2-2014

The Effect Of Malintent On Visual Attention

David Brian Wallace
Graduate Center, City University of New York

THE EFFECT OF MALINTENT ON VISUAL ATTENTION

by

DAVID BRIAN WALLACE

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2014

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Maria Hartwig

Date

Chair of Examining Committee

Maureen O'Connor

Date

Executive Officer

Saul Kassin

Steven Penrod

Jennifer Dysart

Pär-Anders Granhag

Bruno Verschuere

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

THE EFFECT OF MALINTENT ON VISUAL ATTENTION

by

David Brian Wallace

Adviser: Dr. Maria Hartwig

Malintent is a relatively unexplored phenomenon, despite the practical and theoretical interest in its detection. The four studies presented here, informed by the Guilty Knowledge Test and Guided Search model of visual attention, illustrate how eye tracking can (and cannot) be utilized to detect malicious intent. Malintent, induced in two different mock crime paradigms (murder; theft), boosted the frequency and duration of fixations that fell upon objects relevant to their task. The size of the malintent effect ranged from moderate (fixation duration) to large (fixation frequency). The increased visual attention given to task-relevant objects was not unique to malintent; benign intent individuals, given the same task to perform without the transgressive framing, showed similar gaze patterns when they were unaware of the eye-tracking equipment. When made aware that their gaze was being tracked, those with malintent successfully avoided looking at the task-relevant object, while those with benign intent looked at it more often. The large moderating effect of eye-tracking awareness on the malintent effect poses both a challenge and an opportunity to the successful detection of malintent.

Table of Contents

The Effect of Malintent on Visual Attention

The detection of malicious intent, or *malintent*, has taken on additional urgency after the terrorist attacks on September 11th, 2001. American law enforcement and security agencies in particular are increasingly motivated to prevent terrorism before it occurs – and yet failed to do so in two recent events that captured the public's attention: two explosions at the Boston Marathon in April 2013, and the failed Times Square device in May of 2010. In both cases, the explosive devices were constructed from easily obtainable materials placed in a pressure cooker – and yet they caused or had the potential to cause large numbers of casualties.

One of the suspects in the Boston Marathon bombings, Tamerlan Tsarnaev, was interviewed by the FBI two years before the bombing. It may never be known whether he had any malintent at the time, but it illustrates the potential difficulty of determining whether an individual with strong ideological opinions intends to express them via acts of terrorism. Given the ease of constructing a high-impact explosive device, one may harbor malintent for long periods, undetectable by even suspicious law enforcement (as the FBI were or should have been, after a tipoff from Russia's security services). Only in the short period before the act is actually carried out must any physical component to the plan be included – the purchase of a pressure cooker, some fireworks, and nails. At that point, it may be too late in many cases. Accurate malintent detection could play a crucial role in preventing such events.

Detecting malintent is daunting from a psychological point of view, however. Intentions, until acted upon, can be vague and ephemeral. If one intends to murder an individual, but doesn't (yet?) know how or when, can this be considered malintent? Even concrete plans are subject to change as rapidly as the planner modifies their thoughts. To place some boundaries on a potentially broad scope of thoughts which one might arguably label as malintent, we borrow from (Granhag, 2010) definition of intent, which is tailored to forensic psychology research; he constraints the scope of intent to specific acts (with a when, how, and where) planned by an individual who is in custody or otherwise available for interview or interrogation. For our purposes then, *malintent* occurs when an individual has a specific plan to commit a series of actions at a known place and time, which he or she knows is considered transgressive in some way. Most malintent of interest to forensic psychology will concern transgressions of a criminal sort, though malintent could certainly include common social taboos such as adultery (which technically remains a crime in several states; Bronner, 2012).

Having placed some constraints upon the definition of malintent, how might one go about identifying individuals who possess it? In the following pages, we outline existing psychological theory and research that may be relevant to malintent detection, beginning with a discussion of current attempts at detecting malintent, and note the near complete absence of malintent-detection research and malintent-specific theory. We then briefly discuss the more robust field of deception detection, including false intent (lies about future plans), as lie detection may offer some insight into what does and does not make for successful malintent detection – the Guilty Knowledge Test (GKT) and orienting response are especially relevant. We then move on to consider theory of human visual attention, especially the Guided Search (GS) model, and discuss why eye tracking may be a good place to start with detecting malintent. Then we sample the eye-tracking research, presenting relevant non-forensic studies, as well as take a look at how eye tracking is used to measure visual attention in forensic psychology contexts specifically.

After reviewing the psychological literature, we take inspiration from the Guilty Knowledge Test, orienting response, and the Guided Search model to present a theory for the effect of malintent on visual attention patterns: in brief, that the intended interaction with an object (person/place/thing) as part of a plan, as well as knowing that the intent is transgressive will independently produce extra visual salience for said objects. Finally, we present the results of four empirical studies which test various aspects and implications of this theory, using unobtrusive eye-tracking equipment to record measures of visual attention.

Malintent

SPOT. Efforts to detect malintent are especially visible to the general public in security screening settings, such as airports, where multiple new programs have been introduced. One program, the Transportation Security Administration's (TSA) Screening Passengers through Observational Techniques (SPOT) is "a security system based on human observation of suspicious behaviors." (US Department of Homeland Security, 2008, p. 11). SPOT is (as of March 2010) present in over 100 airports, and employs over 3,000 behavior detection officers (BDO), who inspect the behavior of airport passengers as they proceed through security screening, seeking individuals who harbor some form of malintent, whether it be terrorism, or more mundane crimes like smuggling (United States Government Accountability Office, 2010).

Based on work by Paul Ekman, SPOT relies on an emotional approach toward malintent-detection: the core assumption is that individuals who harbor malintent experience emotion related to the event; fear, anxiety, and other affective states can be detected through "leakage" in various channels of nonverbal behavior, such as brief involuntary facial movements known as micro-expressions (Ekman & O'Sullivan, 2006). However, there is no published research to support this claim in the domain of malintent generally, nor SPOT specifically. The idea that micro-expressions can be used to detect emotional concealment is a questionable one, as demonstrated by Porter and ten Brinke (2008). In their study, participants were shown emotionally evocative images and asked to suppress any emotional facial expressions, or to mask them with an emotion inconsistent with what they were being shown (e.g. to look happy while viewing a disgusting image); results showed that untrained participants were highly successful at maintaining a neutral expression to mask a felt emotion, that micro-expressions were rare, and that evaluators showed poor accuracy (60%) in detecting the emotional deception.

Another program, Future Attribute Screening Technology (FAST), while not (yet) implemented, “is an initiative to develop innovative, non-invasive technologies to screen people at security checkpoints” (Department of Homeland Security, 2011). The government’s reveal of FAST has provoked sensationalist media claims such as “Homeland Security is now testing the next generation of security screening — a body scanner that can read your mind” (Barrie, 2008). As with SPOT, there is no peer-reviewed research on the potential efficacy of a program such as FAST. The lack of published research for these two projects is not unique to government work – there is currently no published peer-reviewed research on detecting malintent using any procedure whatsoever.

Aikins, Martin, and Morgan (2010) have come the closest to providing data on the detection of malintent. They hypothesized that, due to the cognitive effort required by deception, individuals with deceptive intent would show decreases in respiratory sinus arrhythmia (RSA) – the variance in heart rate caused by breathing. They fitted participants with a shirt embedded with physiological recording devices (to measure RSA) then assigned them to either performing a mock crime (theft) that they should lie about when questioned, or perform the same task without identifying it as a theft, and to tell the truth when questioned. Before venturing to complete the tasks, participants’ RSA were measured during two cognitive load exercises. The authors found that RSA was lower for the guilty condition during one of the two cognitive load tasks. While the authors interpreted the results as detecting intent to deceive, it is unclear, given their design, whether the RSA differences were attributable to their intention to commit a crime, or their intention to lie about it afterwards, as the two were confounded in the study. No classification accuracy was presented.

False Intent

The next-closest body of research to malintent is a series of recent studies regarding the detection of deception in a statement made by an individual regarding their intentions – lying about what they plan to do in the future. These (false) statements are termed false intent, or deceptive intent. It is important to note that in much of the false intent research, the intent is benign. This distinction between detecting malintent and detecting false intent is not a subtle one -- Granhag (2010) draws a bold line between the two, equating the former to “mindreading” (p. 38). The critical difference between the two is that detecting false intent hinges upon evaluating the statement made by a suspect. Thus, when a suspect is available to give a statement, and if deception-detection techniques are highly accurate, then malintent can be identified via a false intent strategy. As can be seen in the subsequent descriptions of false intent research, high accuracy rates are not common.

Warmelink et al. (2010) tested the utility of thermal imaging to detect false intent from travelers at an airport. Participants were asked to tell the truth, or lie, about their travel plans, and were interviewed about these plans while their skin temperature was recorded with a thermal imaging camera. While there was a significant veracity effect, manifested in skin temperatures for liars increasing as the interview progressed (while truth-tellers’ temperatures remained constant), the differences produced only moderately accurate classification rates, 68% of truth-tellers and 65% of liars. Despite the fact that research generally shows individuals have poor deception-detection skills (Bond & DePaulo, 2006; 2008), the interviewers’ subjective assessments outperformed both chance and the objective measure. Classification models using both the thermal imaging and the interviewers' judgments together were no better than the interviewers alone – implying that the interviewers were capable of picking up on (undetermined) cues which covaried with the thermal imaging data.

