geographical restrictions and subjects did not receive payment for taking part. Subject recruitment continued until we reached a minimum of 100 responses per scene. The design was identical to that of Experiment 2a.

Stimuli and Procedure

We used the same stimuli and program as in Experiment 2a with one exception. We amended the program to allow subjects to adjust the scale of the shadow as well as the rotation. The scale was randomised so that the target shadow initially appeared between 0.5 and 1.5 of its correct scale. Subjects were instructed to "Please use the left and right arrow keys to change the shadow rotation and the up and down arrow keys to change the shadow size, then press enter when you think it is consistent with the other shadows in the scene."

Results and Discussion

An analysis of the response time data suggested that subjects were engaged with the task and spent a reasonable amount of time positioning the target shadow in the scenes. The mean response time per image was 34.7 s (*SD* = 27.3 s) and the median response time 28.6 s (interquartile range: 19.1, 42.5 s).

Overall accuracy on the shadow task

Replicating the result from Experiment 2a, across the four scenes, subjects placed the shadow between +/-1° of the correct shadow position just 7% of the time, 95% CI [4%, 9%]. With the more lenient classification for accuracy on the task, that is positioning the shadow between +/-10° of the correct position, 46% of shadows were positioned accurately (cf. Expt 2a: 51%), 95% CI [41%, 51%]. And 95% of the shadows were positioned between +/-40° of the correct location, 95% CI [93%, 97%]. Of the remaining shadows, 4.8% of shadows were positioned as widely as -124° and +124° of the correct location.[8] Recall that the results from Experiment 2a showed that subjects' accuracy on the shadow task varied by scene; as Figure 4.4 illustrates, we replicated this finding in Experiment 2b. Again, subjects were more accurate in positioning the shadow in Scene 4 than in the other three scenes. In Scene 4, 50% of shadows were positioned between +/-6.7° of the correct location, 95% CI [40%, 60%]. Also replicating the results from Experiment 2a, subjects were least accurate in Scenes 1 and 3, where the target shadow was positioned between +/-6.7° of the correct location by only 25% and 26% of subjects, respectively (Scene 1: 95% CI [17%, 34%]; Scene 3: 95% CI [18%, 35%]). As in Experiment 2a, and in line with previous studies, these results suggest a context effect—the orientation of the objects and resulting configuration of the shadow cues in the scene influenced performance on the task (Tan et al., 2015; Xia et al., 2016). Next, we considered the likelihood that subjects positioned the target shadow to the left or to the right of the consistent position.

[8] One outstanding shadow was positioned at -170°

Preference for shadows to the left or right

Replicating our finding from Experiment 2a, collapsed across the four scenes, subjects positioned 16% more of the shadows to the left of the correct position than to the right, M_{diff} 95% CI [7%, 25%]. Further, we also replicated our finding that this directional preference was context dependent—that is, it varied by scene. As illustrated in Figure 4.4, subjects were more likely to move the shadow to the left of the correct position in Scenes 1 and 3, but more likely to move the shadow to the right of the correct position in Scene 2. In Scene 4, subjects were equally likely to have positioned the shadow to the left or right of its correct location. By replicating our results in Experiment 2a, we offer further support for the view that lighting direction cues in the scene can override any prior assumptions about lighting (Morgenstern et al., 2011, 2014).

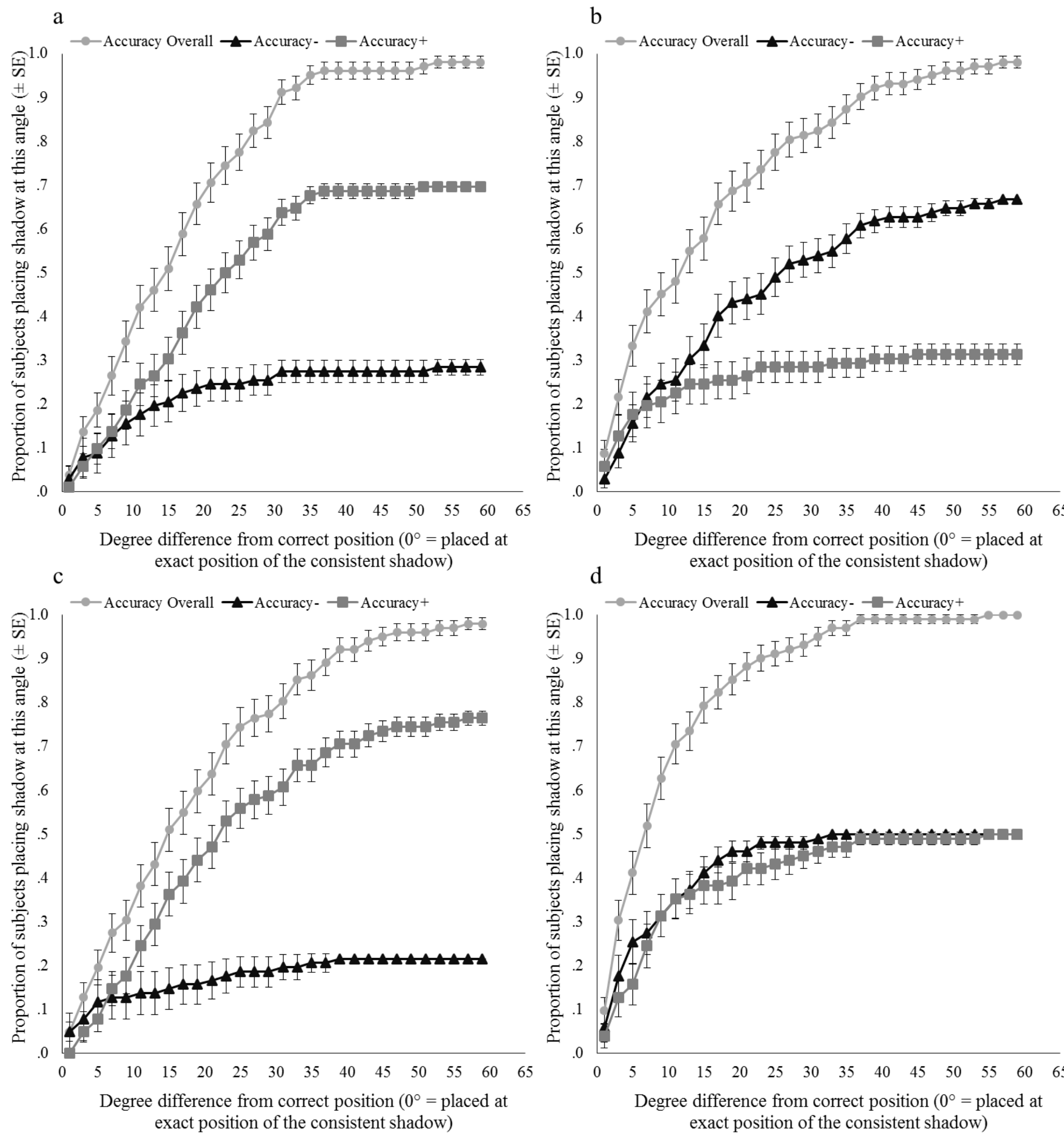


Figure 4.4. Cumulative proportion of responses made by each angle difference level for (a) Scene 1 (b) Scene 2 (c) Scene 3 and (d) Scene 4. The light grey line with circle markers shows the overall proportion of responses made by each angle difference level and therefore includes both negative and positive values—for example, 10° on the graph represents responses between -10° and +10°. The black line with triangle markers shows the proportion of negative angle difference responses (shadow to the right of the consistent position). The dark grey line with square markers shows the proportion of positive angle difference responses (shadow to the left of the consistent position). Error bars represent standard error of the mean.

In sum, allowing subjects to adjust the size of the target shadow in Experiment 2b made virtually no difference to the pattern of results. Therefore, it is possible that being able to change the scale of the target shadow did not prevent subjects using the shape of the shadow as a cue. If so, our results might still overestimate people's ability on the task. To check, we ran a fifth experiment in which we generated different versions of the target shadow that were inconsistent in both position and shape.

Experiment 3

Method

Subjects and Design

A total of 114 subjects (*M* = 25.6 years, *SD* = 9.0, range = 14-52; 48 women, 62 men, and 4 chose not to disclose their gender) completed the task online. Five additional subjects were removed because they experienced technical difficulties. There were no geographical restrictions and subjects did not receive payment for taking part. Subject recruitment continued until we reached a minimum of 100 responses per scene. We used a within-subjects design. Each person viewed a series of four computer-generated images, for each image, subjects used the left and right arrow keys to scroll through 11 possible shadow options for a single lamppost. We measured subjects' accuracy in selecting the shadow that was consistent with the scene light source.

Stimuli

We used the same five original city scenes as in the previous four experiments. This time, however, we created 21 versions of each scene in Maya® (2016; Autodesk, Inc.), each version with the objects in an identical position, but with 21 different light positions. In the consistent version, the target lamppost's shadow was created by the same light source as the rest of the scene. In the other 20 versions of the scene—the inconsistent versions—we created a second light source that only created a shadow for the target lamppost, nothing else in the scene. By changing the position of only the second light source we created 20 versions of the scene in which the shadow for the target lamppost was inconsistent with the shadow configuration for the rest of that scene. Furthermore, these shadows were inconsistent in terms of both position and shape. For 10 of the inconsistent versions of the scene we moved the second light source in 10 equal increments of 200 m to the left of the original light position. For the other 10, we moved the second light source in 10 equal increments of 200 m to the right of the original light position. As a result, we created 21 versions of each of the five scenes: one with consistent lighting for all objects in the scene—including the target lamppost—and 20 with consistent lighting for all objects except the target lamppost. The versions of the scene were numbered from 1 to 21, with the consistent version of the scene always number 11. Versions 10 to 1 were inconsistent, with the target shadow moving incrementally further to the left of the consistent version, while 12 to 21 were the inconsistent versions, with the target shadow moving incrementally further to the right of the consistent version (see Figure 4.5 for a sample of the versions of a scene).

We developed a program in HTML to randomly select one of the 21 versions of the scene to display. As well as this randomly selected version, subjects were able to scroll through a sequence of another 10 consecutive versions of that same scene—crucially, the sequence always included the consistent version. To illustrate, consider, for example, that the program randomly selects version 1, the subject would be able to scroll through versions 1 to 11 of the scene. Or, to consider another example, if the program randomly selects version 15, then the subject will be able to scroll through versions 5 to 15 of the scene. Having generated the sequence, the program randomised which of the 11 versions to display first, thus ensuring that subjects did not always start at the extreme end of a sequence. We programmed the left and right arrow keys on the keyboard to allow subjects to scroll through the 11 versions of the scene.

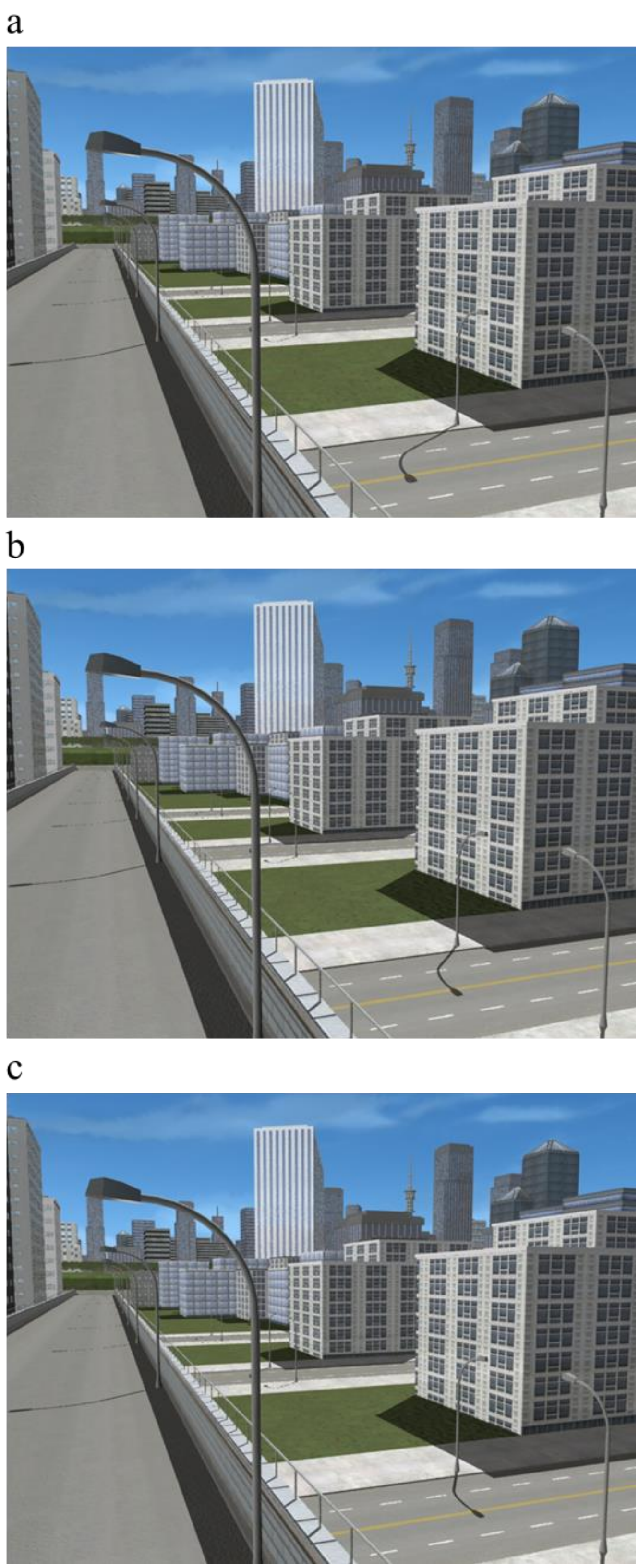


Figure 4.5. Example versions of Scene 3. (a) version 4, inconsistent; (b) version 11, consistent; (c) version 14, inconsistent.

Procedure

The procedure was the same as Experiment 2b, with one exception: subjects scrolled through the 11 versions of each scene rather than moving the shadow at will. We asked subjects to select the version of the scene in which the shadow of the target lamppost was consistent with the other shadows in the scene.

Results and Discussion

An analysis of the response time data suggested that subjects were engaged with the task and spent a reasonable amount of time determining which version of the scene was consistent with a single light source. The mean response time per image was 26.3 s (*SD* = 50.2 s) and the median response time 19.4 s (interquartile range: 12.3, 29.6 s).

Overall accuracy on the shadow task

In the shadow task, subjects scrolled through 11 versions of the scene, each of which showed a different shadow configuration for the target lamppost, and selected the one they judged to be consistent with the scene's lighting. Therefore, as in Experiments 2a and 2b, it was possible to classify subjects' performance on the shadow task in a number of different ways. First, taking a conservative approach we defined an accurate response to be only when subjects selected the consistent version of the scene. Collapsed across the four scenes, the consistent version was selected a mean 25% of the time, 95% CI [21%, 29%]. Second, taking a slightly more lenient approach and defining an accurate response by including one version either side of the consistent shadow position—that is, when versions 10, 11, or 12 were selected—a mean 55% of shadows were positioned correctly, 95% CI [50%, 59%]. Replicating the findings from Experiments 2a and 2b, Figure 4.6 shows that subjects were least accurate in Scene 3, with a mean 40% selecting version 10, 11, or 12 of the scene, 95% CI [31%, 49%]. In contrast to the previous experiments, however, subjects were most accurate in Scene 1, a mean 65% selecting version 10, 11, or 12 of the scene, 95% CI [56%, 74%]. There is not an immediately obvious reason as to why subjects did relatively well on Scene 1 in Experiment 3. Speculatively, it is possible that the shape of the target shadow in Experiments 2a and 2b actually made the inconsistent positions seem more plausible rather than less plausible in Scene 1.

Preference for shadows to the left or right

Across the four scenes, we classified each response according to whether subjects positioned the shadow to the left (versions 10 to 1) or right (versions 12 to 21) of the correct position. In contrast to the results of Experiments 2a and 2b, collapsing across all four scenes, the shadows were equally likely to be positioned to the left or to the right of the correct location, M_{diff} = 0%, 95% CI [-8%, 7%]. Yet as illustrated in Figure 4.6, there was still variation by scene. In line with our previous experiments, subjects were more likely to position the target shadow left of the correct position in Scene 3. And again, in Scene 2, subjects were more likely to position the target shadow right of the correct position. This time, in both Scenes 1 and 4, a similar proportion of subjects selected a target shadow to the left of its correct location as to the right. Again, these findings support the notion that lighting direction cues can override any prior assumptions about lighting (Morgenstern et al., 2011, 2014).

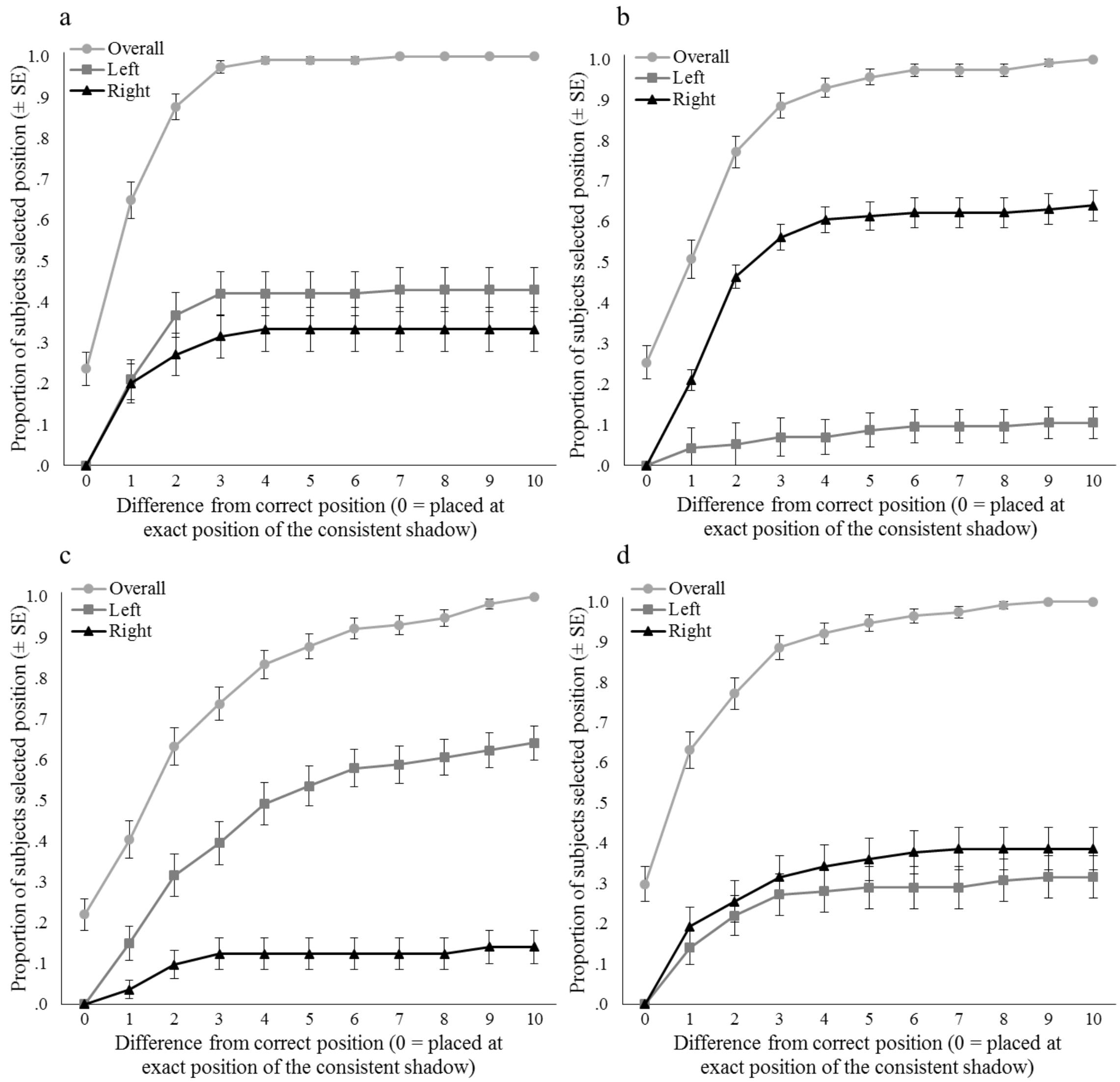


Figure 4.6. Cumulative proportion of responses made by each version of the scene for (a) Scene 1 (b) Scene 2 (c) Scene 3 and (d) Scene 4. The light grey line with circle markers shows the overall proportion of responses made by each inconsistent version of the scene—these are cumulative and therefore a difference of 1 includes subjects selecting versions 10, 11, or 12 of the scene. The black line with triangle markers shows the cumulative proportion of subjects selecting each version of the scene to the right of the consistent shadow position responses. The dark grey line with square markers shows the cumulative proportion of subjects selecting each version of the scene to the left of the consistent shadow position responses Error bars represent standard error of the mean.

Conclusion

Overall the pattern of results across our five shadow experiments was largely consistent. First and foremost, these experiments suggest that people have a limited ability to identify consistent and inconsistent shadows. This finding is surprising considering that subjects viewed scenes in which there was sufficient information to determine the answer objectively. Furthermore, unlike in a real-world scenario, subjects' attention was cued to the target object and its shadow in the scene which meant that subjects only had to check the consistency of this one shadow with the lighting of the scene. When viewing images in real-life, people are unlikely to gain this advantage, rather their task is more difficult than in our experimental scenario because they need to be sceptical of all shadows. There is, however, an important point to consider when interpreting the results of the shadow experiments; although there was information in the scenes that allowed subjects to use the shadow-based analysis technique, it remains possible that they were not aware of this technique. In fact, even with awareness, people might still find it difficult to apply the shadow-based analysis—this would be interesting to explore in future research.

It is possible, then, that a lack of awareness prevented people from using the objective shadow-based analysis. Consider, however, that even without using this objective technique, it should still be possible to identify whether the target shadow is consistent or inconsistent with the other shadows based on a subjective visual inspection of the scene. Put simply, a visual comparison of the position and shape of the target shadow with the other shadows in the scene should help people to make an accurate decision about the consistency or inconsistency of the shadows. In fact, because the human visual system evolved under natural lighting conditions there is good reason to think that people would be sensitive to the appearance of shadows. Yet, our results suggest the opposite: that people are reasonably insensitive to shadow information. To help understand this finding we can look to theoretical accounts of how people process shadow information.

Commonly, researchers have adopted the visual search paradigm to gain insight about how people process shadow information (e.g., Elder, Trithart, Pintilie, & MacLean, 2004; Khuu, Honson, & Challinor, 2016; Lovell, Gilchrist, Tolhurst, & Troscianko, 2009; Porter, Tales, & Leonards, 2010; Rensink & Cavanagh, 2004). In one study, subjects searched for a target item that had a different orientation to the distractor items in the display (Rensink & Cavanagh, 2004). The items in the display were vertically orientated grey rectangle posts with an attached darker grey quadrilateral shape at one end of the post. Based on the implicit light from above assumption—that is, shadows should be below objects (e.g., Adams, 2007; Mamassian & Goutcher, 2001)—when the displays were upright the attached darker region is interpreted as a shadow cast by the post; but when the displays were inverted, the attached darker region is no longer interpreted as a shadow, instead it is simply another part of the object. As such, in the upright trials subjects were searching for an odd shadow amongst other shadows, while in the inverted trials subjects were searching for an odd object amongst other objects. Interestingly, people were slower to detect an odd shadow in a display than an equivalent odd object. Based on these findings, researchers proposed that there is an early stage in visual processing that rapidly identifies and then discounts shadow regions in a scene (Rensink & Cavanagh, 2004).

Indeed this insensitivity to shadows can account for our finding that people struggled to make a subjective judgement about whether or not a target shadow was consistent or inconsistent with the other shadows in the scene. Specifically, if people discard shadow information at an early stage of visual processing, then it follows that they will not have an opportunity to learn about how shadows should appear. Thus, although we directed subjects to use the shadow information in the scenes it is unlikely to have been much help if people do not hold a basic knowledge about how shadows should look. Yet, it is difficult to reconcile some of our other results with the early discounting theory; for example the theory cannot account for our finding that people were more likely to detect inconsistent shadows that were positioned further from the correct position than inconsistent shadows that were positioned closer to the correct position. According to the early discounting theory, all shadows are rapidly identified and discounted; this means that people should be equally likely to incorrectly accept an inconsistent shadow as correct, regardless of the extent of discrepancy with the other shadows in the scene. Therefore, our results suggest that the early discounting theory might not fully account for how people process shadow information.

Other studies showing that shadows can have a profound effect on scene perception also indicate that people do not just ignore shadow information (e.g., Allen, 1999; Castiello, 2001; Kersten, et al., 1996). Such findings have led some researchers to suggest that the notion of an early level mechanism that discounts shadows is too simplistic (Elder et al., 2004; Khuu et al., 2016; Lovell et al., 2009; Porter et al., 2010). An alternative theoretical account of shadow processing is that the visual system adopts a coarse scale analysis of shadow information (Khuu et al., 2016; Lovell et al., 2009; Mamassian, 2004). According to this hypothesis, shadow regions are rapidly identified but not discounted; instead the identified shadows are processed in a coarse manner (Lovell et al., 2009; Mamassian, 2004). Offering support for this coarse scale hypothesis, researchers demonstrated a boundary condition for Rensink and Cavanagh's (2004) finding that people were slower to detect an odd shadow than an equivalent odd object (Lovell et al., 2009). In the original study, the attached darker region of the target item in the display was orientated at a 30° difference to the attached darker region of the distractor items in the display. When this angle difference increased, for example from 30 to 90°, people were as fast to detect the odd shadow as they were to detect the odd object (Lovell et al., 2009). This boundary of the effect illustrates that people can readily detect shadow discrepancies when they are relatively large but not when they are subtle (Lovell et al., 2009; Mamassian, 2004; but see Porter et al., 2010). For subtle shadow discrepancies, an effortful higher-level visual mechanism is invoked to re-evaluate the information, this time taking into account finer details (Lovell et al., 2009).

The coarse scale hypothesis can account for some of our results; namely that people were more likely to detect inconsistent shadows that were positioned further from the correct position than inconsistent shadows that were positioned closer to the correct position. Accordingly, using the coarse scale hypothesis to interpret our results also fits with the notion of a perceptual threshold for detecting lighting inconsistencies (Lopez-Moreno et al., 2010; Tan et al., 2015). That is, there might be a discernible point at which inconsistent shadows are different enough from the consistent position that the inconsistency becomes noticeable—at this point a coarse scale mechanism is effective. For inconsistent shadows that do not reach this threshold a different strategy, perhaps an effortful higher-level mechanism, is used. Although our results neither support nor oppose the existence of a separate higher-level mechanism for processing subtle shadow inconsistencies, our data do indicate that, if such a mechanism does exist, it might not be particularly effective. Indeed, subjects often failed to notice shadow inconsistencies, especially the subtle ones. Furthermore, our finding that subjects were biased to accept the shadow scenes as consistent in Experiments 1a and 1b raises the possibility that people might infrequently employ a more effortful strategy and instead tend to rely on intuitive feelings about shadows.

Unfortunately, these findings do not have overly positive consequences for using shadow information to identify image forgeries in the real world. The results suggest that when an image manipulation creates a relatively large inconsistency in the shadows of the scene, people are likely to notice it and might be able to detect the image as a fake. If the shadow inconsistency is relatively subtle, however, then people are more likely to be fooled by the fake image. What is more, this interpretation might be optimistic; in our experiments, we asked people to check the consistency of the shadows in the scenes—evidence suggests that people give less attention to shadows in more natural viewing conditions (Ehinger et al., 2016; Ostrovsky et al., 2005; Wright, 2005). An interesting follow-up study would be to explore to what extent people can detect shadow inconsistencies when their attention is not directed to the shadows in the scene.

Encouragingly, though, researchers have shown that performance on many types of perceptual tasks can improve as a result of training (e.g., Ellison & Walsh, 1998; Porter et al., 2010; Sireteanu & Rettenbach, 1995, 2000). For example, using the same visual search task as in Rensink and Cavanagh's (2004) study, researchers found that following a training period subjects were as fast to detect an odd shadow in a display as an equivalent odd object (Porter et al., 2010). Moreover, this learning process transferred to other forms of shadow stimuli. Therefore, an interesting avenue for future research would be to examine the influence of training on people's performance on the shadow task. This training could include both showing people how to make use of the objective shadow-based analysis technique as well as perceptual-based learning—for example giving people a number of practice trials before completing the shadow task.

In sum, our findings suggest that people have a limited ability to identify consistent and inconsistent shadows. Yet, the extent of the inconsistency influenced performance: people were more likely to detect inconsistent shadows that were positioned further from the correct position than inconsistent shadows that were positioned closer to the correct position. This result fits with the coarse scale hypothesis of shadow processing—that the visual system typically relies on a coarse representation of shadow information. In real-world situations, then, it is possible that people might be able to use shadow information to help detect fake images if the forger leaves a relatively large shadow inconsistency but not if the inconsistency is only subtle. Seemingly, if the Rathods had chosen to manipulate their fake summit photos to show a timestamp within an hour or so of midday—rather than 6.25am—they might still hold the title of first Indian couple to climb Mount Everest.

Chapter 5 :

Can people identify geometric inconsistencies in reflections?

"…there's the room you can see through the glass—that's just the same as our drawing room, only the things go the other way."

Lewis Carroll (1871)

Introduction

Shortly after the November 2015 Paris terrorist attacks, a doctored photo depicting an innocent man as one of the attackers circulated online and in newspapers (Rawlinson, 2015). The authentic version of the photo featured Veerender Jubbal standing in a bathroom in front of a mirror taking a selfie with his iPad. In the manipulated version of the image, Jubbal is shown wearing a suicide vest and his iPad has been replaced by what appears to be a Qur'an. Although many major news outlets were fooled by the image, the manipulation left several prominent clues that the image was a fake (Butterly, 2015; Rawlinson, 2015). Crucially, when Jubbal photographed his reflection he was standing straight-on to the mirror which means that the camera used to capture the photo must also be visible in the reflection. In the authentic version of the photo, the iPad used to capture the photo can be seen clearly in the reflection. Yet in the manipulated version of the image, the forgers replaced the iPad with a Qur'an thus making it a geometrical impossibility for the image to have been captured. Detecting this basic inconsistency in the geometry of the reflection could have prevented an innocent man becoming a suspect in the attacks. To what extent, then, can people use reflections to help to identify whether photos are authentic or manipulated? In this chapter, we explored this question by examining people's ability to identify whether a target reflection was consistent or inconsistent with the other reflections in the same scene.