Using a similar airport-interview paradigm, (Vrij, Granhag, Mann, & Leal, 2011) asked some travelers to conceal the true purpose of their trip. Importantly, none were experimentally induced to conceal malintent – these were lies about mundane intent. False statements of their travel plans were rated as less plausible than those of truth-tellers, but there were no differences in the level of detail. Classification accuracies were similar between both truth-tellers and liars, at 72% and 74% respectively.

The preceding research on false intent is not redundant with the large body of existing deception detection research (where lies are commonly about past acts, not future plans). Vrij, Leal, Mann, and Granhag (2010) found differences between deception about past events and future intent. In their study, participants were interviewed both before and after a mock crime task. Results showed effects of veracity on the plausibility of both statements about intentions and past activities (in both cases, truth-tellers being much more plausible), but the effects of veracity on the number of details in the statements was limited to past activities – liars and truth-tellers were similarly detailed about their future intentions. Their results underscore not only an important difference between the two types of deception, but one of the problems of assessing intent: neither truth-tellers nor liars had the memory from which to pull sufficient details to describe a future event.

Deception Detection

The difficulty of detecting false intent is unsurprising, given that a large body of deception detection (for a thorough review, see Vrij, 2008) has shown that humans are generally poor at lie-detection. In a context where pure guessing results in a 50% accuracy rate, humans do marginally better: 54% of the time they correctly identify whether an individual is lying or telling the truth (Bond & DePaulo, 2006, 2008). A recent meta-analysis indicates that lie detection is poor not because human lie detectors are using the wrong cues, but because the cues are weak: there is little relationship between lying and objective behavioral cues (Hartwig & Bond, 2011).

There are however several specific strategies that have improved deception-detection rates. Despite the many difference between malintent detection via eye tracking and deception detection (primarily the presence or absence of a potential lie to evaluate), we hope to find commonalities in these deception-detection strategies to help inform a malintent-detection theory and an efficient malintent-detection paradigm. One such lie-detection strategy is commonly found in the polygraph domain.

Guilty knowledge and the polygraph. Polygraphic lie detection ("the polygraph") involves the measurement of physiological variables such as galvanic skin response, heart rate, blood pressure, and breathing displacement in a structured questioning context. In the most common procedure in the United States, the Control Question Test (CQT), an examinee is asked both crime relevant and non-relevant "control" questions, and is judged deceptive if he or she consistently exhibits greater physiological arousal when responding to the crime-relevant questions. The CQT has been criticized for being non-scientific and subject to bias (for a review, see National Research Council, 2003).

The less common procedure, the Guilty Knowledge Test (GKT), or Concealed Information Test (CIT), presents the examinee with a question and several response options – essentially, a multiple choice test. The possible answers contain one correct detail of the crime, and several incorrect distractors. Suspects with guilty knowledge exhibit an elevated orienting response to the correct items but not to those that are incorrect, while naïve suspects show similar patterns to all items see (for a meta-analysis and review of the GKT, see Ben-Shakhar & Elaad, 2003; Verschuere et al., 2011). This orienting response describes a constellation of physiological and behavioral reactions to a novel or meaningful stimulus (Siddle, 1991). Information which is personally relevant provokes larger responses. In the context of a polygraph, the crime-relevant answer in the GKT is elevated above its distractors – but only for individuals familiar with the crime (Lykken, 1974). In a meta-analysis of 22 polygraph studies using the GKT, MacLaren (2001) found that 76% of individuals with concealed knowledge were correctly classified, rising to 82% for those involved in mock crimes; the accuracy rate for innocent examinees was 83%.

A polygraph examination requires a high level of compliance from suspects who must agree to answer questions while sitting still, and they must continue to comply throughout the examination. Additionally, guilty suspects who are motivated to evade detection can often achieve false negative test results by using physical and psychological countermeasures (Honts, Devitt, Winbush, & Kircher, 1996).

Other physiological GKT. Orienting response in a GKT context need not be measured via polygraph. Other psychophysiological techniques have captured similar GKT-elicited orienting responses. The P300 (an event-related potential response measured via EEG, or electroencephalography) has shown increasing use in lie-detection and guilty knowledge-like procedures, with elevated P300 responses occurring for guilty suspects in mock crime paradigms

(Hahm et al., 2009). Several P300 studies have reported both high discrimination accuracies as well as resistance to counter-measures (e.g. Hu, Hegeman, Landry, & Rosenfeld, 2012; Rosenfeld et al., 2008). There is evidence to suggest that P300 GKT procedures can be successful in detecting malintent as well (Meixner & Rosenfeld, 2011).

Studies involving reaction time during a GKT have found that that reaction times to statements relevant to a previously-committed mock crime were higher for guilty individuals (Seymour, Seifert, Shafto, & Mosmann, 2000), and that participants tasked with feigning amnesia showed higher reaction times when reporting they were unfamiliar with personally relevant items (Verschuere, Crombez, Degrootte, & Rosseel, 2010). One reaction-time GKT study established that the orienting response involved in a GKT captures attention (Verschuere, Crombez, & Koster, 2004). Participants were asked to classify the direction of a probe while being shown photographs that were either relevant to guilty knowledge (they were to lie about having seen them in the future), mere knowledge (they had seen them previously), or unseen. Reaction times were higher for guilty knowledge photographs, indicating that the photo was capturing attention away from the task. Reaction-time based GKT procedures have recently begun to investigate intent-detection – Noordraven and Verschuere (2013) found longer reaction times and higher error rates for guilty suspects attempting to conceal knowledge of crime-relevant details.

The Guilty Knowledge Test hinges upon an increased response to one relevant item in a sea of irrelevant ones. This is a strategy which can easily be adapted to an eye-tracking paradigm. Although Verschuere et al. (2004) reported attentional impacts of guilty knowledge during the presentation of visual stimuli, there is currently no known research on using measures of visual attention to detect guilty knowledge. This gives us an opportunity to determine whether the orienting response, observed in several physiological measures, extends to the gaze.

Visual Attention

Rather than implementing one of the above GKT-based procedures to detect malintent, we have opted for a relatively unexplored procedure, eye-tracking to capture visual attention patterns, for five reasons. First, we wish to see to what extent the orienting response extends to gaze. Second, eye-tracking equipment can be unobtrusive, and stimuli presented to suspects without them knowing they are being tracked. Third, eye-tracking requires no contact with a suspect – no EEG cap, no blood pressure, respiratory, skin conductance, or pulse-detecting equipment required. Fourth, unlike traditional GKT procedures, eye tracking requires no questioning of the suspect, nor any response. They are only shown images. Finally, because it requires no response, it is language-neutral – a suspect need only have vision adequate enough to look at a screen.

Bottom-up and top-down. In order to create a GKT-based approach for detecting malintent with eye-tracking, we first must peruse the basic research on what influences visual attention. Research on visual attention has revealed two distinct mechanisms which guide the integration of visual information. The first, bottom-up processing, is involuntary, fast-acting, and is controlled entirely by characteristics of the stimulus. In a static image, salience can be accounted for quite well by considering three visual features: orientation, intensity, and color. Itti, Koch, and Niebur (1998) developed a model of biological vision, using these three feature channels, such that a saliency map (as articulated by Koch and Ullman, 1985) of an image (or video, as a series of images) can be created. The saliency map is a two-dimensional representation of the visual environment, in which each area is assigned a saliency -- the sum of the influence of these three bottom-up modalities. When freely viewing a scene (with no instructions), the location with the largest bottom-up saliency is attended to first, and once attended to, inhibited – resulting in a new highest-saliency location. Itti and Koch (2000) described an algorithm to model this saliency map, and tested it using both common psychological visual search stimuli (in which line segments differ in orientation and/or color) as well as complex high-definition nature scenes. In the nature scenes, the model successfully identified a camouflaged military vehicle, and did so in within simulated time scales similar to humans, given the same task. As the authors note, however, their algorithm of a human visual saliency map incorporates only bottom-up saliency, and is blind to top-down influences.

The second mechanism, top-down processing, is controlled by the viewer's motivation, expectation, experience, and/or task at hand (Isaacowitz, 2006). Viewers may select specific objects from a scene to attend to, regardless of their bottom-up saliency (Tipper, Weaver, Jerreat, & Burak, 1994). For example, even when the viewer is not task-motivated to attend to any one object in a scene, objects inconsistent with the context of an image (e.g. a microscope on a bar) receive more, longer fixations than objects that are consistent (e.g. a glass on a bar) (Henderson, Weeks, & Hollingworth, 1999).

Guided Search. Both bottom-up and top-down influences must be considered in a successful model of a visual search task. Visual search occurs when an individual seeks to identify a target item among a field of distractors – for example, attempting to locate a red pen in a junk drawer. Wolfe, Cave, and Franzel (1989) proposed the Guided Search (GS) model to predict which regions during a visual search would receive attention, and in what order. While the GS has since been revised several times (Wolfe & Gancarz, 1997; Wolfe, 1994, 2007), the component relevant to malintent detection remains the same: GS generally describes that top-down and bottom-up factors each contribute a weight to regions in a visual stimulus. Each region's summed weight determines the probability that the region will be attended to (first) (Wolfe, 1994). For example, the model would predict that for images lacking objects or real-world references (e.g., fractals in Parkhurst, Law, & Niebur, 2002), the bottom-up processes would contribute more strongly to the final weight and be the primary drivers of visual attention. Similarly, a red pen in a junk drawer full of pens would only receive increased attention if the searcher's task (a top-down influence) was to find a red object.

Eye tracking of human gaze has been successfully used as a measure of this final product of visual attention. When our attentional system shifts to a new visual target, gaze follows within milliseconds (Schall & Thompson, 1999). Some research has found that even covert targets of visual attention are followed by such a gaze shift (Henderson, 1992). Gaze may then be a valid measure of attention, whether one is attempting to mask the process or not, although other studies have found that attention can easily be directed to locations other than the current fixation (Driver, 2001).

These gaze patterns are measurable via eye-tracking technology (Henderson, 2003). Due to the connection between attention and gaze, human visual attention has been studied across multiple domains by way of a number of varied eye-tracking paradigms (Findlay & Gilchrist, 2003). While much visual attention research fixates on the bottom-up processes (Yantis, 2005), it is the top-down influences that have a particular relevance for any test of malintent.

While the Guided Search model was designed explicitly for a visual search context, which tasks an individual with finding a specific visual stimulus in a scene, several studies manipulating top-down factors are consistent with the GS's notion that attention is assigned to the largest sum of top-down and bottom-up weights.

Top-down influences. A diverse array of research on visual attention, most using some form of eye tracking, has shown that stimuli containing information that is novel, meaningful, adaptive, and relevant to a viewer (all top-down influences) attract more visual attention. One such top-down influence was demonstrated in a classic eye tracking study by Yarbus (1967). The patterns of fixations and saccades differed considerably when viewing a painting depending on the task the participant was given. For example, when asked to identify the ages of the people in the painting, the participant looked much more at the faces of the people, and less at the objects in the room, as compared to a free examination. Yarbus framed this as an information-yield issue: certain aspects of the photograph commanded greater visual attention because, due to the task at hand, they contained more relevant information for the viewer. Visual attention is not vulnerable only to explicit tasks, however. Existing mindsets, motivations, and attitudes also affect the context in which we process visual information. Research on these various top-down influences have found effects in numerous domains. Smokers deprived of nicotine attend more to smoking-related images (Kwak, Na, Kim, Kim, & Lee, 2007); English-speaking consumers pay more attention to advertisements on the left side of an array, due to left-to-right reading bias (Lam, Chau, & Wong, 2007); Expert chess players look more frequently at pieces involved in the optimal move than amateurs (Charness, Reingold, Pomplun, & Stampe, 2001).

The relationship between visual attention and a simultaneous interview may be relevant for the detection of malintent. Participants who were asked to describe scenes while under eye tracking fixated on an object shortly before verbalizing its name (Griffin & Bock, 2000). In a subsequent study, Griffin and Oppenheimer (2006) asked participants instead to give inaccurate names for the objects in the scene while describing them, and reported two important findings: First, participants fixated, as previously, on the object before naming them – even though the name was an incorrect one. Second, participants fixated for longer on the object before their inaccurate description than when they were not asked to lie.

Habituation. An important question in the effects of visual saliency on attention is whether, and how much, habituation occurs – whether salient objects command less attention over time. If habituation is extremely rapid, then any difference between malintent and innocent suspects may only appear very briefly, before being washed out by habituating to the stimulus.

Participants in one bottom-up habituation study (Donk & van Zoest, 2008) were presented with an array of line segments, all but two of which were oriented horizontally. These two exceptions were either offset by small or large amounts, producing low and high-salience segments, respectively. Participants were asked to locate and look at the most salient line segment as quickly as possible. They found a relationship between the latency of their first fixation (that is, how quickly their first fixation occurred after the stimulus onset) and their accuracy: the faster the fixation, the more likely they were to select the high-salience segment over the low-salience segment. Results were similar when they were asked to physically indicate which segment was most salient – short presentation durations resulted in more accurate identification of the high-salience segment than long presentation durations. They concluded that object salience may be sustained only for a brief period by the brain after perception.

Several follow-up studies conducted by Donk and Soesman (2010) revealed a similar effect, with stronger causal attribution to time. When stimulus presentation was very short, participants responded much more quickly to high-salience segments than low-salience once, and this effect shrunk to zero as stimulus presentation time increased. The effect contraction occurred on a very short time frame, under half a second in all cases. They replicated this under several bottom-up paradigms. Research on how adult top-down visual attention sources are weakened over time is scarce – it is unclear in what time scales would be needed to accomplish a similar attenuation for guilty-knowledge related stimuli. These results suggest that the stimuli presented in a malintent-detection paradigm should be high-frequency, and that long exposures to the same photo (e.g. "please look at this image for the next 10 minutes") may be an inefficient use of presentation time.

Emotion and phobia. As those suspected of harboring malintent (innocent or guilty) may also be experiencing strong emotions relevant to either their plot (if guilty) or to the suspicion placed upon them (if either), it is of interest to be aware of the potential effects of emotional influences on visual attention patterns.

While the emotion of the suspect is often the target of interrogators, Nummenmaa, Hyönä, and Calvo (2006) found that participants' visual attention was captured more quickly by emotional content, and that fixations were more frequent and longer-lasting on emotional pictures than unemotional ones. The bias towards emotional stimuli was present even when participants were explicitly instructed to attend to the neutral images. Their emotional pictures always contained humans, and the emotions were self-contained within the photographs (e.g. smiling baby; woman with knife).

Phobic individuals show unique visual attention patterns. Miltner, Krieschel, Hecht, Trippe, and Weiss (2004) tasked participants with locating a mushroom or a spider hidden in a visual display, and found that when spider-phobic individuals were instructed to locate the mushroom, the presence of the spider interfered with their attentional patterns. They oriented immediately to the spider (more often than non-phobic individuals) and as a result, were delayed in locating the mushroom. Whether this interference is of a bottom-up or a top-down nature, they demonstrated that individual differences in attitudes towards visual content can influence the visual search process.

In a similar study, Rinck and Becker (2006) showed photographs of animals (including spiders) to spider-fearful and non-fearful participants under eye tracking. They found that the spider-fearful participants, instead of avoiding looking at the spider, attended to it more quickly and for longer durations than the non-fearful. However, the attention quickly (at roughly the 1s mark) turned to avoidance, at which point the pattern reversed, and the spider-fearful spent less time looking at the spider photographs. This “attention-avoidance” pattern may provide a good framework for making predictions of guilty suspects’ visual attention patterns when viewing their own crime scene photographs, or when those with malintent view photographs of their intended targets.

Change-detection. Change detection, the ability to identify (usually subtle) differences between two visual scenes, may have utility for testing guilty knowledge. Stirk and Underwood (2007) compared top-down and bottom-up influences in a change-detection paradigm, and found that differences in bottom-up salience (as objectively measured by Itti & Koch, 2000) had no impact on change-detection performance. The top-down influence, as measured by thematically inconsistent vs. consistent objects in a scene, did: inconsistent object were detected more quickly, and more accurately than consistent ones. Thus, if stimuli presented during a malintent-detection procedure were inconsistent with either the scene or with the planned transgression, it could manifest in either all suspects, or only in suspects who *know* that it is inconsistent. For example, photographs of a bomb-making lab, after being searched, may reveal an open drawer with a circuit board – perhaps not inconsistent for a naïve viewer, but a recent user of the lab may think, “I thought I closed that drawer”, and thus attend to it more.

Forensic eye-tracking. The use of eye-tracking in lie-detection, malintent, and guilt-detection research is scarce. In some studies, the visual attention measures are secondary:- Bond (2008) utilized eye tracking technology in the investigation of lie-detection expertise. However, the study tracked the visual attention not of the targets, but of the (would-be) lie detectors. Two (presumed expert) detectors were selected on the basis of their high scores on lie-detection tasks, but their visual attention showed no consistent patterns; visual attention-based inferences about the sources of their skill could not be derived.

Recent forensic studies have explored pupil dilation (not a measure of visual attention, but rather, cognitive load) as a means to detect deception. Webb, Honts, Kircher, Bernhardt, and Cook (2009) measured pupil dilation during a CQT polygraph procedure. The CQT includes relevant questions, which are specific questions about the issue being investigated, and probable lie questions, which are vague and may span long periods of time. Pupil dilation was found to be related to deception, in that innocent individuals showed a pupil diameter change that was larger for probable lie questions than relevant questions, but there were no differences for guilty participants. While pupil diameter area was diagnostic for deception, it provided no additional explanatory power to the traditional polygraph measurements.

A similar study examined pupillary responses to (non-mock-crime) deception for both semantic and episodic memory (Dionisio, Granholm, Hillix, & Perrine, 2001). Their findings were consistent with Webb et al. (2009) – regardless of the memory type being tested, individuals asked to lie and provide confabulated details showed greater pupil diameter changes than truth-tellers.