Reflections are common in our everyday environment yet only a relatively small number of studies have tested what people understand about mirror reflections[9] (Bertamini, Spooner, & Hecht, 2003; Bianchi & Savardi, 2012; Croucher, Bertamini, & Hecht, 2002; Hecht, Bertamini, & Gamer, 2005; Lawson, 2010; Lawson & Bertamini,

[9] Although there are two types of reflection, in this chapter we focus on specular reflection—the reflection of light from a smooth surface, such as a flat mirror or window.

2006; Muelenz, Hecht, & Gamer, 2010). In one study using a bird's-eye view diagram of a room, subjects predicted when a target would first become visible in a mirror (Croucher et al., 2002). In one scenario, for example, subjects indicated the point at which a character who walks across the room on a path parallel to the surface of a mirror on the opposite wall would first see their reflection in that mirror. Across a range of these scenarios, subjects made consistent *early errors*. Put simply, they predicted that the character would be able to see their reflection before reaching the edge of the mirror, which is physically impossible. These findings suggest that people's understanding of reflection is limited and biased.

The bird's-eye view diagrams did not allow people to see the reflections in the mirror which meant that they had to rely on their memory or abstract knowledge about mirror reflections when making their predictions. As such, other research has explored how people perceive reflections, such as subjects' ability to distinguish between correct and incorrect mirror reflections when the reflections are visible in the scene (Bertamini et al., 2003; Farid & Bravo, 2010). In one study, subjects judged whether the reflections in a series of computer-generated scenes of household rooms were correct or incorrect (Bertamini et al., 2003). Some of the scenes contained authentic reflections and others had been manipulated to show an incorrect reflection, for example, a left-right reversal of the mirror image. Even when making these perceptual judgements about mirror images, subjects were surprisingly unlikely to notice distortions. A similar pattern of results was found in a more recent study (Farid & Bravo, 2010). The researchers created a series of computer-generated scenes containing only a red cone and a mirror. In half of the scenes, the cone's reflection was inconsistent with the scene geometry such that the reflection was manipulated to be physically impossible given the position of the cone and the mirror in the scene. In the other half, the cone's reflection was consistent with the scene geometry. Overall, subjects performed only slightly better than chance, correctly identifying just 55.7% of the scenes. In these perceptual tasks, people were surprisingly poor at detecting simple inconsistencies in reflections.

A number of possible explanations have been proposed to account for people's poor performance across these reflection tasks (Croucher et al., 2002). The one that appears to be able to account for the majority of the findings, including the early error, is the virtual world rotation hypothesis (Hecht et al., 2005; Muelenz et al., 2010). This hypothesis proposes that people hold a perceptual outward bias that causes them to mentally rotate the mirror reflection of the world to make it more orthogonal (at a right-angle) with respect to their line of sight. As a result, the mirror reflection appears further away from the observer than it would in reality. To illustrate, imagine an observer is viewing a mirror from the right side, owing to their bias the reflected world in the mirror is perceived as if it has been rotated counter-clockwise causing it to shift left, and away from the observer.

It follows, then, that people would expect to see reflections before it is physically possible to do so. In a test of this hypothesis, subjects completed a real-world localisation task; first subjects viewed an object's reflection in a mirror, then, once the mirror had been concealed and the object removed from the room, they attempted to indicate where the real object had been positioned in the room (Muelenz et al., 2010). Unbeknownst to the subjects, the rotation of the mirror was manipulated across trials: In half of the trials the mirror was positioned with no rotation, in the other half the mirror was rotated 2° away from the subjects' point of view. If subjects have a perceptual outward bias then when the actual mirror is rotated away from their point of view this rotation will compensate for their bias and result in higher accuracy in the rotation than the non-rotation condition. Alternatively, if subjects do not have this bias, then they will more accurately locate the real object in the room when the mirror is not rotated. In support of the virtual world rotation hypothesis, the mean error for localising the position of the real object was smaller in the rotation than in the non-rotation condition.

Although the virtual world rotation hypothesis can account for many of the previous findings, it does not explain all of them (Bianchi, Bertamini, & Savardi, 2015). Furthermore this hypothesis struggles to explain why people make effective use of mirrors in everyday life, for instance when driving or checking their appearance. One possibility is that the virtual world rotation hypothesis is only part of the explanation. Indeed, there is an important difference between using mirrors in the real world and making predictions or judgements about the location of reflections in experimental settings: When people use mirrors in real life the reflections are visible and people can use this perceptual information to help them to successfully use the mirror. Typically, such information is not available in experimental reflection tasks (Bertamini et al., 2003; Croucher et al., 2002). If experimental settings more closely mimicked real-life scenarios, would people be better able to predict when their reflection would appear in the mirror? Research suggests the answer is yes. Using a real-life setting where subjects could see the reflections in a real mirror—as opposed to the mirror being covered—subjects made accurate predictions about when their reflection would become visible (Lawson & Bertamini, 2006). Perhaps, then, the errors people make and the outward bias are artefacts of the experimental conditions but when using real mirrors the availability of perceptual information allows people to use them effectively. Even if that is the case, the results of these studies are still theoretically important because they suggest that people do not store readily accessible information from past experiences with mirrors. Instead, people need the information available each time they use a mirror.

It seems likely that the same rationale would hold for people's ability to predict or recognise the correct pattern of reflection in images. That is, to make accurate judgements people need perceptual information to be available in the image to guide them. One such piece of information that might prove useful is the reflection vanishing point (Montague, 2010). It has been known for centuries that reflections adhere to a basic law of optical physics: a smooth surface will reflect the light at the same angle that it hits the surface (Hecht & Zajac, 1974; Ronchi, 1970). As shown in Figure 5.1a, a result of this optical constraint means that, from a bird's-eye view, imaginary parallel lines connect points on the real object in front of the mirror with the same points on the object's reflection behind the mirror (Farid, 2016; O'Brien & Farid, 2012). Owing to perspective projection, when viewing that same scene but from the viewpoint shown in Figure 5.1b, the lines that connect object points and their corresponding points in the reflection will converge to a single point—the reflection vanishing point. Consequently, a geometric-based analysis can be used to objectively verify where the reflection of an object in the world should appear in a mirror. Furthermore, adding fake reflections into a photo, or manipulating a photo that contains reflections, can create inconsistencies that can be identified using the geometric-based analysis. If a line connecting a point on an object and its corresponding point in the reflection does not intersect the reflection vanishing point it highlights an inconsistency (O'Brien & Farid, 2012). As shown in Figure 5.1c, the geometric analysis reveals that the bus stop's reflection is inconsistent with the reflection vanishing point that is consistent with the rest of the scene, indicating that some manipulation has occurred.

This reasonably simple analysis based on the geometric relationship between objects and their reflections is used in digital image forensics to help identify fakes. What is still unknown, however, is whether people also make use of the reflection vanishing point. To our knowledge, none of the reflection studies to date have used stimuli that would allow people to apply this geometric analysis—studies have used either bird's eye view diagrams or scenes that do not contain enough corresponding object and reflection points to compute the reflection vanishing point. Therefore, previous research might have underestimated people's understanding of mirror reflections. Accordingly, in this chapter we explored people's ability to identify scenes that contained consistent and inconsistent reflections when it was possible to make use of the reflection vanishing point.

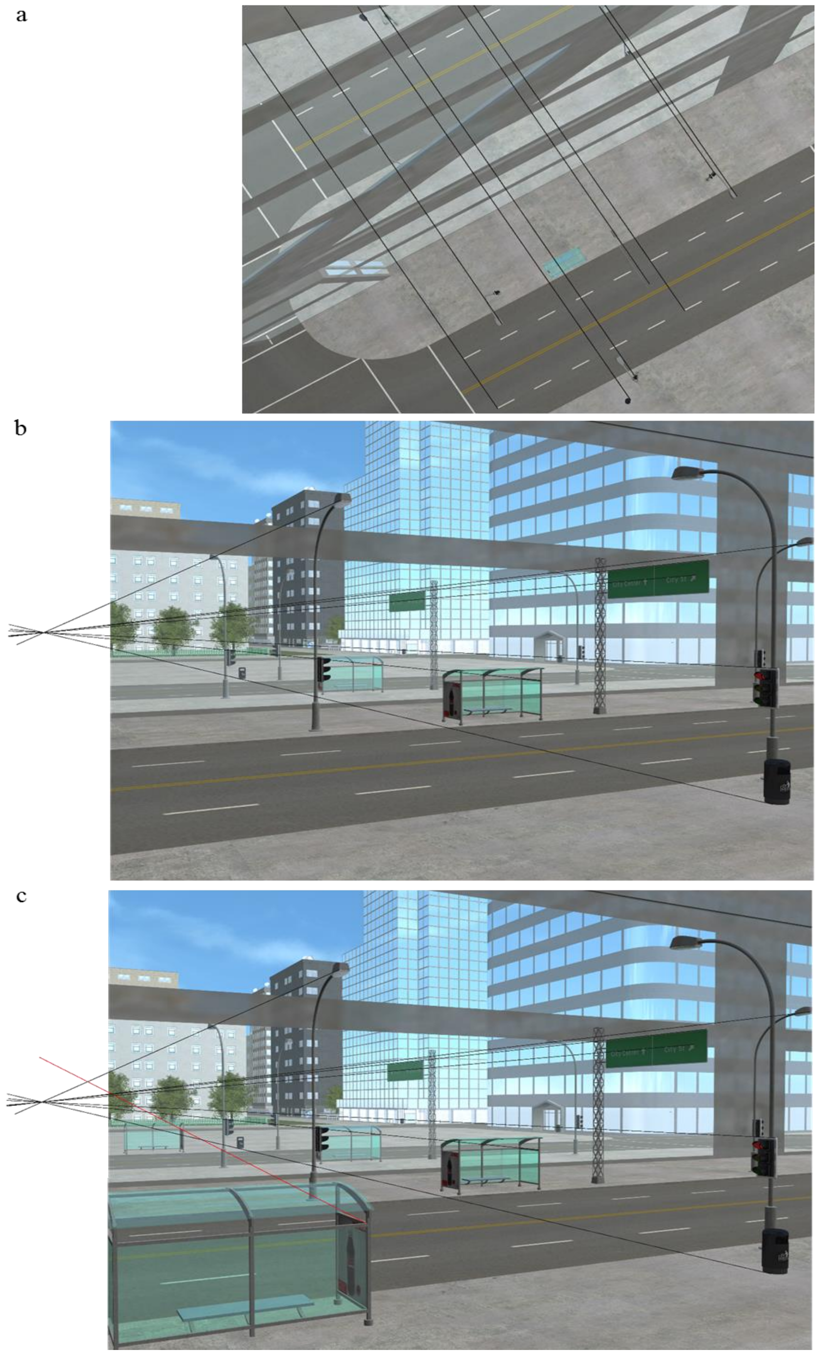


Figure 5.1. Example of the geometrical information available in images: (a) The scene is shown from a bird's-eye view. The geometric relationship between objects and reflections of those objects is constrained by optical law of reflection. As a result, the black lines drawn connecting the objects with their reflections are parallel to one other and perpendicular to the mirror surface. (b) The same scene is shown from a different perspective; here, owing to perspective projection, the black lines are no longer parallel but instead converge to a single vanishing point. (c) Again the same scene is shown but the bus stop on the left and its reflection have been added. The red line connecting the bus stop with its reflection does not intersect the scene's reflection vanishing point and highlights an inconsistency.

Experiment 1

Method

Subjects and Design

A total of 79 subjects (*M* = 26.6 years, *SD* = 10.8, range = 13-68; 39 men, 38 women, and 2 chose not to disclose their gender) completed the study online. Five additional subjects were removed, 4 who experienced technical difficulties, and 1 who had missing response time data for at least one response on the task. There were no geographical restrictions and subjects did not receive payment for taking part, but they did receive feedback on their performance at the end of the task. Subject recruitment continued until we reached a minimum of 20 responses per image (the stimulus set consisted of four consistent reflection images and eight inconsistent reflection images, further details of the images are provided in the following Stimuli section). We used a within-subjects design. Each person viewed a series of four computer-generated images, half of which had consistent reflections, and half of which were manipulated to show inconsistent reflections. We measured people's accuracy in determining whether an image had consistent or inconsistent reflections.

Stimuli

Using Maya® (2016; Autodesk, Inc.), a 3-D animation software, we created five different outdoor city scenes from the same 3-D cityscape model that was used to create the shadow scenes in Chapter 4. All scenes were rendered as TIF files at a resolution of 960 × 720 pixels. Each of the five scenes included a flat, smooth reflective surface. In each scene, the target object was a street sign placed adjacent to the reflective surface so that the sign and its reflection were both visible. In addition, to ensure subjects could use the geometric analysis to locate the reflection vanishing point we made a number of other non-target objects and their corresponding reflections visible in the scene. The scenes were rendered without shadows to ensure that we did not provide subjects with any additional cues that would influence their ability on the reflection task. These five scenes were the originals with consistent reflections.

To create our inconsistent reflection scenes we began by rendering each scene two more times; once with the street sign moved forward relative to the original street sign position (+7 m on the z-axis) and once with the street sign moved backward relative to the original street sign position (-7 m on the z-axis). We then manipulated the images using GNU Image Manipulation Program® (GIMP, Version 2.8). First, we removed the street sign's reflection in the original scene, importantly, the original street sign itself remained in the scene. Second, we cut the reflection of the street sign from one of the other scenes with the street sign moved forwards or backwards. We then overlaid this reflection onto the original scene (see Figure 5.2 for an example of the consistent and inconsistent versions of a scene). We exported the images as PNGs, a lossless format.

Overall, we had three versions of each of the five city scenes, a total of 15 images. The original, non-manipulated version of each of these scenes was used to create our consistent reflection image set. The two manipulated versions of each scene were used to create our inconsistent reflection image set. Subjects saw two consistent reflection and two inconsistent reflection images but always in a different city scene. The fifth city scene was used as a practice.

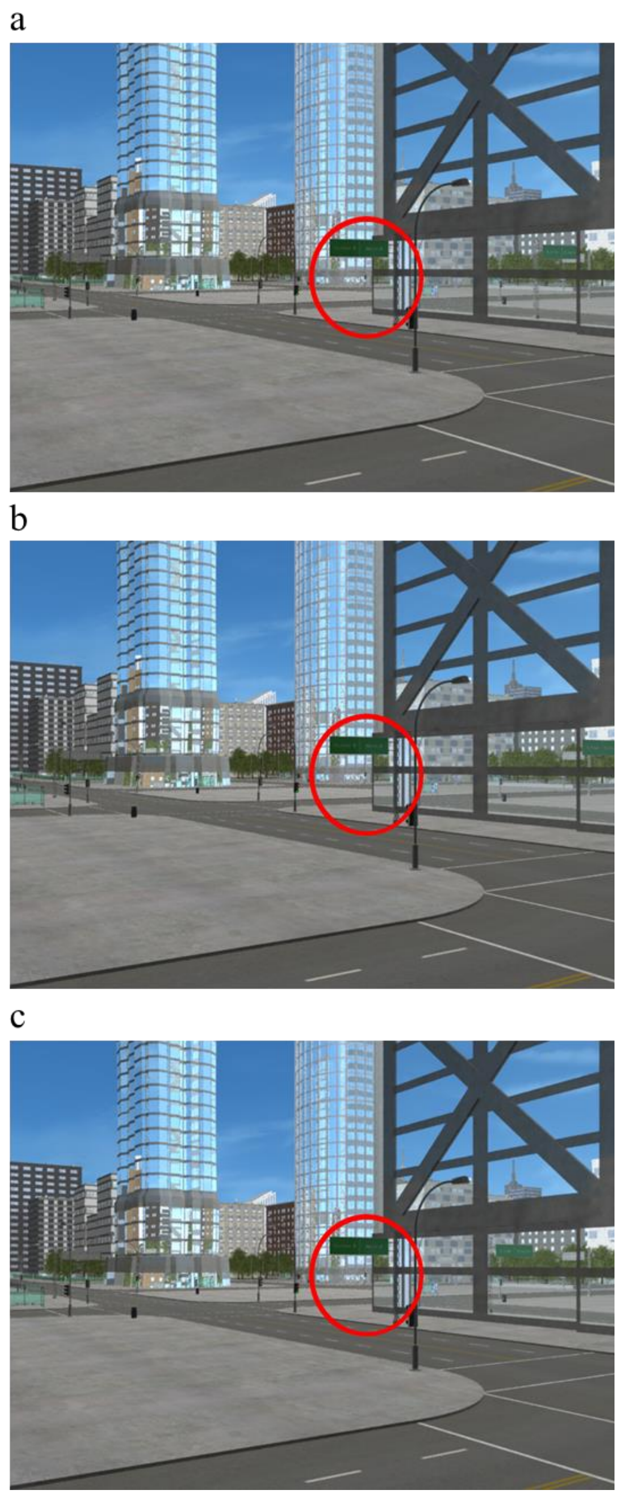


Figure 5.2. Example of the consistent and two inconsistent versions of a scene: (a) original image with consistent reflections, (b) forward inconsistent reflection image with the target street sign reflection moved forward of the consistent position, (c) backward inconsistent reflection image with the target street sign reflection moved backward of the consistent position. Each subject saw this city scene just once, they were randomly shown a, b, or c. In the example images the target street sign is shown in a red circle.

Procedure

We used the same procedure as Experiments 1a and 1b in Chapter 4, with the following two adjustments: We cued subjects' attention to the target street sign on which they need to base their response, and we asked "Is the street sign's reflection consistent or inconsistent with the other reflections in the scene?"

Results and Discussion

An analysis of the response time data suggested that subjects were engaged with the task and spent a reasonable amount of time determining whether the reflections in the scenes were consistent or inconsistent. The mean response time per image was 23.9 s (SD = 24.8 s) and the median response time 17.4 s (interquartile range: 9.9, 25.5 s).

Overall accuracy on the reflection task

We now turn to our primary research question: Can people identify whether scenes contain consistent or inconsistent reflections? Overall, a mean 50% of the scenes were correctly classified, 95% CI [44%, 55%]. Subjects' ability to distinguish between consistent (42% correct, 95% CI [34%, 50%]) and inconsistent (58% correct, 95% CI [50%, 65%]) reflections was not reliably greater than zero, d' = -0.01, 95% CI [-0.23, 0.21]. Furthermore, subjects showed a bias towards saying that reflections were inconsistent, c = -0.15, 95% CI [-0.25, -0.06]. These results indicate that subjects found it extremely difficult to determine whether the reflection of the target object was consistent or inconsistent with the other reflections in the scene. Furthermore, the results suggest that subjects did not make use of the geometrical information in the scene to compute the reflection vanishing point and objectively determine the answer. Instead, in line with previous research, it seems more likely that subjects had incorrect beliefs about reflections and relied on these to make a subjective judgement about the consistency or inconsistency of the reflections in the scene (e.g., Bertamini et al., 2003; Croucher et al., 2002). We next consider whether any individual factors or image metrics are associated with improved ability to identify consistent and inconsistent reflection scenes.

Individual factors and image metrics

To determine whether individual factors play a role in identifying consistent and inconsistent reflections, we gathered subjects' demographic data, as well as details about their interest in photography, and video gaming experience. We also asked subjects to rate their confidence for each of their decisions and recorded their response time. In addition to these individual factors, we checked whether three properties of the image itself affected people's accuracy on the task. One image property was simply whether the reflection had moved forwards or backwards relative to the consistent reflection position. Figure 5.2 shows examples of how the reflection was moved in the forward and backward inconsistent scenes. The second image property was the distance from the centre of the image to the reflection vanishing point. The third image property was an angle measurement. Specifically, we measured the rotation, in degrees, from the scene's reflection vanishing point to the reflection vanishing point for the target object and its inconsistent reflection.

To check how each factor influenced subjects' performance, we conducted two generalized estimating equation (GEE) analyses—one for the inconsistent reflection scenes and one for the consistent reflection scenes. Specifically, we conducted a repeated measures logistic regression with GEE because our dependant variables were binary with both random and fixed effects (Liang & Zeger, 1986). The results of the GEE analyses are shown in Table 5.1.

The GEE analysis revealed that three of the variables had an effect on subjects' ability to accurately identify inconsistent reflection scenes: reflection position, confidence, and gender. Scenes in which the reflection was moved forward of its consistent position were more likely to be identified as inconsistent compared with scenes in which the reflection was moved backwards of its consistent position. This result fits with the virtual world rotation hypothesis and supports the notion of a perceptual outward bias. Indeed, the inconsistent reflections that appeared further away rather than closer to the observers' viewpoint were more likely to be incorrectly accepted as consistent (Muelenz et al., 2010). On the other hand, it is possible that this result is simply an effect of perspective projection transformation from the 3-D world to the 2-D image. Although the street sign was moved equally in the forward and backward conditions in the 3-D environment, owing to perspective projection the same change in the 3-D environment produces a larger change in the foreground than background of the 2-D image. Thus it follows that people will be more sensitive to changes in the foreground than in the background.

We also found a small effect of confidence, such that more confident responses were slightly more likely to be associated with accurate responses than were less confident responses. Finally, females were slightly more likely to correctly identify inconsistent reflection scenes than males. The results of the GEE analyses for the consistent scenes revealed that none of the variables had an effect on subjects' ability to accurately identify consistent reflection scenes.

Table 5.1
Results of the GEE binary logistic models to determine variables that predict accuracy in the reflection task

	Inconsistent			Consistent		
Predictor	*B*	*OR [95% CI]*	*p*	*B*	*OR [95% CI]*	*p*
Reflection position = Forward	0.82	2.26 [1.08, 4.74]	.03	-	-	-
Confidence	0.01	1.02 [1.00, 1.03]	.04	-0.01	0.99 [0.98, 1.01]	.24
Gender = Female	0.71	2.04 [1.01, 4.10]	.05	-0.10	0.91 [0.46, 1.78]	.78
Vanishing point distance	-0.05	0.95 [0.89, 1.01]	.12	0.03	1.03 [0.97, 1.11]	.34
Interest in photography = Interested	-0.61	0.54 [0.25, 1.16]	.12	-0.21	0.81 [0.41, 1.59]	.54
Video gaming = Frequent (at least twice a month)	0.47	1.59 [0.80, 3.18]	.19	0.40	1.49 [0.79, 2.83]	.22
Angle difference	-2.22	0.11 [0.00, 3.96]	.23	-	-	-
Response time	0.00	1.00 [1.00, 1.01]	.41	0.00	1.00 [0.99, 1.01]	.79

Note. CI = confidence interval. *B* and odds ratios (*OR*) estimate the degree of change in accuracy associated with one unit change in the independent variable. An odds ratio of 1 indicates no effect of the independent variable on accuracy, values of 1.5, 2.5, and 4.0 are generally considered to reflect small, medium, and large effect sizes, respectively (Rosenthal, 1996). The category order for factors was set to descending to make the reference level 0. The reference groups are: Reflection position = back, Video game playing = Infrequent (never/less than once a month/about once a month), Gender = Male, Interest in photography = Not Interested. Response time, confidence, reflection vanishing point distance, and angle difference were added as continuous variables. The two subjects who chose not to disclose their gender were excluded from these analyses leaving a total sample of $n = 77$. The reflection position and angle difference predictor variables were not applicable in the consistent reflection scenes.

In Experiment 1, subjects correctly classified a mean 50% of the scenes indicating their performance on the reflection task was no better than would be expected by chance alone. These results suggest that people have an extremely limited ability to identify when the reflections within a scene are consistent and when they are inconsistent: this is somewhat surprising because each scene contained sufficient information to objectively determine the answer. That is, by calculating the location of the scene's reflection vanishing point, the consistency of the target object's reflection with that vanishing point can be objectively verified. There is, however, a possibility that it might have been too difficult to make use of the reflection vanishing points in the four scenes used in Experiment 1. There are two reasons for this possibility. One reason is that the scenes' reflection vanishing point was located a mean 48.9 cm from the image centre, a distance which might have been large enough to make it difficult to use this information. A second reason is that in the inconsistent scenes, the mean angle difference between the target object's reflection vanishing point and the reflection vanishing point for the rest of the scene was just 1°. Perhaps these two factors made it too difficult for subjects to use the geometrical information and locate the reflection vanishing point. For these reasons, and because relatively little research has examined people's perception of reflections, we ran a second experiment with new stimuli. For this new stimuli we decreased the distance of reflection vanishing point from the centre of the image, and increased the angle difference from the scene reflection vanishing point to the vanishing point for the target object and its inconsistent reflection. In Experiment 2, we checked whether changing these image properties influences performance on the task.

Experiment 2

Method

Subjects and Design

A total of 97 subjects (*M* = 25.5 years, *SD* = 9.0, range = 15-57; 58 men, 36 women, and 3 chose not to disclose their gender) completed the study online. Eight additional subjects were removed, 5 who experienced technical difficulties and 3 who failed to understand instructions. There were no geographical restrictions and subjects did not receive payment for taking part, but they did receive feedback on their performance at the end of the task. Subject recruitment continued until we reached a minimum of 20 responses per image. The design was identical to that of Experiment 1.

Stimuli

We created our stimuli following the procedure as in Experiment 1, with two exceptions. First, we changed viewing perspective to bring the reflection vanishing point closer to the centre of the image—the mean distance was 29.8 cm (cf. Expt 1: 48.9 cm). Second, to increase the angle difference between the original, consistent reflection and the manipulated, inconsistent reflection we moved the street sign further from its original position. In Experiment 1, we moved the street sign 7 m on the z-axis relative to the original street sign position, this time we moved it by 10 m. The resulting mean angle difference between the reflection vanishing point for the inconsistent reflections and the reflection vanishing point for the rest of the scene was 3.5° (cf. Expt 1: 1°).

Procedure

The procedure was identical to that used in Experiment 1.

Results and Discussion

An analysis of the response time data indicated that subjects spent a reasonable amount of time determining whether the reflections in the scenes were consistent or inconsistent. The mean response time per image was 23.6 s (*SD* = 17.4 s) and the median response time 18.7 s (interquartile range: 13.4, 29.6 s).

Overall accuracy on the reflection task

Overall subjects correctly classified a mean 73% of the reflection scenes (cf. Expt 1: 50%), 95% CI [68%, 78%]. Subjects showed a reasonably good ability to discriminate between consistent (72% correct, 95% CI [65%, 80%]) and inconsistent (75% correct, 95% CI [68%, 81%]) reflection scenes, d' = 0.91, 95% CI [0.72, 1.10]. Unlike in Experiment 1, these results suggest that subjects have some ability to identify consistent and inconsistent reflections. Perhaps, then, subjects can make use of the reflection vanishing point to objectively judge the consistency of the reflections in a scene, but only in instances where the vanishing point is relatively easy to determine. In addition, in contrast to Experiment 1, subjects did not show a bias towards saying that reflections were inconsistent, c = -0.03, 95% CI [-0.12, 0.07]. This finding offers some support for our suggestion that people might only rely on perceptual biases to make judgements about reflections when there is a lack of information available to make a more informed decision.

Individual factors and image metrics

As in Experiment 1, we conducted two GEE analyses—one for the inconsistent reflection scenes and one for the consistent reflection scenes. Preliminary analyses revealed a variance inflation factor of 11.8 for the angle difference variable, suggesting that this variable was correlated with one or more of the other predictor variables, therefore we removed the angle difference variable from the analyses. Other than removing the angle difference variable, we included the same factors used in the GEE models in Experiment 1. The results of the GEE analyses are shown in Table 5.2. Two of the variables had an effect on subjects' ability to accurately identify inconsistent reflection scenes. Replicating our finding in Experiment 1, more confident responses were slightly more likely to be associated with accurate responses than less confident responses. There was also an effect of distance from the centre of the image to the reflection vanishing point: scenes in which the reflection vanishing point was closer to the image were more likely to be identified as inconsistent compared with scenes in which the reflection vanishing point was further from the image. This time, however, we did not find an effect of reflection position or gender.

Next, considering the consistent reflection scenes, the GEE analysis revealed that only one variable had an effect on subjects' ability to accurately identify consistent reflection scenes—the distance of the reflection vanishing point. As with the inconsistent scenes, when the reflection vanishing point was closer to the centre of the image the scenes were more likely to be identified as consistent compared with when the reflection vanishing point was further from the centre of the scene. It appears, then, that people might be able to make use of the geometrical information provided in the scenes to objectively judge the validity of the reflections when the reflection vanishing point is closer to the centre of the image.

That said, our results warrant a second interpretation that is not based on subjects making use of the reflection vanishing point. Another possibility is that moving the inconsistent reflections further from the consistent position in Experiment 2 than in Experiment 1 made it more visually apparent when the reflections were consistent versus inconsistent. If so, perhaps even based on a visual inspection of the scene, the correspondence between the object and its reflection did not match subjects' subjective expectation of how it should look—including for the backward inconsistent reflections.

Table 5.2
Results of the GEE binary logistic models to determine variables that predict accuracy in the reflection task

	Inconsistent			Consistent		
Predictor	*B*	*OR [95% CI]*	*p*	*B*	*OR [95% CI]*	*p*
Reflection position = Forward	0.04	1.05 [0.53, 2.08]	.90	-	-	-
Confidence	0.02	1.02 [1.01, 1.04]	.002	-0.01	0.99 [0.97, 1.01]	.38
Gender = Female	-0.40	0.67 [0.29, 1.57]	.36	-0.42	0.66 [0.28, 1.55]	.34
Vanishing point distance	-0.17	0.84 [0.71, 1.00]	.04	-0.20	0.82 [0.72, 0.93]	.002
Interest in photography = Interested	0.08	1.08 [0.47, 2.48]	.85	-0.12	0.89 [0.42, 1.91]	.77
Video gaming = Frequent (at least twice a month)	-0.35	0.70 [0.30, 1.64]	.42	0.77	2.16 [0.98, 4.75]	.06
Response time	0.00	1.00 [0.98, 1.02]	.88	0.00	1.00 [0.99, 1.01]	.89

Note. CI = confidence interval. *B* and odds ratios (*OR*) estimate the degree of change in accuracy associated with one unit change in the independent variable. An odds ratio of 1 indicates no effect of the independent variable on accuracy, values of 1.5, 2.5, and 4.0 are generally considered to reflect small, medium, and large effect sizes, respectively (Rosenthal, 1996). The category order for factors was set to descending to make the reference level 0. The reference groups are: Reflection position = back, Video game playing = Infrequent (never/less than once a month/about once a month), Gender = Male, Interest in photography = Not Interested. Response time, confidence, and reflection vanishing point distance were added as continuous variables. The three subjects who chose not to disclose their gender were excluded from these analyses leaving a total sample of n = 94. The reflection position predictor variables was not applicable in the consistent reflection scenes.

So why did subjects perform better on the reflection task in Experiment 2 than in Experiment 1? Given that we made two changes to the stimuli between Experiments 1 and 2 there are two possible reasons. One possibility is that creating scenes with the reflection vanishing point closer to the centre of the image made it easier for people to use the geometric analysis to work out the answer. A second possibility is that the bigger physical distance between the consistent and inconsistent reflection position made it easier to make a subjective judgement about the consistency or inconsistency of the reflections in the scene. To check which of these possible explanations best accounts for people's better performance in Experiment 2 than in Experiment 1 we ran a third experiment in which we changed only one variable. In Experiment 3, the scene reflection vanishing point remained the same as in Experiment 2, but we decreased the distance between the inconsistent and the consistent reflection position to match the distance in Experiment 1.

Experiment 3

Method

Subjects and Design

A total of 120 subjects (*M* = 30.9 years, *SD* = 13.7, range = 14-77; 53 men, 62 women, and 5 chose not to disclose their gender) completed the study online. A further 10 subjects were excluded from the analyses because they experienced technical difficulties. There were no geographical restrictions and subjects did not receive payment for taking part, but they did receive feedback on their performance at the end of the task. Subject recruitment continued until we reached a minimum of 20 responses per image. The design was identical to that of Experiments 1 and 2.

Stimuli

The stimuli remained the same as in Experiment 2 with just one exception: we decreased the distance that we moved the street sign reflection from its consistent position when creating the inconsistent scenes. In Experiment 2, we moved the street sign 10 m on the z-axis relative to the original street sign position, this time we moved it the same distance as in Experiment 1—7 m. The resulting mean angle difference between the reflection vanishing point for the inconsistent reflections and the reflection vanishing point for the rest of the scene was 2.4° (cf. Expt 2: 3.5°).

Procedure

The procedure was identical to that used in Experiments 1 and 2.

Results and Discussion

An analysis of the response time data indicated that subjects spent a reasonable amount of time determining whether the reflections in the scenes were consistent or inconsistent. The mean response time per image was 45.0 s (*SD* = 144.1 s) and the median response time 26.2 s (interquartile range: 18.8, 38.6 s).

Overall accuracy on the reflection task

Overall subjects correctly classified a mean 62% of the reflection scenes (cf. Expt 1: 50%; Expt 2: 73%), 95% CI [57%, 67%]. Subjects had some ability to discriminate between consistent (68% correct, 95% CI [62%, 75%]) and inconsistent (55% correct, 95% CI [49%, 62%]) reflection scenes, d' = 0.46, 95% CI [0.27, 0.65]. Our results show that subjects in Experiment 3 correctly classified a mean 12% more of the reflection scenes as consistent or inconsistent than subjects in Experiment 1, 95% CI [5%, 20%]. This difference in performance suggests that the position of the reflection vanishing point might influence people's ability to distinguish between consistent and inconsistent reflection scenes. On the other hand, subjects in Experiment 3 correctly classified a mean 11% fewer of the scenes as consistent or inconsistent than subjects in Experiment 2, 95% CI [5%, 19%], indicating that the extent of the inconsistency might also have an effect on people's performance on the task. In line with this suggestion, in Experiment 2 the inconsistent reflections were positioned further from the consistent position than in Experiment 3 and we did not find evidence of a response bias. Yet in Experiment 3, subjects showed a bias towards accepting the reflection scenes as consistent, c = 0.12, 95% CI [0.04, 0.21]. Taken together, these results suggest that people might have a relatively conservative criterion for judging that the reflections in a scene are inconsistent. As such, it is possible that people have a perceptual threshold for detecting reflection inconsistencies—that is, there is a point at which the inconsistent reflections are close enough to the consistent position that people will find it extremely difficult to detect the inconsistency; instead, they simply accept the reflection as consistent.

Individual factors and image metrics

As in Experiments 1 and 2, we conducted two GEE analyses—one for the inconsistent reflection scenes and one for the consistent reflection scenes. We included the same factors used in the GEE models in Experiment 2. The results of the GEE analyses are shown in Table 5.3. The GEE analyses revealed that none of the variables had an effect on subjects' ability to accurately identify consistent reflection scenes. Only one variable had an effect on subjects' ability to accurately identify inconsistent reflection scenes: replicating our finding in Experiment 1, we found an effect of reflection position. Scenes in which the reflection was moved forward of its consistent position were more likely to be identified as inconsistent compared with scenes in which the reflection was moved backwards of its consistent position. As mentioned when discussing the finding in Experiment 1, there are two ways to account for this effect of reflection position. On the one hand, this result fits with the virtual world rotation hypothesis and supports the notion of a perceptual outward bias (Muelenz et al., 2010). On the other hand, this result might simply be an effect of perspective projection: the same change in the 3-D environment produces a larger change in the foreground than background of the 2-D image and thus people are more sensitive to changes in the foreground than background of the 2-D scene. Unfortunately, our data do not allow us to determine which of these explanations best accounts for our finding that reflection position has an influence on subjects' ability to correctly identify inconsistent reflections.

Interestingly, however, our finding that there was an influence of reflection position in Experiments 1 and 3, but not Experiment 2, offers support to the notion of a perceptual threshold for detecting when reflections in a scene are inconsistent. Put simply, in Experiment 2 when the inconsistent reflections were moved 10 m from the consistent position we did not find a reliable effect of reflection position on subjects' ability to identify the inconsistent scenes. Yet in Experiments 1 and 3 when the inconsistent reflections were moved a smaller distance (7 m) from the consistent position, we did find a reliable effect of reflection position on performance. These findings suggest that there might be a point at which the inconsistent reflection becomes different enough from its consistent position to make the inconsistency noticeable. That said, it is important to note that adjusting the distance of the inconsistent reflections from the original position also changes the angle difference between the scene reflection vanishing point and the reflection vanishing point for the inconsistent reflection—as the distance increases, so does the angle difference. Thus we are not able to isolate the two factors and test them individually. Although our results appear to support the notion of a perceptual threshold in people's ability to subjectively determine the validity of the reflections based on a visual inspection of the scene, we cannot rule out an alternative explanation. Instead, it remains possible that changes to the angle difference between the scene reflection vanishing point and the reflection vanishing point for the inconsistent reflection affects people's ability to use the geometric information in the scene.

Table 5.3

Results of the GEE binary logistic models to determine variables that predict accuracy in the reflection task

	Inconsistent			Consistent		
Predictor	*B*	*OR [95% CI]*	*p*	*B*	*OR [95% CI]*	*p*
Reflection position = Forward	1.13	3.10 [1.84, 5.23]	<.001	-	-	-
Confidence	0.01	1.01 [0.99, 1.02]	.42	0.01	1.01 [1.00, 1.03]	.12
Gender = Female	0.21	1.23 [0.65, 2.31]	.52	-0.34	0.71 [0.32, 1.57]	.40
Vanishing point distance	-0.10	0.91 [0.80, 1.03]	.13	-0.08	0.93 [0.80, 1.07]	.29
Interest in photography = Interested	-0.03	0.97 [0.52, 1.79]	.92	0.46	1.59 [0.82, 3.10]	.17
Video gaming = Frequent (at least twice a month)	0.36	1.43 [0.75, 2.73]	.28	0.39	1.47 [0.71, 3.06]	.30
Response time	0.00	1.00 [0.99, 1.01]	.59	-0.01	0.99 [0.98, 1.00]	.15

Note. CI = confidence interval. *B* and odds ratios (*OR*) estimate the degree of change in accuracy associated with one unit change in the independent variable. An odds ratio of 1 indicates no effect of the independent variable on accuracy, values of 1.5, 2.5, and 4.0 are generally considered to reflect small, medium, and large effect sizes, respectively (Rosenthal, 1996). The category order for factors was set to descending to make the reference level 0. The reference groups are: Reflection position = back, Video game playing = Infrequent (never/less than once a month/about once a month), Gender = Male, Interest in photography = Not Interested. Response time, confidence, and reflection vanishing point distance were added as continuous variables. The five subjects who chose not to disclose their gender were excluded from these analyses leaving a total sample of n = 115. The reflection position predictor variables was not applicable in the consistent reflection scenes.

Conclusion

In three experiments we examined people's ability to identify whether scenes contained consistent or inconsistent reflections. The results of Experiment 1 suggest that people have an extremely limited ability to identify when reflections in a scene are consistent or inconsistent. Yet, in Experiment 2, when we brought the reflection vanishing point closer to the centre of the image and also moved the inconsistent reflections further from the consistent position, subjects' performance on the task improved. Moreover, in a third experiment we kept the vanishing point position the same as in Experiment 2 but decreased the distance between the inconsistent and the consistent reflection position to match the distance in Experiment 1. Subjects in Experiment 3 correctly classified fewer of the consistent and inconsistent scenes than in Experiment 2 but more than in Experiment 1. Thus, it seems people's understanding of how reflections should appear in images is not straightforward, but rather might depend on various factors, including the location of the reflection vanishing point and extent of the inconsistency.

The results in Experiment 1 further support the notion that people have a limited and biased understanding of reflections (e.g., Bertamini et al., 2003; Croucher et al., 2002; Farid & Bravo, 2010; Muelenz et al., 2010). In line with the virtual world rotation hypothesis, we found an effect of reflection position—subjects accurately identified more of the forward than backward inconsistent reflections which supports the notion of a perceptual outward bias (Muelenz et al., 2010). Conversely, this finding might be more simply accounted for by perspective projection (Farid, 2016; Hartley & Zisserman, 2004). To illustrate, consider the image in Figure 5.3 that shows train tracks receding into the distance. Although in reality there is, of course, a fixed distance between the tracks, the properties of perspective projection mean that the tracks appear closer together in the background than in the foreground. The same principle applies to our scenes: we moved the forward and backward inconsistent reflections the same distance in the 3-D world and accordingly, in the 2-D image, the difference appeared larger in the foreground than in the background. Therefore, people might have more easily noticed the forward than backward inconsistent reflections in the scenes. Overall, the findings in Experiment 1 suggest that people did not take advantage of the geometrical information provided in the scenes to make an objective judgement about the reflections. Instead, our results indicate that people relied on a subjective assessment and in line with previous findings, it seems this assessment was based on incorrect beliefs about reflections (e.g., Bertamini et al., 2003; Croucher et al., 2002). It remains possible, however, that subjects found it too difficult make use of the geometrical information to locate the reflection vanishing point in the scenes used in Experiment 1.

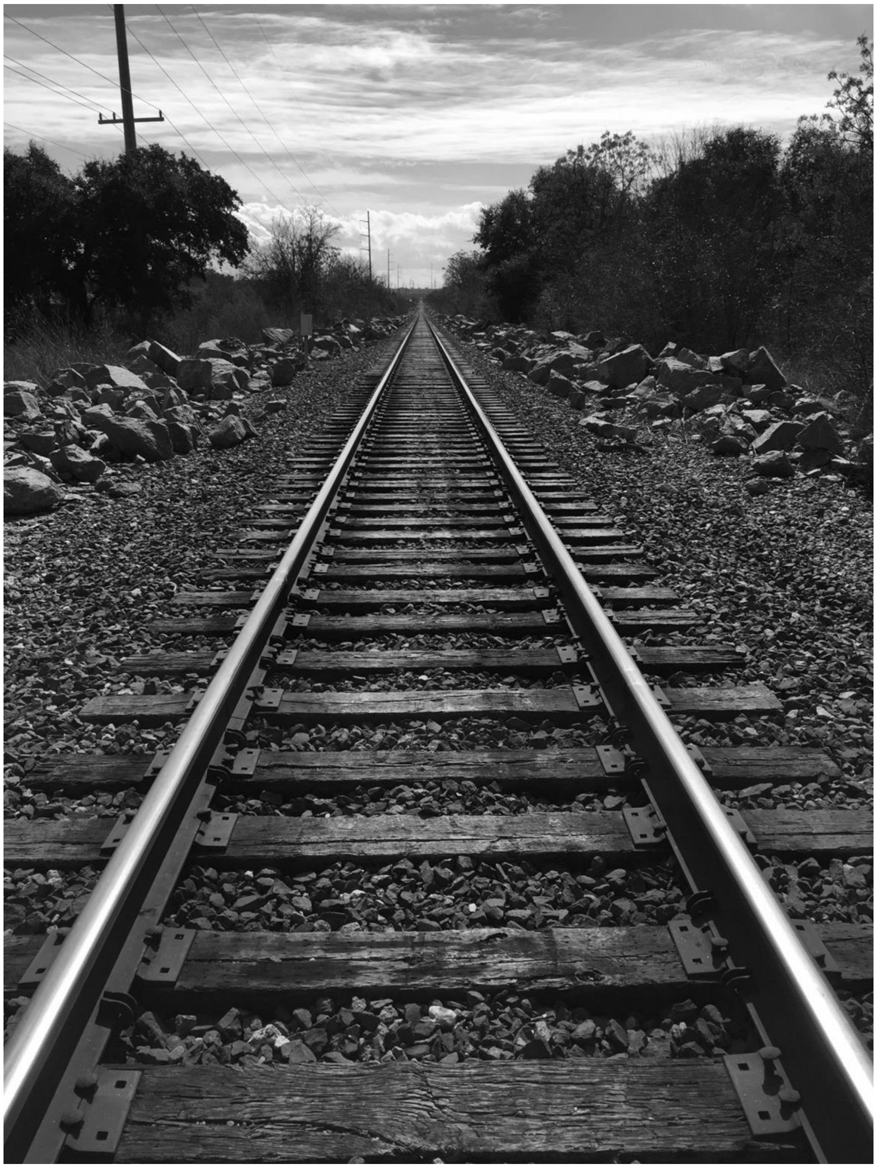


Figure 5.3. Train tracks receding into the distance. Creative Commons license CC0 1.0 / Public Domain.

Indeed, in Experiment 2 when we made it easier for subjects to locate the reflection vanishing point they showed a reasonably good ability to identify whether scenes contained consistent or inconsistent reflections. Furthermore, the results of the GEE analyses revealed that accuracy on the task was associated with distance of the reflection vanishing point from the centre of the image—accuracy was higher when the reflection vanishing point was closer compared with further from the centre of the scene. Seemingly, when the geometrical information in the scene is relatively easy to use, people might rely on it to objectively judge the validity of reflections in a scene. If so, it raises the possibility that people might be able to use this information to help detect manipulated images in the real world. And although our results suggest the ability to make use of this information is limited to situations in which it is easy to apply the geometric analysis, it is possible that with training and practice people's ability will improve.

Yet, we cannot rule out another explanation for subjects' better ability in Experiment 2: moving the inconsistent reflections further from the consistent position in Experiment 2 than in Experiment 1 might simply have made it more visually apparent that the reflections were inconsistent. In fact, the finding that reflection position—forward or backward—had an effect on subjects' accuracy in Experiment 1 but not Experiment 2 offers some support for this explanation. We proposed that this effect of reflection position might be owing to the principles of perspective projection. That is, to give the impression of depth in a 2-D image objects appear smaller and closer together in the background than in the foreground of a scene. Thus it is possible that perspective makes it more difficult for people to make accurate judgements about objects and their reflections when they appear further away than when they are closer. We did not, however, replicate the effect of reflection position in Experiment 2—subjects were as accurate on the backward inconsistent reflections as they were on the forward inconsistent reflections. Perhaps, then, there is a point at which the influence of perspective becomes less important than sensitivity to the absolute distance between the object and its reflection. That is, the results raise the possibility that people have a perceptual threshold for detecting inconsistencies in reflections.

To further test this possibility of a perceptual threshold for detecting inconsistencies in reflections we conducted a third experiment. In this third experiment, the scene reflection vanishing point remained the same as in Experiment 2, but we decreased the distance between the inconsistent reflections and the consistent position to match the distance in Experiment 1. In Experiment 3, subjects correctly classified a mean 11% fewer of the scenes than in Experiment 2. Furthermore, unlike in Experiment 2, subjects showed a bias to accept the reflection scenes as consistent. Also, we found an effect of reflection position—subjects correctly identified more of the forward inconsistent reflections than the backward inconsistent reflections. Together, these findings offer some support to the notion of a perceptual threshold for noticing when the reflections in a scene are inconsistent.

Specifically, decreasing the distance between the inconsistent reflection and the consistent position might simply have made it more difficult to determine whether or not the reflections were inconsistent based on a visual inspection of the scene. It is important to note, however, that adjusting the distance of the inconsistent reflections from the original position also changes the angle difference between the scene reflection vanishing point and the reflection vanishing point for the inconsistent reflection. As such, it remains possible that there is another explanation for these findings—that it is the change in the angle difference that affects people's ability to use the geometric information in the scene. Our results also offer some support for this idea that subjects are using the geometric information to help them to identify the consistent and inconsistent reflections. In particular, in Experiment 3, a mean 12% more of the reflection scenes were correctly classified than in Experiment 1. In both experiments, the distance between the inconsistent reflection position and the consistent position was the same, but the scene vanishing point was closer to the centre of the image in Experiment 3 than in Experiment 1. Therefore, making it easier to use the geometric information in the scene might also influence people's ability to determine when reflections are consistent and when they are inconsistent.

Future research might examine which of these explanations—geometric analysis or a perceptual threshold—best accounts for our findings. Unfortunately, however, it is difficult to isolate the variables of interest: as noted above, adjusting the distance of the inconsistent reflections from the original position also changes the angle difference between the scene reflection vanishing point and the reflection vanishing point for the inconsistent reflection. As such, determining the general strategy subjects rely on to make judgments in the reflection task is not straightforward. Perhaps one fruitful approach would be to examine whether training people to use the geometric analysis would influence performance on the task.

Irrespective of the explanation, our results reveal that people have only a limited ability to detect reflection inconsistencies. Moreover, this ability appears to be dependent upon a number of factors, including the location of the reflection vanishing point and the extent of the inconsistency. Essentially, then, people might have a basic understanding of mirror reflections, but this understanding does not appear to be particularly precise. And, perhaps most importantly, our results imply that in the real world, where people are not prompted to check the geometry of reflections, simple inconsistencies will go undetected. Such was the case when major news outlets were fooled by the fake image that depicted the innocent Veerender Jubbal as a terrorist.

Part Two

Chapter 6 :

Introduction to photos and memory

"Today everything exists to end in a photograph."

Susan Sontag (1977)

Introduction

Snap-happy photographers are used to the quip that they are "taking pictures rather than living in the moment" (Jefferies, 2013). Indeed, some critics suggest that taking photos detracts from an experience and impairs people's ability to engage with the real world (Mols et al., 2015). Nonetheless, people take photos because they believe it helps them to remember (Chalfen, 1998; Harrison, 2002). But there is good reason to think that photography could have the opposite effect. To what extent does taking photos actually hurt our ability to recall or recognise those photographed objects later? That is the question we are interested in here.

It is also the question that Henkel (2014) set out to answer. To examine the effects of taking photos on memory, Henkel took a group of adults on a museum tour. She asked them to photograph some objects, but to simply view others. The next day, everyone took two kinds of memory tests—free-recall and recognition—to determine if their memories were enhanced or impaired by the act of taking a photo. Although subjects performed similarly well on the free-recall test regardless of whether they had photographed that object or not, the recognition test told a different story: Subjects recognised fewer of the objects they had photographed—and fewer details about those objects—relative to objects they merely viewed. Table 6.1 summarises Henkel's results.

Taking photos might hurt memory

These results suggest that photographing objects might hurt memory—findings that make sense in light of the wider research on attention and memory. We know, for example, that attention is key to remembering (Chun & Turk-Browne, 2007; James, 1890), yet, attentional resources are thought to be finite; a bottleneck restricts the amount of information we process at any one time (Levy, Pashler, & Boer, 2006; for a review, see Pashler, 1994). Multitasking divides these finite resources across two or more tasks, impairing people's performance on the primary task (e.g., Broadbent, 1958). In experiments using driving simulators, talking on the phone or texting while driving, for example, impairs people's driving ability (Drews, Yazdani, Godfrey, Cooper, & Strayer, 2009; Strayer & Johnston, 2001). In addition, outside of the laboratory and the driving simulator, we see the same problem in real-world multitasking scenarios (e.g., Hyman, Boss, Wise, McKenzie, & Caggiano, 2010). If taking photos is a similar act of multitasking, then it is reasonable to expect that photographing objects might impair people's ability to attend to, encode, and later remember those objects. Such a process could help us to understand why Henkel (2014) found that when people took photos, it impaired their ability to recognise objects later.

Henkel (2014) proposed a plausible alternative account of her findings: People pay less attention to experiences they capture on camera because, after all, the camera will "remember" for them. The idea that we offload our memories to external storage systems is not new. Couples in long-term relationships often form a shared "transactive" memory system, dividing their knowledge, and accessing that knowledge from each other when necessary (Hewitt & Roberts, 2015; Wegner, 1986). People use digital technology, including computers and smartphones, similarly—to extend their own limited cognition (Sparrow, Liu, & Wegner, 2011; Storm & Stone, 2015). It seems reasonable, then, to surmise that when Henkel's subjects photographed objects in a museum, it encouraged them to offload their memories onto the camera.

Taking photos might not hurt memory

But as much as Henkel's (2014) findings fit with research in the attention and memory literatures, the findings were puzzling in light of other literatures. For instance, the evidence on dual-task interference is inconsistent: multitasking does not always hurt. Numerous studies have shown that in some circumstances people can perform two tasks concurrently at a level comparable to single-task performance (e.g., Humphreys, Watson, & Jolicœur, 2002; Schumacher et al., 2001). So, how can we reconcile these findings with the studies that have found clear evidence of a dual-task cost? Crucially, the answer to that question might depend on various task-related conditions. For example, dual-task performance is highly sensitive to task instructions. In many dual-task studies, subjects are instructed to prioritise one task over the other and the impaired performance is observed for the non-prioritised task (Meyer & Kieras, 1997; Meyer et al., 1995). In one study in which the performance on two simultaneously conducted tasks were given equal emphasis, and subjects were given extensive practice, dual-task costs were found to be minimal (Schumacher et al., 2001).

What do these findings suggest for the effect of taking photos on memory? Consider that when people take a photo, they make a series of subjective decisions (Estrin, 2013; Kurtz, 2015). People ask themselves “Where should I stand?” and “How should I frame this photo?” In doing so, they evaluate the environment to determine what should appear in the photo and what should not. This considered approach should help to prioritise processing of the visual scene and minimise whatever distraction the act of photographing might cause.

The “enactment effect” also makes Henkel’s (2014) findings surprising. Taking photos should make an experience more distinctive; after all, the enactment effect shows that learning-by-doing enhances memory more than learning-by-observing (for a review, see Roediger & Zaromb, 2010). Performing an action draws people’s attention to the specific details of objects, thereby providing a richer encoding that results in a more distinctive memory trace (Engelkamp & Zimmer, 1997; Engelkamp, Seiler, & Zimmer, 2004). Taking photos of an object, then, as opposed to simply viewing it, might also draw people’s attention to details, creating a stronger, more detailed memory trace and ultimately improving memory for that object.

Considered together, the literature supports two predictions: that taking photos will hurt memory, and that taking photos will help memory. Henkel’s (2014) findings offer limited support for the idea that taking photos can hurt memory, however, the size of the effects was rather small. In Henkel’s Experiment 1, subjects recalled only 1.0% more, and recognised only 2.8% more, of the objects they viewed than the objects they photographed. In Experiment 2, there was no free-recall test but subjects recognised just 4.6% more of the objects they viewed than the objects they photographed. Furthermore, Henkel’s results showed that taking photos does not always hurt memory; zooming in to photograph a part of the object had only a trivial effect on memory—subjects recognised just 1.2% more of the objects they simply viewed than the objects they photographed partly. Therefore, Henkel’s findings do not offer good support for one prediction or the other. These findings, when considered against a backdrop of research supporting opposite predictions, prompted us to learn more about the nature of the photo-taking impairment effect in general. Our first experiment, a close replication of Henkel's original study yielded findings in line with the idea that there was plausibly no effect of photo-taking on memory. We conducted the next four experiments to further examine the influence of photo-taking on memory: To improve experimental control, we developed a computer-based analogue of Henkel's procedure. This highly controlled paradigm allowed us to carefully test for a photo-taking memory decrement across a wider range of conditions than has previously been tested. Expanding the "test-space" in turn increased the possibility of detecting any effects of photo-taking on memory as well as allowing us to examine the generalisability of any such effects.

Table 6.1
Mean Difference Scores [95% confidence intervals] from Henkel's (2014) Experiments.

	Experiment 1	Experiment 2	
Test	View vs. Photo-whole	View vs. Photo-whole	View vs. Photo-part
Free-recall	1.0% [-5.8%, 7.8%]	-	-
Name- and photo-recognition	2.8% [-0.1%, 5.7%]	4.6% [0.6%, 8.6%]	1.2% [-1.5%, 3.9%]
Source accuracy	5.3% [-1.5%, 12.2%]	-1.3% [-10%, 7.5%]	-5.4% [-13.4%, 2.7%]
Visual details	9.9% [3.7%, 16.0%]	5.9% [0.5%, 11.3%]	-2.2% [-8.0%, 3.6%]
Location	-	6.8% [1.7%, 11.8%]	29.2% [21.8%, 36.7%]

Note. The 95% confidence intervals are calculated from the planned contrasts of the sample means. In Experiment 1, subjects either viewed (view) or photographed objects as a whole (photo-whole). In Experiment 2, as well as the view and photo-whole conditions, there was a third condition in which subjects photographed a specific part of the object (photo-part). Henkel (2014) used five types of memory test: (a) free-recall—subjects wrote down the names of the objects, (b) name- and photo-recognition—subjects distinguished between objects from the museum and objects they had not seen before, first on a memory test where the name of each object were cues, and then on a test where the name plus a photo of each object served as cues, (c) source accuracy—subjects recalled their action (observe or photograph) for each remembered object, (d) visual details—subjects answered multiple choice questions about visual details of the objects, (e) location-recognition—subjects indicated the location of the objects in the museum. Cells with a dash indicate that data for the test were not collected. A positive mean difference indicates a better performance in the view than photo-whole or photo-part condition. A negative mean difference indicates a better performance in the photo-whole or photo-part than view condition.

Chapter 7 :

To what extent does taking photos affect what people remember?

Experiment 1

In Experiment 1, we aimed to replicate the basic patterns in Henkel's (2014) Experiment 1.

Method

Subjects

A total of 42 subjects (*M* = 23 years, *SD* = 4.5, range = 18–35; 26 women) from Otago University completed the experiment individually and received NZD30. All subjects confirmed that they had not visited the Otago Museum in at least the past year. One additional subject failed to complete the experiment. We aimed to recruit as many people as possible given budgetary constraints.

Design

We used a single-factor (Action: view, photograph) within-subjects design.

Procedure

We followed Henkel's (2014, Experiment 1) procedure, with seven exceptions that served to improve experimental rigor and generalisability; details of these exceptions appear in Appendix B (Table B.1).

Subjects arrived at the museum and received detailed instructions about the experiment. They practiced using the camera function on an iPod Touch (32GB), learning how to zoom and focus correctly. We told them to frame each photo carefully so that the camera captured the whole object. We also asked subjects to pay attention to each of the objects because they would be asked about them the following day. Subjects were told, falsely, that their photos would be available during the second session. We gave subjects this false information to make the process more similar to everyday life—after all, if people really do rely on cameras to remember for them, it is likely that they do so believing they will have access to their photos later.

We selected 30 objects for the tour, including sculptures, pottery, models, tools, and clothing. The objects were distributed across four rooms, and the tour was structured so that subjects passed each object only once. Half the subjects took the tour starting from one point in the museum (A) and ending at another point (B), and the other half took the tour in the reverse order. Subjects viewed half of the objects but did not photograph them; for the remaining objects, they photographed the whole object. We randomly determined the order in which objects were viewed or photographed, and then counterbalanced across this viewed/photographed dimension as well as the direction of the tour (AB or BA). This counterbalancing resulted in four versions of the museum tour, and subjects were randomly assigned to one of these versions.

To begin the tour, subjects read aloud the name of the first object and the researcher led them to that object. After 25 s observing the object, subjects were either told to turn away from the object (for "view" objects), or to photograph the object in its entirety (for "photo" objects). Regardless of the instruction, subjects had 25 s to view the objects; when they took a photo, they received additional time to do so. Subjects then completed a brief distractor (backward counting) task for 10 s before reading aloud the name of the next object. We repeated this procedure for the remaining 29 objects.

The following day, subjects came into the lab, where we tested their memories for the 30 objects using four tests. Test 1 was a free-recall test. Subjects wrote the names of all of the objects they remembered, regardless of whether they had photographed or only viewed them. Test 2 was a name-recognition test in which subjects distinguished among names of objects from the tour and names of objects they had not seen. Subjects were randomly presented with the names of the 30 old objects and 10 new objects, one name at a time. For each object, subjects indicated whether they took a photo of it, observed it, or believed that the object was not part of the tour. Subjects had unlimited time to respond, and the name of each object remained on the computer screen until subjects made a response. Test 3 was a photo-recognition test. This test was similar to the name-recognition test, but subjects viewed photos rather than names of the objects. The final test was a location-recognition test in which subjects examined printed photos of the 30 objects and used a floor plan of the museum to indicate which of the four rooms each object was located.