Three recent studies have used the GKT or procedures similar to the GKT in combination with eye-tracking to attempt to identify deception or guilt. (Cook et al., 2012) utilized a mock-crime paradigm, after which participants responded to questions presented via computer, which were either related or unrelated to the crime. Eye-tracking data (pupillary responses, fixation count, and fixation duration) alongside with reaction times accurately predicted participants' guilt. In a similar study, Peth, Kim, and Gamer (2013) successfully discriminated between guilty and innocent mock crime participants using eye-tracking data (blinks, fixation count, and fixation duration) recorded during a computer-administered GKT presentation. However, they noted that the ocular measures were superfluous when recorded in concert with the traditional physiological measures present in a GKT (heart rate, electro-dermal, and respiratory measure). Using human faces as the stimuli presented under eye-tracking, (Schwedes & Wentura, 2012) found that participants had longer average fixation durations when attempting to conceal recognition of known faces (as compared to unknown faces). None of these studies used procedures in which the eye-tracking device was unobtrusive.

Self-presentation. The self-presentational perspective of deception describes lying as an attempt to influence the impression that others make of them (DePaulo et al., 2003). For example, an individual wishing to be known as intellectual and well-read may lie about recently (or ever) reading Sartre. In order to successfully create such an impression, one must lie convincingly. Thus inherent in most self-presentation is the motivation to appear truthful. In a malintent-detection context, then, we may expect individuals to adopt a similar impression-management strategy if they know they are being observed. One might expect that individuals who intend to commit bad acts make some attempt to appear as if they have only benign plans. Han Solo exemplified such behavior in *Return of the Jedi*, by instructing his co-pilot Chewbacca to "Keep your distance… but don't look like you're trying to keeping your distance… fly casual." For malintent participants, this would manifest in avoidance of objects relevant to their malintent.

Theoretical Convergence

When it comes to an existing theory supporting the detection of malintent with visual attention, there is substantial overlap between two presented here: the Guided Search model and the orienting response. Both predict that objects relevant to a crime should receive an increased response from individuals who are familiar with the crime. Indeed, both Yarbus (1967) and Lykken (1974) framed the issue similarly: relevant objects offer a larger information-yield, and thus evoke a larger attentional response from the viewer. This lays the groundwork for what is, essentially, a visual Guilty Knowledge Test for malintent. Such a test would involve presenting examinees with visual stimuli portraying persons, places, and/or objects relevant to a planned transgression. Stimuli regions that are malintent-relevant would have increased signal value for malintent examinees, and thus produce larger responses; these responses can be captured with eye tracking.

Eye-tracking could be used as a complement to existing credibility assessment procedures. In contrast to the polygraph and traditional interrogation procedures, eye-tracking has the advantage of enabling the collection of information from suspects remotely, without personal contact; compliantly, as minimal cooperation is required – they need only look at a screen; and covertly, without suspects knowing they are being monitored – possibly reducing or eliminating the effectiveness of countermeasures.

In an investigatory context, Granhag (2010) noted that identifying which individual(s) (from a pool of individuals) plans to commit a specific act requires some amount of evidence on the part of the investigators. This is true for the current paradigm, a visual test for malintent.

Even circumstantial or hearsay evidence could be sufficient to produce a set of stimuli for a test. While this act of collecting evidence is likely more effortful than simply asking questions, it also provides information regarding the prior probability of guilt.

Research Required

Two recently published reports have emphasized the urgent need for empirical research on malintent-detection tools. In an article in *Nature*, Weinberger (2010) profiled both the government's SPOT and FAST programs (described above), highlighting the lack of supporting evidence in both cases, and skepticism among scientists regarding their theoretical underpinnings.

In particular, TSA’s SPOT was the subject of a highly critical Government Accountability Office (GAO) report noting that as of March 2010, TSA has deployed 3,000 "Behavioral Detection Officers" to airports all over the country, with more on the way, at an estimated cost of $1.2 billion over the next 5 years. The GAO noted that the “TSA proceeded with deploying SPOT on a nationwide basis before determining whether the list of passenger behaviors and appearances underpinning the SPOT program were scientifically validated…” (p. 15). Beyond failing to validate that their behaviors were actually predictive of malintent, the report stated that, after seven years of SPOT operation, the GAO was “unable to use the SPOT referral data to assess whether any behavior… could be used to reliably predict the final outcome of an incident…”. Not only has SPOT not “ever resulted in the arrest of anyone who is a terrorist, or who was planning to engage in a terrorist-related activity” (p. 46), but the GAO identified twenty-three occasions on which an individual allegedly involved in terrorist plots passed through SPOT airports. None of these twenty-three occasions resulted in the detection of malintent or further investigatory attention.

As current malintent-detection processes are focused towards security screening (e.g. SPOT and FAST), a brief discussion of how classification accuracy and base rates of guilt can be context-sensitive is warranted. FAST, for example, claims a classification accuracy rate in the mid-70s (Weinberger, 2010). Malintent in a screening setting is a rare occurrence, especially for terrorism-related crimes – most travelers have no nefarious plans. Hartsfield-Jackson Atlanta International Airport alone processed almost 90 million passengers in 2010. If every passenger were subjected to FAST and 1% of those passengers (a very high estimate) had some form of malintent, nearly 27 million innocent passengers would be incorrectly flagged as potentially possessing malintent by FAST – a massive number to subject to further screening. Eye tracking as a Guilty Knowledge Test, as described here, is not (in the near-term) appropriate for screening contexts – primarily because some knowledge of the crime (in order to generate stimuli) is necessary. However, if this limitation was surmounted (as may be possible given future research), a false positive would need to be tremendously high in order to be logistically feasible.

The phrase "malintent detection" contains an implicit frame: that the priority is identifying individuals who possess malintent. While this is of course important, the accurate exclusion of malintent from innocents is also of great value. While we continue to frame the phenomenon as malintent detection when describing the studies below, we hope that the reader remains mindful of the importance of reducing suspicion upon the innocent as well.

The Present Studies

The following four studies investigated the extent to which visual attention, captured via unobtrusive eye tracking, could be used to identify malintent. Collectively, the studies had four primary goals. 1) Explore the effect of malintent on visual attention patterns; we predict that malintent will increase visual attention to items that are malintent-relevant when compared to a naïve sample, but that the effect will not extend to malintent-irrelevant items. 2) Examine the effect sizes of malintent (if any) on various measures of visual attention (e.g. mean fixation duration, fixation count) to identify where it is most pronounced. 3) Compare the effects of malintent and benign intent (a plan without the transgressive component) to determine whether the transgression aspect of malintent has additional influence on gaze patterns above a non-transgressive interaction intent; we predict small additional visual attention increases for malintent as compared to benign intent. 4) Test the moderating effect of participants' eye-tracking awareness on the effects of malintent; we expect awareness of eye-tracking and its associated incriminatory implications to attenuate the effect of malintent on their attention towards malintent-relevant objects, while awareness should have no effect on those with benign intent or no intent.

General Method

Overview

A series of four experiments utilized two mock crime paradigms to determine the effect of malintent on visual attention patterns. Unless described otherwise in their respective methods sections, these experiments employed the same stimuli, presented with the same eye-tracking equipment, and utilizing similar data analysis plans – each described in the summary below.

Stimuli and Coding

Stimuli consisted of 29 images and one 30-second silent video of the interior of a windowless office (approximately 4m x 3m). The office contained a filing cabinet, a copy machine, and a desk with a chair. On the desk were several office-typical items, such as a telephone, pencil sharpener, paperclips, a candy bar, and post-it notes. The objects were all consistent with what one might find in a university office. The images differed only in the location and angle from which they were taken; these ranged from close-ups of one half of the desk to wide shots of the entire desk and area surrounding it. No changes were made to the office or arrangement of items on the desk between photographs. Images were taken from heights of 4-6m, and contained at least some portion of the desk and its contents contained in the frame. The video began with camera entering the room, then slowly circling the desk, zooming in on the desk, panning out, then leaving the office – at all times, some part of the desk was in frame, and no humans or other objects not seen in at least one of the photographs was visible. Presenting a large number of photographs at a high frequency was chosen (as opposed to showing one photo of the desk for 30 seconds) in order to yield a potentially large number of orienting responses to a target object.

The contents of each photograph and each frame of the video were coded using SMI BeGaze 3.1 by a research assistant blind to the study design. For every object visible in a photograph or video frame, a polygon was drawn around the visible portion of that object (for example, see Figure 1). Each polygon was drawn to mimic the contours of the object; buffers of no larger than ~20px were allowed on any given side. When objects directly abutted each other, polygons were either drawn to be non-overlapping, or the polygon describing the object in the foreground received precedence by the software. Each of the 122,957 fixations recorded in these studies was tagged by BeGaze with the name of the polygon in which it fell.

Equipment and Software

Participants' visual attention patterns were recorded while they sat at a computer workstation equipped with an SMI IViewX RED eye-tracking system. The system operated at 60Hz, with spatial resolution of 0.1° and gaze position accuracy of 0.4°. Stimuli were presented at 1280 x 1024px on a 17” LCD monitor. The eye-tracking device consists of an unobtrusive black bar mounted underneath the computer screen. The eye-tracking operator was seated to the right and behind the participant (so as not to be in the line of sight), with the participant’s output mirrored on the administrative console.