Results

For all experiments, we followed Cumming's (2012) recommendations and calculated a precise estimate of the actual size of the effects. We also analysed our data using conventional null hypothesis significance testing; those results appear in the Appendix C. The data is also available online: https://osf.io/54pxy/

Free-recall

To determine the effect of photographing objects on subjects' ability to recall them, we calculated the mean accuracy for each subject, and then classified those means according to whether subjects had viewed or photographed each object. Recall that Henkel's (2014) subjects recalled a mean 1.0% more of the objects they viewed than the objects they photographed. We found a different pattern of results: Taking photos led subjects to recall a mean 7.8% more objects than simply viewing (54.1% vs 46.4%), although the confidence interval around the mean difference indicates that there is plausibly only a small benefit of photographing, M_{diff} 95% CI[10] [2.1%, 13.4%].

Recognition

In the name- and photo-recognition tests, we considered "Observed" or "Photographed" responses to be correct for objects that were part of the museum tour. We considered "Not seen before" responses as correct for new objects that were not part of the museum tour. We used signal detection analysis (Green & Swets, 1966; Stanislaw & Todorov, 1999) to calculate the percentage of correctly identified objects from the museum tour (hits) and the percentage of new objects that were incorrectly identified as objects from the museum tour (false alarms). Because both response bias (c) and sensitivity (d') can each illuminate the pattern of responses, we calculated both, and report the results in Table 7.1. As the top row of the table shows, subjects were biased towards saying they had seen objects on the tour, but the confidence interval indicates that this bias was plausibly small. Further, we found that the plausible value range for d' was greater than zero, indicating that subjects were able to discriminate new from old objects.

[10] The 95% confidence intervals are calculated from the planned contrasts of the sample means.

Table 7.1

Response Bias (c) and Sensitivity (d') in the Name- and Photo-Recognition Tests

Experiment	*c*		*d'*	
	M	95% CI	*M*	95% CI
1	-0.23	[-0.37, -.09]	2.88	[2.63, 3.13]
2	0.09	[0.00, 0.17]	2.12	[1.89, 2.35]
3	0.19	[0.08, 0.31]	1.40	[1.17, 1.61]
4	0.08	[-0.04, 0.21]	1.35	[1.09, 1.61]
5	0.07	[-0.04, 0.19]	1.43	[1.13, 1.73]

Note. In Experiments 4 and 5, *c* and *d'* are calculated based on the photo-recognition test data only. The square brackets show the 95% confidence interval for the sample mean. For response bias (*c*), negative values signify a bias to say that objects were seen on the tour, positive values signify a bias to say that objects were not seen on the tour, zero indicates no bias. For *d'*, higher values indicate better discrimination ability, zero signifies no ability.

We now consider the extent to which taking photos affected subjects' ability to recognise which objects were on the museum tour. In Henkel's (2014) Experiment 1, collapsed across the name- and photo-recognition tests, subjects recognised a mean 2.8% more of the objects they viewed than the objects they photographed. By contrast, we found that subjects' ability to recognise objects was the same regardless of the action they took—they correctly recognised 94.3% of objects in the photograph condition and 94.0% of objects in the view condition, M_{diff} = 0.3%, 95% CI [-1.8%, 2.4%].

We also considered subjects' ability to report whether they photographed or simply viewed each object on the tour. To calculate source accuracy, we again used the data from the name- and photo-recognition tests, but classified responses in a different way. For viewed objects we considered an "Observed" response as correct; for photographed objects we considered a "Photographed" response as correct. Henkel (2014) found a small photo-taking impairment effect such that subjects correctly recalled their action (photographing vs. viewing the object) a mean 5.3% more for viewed objects than for photographed objects. We found a slightly larger impairment effect: Subjects correctly recalled their action a mean 22.2% more for viewed than for photographed objects (70.7% vs. 48.5%), M_{diff} 95% CI [13.5%, 31.0%]. Finally, we found that photographing an object had a trivial effect—and plausibly no effect—on subjects' memory for the location of that object: Subjects recognised location at a similar rate for viewed objects and photographed objects (81.3% vs. 80.3%), M_{diff} 95% CI [-3.2%, 5.1%].

The only finding consistent with the idea that taking photos hurts memory was that subjects were markedly more accurate in recalling the action they had performed for the viewed objects than for the photographed objects. But we found no good support for the idea that taking photos of an object does much to one's memory for that object. In Experiment 2, we developed a computer-based analogue for the museum tour so that we could begin to investigate under which conditions (if any), and to what extent, taking photos might hurt our ability to recall or recognise those photographed objects later.

Experiment 2

In Experiment 2, we followed Henkel's (2014, Experiment 2) method. Accordingly, there were three action conditions: (a) "view", as in Experiment 1; (b) "photo-whole", as in Experiment 1; and (c) a "photo-part" condition, in which subjects photographed a specific part of the object. Using a computer-based analogue for the museum tour afforded more control over the procedure and materials.

Method

Subjects

We recruited 42 subjects (M = 23 years, SD = 3.7, range = 17–35; 20 women) from Warwick University who each received £4. We aimed to recruit as many subjects as possible given financial constraints.

Design

We used a single-factor (Action: view, photo-whole, photo-part) within-subjects design.

Procedure

To increase experimental rigor and/or generalisability, we made five amendments to Henkel's (2014) Experiment 2 methodology, details of which appear in Appendix B (Table B.2). Subjects arrived at the lab and received detailed instructions about the experiment. Before the experiment proper began, subjects practiced navigating through a virtual tour of an exhibition at London's Tate Britain. They also practiced each of the three actions—view, photo-whole, photo-part—that they would perform during the experiment.

We used BlitzMax IDE (Version 1.26; Blitz Research Ltd), a videogame programming language, to control all details for each of the stops, including instructions to subjects, and how long they had to view or photograph objects. To take photos, subjects used a virtual camera that appeared on-screen and mimicked a real camera in several ways: Subjects "lined up" their photo by moving the mouse; they could also zoom in on the object by using the "up" and "down" arrow keys on the keyboard. Subjects took the "photo" by clicking a mouse button. Each time subjects took a photo, the program paused for 2 s while displaying the image captured by the camera. Subjects were told to make sure they framed each photo carefully and to take the best shot possible.

Then the experiment proper began. This tour was based on another exhibition at London's Tate Britain, and included stops at 15 objects, five for each of the three action conditions. Figure 7.1a shows some sample objects. The video recording of the museum tour was captured at a resolution of 1920×1080 on an Olympus Stylus SH-50 and was 5 min 27 s long. The objects were located in a single room, and the tour was structured so that subjects passed each object only once. We randomly presented the order that the objects appeared in the view, photo-whole and photo-part action conditions and then counterbalanced the objects across the three conditions. For photo-part objects, we selected two parts of each object that subjects could zoom in on to photograph, and the specified part was also counterbalanced. Our counterbalancing resulted in six versions of the museum tour, and subjects were randomly assigned to one of these versions.

At each of the 15 stops, subjects saw the object name on-screen for 1.5 s and then, regardless of condition, had 5 s to view the object. For "viewed" objects, subjects did not perform any further action, and the museum tour continued. For the photo-whole and photo-part objects, after the 5 s view time a message appeared at the bottom of the screen instructing subjects which action to take: to photograph the whole object or to zoom in and photograph the part of the object as specified in the message. Figure 7.1b shows an example of the virtual camera being used in the photo-whole condition. In both photo conditions, subjects were given an additional 10 s to take the photo. Subjects' photos were saved so that we could verify that they had followed the instructions.

After the tour phase, subjects worked on a 20-min filler task—a crossword puzzle—before completing four memory tests. First, as in Experiment 1, subjects completed the free-recall test. Second, subjects completed the name-recognition test, in which they were required to distinguish between the 15 objects from the tour and five objects they had never seen. Third, subjects completed the visual details test: For each object recognised as old in the name-recognition test, subjects answered two multiple-choice questions about visual details of that object. Finally, subjects completed the photo-recognition test.

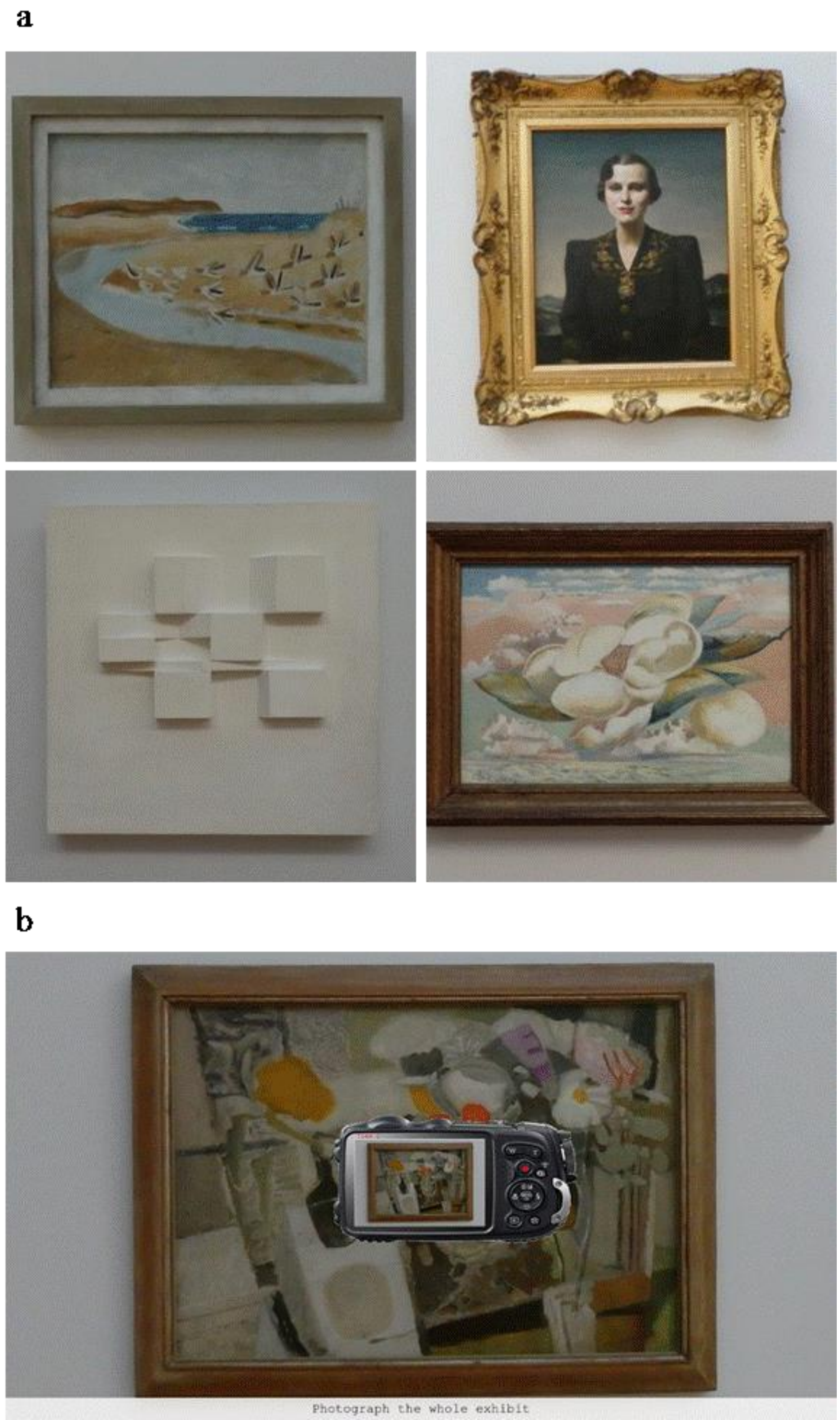


Figure 7.1. Examples of the objects (a) and virtual camera (b) used in Experiments 2 and 3.

Results

We examined subjects' photos to verify that they photographed the whole of the object or the specified part of the object when instructed. All subjects followed instructions.

Free-recall

In line with the pattern of results from Experiment 1, subjects recalled a mean 6.7% more photo-whole objects than view objects (31.0% vs 24.3%), M_{diff} 95% CI [-1.3%, 14.7%]. Further, subjects recalled a mean 4.3% more of the photo-part than view objects (28.6% vs 24.3%), M_{diff} 95% CI [-3.6%, 12.1%]. Yet the confidence intervals around both of these mean differences shows that there was plausibly no benefit of taking a photo on recall.

Recognition

In contrast to Henkel (2014), our subjects recognised a mean 4.0% more photo-whole objects than view objects (84.0% vs. 80.0%), M_{diff} 95% CI [-1.2%, 9.3%]. Further, subjects recognised a mean 5.0% more photo-part objects than view objects (85.0% vs. 80.0%), M_{diff} 95% CI [0.3%, 9.7%]. It is important to note that the confidence intervals for both of these differences include zero; it is therefore plausible that photographing did nothing at all to subjects' ability to recognise the objects they encountered during the tour.

Recall that subjects also answered two multiple-choice questions about the visual details of objects that they recognised as old in the name-recognition test. Our results suggest that photographing had plausibly no effect on subjects' memory of the objects: Subjects correctly answered a mean 4.8% more of the visual detail questions for photo-whole than view objects (35.0% vs. 30.2%), M_{diff} 95% CI [-1.4%, 11.0%], and 3.8% more for photo-part than view objects (34.0% vs. 30.2%), M_{diff} = 3.8%, 95% CI [-1.8%, 9.4%].

Next we examined subjects' ability to report the specific action they performed for each of the objects on the tour. In our first experiment, subjects recalled their action a mean 22.2% more often for viewed than for photographed objects. This time, similar to Henkel's (2014) findings, source accuracy for the view and photo-whole objects was plausibly no different, M_{diff} = 1.9%, 95% CI [-9.2%, 13.0%]. Furthermore, subjects were more likely to remember the action they performed for photo-part objects than for view objects (61.0% vs. 46.9%), M_{diff} = 14.0%, 95% CI [6.7%, 21.4%].

In sum, our first experiment using the computer-based analogue of the museum tour did not offer any good evidence to suggest that taking photos hurts memory. Nonetheless, it is possible that such memory effects only occur (or occur more strongly) sometime after a photo is taken. To examine this possibility, in Experiment 3 we included a 48-hour delay between the tour phase and the test phase.

Experiment 3

Method

Subjects

We initially recruited 42 subjects but replaced 13 who completed the tour phase of the experiment but not the online test emailed to them 48 hours later. Our final sample included 42 subjects (*M* = 32 years, *SD* = 14.9, range = 16–56; 11 men, 24 women, 7 declined to respond) who participated at Warwick University in exchange for £4 or volunteered to participate without payment.

Design

We used a single-factor (Action: view, photo-whole, photo-part) within-subjects design.

Procedure

The procedure was similar to Experiment 2, but we increased the delay between the museum tour and memory test from 20 min to at least 48 hours. We emailed subjects a link so that they could complete the memory test online. The mean delay between the museum tour and memory test was 64 hours (range = 48–148). We also excluded the free-recall test, based on subjects' poor performance in Experiment 2.

Results

Recognition

As in Experiment 2, we found that photographing objects had plausibly no effect on subjects' ability to recognise which objects were on the museum tour. Subjects recognised a mean 2.1% more photo-whole than view objects (69.0% vs. 66.9%), 95%

CI [-4.8%, 9.0%]. Further, subjects recognised 3.1% more photo-part than view objects (70.0% vs. 66.9%), 95% CI [-2.5%, 8.7%].

For the visual details test, our findings were similar to those in Experiment 2: There was plausibly no difference in the mean proportion of correctly answered questions for view and photo-whole objects (24.3% vs. 25.2%; M_{diff} = 1.1%, 95% CI [-5.3%, 7.3%]), or view and photo-part objects (24.3% vs. 23.3%; M_{diff} = 1.0%, 95% CI [-5.5%, 7.4%]).

Finally, we considered source accuracy. Subjects' ability to recall whether they viewed or photographed an object did not differ between the view and photo-whole conditions—a mean 35.0% for source accuracy in the view condition and 34.8% in the photo-whole condition, M_{diff} = 0.2%, 95% CI [-6.7%, 7.2%]. Recall that in Experiment 2, subjects remembered the action performed for 14.0% more photo-part than view objects. This time, we found plausibly no difference, as subjects recalled their action a mean 4.3% more often for photo-part than for view objects (39.3% vs. 35.0%), M_{diff} = 4.3%, 95% CI [-3.6%, 12.2%].

To summarise, in Experiment 3 when we extended the retention interval between the tour phase and the test phase we still found there was plausibly no effect of taking photos on memory. Nonetheless, a critic might counter that although the computer-based paradigm gave us much more control than walking people around a real museum, a video recording of a museum tour still left us with several factors that we could not control. Using the video recording we could not, for example, manipulate the attributes of the target and distractor objects or the size and layout of the museum. To address this criticism, we created a method with even more experimental control, and to minimise the influence of noise from these other factors, we developed a virtual museum.

Experiment 4

Method

Subjects

We recruited 42 subjects (M = 30 years, SD = 15.6, range = 16–67; 27 women) from Warwick University who volunteered to participate without payment.

Design

We used a single-factor (Action: view, photo-whole, photo-part) within-subjects design.

Procedure

The procedure was similar to Experiment 3, with eight exceptions that added experimental rigor and/or generalisability; details of these exceptions appear in Appendix B (see Table B.3). Subjects arrived at the lab and received detailed instructions about the experiment. This time, rather than watching the video of the museum tour at the Tate Britain, subjects were navigated through a virtual museum created using 3D World Studio (Version 5.6; Leadwerks). To simulate a real museum, we included both paintings and sculptures. For the paintings, we downloaded images, permitted for non-commercial re-use with modification, via Google Image search. For the sculptures, we obtained 3-D models from turbosquid.com. Figure 7.2a shows some sample objects. As in Experiments 2 and 3, we used BlitzMax IDE (Version 1.26; Blitz Research Ltd) to control all details for each of the 15 stops on the tour and to create the virtual camera. Figure 7.2b shows an example of the virtual camera being used in the photo-whole condition. Following the tour phase, subjects were given a 2-min filler task before completing the photo-recognition test. In the photo-recognition test, the 15 objects from the museum tour were randomly intermixed with 15 new objects. We piloted the new materials and found that, even with a 2-min delay, subjects' recognition performance was at a level similar to that observed in Experiment 3.

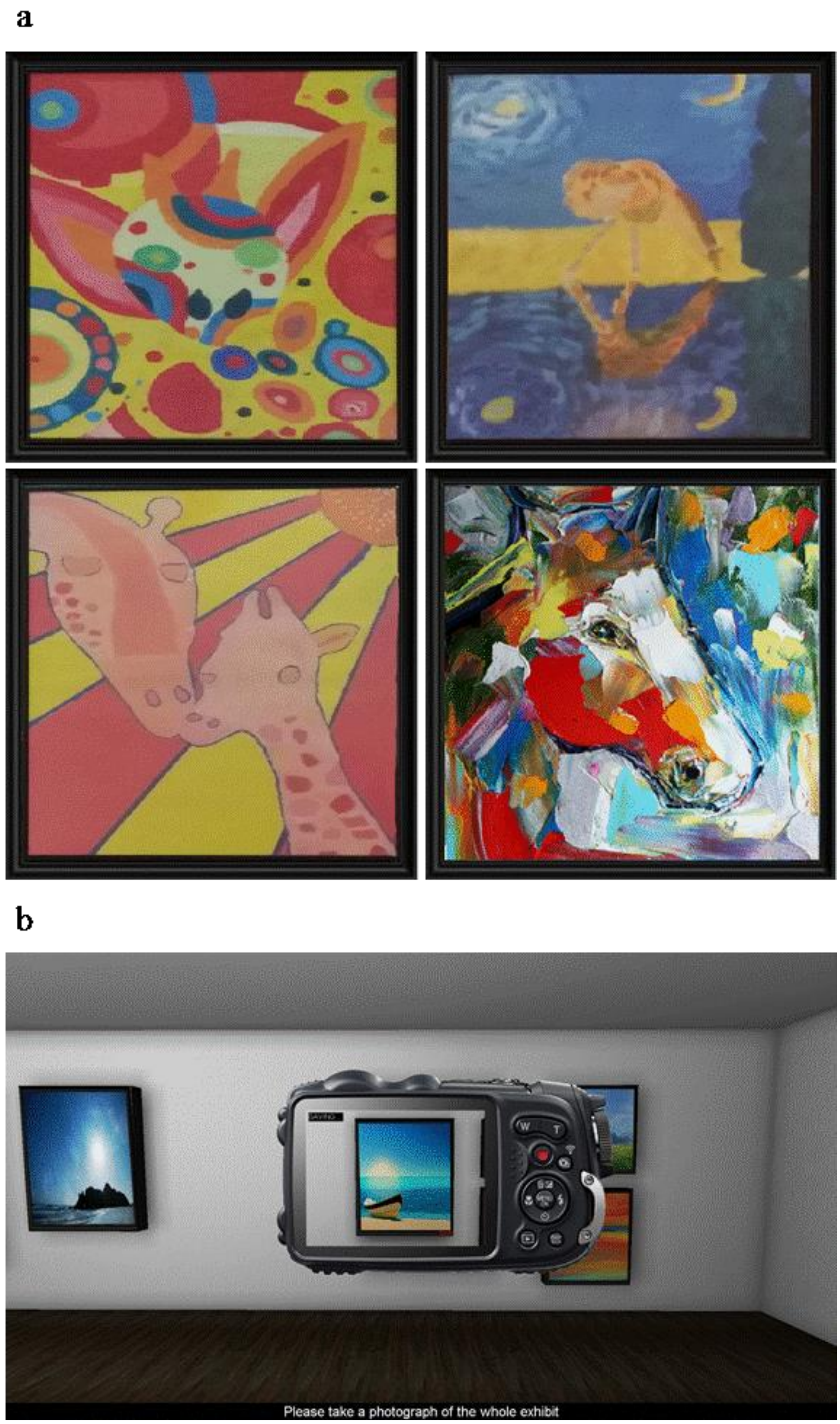


Figure 7.2. Examples of the objects (a) and virtual camera (b) used in Experiments 4 and 5.

Results

Recognition

Once again, we found that taking photos had plausibly no effect on subjects' ability to recognise which objects were on the museum tour. Subjects recognised a mean 5.7% more view than photo-whole objects (70.5% vs. 64.8%), M_{diff} 95% CI [-3.0%, 14.5%]. In addition, subjects recognised 7.1% more photo-part than view objects (77.6% vs. 70.5%), M_{diff} 95% CI [-0.9%, 15.1%]. The one notable finding was that the way in which subjects took photos might have had a small effect on their ability to recognise objects: Subjects recognised a mean 12.9% more photo-part than photo-whole objects (77.6% vs. 64.8%), M_{diff} 95% CI [3.6%, 22.1%].

There are two possible reasons why subjects might be more likely to remember objects they partially—as opposed to wholly—photographed. One reason is that subjects pay more attention to the object when they are asked to locate and then zoom in to photograph a certain portion of it, thereby making it more memorable. A second reason is that by instructing subjects to photograph a specific part of the object we provide an additional memory cue—a verbal description of part of the object. Which of these reasons, then, accounts for subjects' remembering partially photographed objects better than wholly photographed objects? Is it the action of zooming, or is it the verbal description of the object in the instruction that makes these partially photographed objects more memorable? To determine which of these explanations best accounts for subjects' better memory of the objects they photographed partly, in Experiment 5 we changed the wording of the "photo-part" instruction to allow subjects to choose which part of the object to photograph.

Experiment 5

Method

Subjects

We recruited 42 subjects and replaced 2 subjects who failed to follow instructions. The final sample consisted of 42 subjects (M = 21.4 years, SD = 5.7, range = 17-52; 26 women) from Warwick University who received £2 for participating.

Design

We used a single-factor (Action: view, photo-whole, photo-part) within-subjects design.

Procedure

The procedure was similar to Experiment 4. For the photo-part objects, however, rather than informing subjects specifically which part of the object to photograph, subjects were simply asked to “Please take a photograph of a part of the painting.”

Results

Recognition

In line with the findings in Experiments 1 to 4, our results revealed plausibly no effect of photographing on recognition memory. Subjects recognised 6.7% more of the view than photo-whole objects (72.4% vs. 65.7%), M_{diff} 95% CI [-0.2%, 13.5%]. In addition, subjects recognised 7.6% more of the photo-part than view objects (80.0% vs. 72.4%), M_{diff} 95% CI [-0.9%, 16.1%]. Despite our change to the instruction for how to photograph the photo-part objects, subjects still recognised 14.3% more of the photo-part than photo-whole objects (80.0% vs. 65.7%), M_{diff} 95% CI [6.1%, 22.5%]. This finding does not fit with the account that the verbal description of part of the object provided an additional cue that enhanced memory. Rather, it suggests that subjects' better memory for photo-part objects could be due to the additional attention required to locate and zoom in on a certain portion of the object.

Recall that we developed the computer-based museum tour procedure to examine the influence of taking photos on memory under highly controlled conditions. We used this procedure in Experiments 2-5 to explore a range of factors that might influence whether or not photo-taking impairs memory. In each of our four computer-based experiments and also in Experiment 1—a close replication of Henkel’s (2014) study—we found there was plausibly no effect of taking photos on memory. To obtain a more precise estimate of the extent to which taking photos influences memory, we conducted a "mini meta-analysis" including the data from our experiments and Henkel’s original experiments (see Cumming, 2012, 2013).

Mini meta-analysis of the effect of photo-taking on memory

Here, we examined the data from the photo-recognition tests, the only test used in all five of our experiments and in Henkel's (2014) two experiments. For each experiment, we calculated the mean difference in the proportion of objects that subjects correctly recognised as old when they merely viewed the object (view condition) compared to when they photographed the object as a whole (photo-whole condition). We used ESCI (Cumming, 2013) to conduct the analysis and to generate a random effects model meta-analysis on the mean difference data. As Figure 7.3 shows, the result of the meta-analysis is an estimated raw effect size of 0.02, 95% CI [-0.02, 0.05], $z = .99$, $p = .32$. This estimate represents a 2% advantage for viewing over photographing objects, but includes a range of values that plausibly include photo-taking impairment, photo-taking improvement, or no effect of photographing.

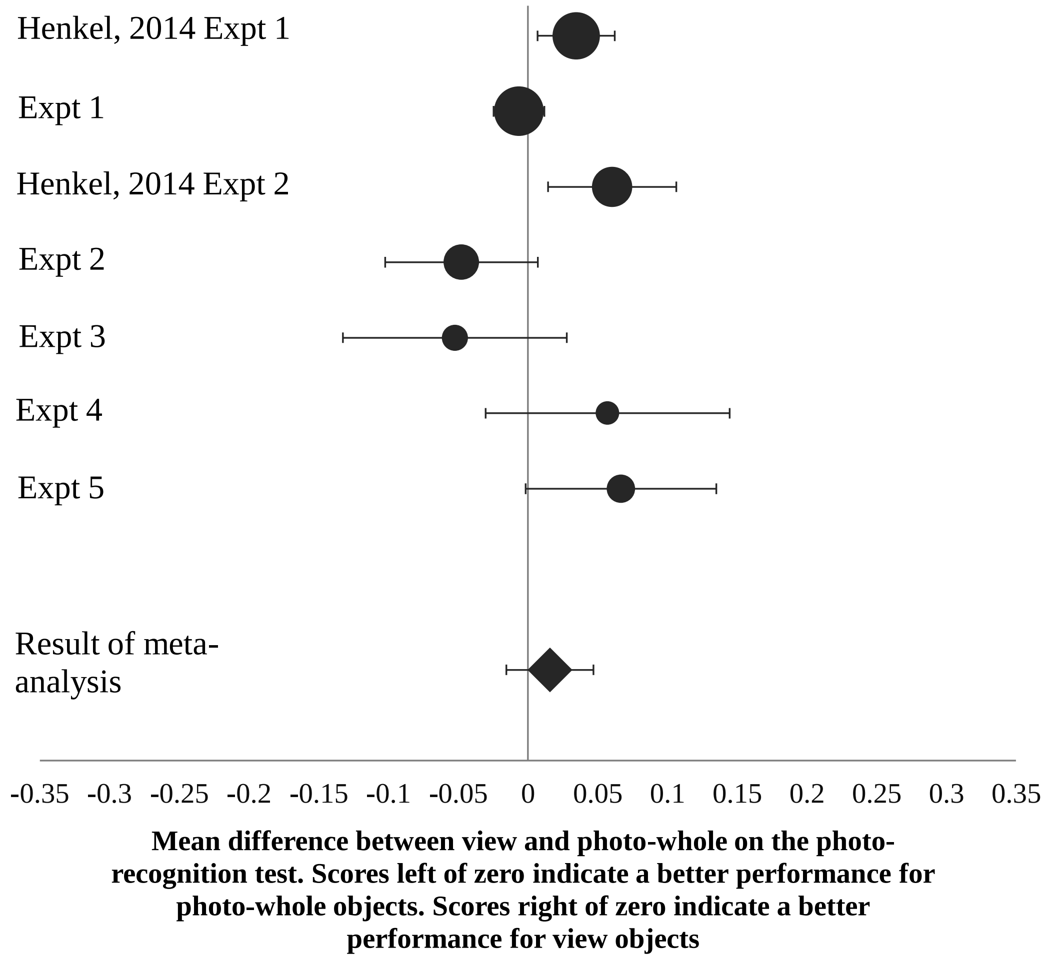


Figure 7.3. Forest plot of effect sizes, including 95% confidence intervals. The location of each circle on the horizontal axis represents the raw effect size—the mean difference in subjects' ability to recognise objects they merely viewed compared to objects they photographed as a whole (calculated using the photo-recognition test data). The size of the circle represents the weighting given to each experiment in the meta-analysis, and is based on sample size and standard deviation. The diamond shape shows the result of the meta-analysis, the estimated effect size. Shapes to the left of zero indicate a photo-taking improvement effect (subjects recognised more objects they photographed as a whole than objects they merely viewed) and shapes to the right of zero indicate a photo-taking impairment effect (subjects recognised more objects they merely viewed than objects they photographed as a whole).

Finally, we calculated the mean difference in the proportion of objects that subjects correctly recognised as old when they merely viewed the object (view condition) compared to when they photographed part of the object (photo-part condition). We calculated these mean differences for each of the five experiments that included the photo-part condition: Henkel's (2014) Experiment 2 and our Experiments 2 to 5. We subjected these data to a random effects model meta-analysis. As Figure 7.4 shows, the result of the meta-analysis is an estimated raw effect size of -0.07, 95% CI [-0.13, -0.01], $z = -2.14$, $p = .03$. This estimate represents a 7% advantage for photographing partly over merely viewing the object, but includes a range of values that indicate there is plausibly only a small benefit.

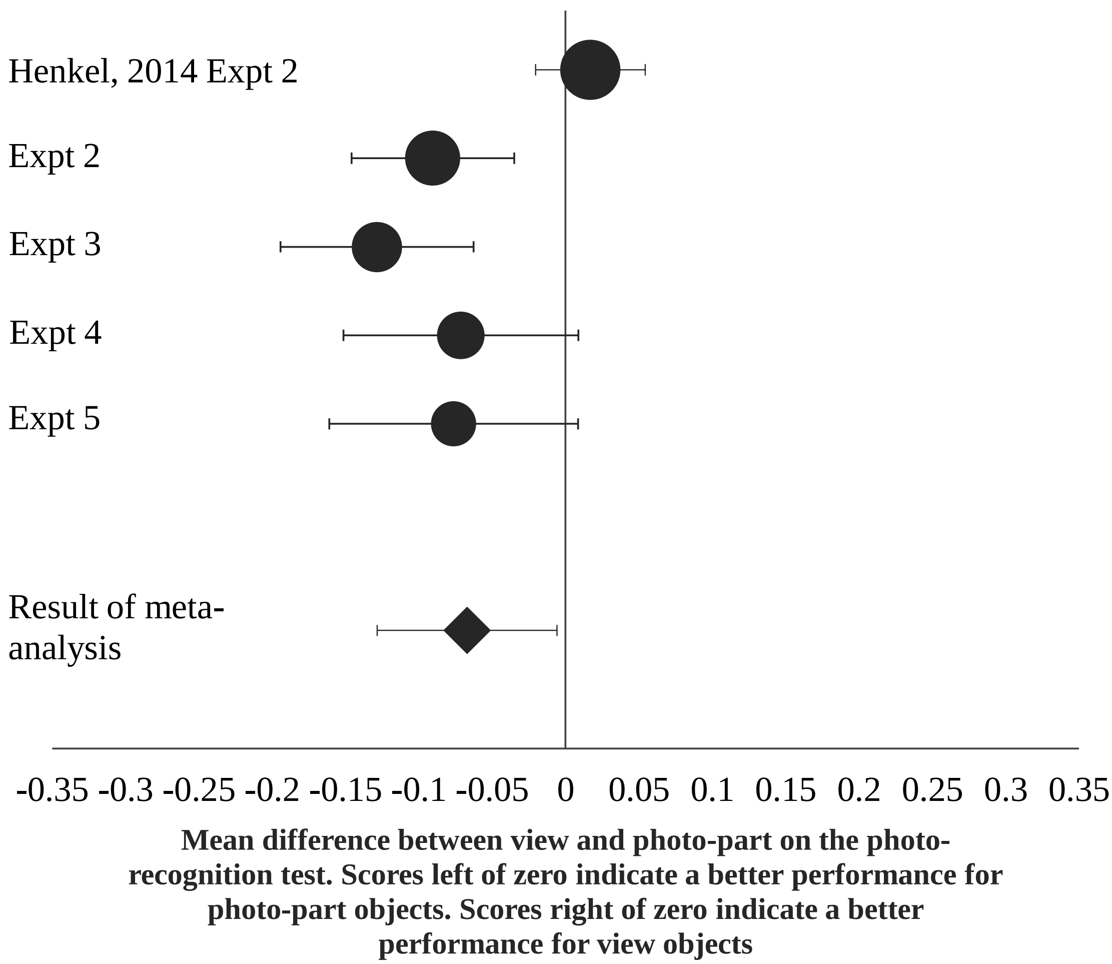


Figure 7.4. Forest plot of effect sizes, including 95% confidence intervals. The location of each circle on the horizontal axis represents the raw effect size—the mean difference in subjects' ability to recognise objects they merely viewed compared to objects they photographed partly (calculated using the photo-recognition test data). The size of the circle represents the weighting given to each experiment in the meta-analysis, and is based on sample size and standard deviation. The diamond shape shows the result of the meta-analysis, the estimated effect size. Shapes to the left of zero indicate a photo-taking improvement effect (subjects recognised more objects they photographed partly than objects they merely viewed) and shapes to the right of zero indicate a photo-taking impairment effect (subjects recognised more objects they merely viewed than objects they photographed partly).

Conclusion

Across five experiments and a mini meta-analysis, we found limited support for the proposal that taking photos impairs people's memories. More specifically, the mini meta-analysis estimated the size of the effect was 2%—but plausibly zero (Cumming, 2012).

How do we reconcile these findings with the extended-cognition literature, which shows that people frequently offload their memories onto external devices, such as smartphones and search engines, instead of committing the information to memory—and that doing so impairs their memory for that information later on (Sparrow et al., 2011; Storm & Stone, 2015)? It seems reasonable to see the act of photographing objects as parallel: Why, then, did we find so little support for this notion that taking photos impairs memory?

One answer might hinge on the type of memories that people offload. People might be willing to offload trivia questions and wordlists, but be more reluctant to offload autobiographical information. After all, autobiographical memories matter: they shape our identity, guide our decisions, and help us to form social relationships (Bluck, Alea, Habermas, & Rubin, 2005). If people are more reluctant to trust their personal memories to external storage, the act of taking photos would not harm memory performance. Future studies might explore this possibility by manipulating the type of information that people are asked to remember to determine the extent to which people are willing to offload semantic rather than personal information.

It is also worth considering the size of the cognitive offloading effect. In light of our findings, we examined Storm and Stone's (2015) recent study which showed that saving previously studied information onto the computer can improve memory for new information. In this study, subjects learned two word lists—A and B. After studying List A, subjects either saved it onto the computer or deleted it, they then studied List B. Finally, after a short delay, subjects were asked to recall the words in List B. Subjects recalled more of the words when they had saved List A than when they had deleted it—seemingly, offloading information onto the computer enhanced subjects' ability to learn and remember new information. The researchers conducted two further experiments to replicate and extend their results. We used the reported data from these three experiments and Cumming's (2013) mini meta-analysis procedure to calculate a more precise estimate of the cognitive offloading effect. The result, an estimated raw effect size of 11%, 95% CI [8%, 15%], suggests a slightly larger effect for offloading (and subsequently failing to report) words than photos. Yet, we cannot know whether this difference is attributable to the medium (words or photos) or the memory (word lists or autobiographical memories) and how that factor affects memory storage.

There are important differences in both Henkel's (2014) original procedure and ours when compared with how real-world people take real-world photos. Even so, these differences might actually mean photos enhance, rather than impair, memory. For example, people usually choose to photograph objects in which they are interested, and therefore might encode these objects better than objects that they are directed to photograph. That interest in turn might well override whatever degree of impairment could be associated with taking a photo. In line with this suggestion, when we asked subjects to choose which aspect of the object to photograph (Experiment 5), they remembered a mean 14.3% (95% CI [6.1%, 22.5%]) more of the objects they photographed partly than the objects they photographed wholly. Given that people's tendency to take so many photos is a relatively new development (Heyman, 2015) there are many outstanding empirical questions to explore, including how taking photos influences our real-world memories.

But considered as a whole, our data support the conclusion that taking photos of an object has little effect on one's memory for that object—and plausibly no effect at all. Our findings also raise new questions about extended cognition in general. With estimates suggesting people will take a record 1.3 trillion photos in 2017 (Heyman, 2015), the possibility that taking photos might have little effect on memory is good news—particularly for the snap-happy among us.

Chapter 8 :
General Discussion

The aims of this thesis were twofold. The aim in Part One was to gain an understanding of people's ability to discriminate between authentic and manipulated images of real-world scenes and to begin to look for ways that might improve this ability. The aim in Part Two was to investigate how taking photos affects people's memory of the events that are photographed. Before discussing the implications of the results, the key findings in each chapter are summarised.

Summary

In Chapter 3, two experiments examined people's ability to detect and locate manipulations within images. The results revealed that people have an extremely limited ability to distinguish between original and manipulated images of real-world scenes. Furthermore, even when subjects correctly detected that an image had been manipulated, they were often unable to locate the manipulation. That said, in Experiment 2, when comparing people's performance on the detection and location tasks with what they should achieve by chance on those tasks, it was surprising to find that people performed better on the location than on the detection task. The results in Experiment 1 offered some support for the prediction that people are better able to identify physically implausible changes than physically plausible ones. This finding, however, was not replicated in Experiment 2, where the results indicated that the amount of change might be more important than the plausibility of the change.

In Chapters 4 and 5 we explored two possible ways that might help people to better identify when images have been manipulated. In Chapter 4, we looked at people's ability to determine the physical consistency of shadow information within a scene illuminated by a single light source. Across five experiments the results showed that people were insensitive to lighting inconsistencies. The results also revealed that the extent of the inconsistency influenced people's ability to detect inconsistent shadows. That is, inconsistent shadows positioned further from the correct position were more likely to be associated with accurate responses than inconsistent shadows positioned closer to the correct position.

In Chapter 5 we explored people's ability to make use of reflection information in scenes. Overall, the results suggested that people had a reasonably limited ability to determine whether the reflections in a scene were consistent or inconsistent. Similar to the results of the shadow experiments, the extent of the inconsistency influenced people's ability to detect manipulations. Specifically, in Experiment 2, when the inconsistent reflections were further from the correct position than in Experiments 1 and 3, people correctly detected more of the inconsistent reflection scenes.

In Chapters 6 and 7, we aimed to learn more about the nature of the photo-taking impairment effect in general. Specifically, we examined the extent that taking photos of objects influenced people's ability to remember those objects later. Across five experiments, taking photos had a small, or plausibly no, effect on memory. Furthermore, the result of a mini meta-analysis including the data from our experiments and Henkel's (2014) original experiments estimated the size of the photo-taking impairment effect to be a trivial 2%.

Theoretical implications

Detecting and locating image manipulations

The research presented in Chapters 3 to 5 has important implications for our understanding of how the visual system processes information. The human visual system is remarkable in many ways. Central to this premise is that people are able to process visual information in a rapid and effortless manner—just a glance at an image is often enough to glean the basic meaning (Potter, 1975; Thorpe et al., 1996). The ease of processing visual information can help to explain why photos have become so popular; for many people, it is easier and quicker to take a message from an image than from a sequence of text. But if people are able to process images so readily, why do they find it difficult to identify manipulations in images? It seems likely that we can account for this finding by considering that there is a limit to the amount of information the visual system can process at any one time. Indeed, continually representing highly detailed information from the visual world in the brain would be cognitively expensive and would overwhelm the system (Liverence & Franconeri, 2015; Tsotsos, 1990). The extent to which the visual detail in the world is represented in visual memory remains a topic for debate.

A number of theories of visual perception exist. The main point that these theories diverge on concerns the amount of information from the visual world that is represented in visual memory. At one extreme is the suggestion that very little of the visual world is processed and that internal representations are sparse, while at the other extreme is the idea that people create a rich and detailed internal representation of the visual world. To better understand these theories it is useful to consider three possible levels of processing within visual perception, as shown in Figure 8.1. The first level involves early and automatic processing of the visual scene, resulting in a low-level representation that is not consciously accessible. At the second level, this low-level representation is coded coarsely; again this happens automatically but the information is now accessible for quick and approximate decisions. The third level involves a separate effortful process that requires focused attention to produce a highly detailed and consciously accessible internal representation of the attended aspect of the scene.

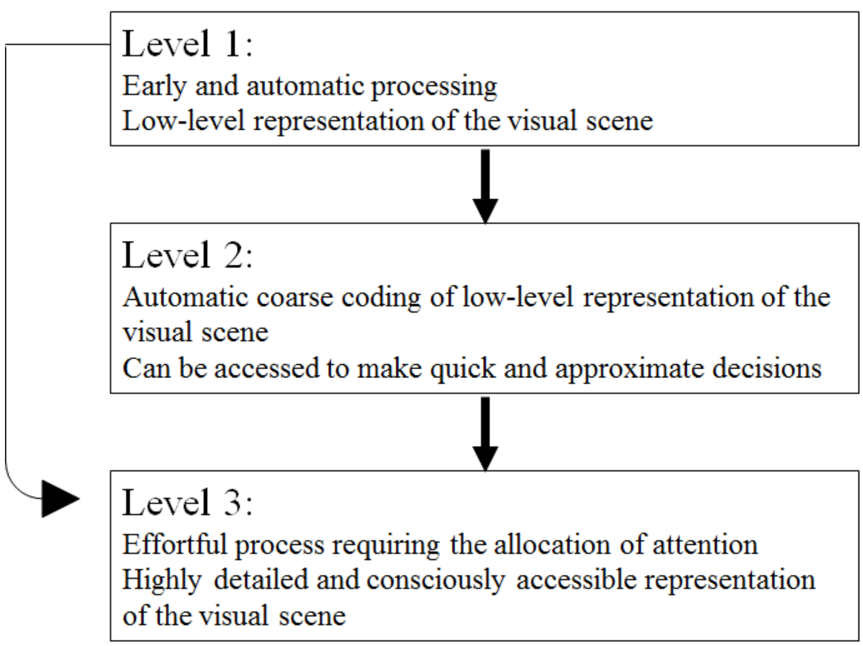


Figure 8.1. Levels of processing approach to visual perception. A number of theories suggest that visual perception involves Levels 1 and 3. Our results in Chapters 4 and 5 indicate that another level—Level 2—might be used when processing shadows and reflections.

The sparse representation theory proposes that the visual world is represented only sparsely in the mind; people might take the gist from a scene but the details are not represented at all, instead the world itself acts as an “external memory” (Dennett, 1991; O’Regan, 1992; Stroud, 1955). Thus, rather than using limited perceptual resources to encode the visual details, this information is accessed directly from the world as and when it is needed. A slightly less extreme version of this sparse representation theory is that visual memory is limited to the currently attended aspect of a scene (e.g., O’Regan, Rensink, & Clark, 1999; Simons & Levin, 1997, 1998). Support for the general notion of incomplete representation of the visual world in the brain comes from change blindness studies in which people are slow to detect even large changes that occur during a real or simulated eye blink. Therefore, although people tend to efficiently encode the aspects of the scene that are important for understanding the gist meaning (Potter, 1975; Schyns & Oliva, 1994), to encode the finer visual details in the scene requires the serial focus of attention on objects through movements of the eyes (Hollingworth & Henderson, 2002; Nelson & Loftus, 1980; O’Regan et al., 1999; Simons & Levin, 1997, 1998). That is, to process finer visual details, and not only the gist of the scene, effortful attention is likely to be required. As such, the incomplete representation theory suggests that visual perception involves a combination of the first and third levels of processing.

The results in Chapter 3 offer some support to this theory. In Experiment 2, subjects performed better in the location task than in the detection task (relative to chance performance for each task). In line with the theories of incomplete visual representation (e.g., O’Regan, 1992; Simons & Levin, 1997), it is possible that the detection task encouraged a more general search of the scene whereby observers encoded the gist but largely ignored the finer visual details. The location task, however, might have encouraged observers to expend greater effort in attending to, and as such encoding, the visual details. Thus the difference in performance on the detection and location tasks might have been a result of people’s tendency to use different strategies in each: The location task might have encouraged an effortful strategy to encode more of the finer visual details in the scene compared to the detection task. Generally then, encoding only the most important aspects of the scene—the gist meaning—is an adaptive strategy that allows the visual and perceptual systems to cope with the continually changing visual input that they receive from the world. Yet neglect of the finer visual details might explain why people often fail to notice when images have been manipulated. Future research might examine whether manipulations that affect the gist meaning of a scene are detected more readily than manipulations that affect the geometric plausibility of the scene.

Moreover, even when people take an effortful approach to attend to the details of a scene, aspects such as shadows and reflections rarely receive attention (Ehinger et al., 2016; Rensink & Cavanagh, 2004; Sareen, Ehinger, & Wolfe, 2015). It is possible therefore that people have little opportunity to learn about how shadows and reflections should look. In fact, we explored specifically whether people could identify if the shadows (Chapter 4) and reflections (Chapter 5) in a scene were consistent or inconsistent and found that, overall, people's ability to do so was extremely limited. This finding supports the idea that these details might not be included in people's internal representation of the visual world and that they infrequently receive effortful attention (Ehinger et al., 2016; O'Regan, 1992; Rensink & Cavanagh, 2004; Sareen et al., 2015). Yet when we looked more closely at the results of our shadow and reflection experiments, we found that the extent of the inconsistency influenced people's performance on the shadow and reflection tasks. Specifically, people were more likely to detect inconsistent shadows and reflections positioned further from the correct position than inconsistent shadows and reflections positioned closer to the correct position. According to the incomplete representation and early discounting theories (e.g., Rensink & Cavanagh, 2004; Simons & Levin, 1997, 1998) such visual details are not usually attended and therefore not "seen"—which means that people should be equally likely to accept an inconsistent shadow or reflection as inconsistent regardless of the extent of the discrepancy.

The difference in people's ability to identify inconsistent shadows and reflections depending on the extent of the inconsistency offers support to the theory that the visual system adopts a coarse scale analysis of fine visual details (Khuu et al., 2016; Lovell et al., 2009; Mamassian, 2004). The coarse scale hypothesis proposes that shadow regions are rapidly identified but not discounted—instead they are represented coarsely. As described above and shown in Figure 8.1, at the second level of visual processing low-level representations are coded coarsely and can be accessed for quick and approximate decisions. This coarse scale analysis means that relatively large shadow discrepancies are fairly easily noticed, but more subtle ones are missed. One possibility is that to detect the more subtle inconsistencies a more effortful higher-level mechanism is required to evaluate the shadow information in the scene (Lovell et al., 2009). Therefore it might be possible that the visual details are encoded by the perceptual system even in the absence of attention. Relatedly, studies in the area of inattentional blindness have revealed that background objects can influence people's judgements even when people claim to be unaware of the objects (C. M. Moore & Egeth, 1997). Yet for visual details to reach an explicit level of internal representation, effortful attention is required.

As well as supporting a coarse scale analysis of shadow information, our results also fit with the notion of a perceptual threshold for detecting lighting inconsistencies (Lopez-Moreno et al., 2010; Tan et al., 2015). It seems possible then, that there is a discernible point at which the inconsistent shadows are different enough from the consistent position for the inconsistency to become noticeable and the coarse scale mechanism to detect it. But for shadow inconsistencies that do not pass this perceptual threshold, a more effortful strategy is required for people to detect the inconsistency. Indeed, our results from the three reflection experiments in Chapter 5 revealed a similar possibility—that people might have a perceptual threshold for noticing inconsistencies in the reflections in a scene. People's ability to identify inconsistent reflections appeared to be dependent upon a number of factors, such as the extent of the inconsistency. In particular, in Experiment 2, when the inconsistent reflections were moved further from the consistent position than in Experiments 1 or 3, people detected the inconsistencies 75% of the time (cf. 58% in Experiment 1 and 55% in Experiment 3). Taken together, our subjects' ability to detect relatively large discrepancies in shadows (Chapter 4) and reflections (Chapter 5) does not have particularly positive consequences for using these visual details to identify image forgeries in the real world. Essentially, the results suggest that when image manipulations create relatively large inconsistencies in the shadows or reflections of the scene, people might be able to detect the image as a fake. Yet, when the inconsistencies are relatively subtle, then people are more likely to be fooled by the fake image.

The findings in Part One of this thesis add to our theoretical understanding of the perception of visual details in scenes. Notably, our results broadly support the notion that people's internal representation of the visual world is incomplete (e.g., Dennett, 1991; O'Regan, 1992; Simons & Levin, 1997, 1998; Stroud, 1955). That said, our findings are difficult to reconcile with the idea that visual details are not perceived at all; instead it is possible that these details might be encoded, even if only coarsely.

The influence of taking photos on memory

The research presented in Chapters 6 and 7 has important implications for our understanding of the workings of human memory. Limits in the visual system mean that people have adapted to encode only the gist of a scene and rely on the world as an external memory source that holds more detailed visual information that can be accessed as and when it is required (O'Regan, 1992; O'Regan & Noë, 2001). In general, the idea behind this theory relates to our research in Chapter 7. In a recent study, Henkel (2014) showed that taking photos of objects on a museum tour influenced people's ability to remember those objects later. Specifically, people recognised fewer of the objects they had photographed—and fewer details about those objects—relative to objects they merely viewed. One explanation for these findings is that people expect the camera to remember for them and therefore pay less attention to content that is captured in photos. Put another way, people might offload their memories onto the camera which acts as an external storage system. The idea that people offload memories onto digital devices is theoretically interesting and could potentially further knowledge about the workings of human memory. For this reason, in Chapter 7 we further explored the impact of taking photos on memory.

Somewhat surprisingly, across five experiments we found little evidence to suggest that people offload their memories onto digital cameras. In fact, when we conducted a mini meta-analysis (see Cumming, 2012, 2013) including the data from our five experiments and Henkel's (2014) original experiments, the estimated size of the photo-taking impairment effect was a trivial 2%—but taking photos could plausibly have no effect at all on memory. Although our results suggest that people might not use digital devices to extend their cognitive capacity, other research in the area of extended-cognition is not concordant with our finding. Specifically, a number of studies have found evidence to suggest that people frequently offload the task of remembering onto external devices, such as smartphones and search engines, instead of committing the information to their own memory. We know that offloading memories can impair people's ability to remember that information later on (Sparrow et al., 2011; Storm & Stone, 2015). At first glance the results of our studies seem difficult to reconcile with the extended-cognition literature in general. Indeed, it seems reasonable to consider that taking photos offers an opportunity to offload memories onto the camera, so why didn't we find any strong evidence to suggest that taking photos impaired memory? Although we are unable to fully answer this question, we can outline two possible explanations.

Before considering the two explanations, however, it is important to note a key difference between our research and other extended-cognition experiments: the nature of the to-be-remembered information. The original photo-taking impairment research (Henkel, 2014) and our five follow-up experiments are the first to look at people's tendency to offload content captured in photos. To our knowledge, all of the other extended-cognition experiments have used generic trivia statements or word lists (e.g., Sparrow et al., 2011; Storm & Stone, 2015). It is possible that the medium (words or photos) or the memory (word lists / trivia statements or autobiographical memories) influence whether or not information will be offloaded.

First, we can draw on research in the area of visual perception to offer one possible explanation for our findings. Because cognitive resources are limited, the perceptual system typically encodes the gist meaning from a scene quickly and effortlessly, but to encode the details requires effortful attention. Put simply, with only a glance at an image, people can glean its basic meaning (Potter, 1975; Thorpe et al., 1996). Therefore, if people find it easier to process visual information than text-based information perhaps in turn this also makes it easier to later recognise visual as opposed to text-based information. If so, it follows that there might be a stronger cognitive offloading effect, and consequently memory impairment, for words than for photos.

Second, our findings might be explained by the difference in the type of memory rather than the medium of the information. Considering that people take photos of their own experiences, it is likely that these photos often contain meaningful autobiographical information. And because autobiographical memories are important—they shape our identity, guide our decisions, and help us to form social relationships (Bluck et al., 2005)—it is possible that people are reluctant to trust these memories to external devices. Yet generic word lists and trivia statements that are not personally relevant provide people with little meaningful information and accordingly people might be more likely to trust that information to an external device. Therefore, prioritising the encoding of personally meaningful experiences might help to explain why research findings suggest less of a cognitive offloading effect for autobiographical memories than word lists and trivia statements. In future research, it would be interesting to test the idea that information type influences people's likelihood to offload, and thus impair memory for, that information. One approach might be to compare the extent to which memory is impaired for autobiographical information in photos versus text.

To summarise, the findings in Part Two of this thesis offer insight into how people perceive and remember their experiences in the world. In general, limits in human cognition mean that people offload information onto external digital devices, but our findings highlight that this might not always be the case. It remains possible that there is something special about the way people process information in photos or autobiographical memories that makes this information less likely to be offloaded—and as a result less likely to impair memory for that information later on. As such, our findings raise new questions about the potential differences in the way that information is offloaded onto external devices. Specifically, it seems possible that the medium or the type of memory that the information relates to influences the likelihood that it is offloaded.

Practical implications

Taken together, the studies presented in this thesis provide the first empirical tests of people's ability to detect image manipulations in real-world scenes and also explore ways that might help people to improve this ability. Accordingly, the results of this research have practical implications for a number of areas where image manipulation could pose problems.

Image manipulation and legal processes

The findings from Part One of this thesis have practical implications for people working within criminal justice settings. It has become commonplace for digital photos to be used as evidence in the criminal justice system, however the rules and guidelines for the admission of photographic evidence in legal cases has not been adequately updated to reflect the unique challenges of the digital age (Johnson, 2012). Instead, digital photos are admissible as evidence based on largely the same grounds as analogue photos, which means that photographic evidence is often admitted based on a witness testimony that the image is authentic and an accurate representation of events (Facciola & Barrett, 2016; Federal Rules of Evidence, 1975; Galves, 2000; Parry, 2009). The findings in Chapter 3 highlight potential pitfalls of such a strategy. Namely, people find it incredibly difficult to determine when an image of a real-world scene is the original, unaltered photo, or when it has been manipulated. Furthermore, even when people are able to detect that an image has been altered, our findings suggest they often fail to locate precisely what has changed.

So what can be done to improve the current situation where digitally altered photos might too easily be admitted into evidence in legal cases? One possibility is to conduct research to better understand why people find it difficult to identify when images have been manipulated and to then use this information to help people to improve at this task. In Chapters 4 and 5, we looked at the extent that people could identify inconsistencies in shadows and reflections. In particular, we explored people's ability to identify inconsistencies when there was enough information in the scene to use two types of geometric analysis based on shadow (Chapter 4) and reflection (Chapter 5) information. These types of analyses form the basis of some of the digital image forensic computer programs that can help to verify the authenticity of images (e.g., Kee et al., 2013; O'Brien & Farid, 2012). Unfortunately, however, our results suggest that people are reasonably insensitive to inconsistencies in shadows and reflections which indicates that they might struggle to use these geometric-based strategies to help them to judge the authenticity of images. Yet it remains possible that, through training, people could learn to use these techniques. In addition, future research could examine whether raising awareness of the extent and possibilities of image manipulation, and whether various types of training, can improve people's ability to distinguish fake images from real ones.

Consequently, then, our findings suggest that relying on people to testify about the authenticity of images in legal settings is potentially problematic. Therefore it is reasonable to suggest that the current rules and guidelines that govern the admissibility of photographic evidence in legal cases should be updated to better account for the ease of digitally altering photos and people's limited ability to detect such alterations. Given that our research is the first to show that people have a limited ability to identify manipulations in real-world scenes, we unfortunately do not yet know enough to make strong recommendations about how to amend the rules and guidelines. At this stage our findings suggest that something needs to change, but to give sound advice on how best to update the policies more research is required. Ideally, psychological scientists, digital forensic experts, legal scholars, and policy makers might work together to conduct further research and develop research-led policies.

In the meantime, one sensible strategy might be to place more reliance on digital image forensics to verify images that are used in legal cases—current guidelines do not request this, thus highlighting a potential area for change. Following the controversy in the 2015 World Press Photo competition, the organisers decided to introduce new safeguards for detecting manipulated photos, including a computerised photo-verification test (World Press Photo, n.d.). A similar safeguarding strategy in the legal arena might be sensible. As it stands, digital image forensic experts are typically approached independently by the defence or prosecution to verify or discredit photo evidence. The lack of an official procedure for using digital forensic tools to verify the authenticity of images means that those who are not aware of the problems associated with digital photography, or of the field of digital image forensics, do not access such tools. Furthermore, independently calling on digital image experts can be expensive. Thus, the most powerful and promising digital forensic tools are often reserved for a small number of cases.

Yet it is important to note that digital image forensics are not infallible. Many researchers have warned that forgers might work out ways to get around the digital image forensic software and beat the system (Böhme & Kirchner, 2012; Gloe, Kirchner, Winkler, & Böhme, 2007). Furthermore, digital image forensics are not at the stage where users can simply upload an image and at the click of a button get back an unequivocal answer as to whether it is authentic or not (Wen, 2017). Although developing such technology is the aim of one of the research projects being conducted by the Defense Advanced Research Projects Agency (DARPA) in the US, in the current situation using the suite of forensic tools to analyse images requires expertise and time. Therefore a promising research investment is to continue to explore ways to improve people's ability to detect when photos have been altered. As shown in Chapters 4 and 5 of this thesis, we can borrow techniques from the field of digital image forensics to discover ways to better equip people to authenticate images on their own (Horaczek, 2017). Of course, by teaching people ways to better detect manipulations we might also inadvertently help forgers to create even more convincing fake images—a paradox that is difficult to avoid. On a positive note, even though serious image forgers might make the effort to correct all potential tell-tale signs that their edits could create in the image, many laypeople are unlikely to make such a concerted effort. Thus, providing tips for detecting fake images should at least make it harder for many people to fool others with manipulated images.