Participants were calibrated on the eye-tracking device using varying cover stories; in all cases, calibration involved following a black dot on the screen to four or nine locations. The device provided feedback to the eye-tracking operator regarding the quality of the calibration. If participants’ first or second calibration was inadequate (as judged by the recorded fixation locations falling outside of the defined circle located on the screen on the screen), a second or third calibration was performed. Participants who failed calibration thrice were terminated and released from the study.

SMI’s Experiment Center was used to present the stimuli and on-screen instructions to participants. SMI’s BeGaze 3.1 was used to code the stimuli and collect the raw fixation data. Data cleaning and most analyses were conducted using R 2.11-2.15.3, with additional analyses performed in SPSS 18-20.

General Procedure

While two different paradigms were utilized for this series of experiments, the eye-tracking phase of the procedure contained a high level of commonality. After receiving some instructions and/or performing a task, participants sat in a chair in front of a screen equipped with the eye-tracking device. Participants had no physical interaction with the eye-tracking equipment; they only looked at the screen. After completing the calibration, participants were first shown the 29 photographs in a randomized order, then the 30-second video. After exposure to the stimuli, participants were asked to complete a questionnaire, after which they were debriefed.

Analysis Notes

Much of the sample data in the following studies possessed two characteristics that presented potential problems for proper parametric analysis and interpretation. First, several of the visual attention measures were non-normally distributed. Normality was checked by calculating the *Z*-score of the skew and kurtosis for each of an experiments' cells and comparing them against a threshold value of $Z \pm 3$. Variables exceeding the threshold were considered to be significantly skewed and/or kurtotic – an issue of mild concern, given that some groups were below or at the size threshold for ensuring that assumption of a normal sampling distribution was met. The second issue was significantly heterogeneous variance between experimental conditions. Homogeneity of variance was tested with Levene's test, and several measures in each of the following studies exhibited statistically significant differences in variance between cells. These issues can both result in inflated type I errors as well as poor estimates of effect sizes.

In order to minimize any interpretive errors resulting from the shape of the sample distributions, all relevant parametric analyses (predominantly ANOVA) were accompanied by condition-stratified bootstrapping analyses which generated hypothesis-relevant pairwise comparisons of cell means, their associated bias-corrected and accelerated confidence intervals (using the Efron, 1987 method), and effect sizes. For example, a custom Cohen's *d* function was written to use with the *boot* package in R to generate bootstrapped effect sizes for pairwise comparisons. These bootstrap-generated results are preferable to both the omnibus F-tests and traditional post-hoc comparisons, which can be unreliable in cases where there are assumption violations, as observed in these data. For these reasons, while hull hypothesis significance testing

(NHST) is performed and reported throughout (generally in the form of an F-ratio and associated p-value), bootstrapped results, including effect sizes, are given preferential interpretive attention throughout (Carver, 1978, 1993).

These data followed a nested pattern: fixations within images within participants within experimental conditions. However, no 3- or 4-level hierarchical analyses are reported here. Rather, the data have been aggregated at the subject level, to retain the focus on the experimental condition-level differences, and allow for subject-based classification procedures. One example of the aggregation process follows in the measures section below.

Visual Attention Measures

Several variables intended to capture visual attention were calculated from the fixation data and reported in these experiments. One such measure, *fixation count*, is described here to outline the aggregation process, which was similar for each visual attention measure. *Fixation count* describes to the number of times a participant looked at an object (as defined by the polygon coding process), across the set of stimuli. To control for individual-level differences, and to correct for participants for whom only partial data was collected, fixation count is analyzed and reported as a percentage of each participant's total number of fixations. For example, eye-tracking recorded 541 fixations for subject #1 in experiment 1, 21 of which were looking at the post-it notes, for a fixation count of 3.9%. Condition means and standard deviations were then calculated from these subject-level values.

Study 1

Using an eye-tracking paradigm, Study 1 served as a proof of concept – successfully identifying unique patterns of visual attention in individuals with malintent (as compared to oblivious individuals) when they were shown images of the (future) crime scene. Malintent participants were told they would poison one of two objects on a desk; others were innocent and unaware of any planned transgression. Before malintent participants went to execute the task, their visual attention was recorded via covert eye-tracking equipment. According to both the Guided Search model and the orienting response, the object which malintent individuals planned to interact with should receive higher visual saliency. Thus, we predicted that malintent participants would look at these objects made more relevant due to their planned crime more often, and for longer, than innocent participants.

Method

Participants and Design

Sixty-one undergraduate students participated in exchange for class credit, and were randomly assigned to one of three conditions: malintent (key), malintent (post-it), or innocent. Eleven participants were excluded due to poor calibration on the eye-tracking device. The final sample thus consisted of N = 50 participants: 18 malintent (key), 15 malintent (post-it), and 17 innocent.

Procedure

After providing informed consent, participants sat at a computer workstation equipped with an SMI IViewX RED eye-tracking system. Participants were not informed that their eye movements would be recorded.

Malintent conditions. Participants in malintent conditions were instructed to imagine that they had been recruited as a secret agent to undertake a mission for the "Organization." Participants were then told that their identity would have to be confirmed with an iris scan, which served as a cover for the eye-tracking calibration procedure. After calibration, participants were told that an "enemy agent" had been located in the vicinity and that their mission was to eliminate that agent. Before proceeding, they were asked to confirm their willingness to undertake the mission; all participants did so.

Participants were instructed that the enemy agent was posing as university professor and that they were to sneak into his empty office, place a contact poison on an item on the desk, and then exit without leaving a trace of evidence. Participants were told that the target need only touch the poisoned object to be affected. The Organization's analysts then recommended the object to poison. *Malintent (post-it)* participants were told to poison a post-it note pad; *malintent (key)* participants were told to poison a key. Randomly assigning malintent participants to poison one of these two specific objects was performed for several important reasons. First, two objects (rather than one or three) were chosen in order to compare the gaze patterns of the malintent conditions – it is of interest to determine how much consistency (if any) exists across objects made salient by malintent. Second, for stimulus generalizability, it is important to ensure that the visual attention patterns of the malintent were not unique to a specific object. Finally, the key and post-it were selected due to their relative small size on the screen, which would make any additional attention given to them more detectable.

Participants were told that they would be shown photographs of the target's office so that they could become familiar with the site in anticipation of their mission. They were then shown photographs followed by a video recording (see below) on the LCD monitor during which their eye movements were covertly recorded. Afterward, participants were informed that their mission would not be completed. They were then debriefed and thanked for their participation.

Innocent condition. After providing informed consent, participants in the innocent condition were calibrated with the same cover story given to malintent participants – that it was necessary to confirm their identity before proceeding. However, there was no mention of an Organization or secret agents. Once calibrated, participants were asked to view the photographs, followed by the video, while their eye movements were recorded. They were merely instructed that they would be asked questions about the stimuli afterwards. Participants were then debriefed and thanked.

Stimuli

Twenty-nine photographs of an office were presented for 2 seconds each, in randomized order, followed by one 30-second video of the same office. The video recording portrayed a slow "scan" of the room and its contents, at times zooming in on portions of the desk, which was central for most of the 30 seconds.

Dependent Measures

Four measures of visual attention were constructed for each of the 2 target objects, yielding eight dependent variables. These DVs all involved fixations – the participant halting and maintaining their gaze on a single location for longer than 99ms. Saccades (the rapid movement of the eyes between fixations) were not analyzed, as neither the Guided Search model nor the orienting response offer guidance in making saccade-based predictions (except, of course, that early saccades should be directed toward the target object, a behavior captured equally well by the fixation at the saccade's terminus). These DVs were chosen because this is an exploratory study, and they were the most intuitive and easily measured fixation-based measures. Additionally, neither theory nor known existing eye-tracking research provides any insight as to which would be most likely to demonstrate malintent-related patterns.

Fixation count refers to the number of times a participant looks at an object. To control for individual-level differences, and to correct for participants for whom only partial data was collected, fixation count is expressed as a percentage of each participant's total number of fixations.

Fixation duration mean, or dwell time, measures how long (in ms) the average fixation on an object lasts. Fixation duration mean is independent of fixation count, as a participant may look at an object infrequently but dwell on it for long periods of time when they do so (or vice-versa).

Fixation duration variance, derived from each participant's distribution of fixation durations on an object, reflects an individual's consistency in fixation duration. High fixation duration variances would indicate a pattern of both very brief and very long fixations, rather than fixations that are constant in duration – a difference that may not be reflected in the fixation duration mean. It is possible that malintent produces such a pattern: many quick automatic glances combined with longer periods of visual study.

Fixation time indicates the percentage of a participant's total fixation time spent looking at a given object. Fixation time was constructed by summing the duration of fixations for each object then dividing by the total fixation time across all fixations. Although fixation count, duration mean, and duration variance are (potentially) independent, fixation time is by definition covariant with the others (e.g., an increase in count or duration mean would also increase fixation time).