Image manipulation and well-being

Another practical implication of the studies described in Chapter 3 concerns the media's use of digital-editing software to alter a person's appearance. It has become standard for advertisers and magazine editors to present retouched images of celebrities and models that align with societal views of ideal beauty (e.g., Grabe et al., 2008; Groesz et al., 2002; Kee & Farid, 2011; Sheehan, 2014). The portrayal of such difficult-to-achieve, if not impossible, standards for beauty and thinness is worrying because research shows that continual exposure to these beauty "ideals" can lead to psychological problems as well as put people at risk of engaging in dangerous eating and exercise behaviour (e.g., Fallon, 1990; Heinberg, 1996; Morry & Staska, 2001; Owen & Spencer, 2013; Thompson & Stice, 2001). Chapter 3 shows that one reason manipulated images can have such a negative effect is because people frequently fail to realise that these images have been manipulated—indeed, only 40% of the airbrushing manipulations in Experiment 2 were detected (10% below chance performance). If people think that these images are the truth and that people really do look the way they are portrayed in the media then it might seem more reasonable for them to aspire to these unrealistic beauty standards.

The negative impact of ubiquitous idealised and unrealistic representations of physical beauty is being recognised and several countries have implemented formal legislation in an attempt to challenge the permeation of these unrealistic beauty standards (Gladstone, 2016; Wallwork, 2015). Unfortunately, however, the strategies being implemented are not based on a solid body of scientific evidence about what works best. One strategy, for instance, is to add disclaimer labels to images that indicate to viewers that the photos have been retouched. Yet research shows that such disclaimer labels are not a sufficient solution; labels not only fail to offer protective effects against exposure to unrealistic standards of physical beauty (e.g., Tiggemann et al., 2013) but can even have a harmful impact by increasing accessibility to negative thoughts (Selimbegović & Chatard, 2015). Essentially, adding disclaimer labels to retouched images might only prove useful if people are able to work out specifically what has been airbrushed in the image. Recall, our findings suggest that people are rarely able to do so. In Experiment 2, just 13% of people accurately detected the airbrushing manipulations and also then went on to locate the change in the image. To put this result into perspective, consider that more than three times as many people (44%) failed to detect or locate the airbrushing manipulations. And twice as many people (27%) managed to detect that something in the airbrushed image had been altered but failed to correctly select the specific area that had changed. Thus, even when people thought that something had been altered, they often could not tell precisely what had changed. Our results add to the literature suggesting that current "solutions" for the negative impact of idealised and unrealistic representations of physical beauty might not suffice.

An interesting finding is that people's perception of how much an image has been retouched correlates with a quantitative measure of the amount of physical change that has been made to the image (Kee & Farid, 2011). In Chapter 3, we considered how the amount of change made to an image influenced people's ability to detect and locate that change. In particular, Experiment 2 revealed a positive correlation between the amount of change and people's ability to both detect and locate manipulations. On average, however, the airbrushing manipulations created a smaller amount of change to the manipulated images than the other manipulation types—that is, the airbrushing manipulations in our research were relatively subtle. Future research might explore a wider range of airbrushing manipulations—from subtle to extreme—to determine whether the relationship between the extent of the change and people's ability to detect and locate that change persists. Indeed, using an objective measure to inform people how much an image has been retouched could be a more helpful approach than simply adding a disclaimer label. Perhaps providing more information about the extent of retouching in images could serve to better protect people from the negative consequences of exposure to unrealistic representations of physical beauty. In addition, development of such an objective measure could serve as a useful tool for the industry as a whole; organisations and professional photo editors could use this as a guide to minimise or prevent extreme retouching to images (Kee & Farid, 2011).

Future research

The research presented in Part One of this thesis represents the first empirical test of people's ability to identify when images of real-world scenes have been manipulated and explores ways to help people to better detect photo forgeries. The findings have important theoretical and practical implications but given the newness of this research area there are numerous outstanding questions for future research. In fact, as digital technology continues to advance at an incredible speed, the frequency and sophistication of image forgeries is growing. Thus it is important that the relatively neglected area of exploring people's ability to detect photo forgeries continues to receive research interest.

In future research it would be interesting to examine whether raising awareness of image manipulation and whether various types of training can improve people's ability to distinguish fake images from real ones. At first glance, it might seem obvious—better awareness and appreciation of image manipulation will help people to detect when images have been altered. It is possible, however, that people will not apply awareness or knowledge gained from a training scenario to images that they see in the real world. In addition, there is a potential cost associated with raising awareness or training people to look for manipulations: that they lose trust in authentic images. Therefore, any research that looks at the effectiveness of awareness and training should not only consider people's ability to detect when a photo has been altered, but also their ability to correctly identify when a photo has *not* been altered. Signal detection theory offers the necessary measures to do so; to consider any awareness or training strategy truly beneficial then it is important to see an improvement in d' for the experimental group compared with the control group—as opposed to only considering the effect of that strategy on the overall percentage of manipulated photos that are correctly detected. The results in Chapter 3 demonstrate how easy it might be to sway people's tendency to simply say images are manipulated. In Experiment 1 we asked subjects to attempt to locate the manipulation only in photos that they had previously identified as manipulated. We made a small change to the method in Experiment 2 so that subjects were asked to make a guess about the manipulated region regardless of their answer in the detection task. In Experiment 1 people showed a bias to accept the images we showed them as real but we did not find this bias in Experiment 2. Put simply then, a small change to the method between Experiments 1 and 2 appeared to influence people's bias to accept the images we showed them as real.

Another avenue for future research would be to see whether it is possible to train people to more effectively use the shadow and reflection information in scenes. Encouragingly, researchers have shown that performance on many types of perceptual tasks can improve as a result of training (e.g., Ellison & Walsh, 1998; Porter et al., 2010; Sireteanu & Rettenbach, 1995, 2000). Furthermore, it would be interesting to compare how effective a range of different training methods are in improving people's ability to identify consistent and inconsistent shadows and reflections. One method might involve training people to make use of the objective shadow- and reflection-based geometric analysis techniques. Another method might be to use perceptual-based learning—for example giving people a number of practice trials before completing the shadow or reflection task. As well as testing the influence of such training methods on immediate performance on the tasks, if training does improve people's ability to identify consistent and inconsistent scenes, it is important to check whether the improvement persists over time.

The research presented in Part Two of this thesis extends the relatively limited body of research looking at the effect of photography on memory. The results from the five experiments in Chapter 7 not only suggest that the influence of taking photos on memory might not be particularly robust, but also raise new questions about people's willingness to offload cognitive tasks, like remembering, onto external devices. Future research could examine whether people might be more reluctant to offload autobiographical memories than word lists and trivia statements, or alternatively whether the medium of the information influences the likelihood that information will be offloaded.

In addition, to further explore the effect of taking photos on memory it would be interesting to run a series of experiments that consider how people take real-world photos. Based on Henkel's (2014) original procedure, in our five replication experiments we instructed subjects to photograph particular objects on the museum tour. Instructing subjects about what to photograph marks an important difference to how photos are taken in the real world—people typically choose what to photograph and often they photograph things that they are interested in. Furthermore, in the real world, people are not usually given a time limit to capture a photo, and they often have the opportunity to review the photos they take. Therefore, it would be interesting to conduct further research that looks at the impact of taking photos on people's memory for the photographed content under conditions that are more similar to people's experiences in the real world (Barasch, Diehl, Silverman, & Zauberman, 2017). Aligning the experimental conditions more closely with real-life experiences might actually mean that taking photos enhances, rather than impairs, memory.

Concluding remarks

The aims in Part One of this thesis were to gain an understanding of people's ability to discriminate between authentic and manipulated images of real-world scenes and to begin to look for ways that might improve this ability. The aim in Part Two was to investigate how taking photos affects people's memory of the photographed content. The findings in Part One suggest that people have an extremely limited ability to identify manipulations in images of real-world scenes. One possible reason for people's limited ability to detect manipulations is that when people process images, they typically encode the gist meaning but not necessarily the details. Accordingly, people might frequently neglect information that could help them to accurately determine whether images are real or fake. It is clear that image manipulation is not going away; as digital technology improves forgeries are only going to become more visually compelling and thus more difficult to detect. The challenge now is to try to find ways to prevent people being fooled by manipulated photos. The findings in Part Two suggest that taking photos might have only a small effect on people's memories. With projections indicating that the number of photos captured each year will continue to grow exponentially (Heyman, 2015; Lee 2016), perhaps the finding that taking photos plausibly has little effect on memory is good news. That said, our research raises new questions for the area of extended-cognition in general. Further exploring the possible nuances concerning the effects of taking photos on memory will prove useful for furthering theoretical understanding of the workings of human memory.

Chapter 9 :

References

Adams, W. J. (2007). A common light-prior for visual search, shape, and reflectance judgments. *Journal of Vision, 7,* 1–7. doi:10.1167/7.11.11

Allen, B. P. (1999). Shadows as sources of cues for distance of shadow-casting objects. *Perceptual and Motor Skills, 89*, 571–584. doi:10.2466/pms.1999.89.2.571

Altimari, D., & Owens, D. (2017, March 9). Convicted Killer, Suspect In Slayings Of 4 Other Hartford Women, Could Soon Walk Free. *Hartford Courant*. Retrieved from http://www.courant.com/news/connecticut/hc-hartford-swinton-murder-exonerate-20170309-story.html

Amsler, M. (1953). Earliest symptoms of diseases of the macula. *British Journal of Ophthalmology, 37,* 521–537. PMCID:PMC 1324189

Ata, R. N., Thompson, J. K., & Small, B. J. (2013). Effects of exposure to thin-ideal media images on body dissatisfaction: Testing the inclusion of a disclaimer versus warning label. *Body Image, 10*, 472–480. doi:10.1016/j.bodyim.2013.04.004

Autodesk (2016). Depth map and raytraced shadows [Online Maya user guide]. Retrieved from https://knowledge.autodesk.com/support/maya/learn-explore/caas/CloudHelp/cloudhelp/2016/ENU/Maya/files/GUID-BC9C9AFE-D90E-4B54-8A2E-22ED1C89A06C-htm.html

Bakhshi, S., Shamma, D. A., Kennedy, L., & Gilbert, E. (2015). Why We Filter Our Photos and How It Impacts Engagement. *Proceedings of the International AAAI Conference on Web and Social Media 9*, 12–21. Retrieved from http://icwsm.org/2015/

Barasch, A., Diehl, K., Silverman, J., & Zauberman, G. (2017). Photographic Memory: The Effects of Volitional Photo Taking on Memory for Visual and Auditory Aspects of an Experience. *Psychological Science, 28,* 1056–1066. doi:10.1177/0956797617694868

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W.A. Rosenblith (Ed.), *Sensory Communication* (pp. 217–34). Cambridge, MA: MIT Press.

Baron, C. (2008). *Adobe Photoshop Forensics: Sleuths, Truths, and Fauxtography*. Boston, MA: Thomson Course Technology.

Bartlett, F. C. (1932). *Remembering: a study in experimental and social psychology*. Cambridge, England: Cambridge University Press.

Baxandall, M. (1995). *Shadows and Enlightenment*. New Haven, CT: Yale University Press.

Bertamini, M., Spooner, A., & Hecht, H. (2003). Naïve optics: Predicting and perceiving reflections in mirrors. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 982–1002. doi:10.1037/0096-1523.29.5.982

Bex, P. J. (2010). (In) sensitivity to spatial distortion in natural scenes. *Journal of Vision*, *10*, 1–15. doi:10.1167/10.2.23.

Bex, P. J., Solomon, S. G., & Dakin, S. C. (2009). Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure. *Journal of Vision, 9*, 1–19. doi:10.1167/9.10.1

Bianchi, I., Bertamini, M., & Savardi, U. (2015). Differences between predictions of how a reflection behaves based on the behaviour of an object, and how an object behaves based on the behaviour of its reflection. *Acta Psychologica, 161*, 54–63. doi:10.1016/j.actpsy.2015.08.008

Bianchi, I., & Savardi, U. (2012). What fits into a mirror: Naïve beliefs about the field of view. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 1144–1158. doi:10.1037/a0027035

Bissell, K. L. (2006). Skinny like you: Visual literacy, digital manipulation, and young women's drive to be thin. *Studies in Media & Information Literacy Education, 6*, 1–14. doi:10.3138/sim.6.1.002

BlitzMax IDE (Version 1.26) [Programming language]. Blitz Research Ltd. Retrieved from http://www.bmx-ng.com/main/

Bluck, S., Alea, N., Habermas, T., & Rubin, D. C. (2005). A tale of three functions: The self-reported uses of autobiographical memory. *Social Cognition*, *23*, 91–117. doi:10.1521/soco.23.1.91.59198

Böhme, R., & Kirchner, M. (2013). Counter-forensics: Attacking image forensics. In H. T. Sencar & N. Memon (Eds.), *Digital Image Forensics* (pp. 327–366). New York, NY: Springer.

Bonfiglioli, C., Pavani, F., & Castiello, U. (2004). Differential effects of cast shadows on perception and action. *Perception, 33*, 1291–1304. doi:10.1068/p5325

Boone, J. (2016, June 29). Indian couple accused of faking photo of summit at Mount Everest. *The Guardian*. Retrieved from https://www.theguardian.com/world/2016/jun/29/mount-everest-summit-india-couple-morphed-photos

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*, 387–398. doi:10.1016/j.actpsy.2008.09.005

Bouwens, M. D., & Meurs, J. C. (2003). Sine Amsler Charts: A new method for the follow-up of metamorphopsia in patients undergoing macular pucker surgery. *Graefes Archive for Clinical and Experimental Ophthalmology, 241*, 89–93. doi:10.1007/s00417-002-0613-5

Braje, W. L., Legge, G. E., & Kersten, D. (2000). Invariant recognition of natural objects in the presence of shadows. *Perception, 29*, 383–398. doi:10.1068/p3051

Braun, K. A., Ellis, R., & Loftus, E. F. (2002). Make my memory: How advertising can change our memories of the past. *Psychology & Marketing, 19*, 1–23. doi:10.1002/mar.1000

Broadbent, D. E. (1958). *Perception and communication*. Oxford, England: Pergamon Press. doi:10.1037/10037-000

Butterly, A. (2015, November 16). Sikh man Veerender Jubbal photoshopped to look like a Paris attacker. *BBC Newsbeat*. Retrieved from http://www.bbc.co.uk/newsbeat/article/34830661/sikh-man-veerender-jubbal-photoshopped-to-look-like-a-paris-attacker

Camera & Imaging Products Association (2017). *CIPA Total Shipments of Digital Still Cameras / Film Cameras*. Retrieved from http://www.cipa.jp/stats/report_e.html

Casati, R. (2004). The shadow knows: a primer on the informational structure of cast shadows. *Perception, 33*, 1385–1396. doi.org/10.1068/p5281

Castiello, U. (2001). Implicit processing of shadows. *Vision Research, 41*, 2305–2309. doi.org/10.1016/S0042-6989(01)00141-9

Cavanagh, P. (2005). The artist as neuroscientist. *Nature*, *434*, 301-307. doi:10.1038/434301a

Cavanagh, P., & Leclerc, Y. G. (1989). Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 3–27. doi.org/10.1037/0096-1523.15.1.3

Chalfen, R. (1987). *Snapshot versions of life*. Bowling Green, OH: Bowling Green University Popular Press.

Chalfen, R. (1998). Family photograph appreciation: Dynamics of medium, interpretation and memory. *Communication & Cognition: Monographies, 31*, 161–178. Retrieved from https://philpapers.org

Cherrington, R. (2016, June 29). Samsung's 'Beauty' Filter Is Reportedly Automatically Airbrushing Your Selfies. *Huffington Post UK*. Retrieved from http://www.huffingtonpost.co.uk/entry/samsung-beauty-face-camera_uk_57737d18e4b081f48abaa033

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology, 17*, 177–184. doi:10.1016/j.conb.2007.03.005

Clark, A. (2016, September 4). Fashion industry told to end its quest for 'unattainable thinness'. *The Observer*. Retrieved from https://www.theguardian.com/fashion/2016/sep/04/minimum-bmi-checks-uk-models-womens-equality-party

Connecticut v. Swinton, 268 Conn. 781 (2004).

Craik, K. (1943). *The nature of exploration*. Cambridge, England: Cambridge University Press.

Croucher, C. J., Bertamini, M., & Hecht, H. (2002). Naive optics: understanding the geometry of mirror reflections. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 546–562. doi:10.1037/0096-1523.28.3.546

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.

Cumming, G. (2013). *The new statistics: Estimation for better research*. Retrieved from www.thenewstatistics.com

Dee, H. M., & Santos, P. E. (2011). The perception and content of cast shadows: An interdisciplinary review. *Spatial Cognition & Computation, 11*, 226–253. doi.org/10.1080/13875868.2011.565396

Deffenbacher, K. A., Bornstein, B. H., & Penrod, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior, 30*, 287–307. doi:10.1007/s10979-006-9008-1

Dennett, D. C. (1991). *Consciousness Explained*. Boston, MA: Little, Brown and co.

Djudjic, D. (2017, March 2). Camera sales report for 2016: Lowest sales ever on DSLRs and mirrorless. *DIY Photography*. Retrieved from http://www.diyphotography.net/camera-sales-report-2016-lowest-sales-ever-dslrs-mirrorless/

Drews, F. A., Yazdani, H., Godfrey, C. N., Cooper, J. M., & Strayer, D. L. (2009). Text messaging during simulated driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 51*, 762–770. doi:10.1177/0018720809353319

Ehinger, K. A., Allen, K., & Wolfe, J. M. (2016). Change blindness for cast shadows in natural scenes: Even informative shadow changes are missed. *Attention, Perception, & Psychophysics, 78*, 978-987. doi:10.3758/s13414-015-1054-7

Elder, J. H., Trithart, S., Pintilie, G., & MacLean, D. (2004). Rapid processing of cast and attached shadows. *Perception, 33*, 1319–1338. doi:10.1068/p5323

Ellison, A., & Walsh, V. (1998). Perceptual learning in visual search: some evidence of specificities. *Vision Research, 38*, 333–345. doi:10.1016/S0042-6989(97)00195-8

Engelkamp, J., Seiler, K. H., & Zimmer, H. D. (2004). Memory for actions: Item and relational information in categorized lists. *Psychological Research, 69*, 1–10. doi: 10.1007/s00426-003-0160-7

Engelkamp, J. & Zimmer, H. D. (1997). Sensory factors in memory for subject-performed tasks. *Acta Psychologica, 96*, 43–60. doi:10.1016/S0001-6918(97)00005-X

Enns, J. T., & Rensink, R. A. (1990). Influence of scene-based properties on visual search. *Science, 247*, 721–723. doi:10.1126/science.2300824

Estrin, J. (2013, October). The visual village. *National Geographic*. Retrieved from http://ngm.nationalgeographic.com/2013/10/digital-village/estrin-text

Facciola, J. M., & Barrett, L. (2016, November). Law of the Foal: Careful Steps Towards Digital Competence in Proposed Rules 902(13) and 902(14). *Georgetown Law Technology Review, 1*, 6–15. Retrieved from https://perma.cc/Z9NN-ZY95

Fallon, A. E. (1990). Culture in the mirror: sociocultural determinants of body image. In T. F. Cash & T. Pruzinsky (Eds.), *Body images: Development, deviance, and change* (pp. 80–109). New York, NY: Guilford.

Farid, H. (2006). Digital doctoring: how to tell the real from the fake. *Significance, 3*, 162-166. doi:10.1111/j.1740-9713.2006.00197.x

Farid, H. (2009a). Image forgery detection. *Signal Processing Magazine, IEEE, 26*, 16–25. doi:10.1109/MSP.2008.931079

Farid, H. (2009b). Seeing is not believing. *IEEE Spectrum, 46*, 44–51. doi:10.1109/MSPEC.2009.5186556

Farid, H. (2016). *Photo Forensics*. Cambridge, MA: The MIT Press.

Farid, H., & Bravo, M. J. (2010). Image forensic analyses that elude the human visual system. *Proceedings of SPIE, 7541*, 1–10. doi:10.1117/12.837788

Federal Rules of Evidence, Pub. L No. 93-595, §1, 88 Stat. 1945 (1975). Retrieved from http://www.uscourts.gov/rules-policies/current-rules-practice-procedure

Felber, P. (2002). *Charge-coupled devices: a literature study as a project for ECE 575*. Retrieved from http://www.ece.iit.edu/~pfelber/ccd/project.pdf

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*, 850–855. doi: 10.1111/j.1467-9280.2007.01990.x

Fenn, E., Newman, E. J., Pezdek, K., & Garry, M. (2013). The effect of nonprobative photographs on truthiness persists over time. *Acta Psychologica, 144*, 207–211. doi:10.1016/j.actpsy.2013.06.004

Fister, S. M., & Smith, G. T. (2004). Media effects on expectancies: exposure to realistic female images as a protective factor. *Psychology of Addictive Behaviors*, *18*, 394–397. doi:10.1037/0893-164X.18.4.394

France bans extremely thin models. (2017, May 6). *BBC News*. Retrieved from http://www.bbc.co.uk/news/world-europe-39821036

Frenda, S. J., Knowles, E. D., Saletan, W., & Loftus, E. F. (2013). False memories of fabricated political events. *Journal of Experimental Social Psychology, 49*, 280–286. doi:10.1016/j.jesp.2012.10.013

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological sciences, 360*, 815-836. doi:10.1098/rstb.2005.1622

Galassi, P. (1981). *Before Photography*. New York, NY: The Museum of Modern Art.

Galves, F. (2000). Where the not-so-wild-things are: Computers in the courtroom, the Federal Rules of Evidence, and the need for institutional reform and more judicial acceptance. *Harvard Journal of Law & Technology, 13*, 161-301. Retrieved from http://jolt.law.harvard.edu/articles/v13.php

Garcia, F. (2016, October 29). Pro-Trump Twitter bot 'Neil Turner' posts fake image of immigration officer arresting Latino voter. *The Independent*. Retrieved from http://www.independent.co.uk/news/world/americas/twitter-bot-immigration-arrest-latino-voter-neil-turner-hoax-us-election-2016-a7386061.html

Gardner-Medwin, A. R., & Barlow, H. B. (2001). The limits of counting accuracy in distributed neural representations. *Neural Computation*, *13*, 477–504. doi:10.1162/089976601300014420

Garry, M., Strange, D., Bernstein, D. M., & Kinzett, T. (2007). Photographs can distort memory for the news. *Applied Cognitive Psychology*, *21*, 995–1004. doi: 10.1002/acp.1362

Gatton, M. (2009). First light: inside the palaeolithic camera obscura. In A. Kaniari & M. Wallace (Eds.), *Acts of Seeing: Artists, Scientists and the History of the Visual—a volume dedicated to Martin Kemp* (pp. 146–154). London, England: Artakt and Zidane Press.

Gernsheim, H. (1977). The 150th anniversary of photography. *History of Photography*, *1*, 3-8. doi:10.1080/03087298.1977.10442876

Gernsheim, H. (1986). *A Concise History of Photography* (3rd ed.). Mineola, NY: Dover Publications.

Ghosh, P. (2017, February 1). £1m Queen Elizabeth Prize: Digital camera tech lauded. *BBC News*. Retrieved from http://www.bbc.co.uk/news/science-environment-38815948

Gillin, J. (2017, March 9). Image of Obama in handcuffs for wiretapping Trump is fake photo. *PunditFact*. Retrieved from http://www.politifact.com/punditfact/statements/2017/mar/09/blog-posting/image-obama-handcuffs-wiretapping-trump-fake-photo/

Gladstone, J. (2016). The Skinny on BMI-Based Hiring: An Assessment of the Legality and Effectiveness of Israel's Weight Restriction Law. *Wash. U. Global Stud. L.*

Rev., 15, 495–512. Retrieved from http://openscholarship.wustl.edu/law_globalstudies/

Gloe, T., Kirchner, M., Winkler, A., & Böhme, R. (2007). Can we trust digital image forensics? *Proceedings of the 15th ACM international conference on Multimedia, 15*, 78–86. doi:10.1145/1291233.1291252

GNU Image Manipulation Program (GIMP) (Version 2.8) [Computer software]. Retrieved from https://www.gimp.org/

Grabe, S., Ward, L. M., & Hyde, J. S. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. *Psychological Bulletin, 134*, 460–476. doi:10.1037/0033-2909.134.3.460

Graham, R. (1984). The early history of photography. *History of Photography, 8*, 231–235. doi:10.1080/03087298.1984.10442226

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534–537. doi:10.1038/nature01647

Green, C. S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1465–1478. doi:10.1037/0096-1523.32.6.1465

Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science, 18*, 88–94. doi:10.1111/j.1467-9280.2007.01853.x

Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology, 22*, R197–R206. doi:10.1016/j.cub.2012.02.012

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, England: Wiley.

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

Grinley, M. J. (2002).*Effects of advertising on semantic and episodic memory.* (Unpublished master's thesis). University of Washington, Seattle, Washington.

Groesz, L. M., Levine, M. P., & Murnen, S. K. (2002). The effect of experimental presentation of thin media images on body satisfaction: A meta-analytic review. *International Journal of Eating Disorders, 31*, 1-16. doi:10.1002/eat.10005

Guilshan, C. A. (1992). A picture is worth a thousand lies: Electronic imaging and the future of the admissibility of photographs into evidence. *Rutgers Computer & Tech. Law Journal, 18,* 365–380. Retrieved from https://home.heinonline.org/titles/Law-Journal-Library/Rutgers-Computer-and-Technology-Law-Journal/

Guthrie, C., & Mitchell, B. (2007). The Swinton Six: The impact of State v. Swinton on the authentication of digital images. *Stetson L. Rev., 36*, 661–726. Retrieved from http://heinonline.org/HOL/Index?index=journals/stet&collection=journals

Hadland, A., Campbell, D., & Lambert, P. (2015). *The state of photography: The lives and livelihoods of photojournalists in the digital age*. Retrieved from http://reutersinstitute.politics.ox.ac.uk/publication/state-news-photography-lives-and-livelihoods-photojournalists-digital-age

Hand, M. (2012). *Ubiquitous photography*. Cambridge, England: Polity Press.

Hannon, M. J. (2014). An Increasingly Important Requirement: Authentication of Digital Evidence. *Journal of the Missouri Bar, 70,* 314–323. Retrieved from http://www.mobar.org/journal/novdec2014.htm

Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. *Proceedings of Neural Information Processing Systems (NIPS), 19*, 545-552. Retrieved from http://papers.nips.cc/

Harrison, B. (2002). Photographic visions and narrative inquiry. *Narrative Inquiry, 12*, 87–111. doi:10.1075/ni.12.1.14har

Harrison, K., & Hefner, V. (2014). Virtually perfect: Image retouching and adolescent body image. *Media Psychology, 17*, 134–153. doi:10.1080/15213269.2013.770354

Hartley, R., & Zisserman, A. (2003). *Multiple view geometry in computer vision* (2nd ed.). Cambridge, England: Cambridge University Press

Hecht, E., & Zajac, A. (1974). *Optics*. Reading, MA: Addision-Wesley.

Hecht, H., Bertamini, M., & Gamer, M. (2005). Naive optics: acting on mirror reflections. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1023–1038. doi:10.1037/0096-1523.31.5.1023

Heinberg, L. J. (1996). Theories of body image disturbance: Perceptual, developmental, and sociocultural factors. In J. K. Thompson (Ed.), *Body image, eating disorders,* *and obesity: An integrative guide for assessment and treatment* (pp. 27–47). Washington, DC: American Psychological Association.

Henkel, L. A. (2011). Photograph-induced memory errors: When photographs make people claim they have done things they have not. *Applied Cognitive Psychology, 25*, 78–86. doi:10.1002/acp.1644

Henkel, L. A. (2014). Point-and-shoot memories: The influence of taking photos on memory for a museum tour. *Psychological Science, 25*, 396-402. doi:10.1177/0956797613504438

Hewitt, L. Y., & Roberts, L. D. (2015). Transactive memory systems scale for couples: development and validation. *Frontiers in Psychology, 6*(516), 1-8. doi: 10.3389/fpsyg.2015.00516

Heyman, S. (2015, July 29). Photos, Photos Everywhere. *The New York Times*. Retrieved from https://www.nytimes.com/2015/07/23/arts/international/photos-photos-everywhere.html

Hillen, B. (2017, March 6). 2016 CIPA data shows compact digital camera sales lower than ever. *Digital Photography Review*. Retrieved from https://www.dpreview.com/news/9398648371/2016-cipa-data-shows-compact-digital-camera-sales-lower-than-ever

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 113–136. doi:10.1037/0096-1523.28.1.113

Hooton, J. (2016, September 6). How The Women's Equality Party #NoSizeFitsAll Campaign Demonises Too Thin But Ignores Too Fat. *Huffington Post UK*. Retrieved from http://www.huffingtonpost.co.uk/jacqueline-hooton/how-the-womens-equality-p_b_11865292.html

Horaczek, S. (2017, July 21). Spot faked photos using digital forensic techniques. *Popular Science*. Retrieved from http://www.popsci.com/use-photo-forensics-to-spot-faked-images

Howe, C. Q., & Purves, D. (2005). *Perceiving geometry: Geometrical illusions explained by natural scene statistics.* New York, NY: Springer.

Hulleman, J., & Olivers, C. N. L. (2015). The impending demise of the item in visual search. *Behavioral and Brain Sciences, 17*, 1–76. doi:10.1017/S0140525X15002794

Humphreys, G. W., Watson, D. G., & Jolicœur, P. (2002). Fractionating the preview benefit in search: Dual-task decomposition of visual marking by timing and modality. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 640–660. doi: 10.1037//0096-1523.28.3.640

Hyman, I. E., Jr., Boss, S. M., Wise, B. M., McKenzie, K. E., & Caggiano, J. M. (2010). Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology, 24*, 597–607. doi:10.1002/acp.1638

Hyman, I. E., Jr., Husband, T. H., & Billings, F. J. (1995). False memories of childhood experiences. *Applied Cognitive Psychology, 9*, 181–197. doi:10.1002/acp.2350090302

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506. doi:10.1016/S0042-6989(99)00163-7

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*, 1254–1259. doi:10.1109/34.730558

James, W. (1890). *The principles of psychology*. New York, NY: Henry Holt.

Jefferies, S. (2013, December 13). The death of photography: are camera phones destroying an artform? *The Guardian*. Retrieved from https://www.theguardian.com/artanddesign/2013/dec/13/death-of-photography-camera-phones

Johnson, E. (2012, January 23). Can Your Digital Images Withstand A Court Challenge? [Blog post]. Retrieved from https://articles.forensicfocus.com/author/johnsoneric/

Johnston, E. B. (2001). The repeated reproduction of Bartlett's Remembering. *History of Psychology, 4*, 341-366. doi:10.1037/1093-4510.4.4.341

Kajiya, J.T. (1986). The Rendering Equation. *Proceedings of the Annual Conference on Computer Graphics and Interactive Techniques 13*, 143–150. doi:10.1145/15922.15902

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A., & La Fon, D. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior, 31*, 381–400. doi:0.1007/s10979-006-9073-5

Kee, E., & Farid, H. (2011). A perceptual metric for photo retouching. *Proceedings of the National Academy of Sciences, 108*, 19907–19912. doi:10.1073/pnas.1110747108

Kee, E., O'Brien, J. F., & Farid, H. (2013). Exposing photo manipulation with inconsistent shadows. *ACM Transactions on Graphics (ToG), 32*, 1–12. doi: 10.1145/2487228.2487236

Kelly, J. E., & Nace, D. (1994). Digital imaging and believing photos. *Visual Communication Quarterly, 1*, 4–18. doi:10.1080/15551393.1994.10387481

Kersten, D., Knill, D. C., Mamassian, P., & Bülthoff, I. (1996). Illusory motion from shadows. *Nature, 379*, 31. doi.org/10.1038/379031a0

Kersten, D., Mamassian, P., & Knill, D. C. (1994). *Moving cast shadows and the perception of relative depth* (Technical Report 6). Retrieved from Max Planck Institute for Biological Cybernetics website: http://www.kyb.mpg.de/fileadmin/user_upload/files/publications/pdfs/pdf1461.pdf

Khang, B. G., Koenderink, J. J., & Kappers, A. M. (2006). Perception of illumination direction in images of 3-D convex objects: Influence of surface materials and light fields. *Perception, 35*, 625–645. doi.org/10.1068/p5485

Khuu, S. K., Honson, V. J., & Challinor, K. L. (2016). The influence of cast shadows on the detection of three-dimensional curved contour structure. *Perception, 45*, 425–442. doi:10.1177/0301006615622323

King, H. (2015, April 24). Facetune makes your ugliest selfies look beautiful. *CNET Tech.* Retrieved from http://money.cnn.com/2015/04/24/technology/facetune-mobile-photoshop/index.html

Koenderink, J. J., van Doorn, A. J., & Pont, S. C. (2004). Light direction from shad (ow) ed random Gaussian surfaces. *Perception, 33*, 1405–1420. doi.org/10.1068/p5287

Kurtz, J. L. (2015). Seeing through new eyes: An experimental investigation of the benefits of photography. *Journal of Basic & Applied Sciences, 11*, 354–358. doi:10.6000/1927-5129.2015.11.51

Langer, M. S., & Zucker, S. W. (1997). What is a light source? *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 8*, 172–178. doi:10.1109/CVPR.1997.609316

Lansdale, M., Underwood, G., & Davies, C. (2010). Something Overlooked? How experts in change detection use visual saliency. *Applied Cognitive Psychology*, *24*, 213-225. doi:10.1002/acp.1552

Larsen, J. (2008). Practices and flows of digital photography: An ethnographic framework. *Mobilities, 3*, 141-160. doi: 10.1080/17450100701797398

Larsen, J., & Sandbye, M. (Eds.). (2014). *Digital snaps. The new face of photography*. London, England: I.B. Tauris.

Lawson, R. (2010). People cannot locate the projection of an object on the surface of a mirror. *Cognition, 115*, 336–342. doi:10.1016/j.cognition.2009.12.013

Lawson, R., & Bertamini, M. (2006). Errors in judging information about reflections in mirrors. *Perception, 35*, 1265–1288. doi:10.1068/p5498

Lee, E. (2016, April 14). Fun Facts About the Photography Market. [Blog post]. Retrieved from http://blog.infotrends.com/?p=21310

Lee, P., & Stewart, D. (2016). *Photo sharing: trillions and rising* (TMT Predictions 2016). Retrieved from https://www2.deloitte.com/global/en/pages/technology-media-and-telecommunications/articles/tmt-pred16-telecomm-photo-sharing-trillions-and-rising.html#full-report

Leek, E. C., Davitt, L. I., & Cristino, F. (2015). Implicit encoding of extrinsic object properties in stored representations mediating recognition: Evidence from shadow-specific repetition priming. *Vision Research, 108*, 49–55. doi.org/10.1016/j.visres.2015.01.011

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, *4*, 501–506. doi:10.3758/BF03214339

Levush, R. (2012, April 11). A Matter of Size: Israel's New Legislation against Too-Skinny Models [Blog post]. Retrieved from https://blogs.loc.gov/law/2012/04/a-matter-of-size-israels-new-legislation-against-too-skinny-models/

Levy, J., Pashler, H., & Boer, E. (2006). Central interference in driving is there any stopping the psychological refractory period? *Psychological Science, 17*, 228–235. doi:10.1111/j.1467-9280.2006.01690.x

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73* 13–22. doi:10.2307/2336267

Lindsay, D. S., Hagen, L., Read, J. D., Wade, K. A., & Garry, M. (2004). True photographs and false memories. *Psychological Science, 15*, 149–154. doi:10.1111/j.0956-7976.2004.01503002.x

Liverence, B. M., & Franconeri, S. L. (2015). Resource Limitations in Visual Cognition. In R. Scott & S. Kosslyn (Eds.), *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource* (pp. 1–15). doi:10.1002/9781118900772.etrds0287

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*, 560–572. doi:10.1016/0010-0285(75)90023-7

Loftus, E. F. (2003). Make-believe memories. *American Psychologist, 58*, 867–873. doi:10.1037/0003-066X.58.11.867

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*, 361–366. doi:10.1101/lm.94705

Loftus, E. F. (2013, June). *Elizabeth Loftus: The fiction of memory* [Video file]. Retrieved from http://www.ted.com/talks/elizabeth_loftus_the_fiction_of_memory/

Loftus, E. F. (2017). Eavesdropping on Memory. *Annual Review of Psychology, 68*, 1–18. doi:10.1146/annurev-psych-010416-044138

Loftus, E. F., & Hoffman, H. G. (1989). Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General, 118*, 100–104. doi:10.1037/0096-3445.118.1.100

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19–31. doi:10.1037/0278-7393.4.1.19

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25*, 720–725. doi:10.3928/0048-5713-19951201-07

Lopez-Moreno, J., Sundstedt, V., Sangorrin, F., & Gutierrez, D. (2010). Measuring the perception of light inconsistencies. *Proceedings of the 7th ACM Symposium on Applied Perception in Graphics and Visualization 7*, 25–32. doi:10.1145/1836248.1836252

Lovell, P. G., Gilchrist, I. D., Tolhurst, D. J., & Troscianko, T. (2009). Search for gross illumination discrepancies in images of natural objects. *Journal of Vision, 9*, 1–14. doi:10.1167/9.1.37

Mamassian, P. (2004). Impossible shadows and the shadow correspondence problem. *Perception, 33*, 1279–1290. doi:10.1068/p5280

Mamassian, P. (2008). Ambiguities and conventions in the perception of visual art. *Vision Research, 48*, 2143–2153. doi:10.1016/j.visres.2008.06.010

Mamassian, P., & Goutcher, R. (2001). Prior knowledge on the illumination position. *Cognition, 81*, B1–B9. doi:10.1016/S0010-0277(01)00116-0

Mamassian, P., Knill, D. C., & Kersten, D. (1998). The perception of cast shadows. *Trends in Cognitive Sciences, 2*, 288–295. doi.org/10.1016/S1364-6613(98)01204-2

Marszal, A. (2015, December 4). Narendra Modi 'photoshopped' image of Chennai floods visit goes viral. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/news/worldnews/asia/india/12032546/Narendra-Modi-photoshopped-image-of-Chennai-floods-visit-goes-viral.html

Maya (2016) [Computer software]. San Rafael, CA: Autodesk, Inc.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of Psychological Refractory-Period Phenomena. *Psychological Review, 104*, 749–791. doi:10.1037//0033-295X.104.4.749

Meyer, D. E., Kieras, D. E., Lauber, E., Schumacher, E. H., Glass, J., Zurbriggen, E., ... & Apfelblat, D. (1995). Adaptive executive control: Flexible multiple-task performance without pervasive immutable response-selection bottlenecks. *Acta Psychologica, 90*, 163–190. doi:10.1016/0001-6918(95)00026-Q

Mnookin, J. L. (1998). The image of truth: Photographic evidence and the power of analogy. *Yale Journal of Law & the Humanities, 10*, 1–74. Retrieved from http://digitalcommons.law.yale.edu/yjlh/vol10/iss1/1

Mols, I., Broekhuijsen, M., van den Hoven, E., Markopoulos, P., & Eggen, B. (2015). Do we ruin the moment? Exploring the Design of Novel Capturing Technologies. *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction 27*, 653–661. doi:10.1145/2838739.2838758

Montague, J. (2013). *Basic Perspective Drawing* (6th ed.). Hoboken, NJ: John Wiley and Sons.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology-Human Perception and Performance*, *23*, 339–352. doi:10.1037//0096-1523.23.2.339

Moore, J. L. (2010). Time for an Upgrade: Amending the Federal Rules of Evidence to Address the Challenges of Electronically Stored Information in Civil Litigation. *Jurimetrics, 50*, 147–193. Retrieved from https://web.law.asu.edu/jurimetrics/JurimetricsJournal/ArticlesIssues.aspx

Morgenstern, Y., Geisler, W. S., & Murray, R. F. (2014). Human vision is attuned to the diffuseness of natural light. *Journal of Vision, 14,* 1–18. doi:10.1167/14.9.15

Morgenstern, Y., Murray, R. F., & Harris, L. R. (2011). The human visual system's assumption that light comes from above is weak. *Proceedings of the National Academy of Sciences 108*, 12551–12553. doi:10.1073/pnas.1100794108

Morry, M. M., & Staska, S. L. (2001). Magazine exposure: Internalization, self-objectification, eating attitudes, and body satisfaction in male and female university students. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 33*, 269–279. doi:10.1037/h0087148

Moses, M. (2002, August 17). Readers consume what they see. *Poynter*. Retrieved from http://www.poynter.org/2002/readers-consume-what-they-see/1875/

Muelenz, C., Hecht, H., & Gamer, M. (2010). Testing the egocentric mirror-rotation hypothesis. *Seeing and Perceiving, 23*, 373–383. doi:10.1163/187847510X540000

Mullins, C. (2013). *The world on a plate: the impact of photography on travel imagery and its dissemination in britain, 1839-1888*. (Doctoral dissertation). Retrieved from http://sro.sussex.ac.uk/47201/

Nash, R. A., & Wade, K. A. (2009). Innocent but proven guilty: Eliciting internalized false confessions using doctored video evidence. *Applied Cognitive Psychology, 23*, 624–637. doi:10.1002/acp.1500

Nash, R. A., Wade, K. A., & Lindsay, D. S. (2009). Digitally manipulating memory: Effects of doctored videos and imagination in distorting beliefs and memories. *Memory & Cognition, 37*, 414–424. doi:10.3758/MC.37.4.414

Nelson, W. W., & Loftus, G. R. (1980). The functional visual field during picture viewing. Journ*al of Experimental Psychology: Human Learning and Memory, 6*, 391–399. doi:10.1037/0278-7393.6.4.391

Newhall, B. (1982). *The History of Photography* (5th ed.). London, England: Secker & Warburg.

Newman, E. J., & Feigenson, N. (2013). The truthiness of visual evidence. *The Jury Expert, 25*(5), 9–14. Retrieved from http://www.thejuryexpert.com/2013/11/the-truthiness-of-visual-evidence/

Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J., & Lindsay, D. S. (2012). Nonprobative photographs (or words) inflate truthiness. *Psychonomic Bulletin & Review, 19*, 969–974. doi:10.3758/s13423-012-0292-0

O'Brien, J. F., & Farid, H. (2012). Exposing photo manipulation with inconsistent reflections. *ACM Transactions on Graphics, 31*, 1–11. doi:10.1145/2077341.2077345

Olshausen, B. A., & Field, D. J. (2000). Vision and the coding of natural images. *American Scientist, 88*, 238–245. doi:10.1511/2000.3.238

Oosterhoff, D. (2015, January 8). Fakes, Frauds, and Forgeries: How to Detect Image Manipulation. *Envato Tuts+*. Retrieved from http://photography.tutsplus.com/articles/fakes-frauds-and-forgeries-how-to-detect-imagemanipulation--cms-22230

O'Regan, J. K. (1992). Solving the" real" mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 46*, 461–488. doi:10.1037/h0084327

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*, 939–973. doi:10.1017/S0140525X01000115

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature, 398,* 34. doi:10.1038/17953

O'Shea, J. P., Agrawala, M., & Banks, M. S. (2010). The influence of shape cues on the perception of lighting direction. *Journal of Vision, 10*, 1–21. doi:10.1167/10.12.21

Ostrovsky, Y., Cavanagh, P., & Sinha, P. (2005). Perceiving illumination inconsistencies in scenes. *Perception, 34*, 1301–1314. doi:10.1068/p5418

Owen, R., & Spencer, R. M. C. (2013). Body ideals in women after viewing images of typical and healthy weight models. *Body Image, 10*, 489–494. doi:10.1016/j.bodyim.2013.04.005

Parry, Z. B. (2009). Digital manipulation and photographic evidence: defrauding the courts one thousand words at a time. *Journal of Law, Technology & Policy, 2009*, 175–202. Retrieved from http://illinoisjltp.com/journal/

Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics, 44*, 369–378. doi:10.3758/BF03210419

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244. doi:10.1037/0033-2909.116.2.220

Paul, G. L. (2008). *Foundations of Digital Evidence*. Chicago, IL: American Bar Association.

Pearson, H. (2006). Forensic software traces tweaks to images. *Nature, 439*, 520–521. doi:10.1038/439520b

Peters, J. W. (2010, July 5). On The Economist's Cover, Only a Part of the Picture [Blog post]. Retrieved from https://mediadecoder.blogs.nytimes.com/2010/07/05/on-the-economists-cover-only-a-part-of-the-picture/?_r=0

Peterson, K. F. (2010, January). Enhanced persuasion: Effective use of demonstrative evidence at trial. *Journal of Consumer Attorneys Associates for Southern California Advocate, 37*, 1–5. Retrieved from http://www.theadvocatemagazine.com/Advocate%20Archives.html

Porter, E. G. (2014). Taking Images Seriously. *Colum. L. Rev., 114*, 1687–1782. Retrieved from http://heinonline.org/HOL/Index?index=journals/clr&collection=journals

Porter, G., Tales, A., & Leonards, U. (2010). What makes cast shadows hard to see? *Journal of Vision, 10*, 1–18. doi:10.1167/10.3.13

Porter, S., Yuille, J. C., & Lehman, D. R. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: implications for the recovered memory debate. *Law and Human Behavior, 23*, 517–537. doi:10.1023/A:1022344128649

Potter, M. C. (1975). Meaning in visual search. *Science, 187*, 965–966. doi:10.1126/science.1145183

Pringle, H. L., Irwin, D. E., Kramer, A. F., & Atchley, P. (2001). The role of attentional breadth in perceptual change detection. *Psychonomic Bulletin & Review, 8,* 89–95. doi:10.3758/BF03196143

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*, 79–87. doi:10.1038/4580

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review, 85*, 59–108. doi:10.1037/0033-295X.85.2.59

Rawlinson, K. (2015, November 16). Canadian pictured as Paris terrorist in suspected Gamergate smear. *The Guardian*. Retrieved from https://www.theguardian.com/world/2015/nov/16/canadian-pictured-as-paris-terrorist-in-suspected-gamergate-smear

Rensink, R. A., & Cavanagh, P. (2004). The influence of cast shadows on visual search. *Perception, 33*, 1339–1358. doi:10.1068/p5322

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*, 368–373. doi:10.1111/j.1467-9280.1997.tb00427.x

Robertson, A. R. (1977). The CIE 1976 color-difference formulae. *Color Research and Application, 2*, 7–11. doi:10.1002/j.1520-6378.1977.tb00104.x

Robinson, E. M. (2013). *Introduction to Crime Scene Photography*. Amsterdam, The Netherlands: Academic Press.

Roediger, H. L., III, & Zaromb, F. M. (2010). Memory for actions. In L. Backman & L. Nyberg (Eds.), *Memory, aging and the brain* (pp. 24–52). New York, NY: Psychology Press.

Ronchi, V. (1970). *The nature of light: An historical survey* (V. Barocas, Trans.). London, England: Heinemann.

Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research, 21,* 37–59. doi:10.1300/J079v21n04_02

Sacchi, D. L. M., Agnoli, F., & Loftus, E. F. (2007). Changing history: doctored photographs affect memory for past public events. *Applied Cognitive Psychology, 21*, 1005–1022. doi:10.1002/acp.1394

Safi, M. (2016, August 30). Indian couple banned from climbing after faking ascent of Everest. *The Guardian*. Retrieved from https://www.theguardian.com/world/2016/aug/30/indian-couple-banned-from-climbing-after-faking-ascent-of-everest

Sareen, P., Ehinger, K. A., & Wolfe, J. M. (2015). Through the looking-glass: Objects in the mirror are less real. *Psychonomic Bulletin & Review*, *22*, 980–986. doi:10.3758/s13423-014-0761-8

Saul, H. (2015, October 22). Kate Winslet explains why she has a contract in place to stop Lancome adverts being airbrushed. *The Independent*. Retrieved from http://www.independent.co.uk/news/people/kate-winslet-explains-why-she-has-a-contract-in-place-to-stop-lancome-adverts-being-airbrushed-a6703921.html

Schacter, D. L. (2013). Memory: sins and virtues. *Annals of the New York Academy of Sciences, 1303*, 56–60. doi:10.1111/nyas.12168

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science, 12*, 101–108. doi:10.1111/1467-9280.00318

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5*, 195–200. doi:10.1111/j.1467-9280.1994.tb00500.x

Scott-Brown, K. C., Baker, M. R., & Orbach, H. S. (2000). Comparison blindness. *Visual Cognition, 7*, 253–267. doi:10.1080/135062800394793

Selimbegović, L., & Chatard, A. (2015). Single exposure to disclaimers on airbrushed thin ideal images increases negative thought accessibility. *Body Image, 12*, 1–5. doi:10.1016/j.bodyim.2014.08.012

Sheehan, T. (2014). Retouch yourself: the pleasures and politics of digital cosmetic surgery. In J. Larsen & M. Sandbye (Eds.), *Digital snaps. The new face of photography* (pp. 179–203). London, England: I.B. Tauris.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science, 7*, 301–305. doi:10.1111/j.1467-9280.1996.tb00378.x

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*, 261–267. doi:10.1016/S1364-6613(97)01080-2

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review, 5*, 644–649. doi:10.3758/BF03208840

Sireteanu, R., & Rettenbach, R. (1995). Perceptual learning in visual search: Fast, enduring, but non-specific. *Vision Research, 35*, 2037–2043. doi:10.1016/0042-6989(94)00295-W

Sireteanu, R., & Rettenbach, R. (2000). Perceptual learning in visual search generalizes over tasks, locations, and eyes. *Vision Research, 40*, 2925–2949. doi:10.1016/S0042-6989(00)00145-0

Slater, D. (1991). Consuming Kodak. In: J. Spence & P. Holland (Eds.), *Family Snaps: The Meaning of Domestic Photography* (pp. 49–59). London, England: Virago.

Smith, C. (2013, September 18). Facebook Users Are Uploading 350 Million New Photos Each Day, *Business Insider*. Retrieved from http://www.businessinsider.com/facebook-350-million-photos-each-day-2013-9?IR=T

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, *333*, 776–778. doi:10.1126/science.1207745

Spotorno, S., & Faure, S. (2011). Change detection in complex scenes: hemispheric contribution and the role of perceptual and semantic factors. *Perception*, *40*, 5–22. doi:10.1068/p6524

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*, 137–149. doi:10.3758/BF03207704

Steadman, P. (2001). *Vermeer's Camera*. Oxford, England: Oxford University Press.

Stice, E., Schupak-Neuberg, E., Shaw, H. E., & Stein, R. I. (1994). Relation of media exposure to eating disorder symptomatology: An examination of mediating mechanisms. *Journal of Abnormal Psychology, 103*, 836–840. doi:10.1037/0021-843X.103.4.836

Stice, E., & Shaw, H. E. (1994). Adverse effects of the media portrayed thin-ideal on women and linkages to bulimic symptomatology. *Journal of Social and Clinical Psychology, 13,* 288–308. doi:10.1521/jscp.1994.13.3.288

Stirk, J. A., & Underwood, G. (2007). Low-level visual saliency does not predict change detection in natural scenes. *Journal of Vision, 7,* 1-10. doi:10.1167/7.10.3

Storm, B. C., & Stone, S. M. (2015). Saving-Enhanced Memory The Benefits of Saving on the Learning and Remembering of New Information. *Psychological Science, 26*, 182–188. doi:10.1177/0956797614559285

Strange, D., Garry, M., Bernstein, D. M., & Lindsay, D. (2011). Photographs cause false memories for the news. *Acta Psychologica, 136*, 90–94. doi:10.1016/j.actpsy.2010.10.006

Strange, D., Sutherland, R., & Garry, M. (2006). Event plausibility does not determine children's false memories. *Memory, 14*, 937–951. doi:10.1080/09658210600896105

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. *Psychological Science, 12*, 462–466. doi:10.1111/1467-9280.00386

Stroud, J. M. (1955). The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology: problems and methods* (pp. 174–207). Glencoe, IL: Free Press.

Sturken, M., & Cartwright, L. (2009). *Practices of looking: an introduction to visual culture* (2nd ed.). Oxford, England: Oxford University Press.

Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience, 1*, 183–184. doi:10.1038/630

Symons, L. A., Cuddy, F., & Humphrey, K. (2000). Orientation tuning of shape from shading. *Perception & Psychophysics, 62,* 557-568. doi:10.3758/BF03212107

Tan, M., Lalonde, J. F., Sharan, L., Rushmeier, H., & O'Sullivan, C. (2015). The perception of lighting inconsistencies in composite outdoor scenes. *ACM Transactions on Applied Perception (TAP), 12*, 1–18. doi:10.1145/2810038

Tarr, M. J., Kersten, D., & Bülthoff, H. H. (1998). Why the visual recognition system might encode the effects of illumination. *Vision research*, *38*, 2259–2275. doi.org/10.1016/S0042-6989(98)00041-8

Thompson, J. K., & Heinberg, L. J. (1999). The media's influence on body image disturbance and eating disorders: We've reviled them, now can we rehabilitate them? *Journal of Social Issues*, *55*, 339–353. doi:10.1111/0022-4537.00119

Thompson, J. K., Heinberg, L. J., Altabe, M., & Tantleff-Dunn, S. (1999). *Exacting beauty: Theory, assessment, and treatment of body image disturbance.* Washington, DC: American Psychological Association.

Thompson, J. K., & Stice, E. (2001). Thin-ideal internalization: Mounting evidence for a new risk factor for body-image disturbance and eating pathology. *Current Directions in Psychological Science, 10*, 181–183. doi:10.1111/1467-8721.00144

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520–522. doi:10.1038/381520a0

3D World Studio (Version 5.6) [Computer software]. Sacramento, CA: Leadwerks.

Tiggemann, M., & McGill, B. (2004). The role of social comparison in the effect of magazine advertisements on women's mood and body dissatisfaction. *Journal of Social and Clinical Psychology, 23*, 23–44. doi:10.1521/jscp.23.1.23.26991

Tiggemann, M., Slater, A., Bury, B., Hawkins, K., & Firth, B. (2013). Disclaimer labels on fashion magazine advertisements: Effects on social comparison and body dissatisfaction. *Body Image*, *10*, 45–53. doi:10.1016/j.bodyim.2012.08.001

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*, 189–208. doi:10.1037/h0061626

Tooth, R. (2011, September 4). The head of photography on… picture manipulation and trust in news imagery. *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2011/sep/04/picture-manipulation-news-imagery-photoshop

Trenholm, R. (2007, November 5). Photos: The history of the digital camera. *CNET News.* Retrieved from https://www.cnet.com/uk/news/photos-the-history-of-the-digital-camera/

Tsotsos, J. K. (1990). Analyzing vision at the complexity level. *Behavioral and Brain Sciences*, *13*, 423–445. doi:10.1017/S0140525X00079577

Turley, J. (2012, February 17). Arizona Legislator Moves To Bar Airbrushing Of Models [Blog post]. Retrieved from https://jonathanturley.org/2012/02/17/arizona-legislator-moves-to-bar-airbrushing-of-models/

Turner, C. (2016, May 11). Meghan Trainor pulls music video that digitally altered her waistline. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/news/2016/05/11/meghan-trainor-pulls-music-video-that-digitally-altered-her-wais/

Wade, K. A., Garry, M., Read, J. D., & Lindsay, D. S. (2002). A picture is worth a thousand lies: Using false photographs to create false childhood memories. *Psychonomic Bulletin & Review, 9*, 597–603. doi:10.3758/BF03196318

Wade, K. A., Green, S., & Nash, R. A. (2010). Can fabricated evidence induce false eyewitness testimony? *Applied Cognitive Psychology, 24*, 899–908. doi:10.1002/acp.1607

Wallach, H., & Lewis, C. (1966). The effect of abnormal displacement of the retinal image during eye movements. *Attention, Perception, & Psychophysics, 1*, 25–29. doi:10.3758/BF03207816

Wallwork, E. (2015, March 18). France To Ban Excessively Skinny Models: Fashion Insiders Weigh In. *Huffington Post UK*. Retrieved from http://www.huffingtonpost.co.uk/2015/03/18/france-to-ban-skinny-models-fashion-insiders-comment_n_6894048.html

Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York, NY: Springer-Verlag.

Wen, T. (2017, June 30). The hidden signs that can reveal a fake photo. *BBC Future*. Retrieved from http://www.bbc.com/future/story/20170629-the-hidden-signs-that-can-reveal-if-a-photo-is-fake

Whittaker, S., Bergman, O., & Clough, P. (2010). Easy on that trigger dad: a study of long term family photo retrieval. *Personal and Ubiquitous Computing, 14*, 31–43. doi:10.1007/s00779-009-0218-7

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of vision, 4*, 1120–1135. doi:10.1167/4.12.11

World Press Photo. (n.d.). Photo contest code of ethics. Retrieved from http://www.worldpressphoto.org/activities/photo-contest/code-of-ethics

Worthington, P. (2014, December 11). One trillion photos in 2015. *Mylio*. Retrieved from http://mylio.com/true-stories/tech-today/one-trillion-photos-in-2015-2

Wright, M. J. (2005). Saliency predicts change detection in pictures of natural scenes. *Spatial Vision, 18*, 413–430. doi:10.1163/1568568054389633

Wright, T. (2016). *The photography handbook* (3rd ed.). Abingdon, England: Routledge.

Xia, L., Pont, S. C., & Heynderickx, I. (2016). Effects of scene content and layout on the perceived light direction in 3D spaces. *Journal of Vision, 16*, 1–13. doi:10.1167/16.10.14

Young, K., & Silverman, R. (2014, October 27). Dame Helen Mirren: don't retouch my image, L'Oreal. *The Telegraph*. Retrieved from http://www.telegraph.co.uk/news/celebritynews/11190510/Dame-Helen-Mirren-dont-retouch-my-image-LOreal.html

Appendices

Appendix A: Saliency Analyses in Chapter 3

Saliency analyses

The extent to which certain parts of a scene stand out relative to other parts (object or region saliency) can affect the distribution of visual attention. We ran two saliency analyses to check whether our manipulations influenced the salience of the region where the manipulation had been made.

We chose two of the various models available to compute bottom-up predictions of visual salience; the classic Itti-Koch model (Itti & Koch, 2000; Itti, Koch, & Niebur, 1998) and the Graph-Based Visual Saliency (GBVS) model (Harel, Koch, & Perona, 2006). For simplicity, we will refer to these models as IK and GBVS.

Experiment 1

IK

Using Matlab, we ran the IK model on the six original images and the 30 manipulated versions of these images. For each image, the model created a saliency map with a saliency value for each individual pixel. Recall that in the location task of Experiment 1, subjects saw the image with a 3 × 3 grid overlaid and were asked to select the region that they believed had been manipulated. Therefore, to quantify the saliency values we calculated the mean saliency for each of the nine regions for each the 36 images. We then checked whether our manipulations of the images had affected the saliency of the manipulated region compared with the same region in the original image. That is, did our manipulations make the regions any more or less salient than they were in the original image? Figure A.1 shows the mean saliency of the *manipulated* region alongside the mean saliency of that same region in the *original* image. As shown, our manipulations had no systematic effect.