Results

Group Comparisons

To test for the hypothesized effects of malintent on visual attention patterns, we performed one-way ANOVAs on each of the eight variables described above (derived from the 24,062 total fixations recorded), with malintent (key), and malintent (post-it), and innocent as between-subjects experimental conditions. For each, we report both the omnibus F-test, as well as comparisons between each malintent condition and the innocent control. We predicted that on measures involving key fixations, the malintent (key) condition but not the malintent (post-it) condition would yield larger values than the innocent control. Similarly, for measures involving post-it fixations, we predicted that only the malintent (post-it) condition would score higher than the innocent control. In seven of the eight comparisons below, the results demonstrated statistically significant, large effect sizes in the hypothesized direction (see Table 1).

Fixation count (key). There was a significant effect of malintent on the number of fixations on the key, $F(2{,}47) = 12.57, p < .001, \eta^2 = .348$. Participants who planned to poison the key had a larger percentage of their fixations ($M = 1.7\%, SD = 1.3$) fall on the key than the innocent participants ($M = 0.6\%, SD = 0.3$), $t(33) = 3.53, p = .001, d = 1.12$. Those who planned to poison the post-it, however, did not look at the key more often ($M = 0.4\%, SD = 0.5$) than innocent participants, $t(30) = -1.40, p = .171, d = -0.19$.

Fixation count (post-it). There was also a significant effect of malintent on number of fixations on the post-it, $F(2{,}47) = 8.12, p < .001, \eta^2 = .257$. Showing the opposite pattern of key fixations, malintent (post-it) participants looked at the post-it significantly more frequently ($M = 5.5\%, SD = 3.8$) than innocent participants did ($M = 2.8\%, SD = 1.1$),

$t(30) = 2.79, p = .009, d = 1.09$. Malintent (key) participants did not look at the post-it more often ($M = 2.7\%, SD = 0.6$) than innocent participants, $t(24.7) = -0.32, p = .754, d = -0.11$.

Fixation duration mean (key). The malintent manipulation had a significant effect on the length of fixations on the key, $F(2,47) = 7.95, p = .001, \eta^2 = .253$. Participants who planned to poison the key ($M = 277ms, SD = 68$) did not have significantly longer key fixations than innocents ($M = 227ms, SD = 101$), though the effect size was large, $t(33) = 1.73, p = .093, d = 0.71$. Malintent (post-it) participant fixations on the key were significantly shorter ($M = 141ms, SD = 122$) than those in the innocent condition, $t(30) = -2.17, p = .038, d = -0.81$.

Fixation duration mean (post-it). There was no significant omnibus effect for malintent on the length of post-it fixations, $F(2,47) = 0.91, p = .441, \eta^2 = .037$. Malintent (key) participants ($M = 254ms, SD = 52$) did not look at the post-it for longer than innocent participants ($M = 256ms, SD = 70$), $t(33) = -0.11, p = .911, d = -0.05$. Nor did malintent (post-it) participants ($M = 289ms, SD = 114$) look longer, $t(30) = 0.98, p = .336, d = 0.32$.

Fixation duration variance (key). There was a significant effect of malintent on the standard deviation of fixation durations on the key, $F(2,47) = 7.84, p = .001, \eta^2 = .250$. The average standard deviation for malintent (key) participants was significantly larger ($M = 140, SD = 59$) than for innocent participants ($M = 71, SD = 56$), $t(33) = 3.56, p = .001, d = 1.31$. There was no significant difference between malintent (post-it) ($M = 51, SD = 89$) and innocent participants, $t(30) = -0.74, p = .466, d = -0.36$.

Fixation duration variance (post-it). There was a significant effect of malintent on the standard deviation of fixation durations on the post-it, $F(2,47) = 3.59, p = .036, \eta^2 = .132$. Specific comparisons showed that malintent (post-it) participants showed more variance in their post-it fixation durations ($M = 176, SD = 92$) than innocents ($M = 121, SD = 56$), $t(30) = 2.08, p = .046, d = 0.76$. In contrast, malintent (key) participants ($M = 121, SD = 49$) did not differ from innocents on this measure, $t(33) = 0.01, p = .991, d = 0.01$.

Fixation time (key). There was a significant effect of malintent on the percentage of total fixation time spent dwelling on the key, $F(2,47) = 13.64, p < .001, \eta^2 = .367$ (see Figure 2). Participants who planned on poisoning the key spent more total fixation time ($M = 2.0\%, SD = 0.3$) looking at the key than innocent participants did ($M = 0.6\%, SD = 0.3$), $t(17.8) = 3.78, p = .001, d = 1.44$. Malintent (post-it) participants ($M = 0.4\%, SD = 0.4$) did not significantly differ from innocent participants, $t(30) = -1.53, p = .137, d = -0.60$.

Fixation time (post-it). The malintent manipulation had a significant effect on time spent fixating on the post-it (see Figure 2), $F(2,47) = 8.01, p = .001, \eta^2 = .254$. Post-it poisoners spent more than twice as much time ($M = 6.4\%, SD = 5.1$) fixating on the post-it than did innocent participants ($M = 2.9\%, SD = 1.3$), $t(30) = 2.72, p = .011, d = 1.07$. Key-poisoners ($M = 2.7\%, SD = 0.6$) did not significantly differ from innocents on this measure, $t(22.9) = -0.64, p = .538, d = -0.24$.

As predicted, malintent increased visual attention for the transgression-relevant object, but not for the neutral paired object. The largest effects were observed in the composite variable, percentage of fixation time spent looking at the object (average d = 1.26), and the weakest were observed in the average fixation duration (average d = 0.52). There was also remarkable consistency in the malintent effect (as compared to the innocent group) in each of the four dependent measures – effect sizes were highly similar across the two target objects (see Table 1). This is encouraging, given the wide variety of potential stimulus characteristics (e.g. size, color, shape) which are known to influence bottom-up saliency. This suggests that the effect of malintent on these four dependent measures could be relatively independent of these forms of stimulus characteristics.

Other objects. The four visual attention measures (fixation count, duration mean, duration variance, and time) were also calculated for each of the thirteen remaining objects visible on the desk. None of these objects were made salient to either malintent condition. Fifty-two ANOVAs (13 objects x 4 variables) and one hundred and four planned contrasts (2 for each omnibus ANOVA comparing each malintent condition to the innocent condition) were performed to confirm the uniqueness of the visual attention patterns produced by malintent. As expected, zero of the one hundred and four planned contrasts showed a malintent condition with significantly greater visual attention than the innocent control[1].

Principal Components Analysis

While there were medium to very large effects observed for each of the measures of visual attention described above, it is important to determine how much redundancy was present in those variables. Given that fixation time is a composite of fixation duration and fixation count, we expect at least a moderate relationship between the two. We performed a Principal Components Analysis (PCA) on the eight measures of visual attention to determine the underlying structure, if any. The sample of N = 50 was just adequate for a PCA of this nature, KMO = .587. Bartlett's Test indicated that the variables were related enough to justify extracting

[1] No correction for Type I error was made for these 104 tests. A liberal α of .05 was intentionally chosen, to demonstrate the very specific scope of the effects of the malintent manipulation. The five Type-I error instances one might expect to occur from chance alone were not observed.

components, $\chi^2(28) = 461.2, p < .001$. We used Kaiser's rule for factor retention: all components with an eigenvalue of over 1 were retained. This resulted in extracting two components, which accounted for 76.3% of the variance of the imputed variables. Varimax-rotated factor loadings describe a clear pattern (see Table 2): the four variables describing post-it attention all load highly on component 1, and the four variables describing key attention all load highly on component 2.

Classification

A linear discriminant analysis was conducted to determine how well measures of visual attention can classify participants as possessing malintent, or innocent. The model included all four variable-pairs analyzed previously (fixation count, duration mean, duration variance, and time, for both objects). For a three-group classification model, chance rate of classification is 33%, though in the analysis, prior probabilities were not assumed to be equal. The model significantly discriminated between conditions, $Wilks'\lambda = .326, p < .001$, and produced an 80% overall classification accuracy. False positives were rare (6%); innocents were correctly classified 94% of the time. Malintent participants were accurately classified 78% and 67% of the time, when intending to poison the key and post-it, respectively. Seven of the eight dependent variables (all but fixation duration on the post-it) contributed significantly to the discriminant function, $Wilks'\lambda s = .633 - .868$.

Discussion

Study 1 provided strong initial evidence for the theory that malintent increases visual attention for malintent-relevant items, and the possibility of successful malintent detection. Induced malintent, even within a role-playing laboratory procedure, produced considerable increases in visual attention directed at objects relevant to the malintent, when compared to an innocent control group: Those with malintent looked at malintent-relevant salient more often, their fixations on it were longer and had larger variances, and they spent more time overall looking at the target object. Consequently, 80% of all participants were classified correctly, with accuracy rates of 73% (sensitivity) in the malintent conditions and 94% (specificity) in the innocent condition. Given the short stimuli exposure period and the lack of incentive, the low rate of false positives (6%) observed are encouraging - the identification of innocence being an important objective. Additionally, Study 1 directed participants specifically which object to interact with as part of their malintent task, rather than allowing them to choose.

Several further theoretical and practical questions issues remain to be addressed. One concern is the extent to which benign intent would have the same effect as malintent – that is, how much of the malintent effect observed here can be attributed to the transgressive nature of the act, and how much to the plan to simply interact with it? Also relevant is the question of whether awareness of the eye-tracking procedure affects visual attention patterns. Knowing that attending to relevant regions in a stimulus array could reveal their intentions, examinees may exhibit a pattern similar to that observed in phobic individuals. Studies 2, 3, and 4 were designed to replicate and extend these initial findings.

Study 2

Study 1 served as a proof-of-concept that malintent can influence visual attention. Study 2 aimed to replicate Study 1 with a small number of critical changes, to make classification accuracy more challenging (and more externally valid) and to explore participants' insight into their own gaze patterns.

Post-hoc classification accuracy when the two options are known (as was the case in Study 1) showed that, for the two target objects selected as options in the protocol, the effect sizes seemed to be stable across targets. Study 2 investigated not just the stability of the effect size of malintent across a much larger pool of objects, but how well malintent can be classified when all options are on the table (if you will) – when any one of the objects in the scene may be malintent-relevant (not just two). In terms of the Guided Search model, Study 2 extended the number of potential target objects, and thus the range of bottom-up visual saliency baselines. This larger range allows for determining whether the top-down influences of malintent add the same attentional weights regardless of the bottom-up strength.

It is also important to determine how much of the increased visual attention resulting from malintent is the results of a conscious process or an unconscious one. That is, do participants know that they are spending two to three times as much time fixating on the objects they intend to interact with? If entirely conscious, for example, individuals with malintent may be aware that their visual attention patterns are incriminatory, and countermeasures would be easier to deploy. In this study, we recorded participants' retrospective evaluation of what they looked at most, in order to compare it to both their targets (for malintent participants) and the objects they actually looked at most.

Method

Participants and Design

One hundred and sixty two undergraduate students were recruited for Study 2. Ten participants were dropped from the sample due to one or more of the following issues: failure to calibrate properly on the eye-tracking device, a lack of fixation data from being outside of the detection zone post-calibration, or failure to complete the questionnaire, resulting in a final sample of N = 152, randomly, but unequally, assigned to one of two conditions: malintent (n = 119, or innocent (n = 33).

Procedure

The procedure for Study 2 was identical to that of Study 1, with one change in the instruction phase for malintent participants. Instead of being told that they should poison a specific object, malintent participants were shown a list of every object displayed in the stimuli, and asked to pick one to poison before continuing. Participants were given as much time as they needed to peruse the list and make a selection before experiment continued.

After viewing the 90 seconds of photographs and video, participants were asked to report how anxious they were (on a percentage scale, 0-100%) while looking at the stimuli, and to name the object they looked at the most: “what part of the photographs and video did you spend the most time looking at?”.

Results

Analysis Notes

In Study 1, fixation time percentage yielded the largest effect size for malintent, and was an effective composite of two other visual attention measures (fixation duration and fixation count). Thus, in both the following group mean comparisons and linear discriminant analysis, we chose to focus on fixation time percentage as our measure of visual attention.

Object selection

Of the ten options presented to participants at the outset, 9 were chosen as poison targets by at least two participants. The most commonly selected object was the phone (N = 42), followed by the coffee mug (N = 17), key (N = 17), and water bottle (N = 16). See Table 3 for object selection frequencies. These four objects yielded a large enough n to justify mean comparisons with the oblivious reference group.

Although the baseline of visual attention (as measured by the oblivious participants' percentage of time spent fixating on the object) for each of the objects varied considerably (M = 2% for the key, M = 10% for the phone), the effect size of malintent's influence on visual attention was consistently large, as described in detail for the following four high-n objects.

Fixation Time Percentage for Top Four Objects

Participants who chose to poison the phone gazed at it about twice as often (M = 18.21, SD = 11.82) as both oblivious participants (M = 10.47, SD = 2.58), $t(45.9) = 4.13, p < .001, M_{diff} = 7.75$, 95% CI $[4.13, 11.72], d = 0.88$ (see Table 4), and the non-phone-poisoning malintent participants (M = 8.96, SD = 3.54), $t(45.1) = 4.95, p < .001, d = 1.25$.

Key-poisoners looked at the key ($M = 6.18, SD = 4.13$) more than three times as much as the oblivious reference group ($M = 2.08, SD = 1.47$), $t(15.6) = 3.80, p = .002, M_{diff} = 4.17$, 95% CI $[2.28, 6.34], d = 1.69$. The effect was even larger when comparing the key-poisoners to those who planned to poison other objects ($M = 1.87, SD = 1.33$), $t(14.4) = 4.01, p = .001, d = 2.34$.

Mug-poisoning participants' (M = 18.28, SD = 13.30) mug-related visual attention was significantly greater than both the oblivious participants (M = 5.56, SD = 2.59), $t(15.6) = 3.79, p = .002, M_{diff} = 12.72, 95\%$ CI $[6.87, 19.27], d = 1.68$, and the non-mug-poisoning malintent participants (M = 5.52, SD = 3.49), $t(15.3) = 3.82, p = .002, d = 2.25$.

Finally, participants who planned to poison the water bottle (M = 15.88, SD = 9.84) looked at it significantly more than the oblivious participants (M = 8.08, SD = 3.53), $t(16.9) = 3.08, p = .007, M_{diff} = 7.80, [3.19, 12.68], d = 1.29$, as well as the other malintent participants (M = 8.20, SD = 3.42), $t(15.6) = 3.09, p = .007, d = 1.67$.

Classification

We conducted a linear discriminant analysis (LDA), with leave-one-out cross-validation to attempt to classify participants as a member of one of ten groups: innocent, and nine malintent groups – reflecting the nine objects which were selected at least twice by malintent participants. Classifiers variables were the percentages of time spent on each of the objects on the desk.

For ten categories, the expected chance rate of prediction is 10.0%. The LDA produced a classification accuracy five times as high, at 48.3% correct classification, significantly discriminating between the various malintent targets and innocence, $Wilks'\lambda = .078, p < .001$. As shown in Table 3, classification accuracy varied substantially between categories. The highest accuracies were observed in the largest groups, with 69.7% of the n = 33 innocent participants correctly classified, and 54.8% of the n = 42 phone-poisoners correctly classified. The poorest classification rates were found for the n = 6 post-it poisoners, n = 2 sharpener poisoners, and n = 7 candy poisoners, each yielding a 0% accuracy rate. As expected, we observed a moderately strong correlation between condition classification accuracy and condition size, $r(8) =$

$.442, p = .201$; the objects which few participants selected gave the discriminant functions little data to work with, making classification for those objects difficult.

The above LDA was conducted with prior probabilities of class membership equal to the sample distribution, which, for a phenomenon as rare as malintent, is unrealistic. We conducted a second LDA, using the same variables, but with altered prior probabilities of group membership. We assumed the innocent group to be 1000 times as likely to occur as each of the malintent groups (arguably still a vast underestimate), rather than have likelihood driven by sample distributions. Malintent groups were given equal weights to each other. Here, the relevant chance rate of classification reference is 22.1% (33/149) – reflecting the overall accuracy in this sample if we were to simply label everyone as innocent. With the increased innocent prior probabilities, the overall correct classification rates was lower, as expected, at 30.9% correct, but still an improvement over the relevant chance rate.

Participant Gaze Insight

Three variables were relevant to determining the extent to which participants knew which object captured the plurality of their visual attention: 1) their self-report of which object they spent the most time looking at; 2) the object they actually spent the most time looking at, as captured by the eye-tracking equipment; and 3) the object that they, at the outset, choose to poison[2]. These three variables could be entirely independent – for example, a maximally incongruous participant may have first chosen to poison the telephone, then gazed predominantly at the water bottle, and finally reported that he thought he had looked at stapler more than anything else. We calculated the percentage agreement for each of the three pairs of these

[2] This measure is only relevant for those in the malintent condition. Innocent/oblivious participants were ignored in comparisons made with this variable.

variables as well as the combination of all three, and reported the agreement percentage with its associated bootstrapped 95% confidence interval.

Fewer than half of all participants correctly reported the object that they spent the most time looking at, $M = 40.7\%, 95\%\ \mathrm{CI}\ [32.0, 47.3]$. The remaining comparisons are only relevant for malintent participants, as innocent participants did not select an object to poison at the outset.

Malintent participants were highly consistent between the object they selected to poison at the outset and the object they reported having spent the most time looking at, agreeing over two-thirds of the time, $M = 69.3\%, 95\%\ \mathrm{CI}\ [59.5, 76.1]$. The target they chose, however, only received the highest amount of fixation attention half of the time, $M = 49.7\%, 95\%\ \mathrm{CI}\ [40.2, 57.3]$. Of the malintent participants, 41% were consistent between the three variables – that is, they correctly self-reported that they spent the most time looking at the same object that they had chosen to poison at the beginning of the study, $M = 41.0\%, 95\%\ \mathrm{CI}\ [31.6, 49.6]$.