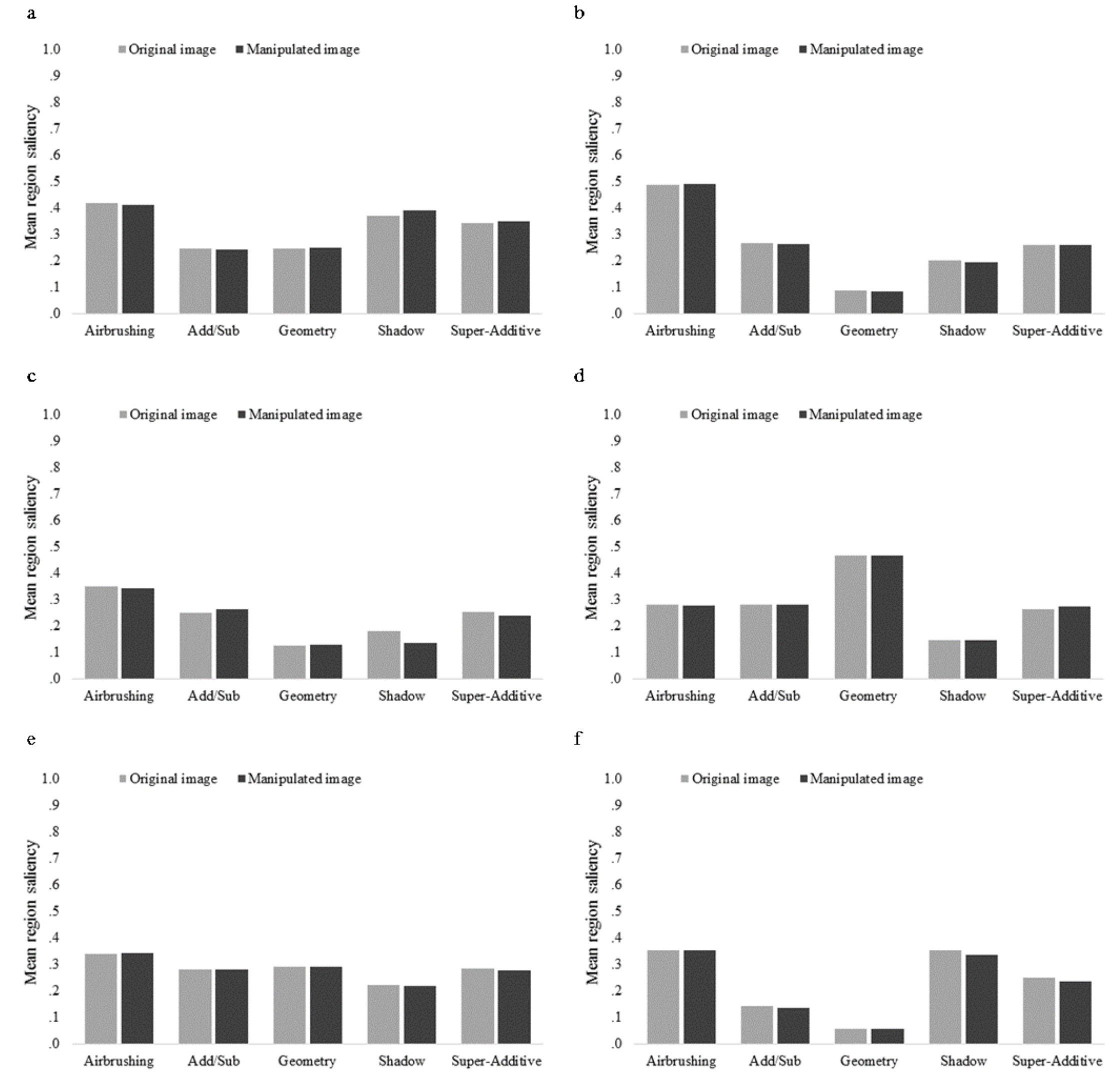


Figure A.1. Mean saliency of the manipulated region in the manipulated image and mean saliency of the same region in the original image, where higher values indicate a more salient region. Data is shown for each of the 6 images, a-f, and for all five manipulation types, computed using the IK model.

GBVS

Next, in Matlab we ran the same analysis using the GBVS model to determine whether our manipulations had influenced region salience. Replicating the results from the IK model, Figure A.2 shows that our manipulations made little difference to the salience of the image regions.

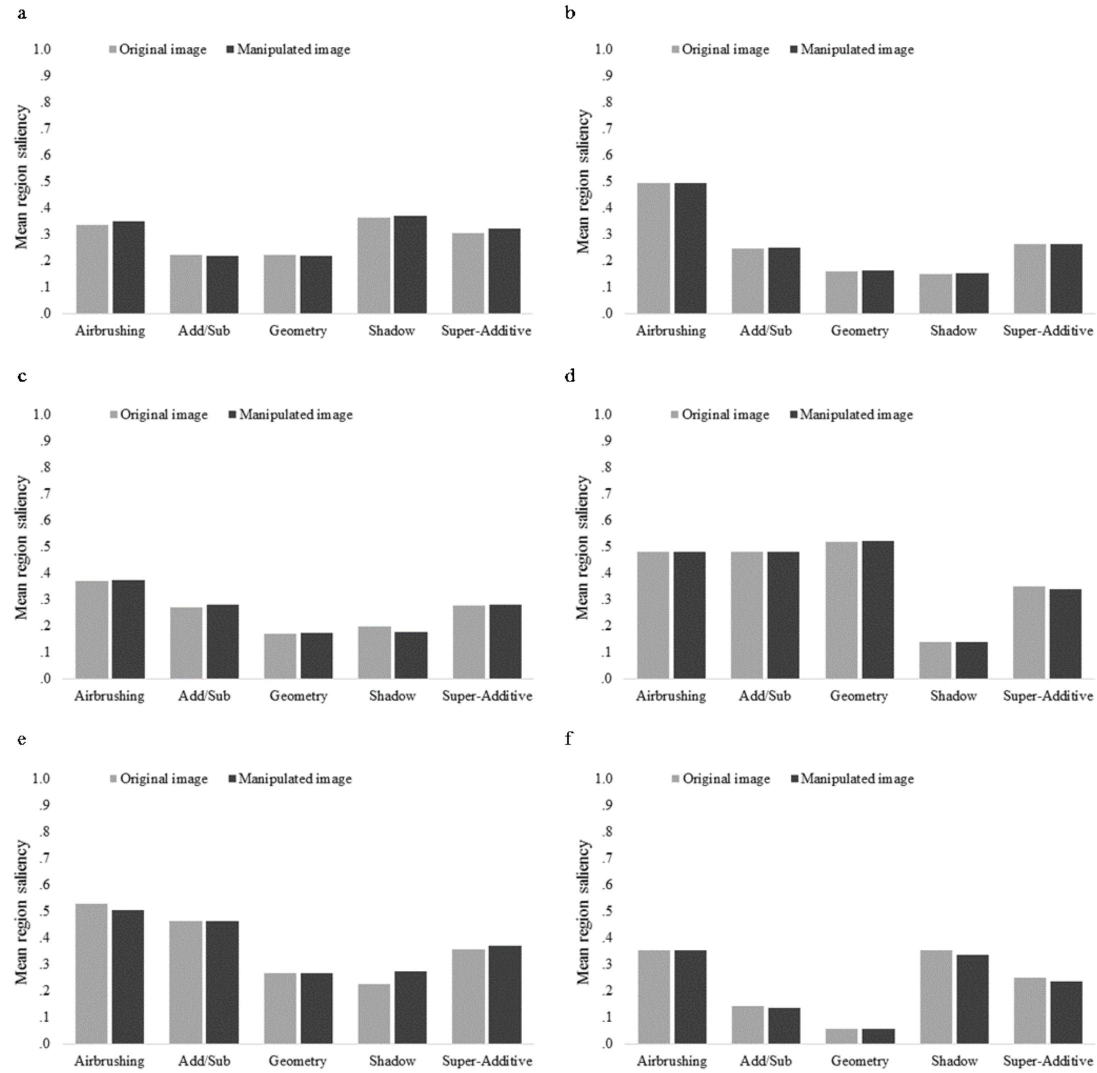


Figure A.2. Mean saliency of the manipulated region in the manipulated image and mean saliency of the same region in the original image, where higher values indicate a more salient region. Data is shown for each of the 6 images, a-f, and for all five manipulation types, computed using the GBVS model.

Experiment 2

IK

We ran the IK model on the six original images and the 30 manipulated versions of these images. For each image, the model created a saliency map with the saliency value for each individual pixel. Recall that in Experiment 2 the location task used a 4 × 3 overlaid grid. Therefore, to quantify the saliency values, we calculated the mean saliency for each of the 12 regions. We calculated these mean values for each region in each of the 36 images.

As in Experiment 1, we considered whether our manipulations of the images had made any difference to the saliency of the manipulated region compared with that same region in the original image. Figure A.3 shows a similar pattern of results to those observed in Experiment 1—our manipulations had little or no systematic effect on region salience.

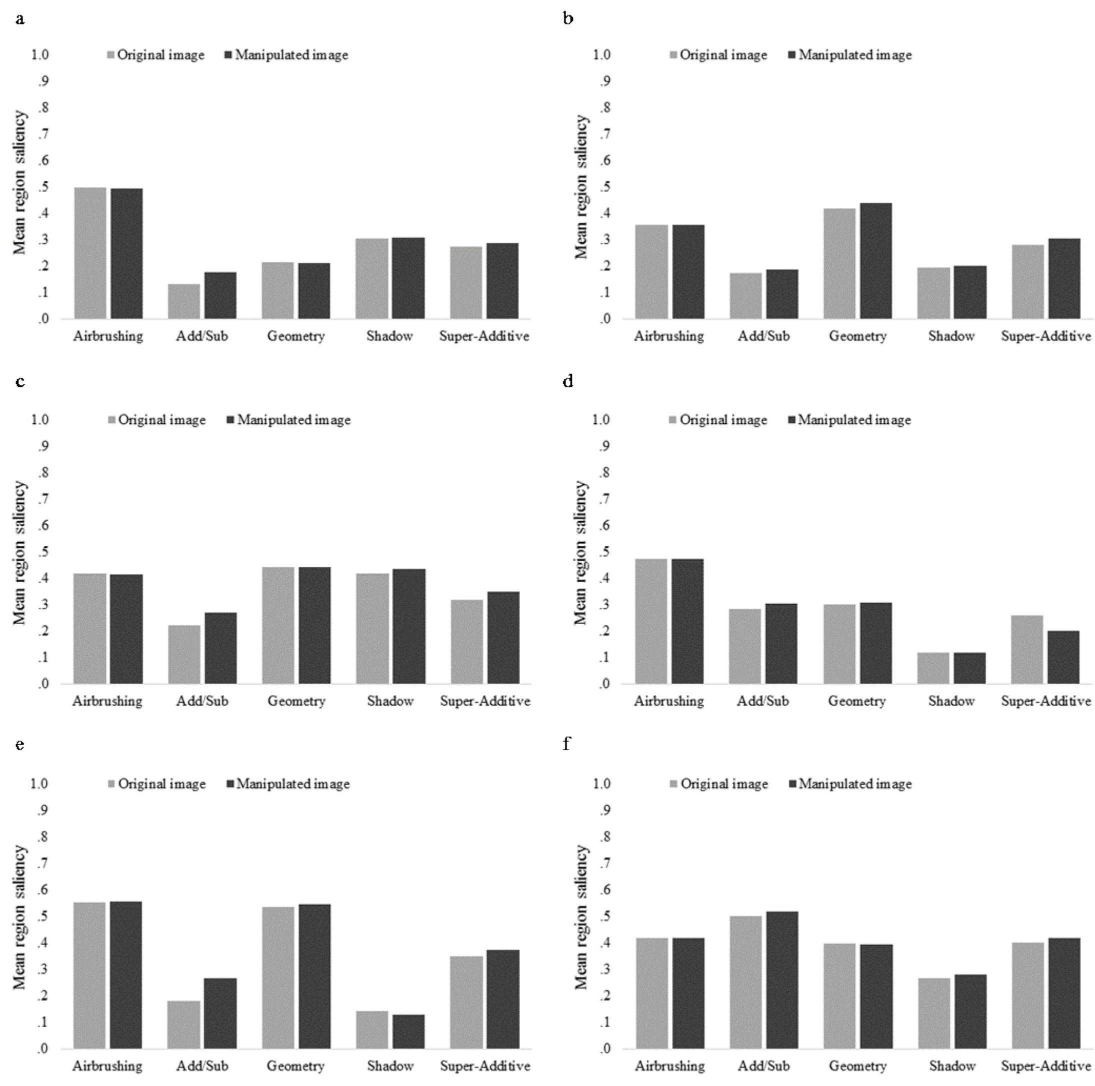


Figure A.3. Mean saliency of the manipulated region in the manipulated image and mean saliency of the same region in the original image, where higher values indicate a more salient region. Data is shown for each of the 6 images, a-f, and for all five manipulation types, computed using the IK model.

GBVS

Next we ran the GBVS model on all six original and the 30 manipulated images. Again, we considered whether our manipulations made the regions any more or less salient than they were in the original image. Replicating our previous results, Figure A.4 shows that our manipulations made little difference to the salience of the images. In sum, as in Experiment 1, our saliency analyses indicate that we did not influence people's accuracy on the detection and location tasks by inadvertently changing the salience of the manipulated regions.

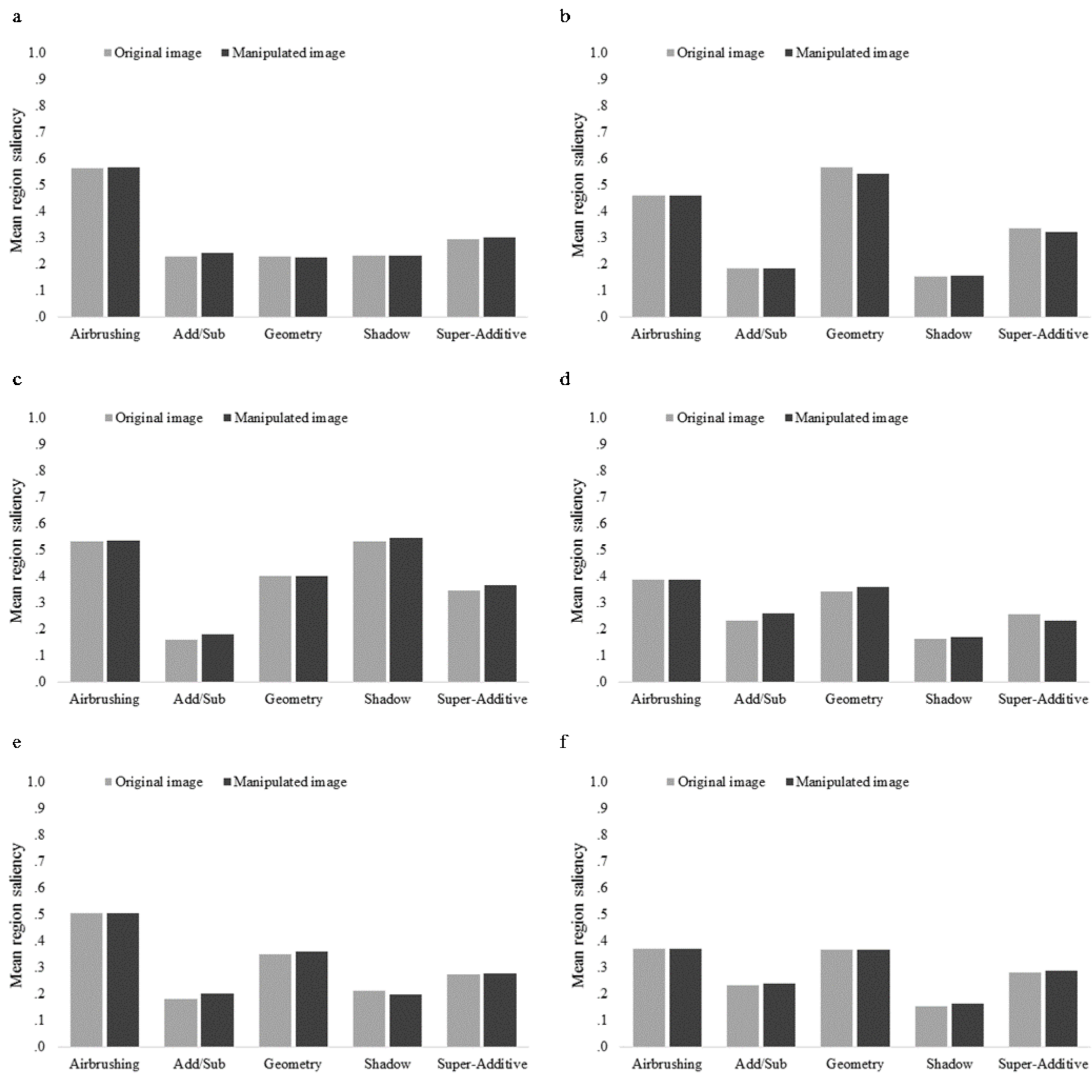


Figure A.4. Mean saliency of the manipulated region in the manipulated image and mean saliency of the same region in the original image, where higher values indicate a more salient region. Data is shown for each of the 6 images, a-f, and for all five manipulation types, computed using the GBVS model.

Appendix B: Exceptions to procedure in Chapter 7

All exceptions were made to add experimental rigor and/or generalisability.

Experiment 1

Table B.1 provides details of the exceptions.

Table B.1
Exceptions to Henkel's (2014, Experiment 1) procedure

Exception	Details and Rationale
Museum location	The experiment took place at a different museum: The Otago Museum in New Zealand.
Camera type	Henkel's subjects used a digital camera to take photos but we used an iPod touch because subjects routinely use similar devices to take photos.
Photo review	Henkel did not explicitly tell subjects that they would get an opportunity to see the photos the next day, but we did. We gave subjects this false information to make the process more similar to everyday life—after all, if people really do rely on cameras to remember for them, it is likely that they do so believing they will have access to their photos later.
Tour direction	Henkel used only one direction of the tour (AB), we also included the reverse (BA) to account for the order in which subjects encountered objects.
Timing for viewing and photographing objects	In Henkel's Experiment 1, subjects had 30 s with each object, regardless of whether they were only viewing or photographing the objects. For the view objects, subjects had the full 30 s to simply view the object. For the photograph objects, however, subjects only had 20 s uninterrupted to view the object and then the remaining 10 s were used to take the photo. Henkel noted the

	time difference as a limitation and changed the timings for Experiment 2; we used her Experiment 2 timings.
Distractor task	The distractor task was included as a mask to stop immediate rehearsal of the objects.
Test phase	We included the location-recognition test from Henkel's Experiment 2. But we did not ask subjects to answer questions about details of the objects. In Henkel's experiments, the visual details test involved subjects answering questions about details of only the objects that were recognised as old in the name-recognition test. It is possible that the additional information provided in the visual detail questions might affect subjects' responses in the photo-recognition test. For the location-recognition test, subjects indicated the location of all 30 objects. In Henkel's Experiment 2, subjects only indicated the location of the objects they recognised as old in the photo-recognition test.

Experiment 2

Table B.2 provides details of the exceptions.

Table B.2

Exceptions to Henkel's (2014, Experiment 2) procedure

Exception	Details and Rationale
Museum tour	To convert the experiment to a lab-based procedure, subjects watched a recording of a museum tour on a computer instead of actively walking around a museum.
Number of objects	Henkel's museum tour included stops at 27 objects, ours included stops at 15 objects.

Camera type	Henkel's subjects used a digital camera to take photos but we used a virtual camera. The virtual camera appeared on-screen when subjects were instructed to photograph an object. If subjects were still unsure about how to use the virtual camera after completing the practice, they were allowed to do the practice again (only one subject asked to do the practice a second time).
Timing for viewing and photographing objects	One reason for developing the lab-based analogue for the museum tour procedure was to remove the external distractions from the museum environment. Using the same timing restraints as in Henkel's Experiment 2, but without external distractions in the museum environment, would likely improve subjects' overall performance on the memory tests, potentially leading to a ceiling effect. For this reason, we conducted several pilot tests with different manipulations of the timing restraints for viewing and photographing the objects in the museum. Subjects' overall performance on the name- and photo-recognition tests was similar to Henkel's Experiment 2 with a 5 s view time and additional 10 s to take a photo.
Test phase	Henkel removed the free-recall test in her Experiment 2, but we included it to see whether we could replicate our finding from our Experiment 1. We did not include the location-recognition test because our subjects did not actively walk around the museum and because the objects were located in a single room.

Experiment 4

Table B.3 provides details of the exceptions to our Experiment 3 procedure.

Table B.3

Exceptions to the Experiment 3 procedure

Exception	Details and Rationale
Museum tour	To gain more experimental control, subjects were navigated through a virtual museum instead of watching a video recording of a museum tour.
Target objects	One reason for moving to use a virtual museum was to have more control over the stimuli. As such, we selected objects with similar attributes which allowed us to create less variation across the objects. We carried out pilot tests to check that each of the 15 objects we selected for our stimulus set were recognised at a similar level in the memory test phase. Subjects' performance on two of the objects was perfect across all three action conditions so we replaced these objects as it seemed likely that they were more salient than the other 13 objects included in the tour.
Distractor objects	We placed three distractor objects near each of the target objects. Therefore, when subjects stopped at each of the 15 target objects in the museum they would see the target object and three distractor objects.
Museum layout	We spaced the objects across four different rooms in the virtual museum. Although this marks a change from our Experiments 2 and 3, it is more similar to the museum layouts used in Henkel's (2014) experiments and our Experiment 1.
Virtual camera	We made a change to the virtual camera to further ensure that subjects were aware their photos were being stored. Each time subjects took a photo, in addition to displaying the shot they had

	captured on the screen of the virtual camera for 2 seconds, we included a message at the top of the screen that read “saving.”
Timing for viewing and photographing objects	We conducted several pilot tests with different manipulations of the timing restraints for viewing and photographing the objects in the museum. Subjects’ overall performance on the photo-recognition test was similar to Henkel’s (2014, Experiment 2) with a 10 s view time and additional 10 s to take a photo.
Length of time for filler task	We also used the pilot tests to determine the length of time for the filler task—we reduced the duration of the filler task to 2 min.
Test phase	We excluded the name-recognition and visual details tests. Whether subjects viewed or photographed objects made a trivial difference to performance on either of these tests, therefore we included only the photo-recognition test.

Appendix C: Null Hypothesis Significance Testing Approach for the results in Chapter 7

We also ran our analyses in null hypothesis significance testing (NHST) terms.

Experiment 1

Free recall

A paired-samples t test revealed that subjects were more likely to recall the photographed objects than the merely viewed objects: M_{view} 46%; $M_{photograph}$ = 54%; $t(41) = 2.78$, $p = .01$, $d = 0.44$.

Recognition

A 2 (Action: view, photograph) × 2 (Retrieval Cue: name-cue, photo-cue) repeated measures analysis of variance (ANOVA) on subjects' correct old recognition showed no effect of action, $F(1, 41) = 0.07$, $p = .79$, $\eta_p^2 = .002$, no effect for the retrieval cue, $F(1, 41) = 1.22$, $p = .28$, $\eta_p^2 = .03$, and no interaction between action and retrieval cue, $F(1, 41) = 0.09$, $p = .77$, $\eta_p^2 = .002$. The results show that there was no difference in subjects' ability to recognise objects as old in the two action conditions and that their recognition performance did not vary as a function of the type of retrieval cue.

For source accuracy, we conducted an Action (view, photograph) × Retrieval Cue (name-cue, photo-cue) repeated measures ANOVA and found a main effect of action, $F(1, 41) = 25.94$, $p < .001$, $\eta_p^2 = .39$. Subjects attributed more of the objects they viewed to the accurate source than objects they photographed, M_{view} 71%; $M_{photograph}$ = 49%. There was also an effect of retrieval cue, $F(1, 41) = 14.18$, $p = .001$, $\eta_p^2 = .26$, in which subjects attributed more objects to the correct source when their memories were cued by photos than by names, $M_{name\text{-}cue}$ = 56%; $M_{photo\text{-}cue}$ = 63%. There was no interaction.

For location accuracy, a paired-samples t test revealed that there was no effect of action on the proportion of objects recognised in the correct location, M_{view} 81%; $M_{photograph}$ = 80%; $t(41) = 0.47$, $p = .64$, $d = 0.07$.

Experiment 2

Free recall

A one-way repeated measures ANOVA revealed no effect of action on the proportion of objects correctly recalled, $F(2, 82) = 1.28$, $p = .28$, $\eta_p^2 = .03$.

Recognition

In NHST terms, for the recognition tests, a 3 (Action: view, photo-whole, photo-part) × 2 (Retrieval Cue: name-cue, photo-cue) repeated measures ANOVA on subjects' correct old recognition revealed no effect of action, $F(2, 82) = 2.48$, $p = .09$, $\eta_p^2 = .06$. There was an effect of retrieval cue, $F(1, 41) = 64.98$, $p < .001$, $\eta_p^2 = .61$—subjects correctly recognised markedly more objects as old in the photo-recognition test than in the name-recognition test, $M_{photo\text{-}cue} = 93\%$, $M_{name\text{-}cue} = 73\%$. There was no interaction between action and retrieval cue, $F(2, 82) = 1.65$, $p = .20$, $\eta_p^2 = .04$.

For accuracy on the visual details test, a one-way repeated measures ANOVA revealed no effect of action, $F(2, 82) = 1.63$, $p = .20$, $\eta_p^2 = .04$. Finally, to explore source accuracy, we conducted an Action (view, photo-whole, photo-part) × Retrieval Cue (name-cue, photo-cue) repeated measures ANOVA and found a main effect of action, $F(2, 82) = 6.61$, $p = .002$, $\eta_p^2 = .14$; a main effect of retrieval cue, $F(1, 41) = 54.29$, $p < .001$, $\eta_p^2 = .57$; and an interaction between action and retrieval cue, $F(2, 82) = 5.64$, $p = .01$, $\eta_p^2 = .12$. This interaction indicates that the type of retrieval cue had a different effect, depending on the action subjects took. To explore this interaction further, we first looked at the effect of action on source accuracy in the name-recognition test. A one-way repeated measures ANOVA revealed no effect of action, $F(2, 82) = 2.1$, $p = .14$, $\eta_p^2 = .05$ indicating that subjects were equally likely to attribute view, photo-whole and photo-part objects to the accurate source when their memories were cued by names. Next we ran a one-way repeated measures ANOVA to look at the effect of action on source accuracy in the photo-recognition test and found a main effect, $F(2, 82) = 9.53$, $p < .001$, $\eta_p^2 = .19$. Post hoc Bonferroni-corrected comparisons showed that subjects were more accurate in recognising which action they had performed for photo-part (74%) than photo-whole objects (47%) or view objects (57%): photo-part and photo-whole, $t(41) = 3.96$, $p < .001$, $d = 1.07$; photo-part and view, $t(41) = 3.77$, $p = .001$, $d = 0.59$.

Experiment 3

Recognition

For the name- and photo-recognition tests, a 3 (Action: view, photo-whole, photo-part) × 2 (Retrieval Cue: name-cue, photo-cue) repeated measures ANOVA on subjects' correct old recognition showed no effect of action, $F(2, 82) = 0.50$, $p = .61$, $\eta_p^2 = .01$.

There was an effect for the retrieval cue, $F(1, 41) = 80.36$, $p < .001$, $\eta_p^2 = .66$, and an interaction between action and retrieval cue, $F(2, 82) = 7.86$, $p = .001$, $\eta_p^2 = .16$. To explore the interaction further, we ran a one-way repeated measures ANOVA for the name-recognition test and for the photo-recognition test. There was no main effect of action on the name-recognition test, $F(2, 82) = 1.28$, $p = .28$, $\eta_p^2 = .03$—subjects correctly recognised a similar proportion of view, photo-whole and photo-part objects. In contrast, there was a main effect of action on the photo-recognition test, $F(2, 82) = 6.60$, $p = .002$, $\eta_p^2 = .14$. But the pattern of results does not fit with photo-taking impairment effect. Post hoc Bonferroni-corrected comparisons revealed that when subjects' memories were cued with photos, they correctly recognised more of the photo-part than view objects: $M_{photo\text{-}part} = 87\%$, $M_{view} = 74\%$, $t(41) = 3.95$, $p < .001$, $d = 0.58$. Subjects recognised a similar proportion of the view and photo-whole objects.

Next, we considered the effect of photographing on subjects' memory for the visual details of the objects. We replicated our finding from Experiment 2—subjects correctly answered a similar proportion of visual detail questions for the view, photo-whole and photo-part objects, $F(2, 82) = 0.19$, $p = .83$, $\eta_p^2 = .01$.

Finally, to explore subjects' ability to remember which objects they photographed and which they viewed, we ran a 3 (Action: view, photo-whole, photo-part) × 2 (Retrieval Cue: name-cue, photo-cue) repeated measures ANOVA on the proportion of objects that subjects correctly attributed to their source. This analysis revealed no main effect of action, $F(2, 82) = 1.01$, $p = .37$, $\eta_p^2 = .02$, but a main effect of retrieval cue, $F(1, 41) = 10.23$, $p = .003$, $\eta_p^2 = .20$, and an interaction, $F(2, 82) = 4.49$, $p = .01$, $\eta_p^2 = .10$. As was the case in Experiment 2, this interaction means that the type of retrieval cue had a different effect depending on the action taken. To investigate further, we ran a one-way repeated measures ANOVA to look at the effect of action on source accuracy in the name-recognition test. The analysis revealed no effect of action, $F(2, 82) = 0.52$, $p = .60$, $\eta_p^2 = .01$, indicating that subjects were equally likely to attribute view, photo-whole and photo-part objects to the correct source when given name cues. A one-way repeated measures ANOVA exploring the effect of action on source accuracy in the photo-recognition test revealed a main effect, $F(2, 82) = 3.88$, $p = .03$, $\eta_p^2 = .09$. As in Experiment 2, post hoc Bonferroni-corrected comparisons showed that subjects attributed more photo-part objects than view objects to the accurate source: $M_{photo\text{-}part} =$

49%, M_{view} = 36%, $t(41) = 2.62$, $p = .01$, $d = 0.51$. Subjects attributed a similar proportion of the view and photo-whole objects to the correct source.

Experiment 4

Recognition

A one-way repeated measures ANOVA revealed that action had an effect on the proportion of objects correctly recognised as old, $F(2, 82) = 4.51$, $p = .01$, $\eta_p^2 = .10$. Post hoc Bonferroni-corrected comparisons indicated one reliable difference: Subjects correctly recognised a greater proportion of the photo-part objects than the photo-whole objects as having been part of the museum tour: $M_{photo\text{-}part}$ = 78%, $M_{photo\text{-}whole}$ = 65%, $t(41) = 2.82$, $p = .01$, $d = 0.54$.

Experiment 5

Recognition

A one-way repeated measures ANOVA on subjects' correct old recognition showed an effect of action, $F(2, 82) = 6.74$, $p = .002$, $\eta_p^2 = .14$. Post hoc Bonferroni-corrected comparisons indicated one reliable difference: As in Experiment 4, subjects correctly recognised a greater proportion of the photo-part objects than the photo-whole objects as having been part of the museum tour: $M_{photo\text{-}part}$ = 80%, $M_{photo\text{-}whole}$ = 66%, $t(41) = 3.53$, $p = .001$, $d = 0.63$.