Discussion

The effect of malintent on visual attention patterns shown in Study 1 was replicated here, with large effect sizes observed when comparing the percentage of fixation time spent on a particular object between innocent viewers and those who intended to poison the object. Malintent participants, who all chose their own target object, looked at their chosen object two to three times more often than oblivious participants did, regardless of the base rate of attention.

Perhaps less promising, as shown by the discriminant analysis, is the idea that investigators with little knowledge would be able to tell with high accuracy rates which object is relevant to a participant's malintent – or if they indeed have any. Classification accuracies of 30-

50%, while still a vast improvement over the relevant baselines, are an indication that the increased visual attention from malintent was in several cases lost in each individual's unique and varied gaze patterns. The inaccuracy of this sort of attempt to classify an individual as innocent or malintent *generally* implies that procedures similar to this may not be suitable as a screening tool. That is, when investigators know little about a potential plot, or have gathered no usable information from a suspect, the risk of misclassification could be quite high.

In this study, participants were asked immediately after their eye-tracking experience to recall what they looked at most – and 60% of them failed to do so correctly. It seems that, for malintent participants at least, they were not relying on their memory of the stimulus viewing experience, but rather making a guess based on what objects were more important to them – and thus naming the object they chose to poison rather than the object they looked at most. They were often incorrect – only 50% of the time did they actually spend the most time on their selected object; a number not coincidentally similar to the overall classification accuracy observed.

Study 3

Studies 1 and 2 established that malintent does indeed produce detectable patterns of visual attention that strikingly differ from innocents. The effect was strong and observed across several measures of visual attention. The experimental conditions of Studies 1 and 2 arguably reflect the two most relevant groups from an external validity perspective: suspects possessing malintent (whom investigators are interested in identifying) and oblivious innocents who know nothing (whom can be easily identified and used as a reference group both in studies as well as in a practical setting). However, these studies have not properly isolated *malintent* from a more generic intention to interact with an object – which may involve no malicious intent whatsoever. Thus, Studies 3 introduced a third condition within the guilt manipulation: benign intent. Benign intent individuals represent innocent suspects who find regions of the visual stimuli to be more salient than the innocent controls, but *not* because they intend to commit a bad act. For example, a garbage can may attract additional attention from both one who intends to plant a bomb in it (malintent) as well as someone who intends to salvage bottles for the recycling deposit (benign intent).

Study 3 introduced an additional manipulation of participant awareness of the eye-tracking device. In Studies 1 and 2, participants were naïve to the presence of the eye-tracking equipment, and thus did not have the opportunity or motive to employ any form of reactivity or countermeasures. As the top-down portion of our visual attention is to some extent under conscious control, it is crucial to determine whether simply knowing that one's visual attention is being monitored will influence gaze behavior.

Study 3 had three goals. First, Study 3 examined whether there are differences in attention patterns between participants with malintent and benign intent. It is both theoretically and practically of interest to determine if there is a difference in visual attention patterns based on the criminality of the task. Theoretically, the Guided Search model only states that top-down influences increase saliency – applying to both malintent and benign intent, each with some unknown weight. Here, we seek identify whether the transgressive nature of the act results in additional weight beyond merely intending to interact with an object. Practically, the benign intent and oblivious conditions allow us to model two crucial real-world comparisons with the malintent group. The benign intent participant represents the worst-case scenario for an innocent: they are planning to perform a task which overlaps in many ways with what a malintent individual intends. The oblivious participant reflects the more-common event in which a suspect knows nothing at all, and has no malintent-overlapping plans. Importantly for potential future applications of this procedure, oblivious participants are easily acquired by any investigative body: members of the general population (non-suspects) who are unfamiliar with the crime scene.

Second, this study investigated the effect of participants' awareness of the presence of eye-tracking equipment. Under the Guided Search model, awareness of the eye-tracking equipment should produce an additional top-down influence for malintent individuals only: the motivation to avoid the object in order to mask their malintent. While GS implies that the weights producing final visual attention would be positive, it does not explicitly preclude the possibility of negative weights, which we expect to observe for malintent participants. Orienting response theory, meanwhile, with a focus on automatic responses, would predict that malintent individuals would still exhibit an increased and immediate response to the relevant stimulus regions. The practicality of using visual attention measures to detect malintent depends on whether awareness of the process, which may provoke countermeasures, moderates the effect size(s).

Finally, Study 3 expanded the crime-generalizability of the malintent effect by varying the nature of the task – this time, with a considerably more mundane crime than the murder modeled in Studies 1 and 2: theft.

Method

Participants and Design

Undergraduate participants (N = 171) were randomly assigned to one of six cells in a 3 (guilt: malintent, benign intent, or innocent) x 2 (eye-tracking awareness: aware vs. unaware) between-subjects design. Participants were young ($M = 19.7, SD = 3.2$), and predominantly female (56.7%, n = 97).

Procedure

After providing informed consent, participants were told that their assistance was needed to help test a new security system, and given one of three roles to play (defining the malintent, benign intent, and oblivious conditions). They were told that a new security surveillance system was being installed at the university, and that their help was needed to test it. They were given no details regarding the nature or scope of the security system, except that there were security agents deployed around campus, and that if they were approached by anyone identifying themselves as from security, they should comply with their requests.

Malintent participants were asked to commit a mock theft. Their instructions were to find and enter an unlocked office, take a key from on top of a desk, and use it to open a safe containing $50 – which they were told they could keep if they successfully complete the task.

Additionally, they were told that, if apprehended at any point, they should simply say they were asked to retrieve the key and to deny any involvement in a crime (though a cover story was not strictly necessary, participants were not questioned).

Benign intent participants were told that they would run an errand, which was a task similar to the malintent participants – but with the transgression framing removed. They were asked to go to an office, open a safe, and return with an envelope to give to the researcher. No mention of incentives for not being caught, theft, nor a safe were made. Both the benign intent and malintent participants were shown a map to the office along with a photocopied image of the key that they needed to complete their task.

Oblivious participants were not made aware of the existence of a key, or any crime, and were asked to simply walk to a specific location and wait.

After the participants had read the written instructions, the researcher delivered the instructions verbally to ensure comprehension. Participants then departed to ostensibly complete their task. Before reaching their destination, they were intercepted by a security officer (a researcher blind to malintent condition), who asked them to accompany them to a nearby room (the room containing the eye-tracking equipment).

The security officer informed the participant that they had reason to suspect that a crime was currently in progress and had some questions for the participant, but before asking them, the participant should turn their attention to the screen in front of them. The computer (with the eye-tracking equipment) displayed a series of instructions. The first page reminded participants that security was suspicious that a crime was going to occur, and also informed them that they would be interrogated to find out what they knew about the crime (no interrogation occurred).

Participants were then given a “vision test”, and told to follow the dot on the screen – actually the eye-tracking calibration procedure. After calibration was complete, participants were told that they would be shown photographs from the suspected crime scene. Additionally, participants in the *aware* condition were told that, “using eye-tracking technology, we will be monitoring and recording what you look at in the photographs. What you look at may determine whether greater suspicion is placed upon you.” Participants in the *unaware* condition received no such information.

Participants were then shown the same stimuli as used in Studies 1 and 2 while their visual attention patterns were recorded. Twenty-nine photographs of an office featuring a desk with several objects on it were displayed under for 3 seconds each (order randomized for each participant) followed by a 30-second video of the same office.

The photographs were not explicitly identified as being the office which the malintent and benign intent individuals were tasked to enter.

After being exposed to the stimuli, participants were told that their task was complete and asked to fill out a questionnaire. The questionnaire included manipulation checks regarding task (“what were you asked to do today?”) and eye-tracking awareness (“Were you aware that your eye movements were being tracked while you were at the computer?”), as well as self-report measures of stress (“How stressed were you?”), retroactive confidence that they could have completed the task (“Immediately after the instructions, how confident were you that you would be able to complete your assigned task?”), and task-completion motivation (“How motivated were you to complete your task today?”). The stress, confidence, and motivation items were requested as a percentage, with two anchors given: 0% meaning not at all stressed/confident/motivated, and 100% meaning completely stressed/confident/motivated.

Additionally, participants were asked whether and how any awareness of eye-tracking changed how they looked at the photographs. After completing the questionnaire, participants were debriefed.

Measures

From the 46,644 fixations recorded during the stimulus-presentation stage in the study (i.e. not during calibration or instruction-presentation phases), several measures of visual attention were calculated for each participant. These measures included those previously introduced in Studies 1 and 2: fixation count (percentage), fixation time (percentage), fixation duration mean; two new visual attention measures were added: fixation distance and fixation latency.

Fixation distance is a measure of how close each participant's fixations were to the key. Fixation distance was calculated as the number of pixels between the Cartesian coordinates of the centroid of the key (in each image) and the location of the fixation. Larger average distances indicate fewer fixations near the object, and could indicate intentional avoidance of the key, as would be expected when participants are aware of the eye-tracking device. If negative visual attention weights are possible, the Guided Search model might predict that malintent participants would produce larger average fixation distances from the key when aware of the eye-tracking device.

Fixation latency reflects how quickly after the onset of each image the participant fixated on the key. Lower average latencies (measured in ms) are indicators of more rapid attention. Orienting response theory would predict that regardless of eye-tracking awareness, malintent and benign intent participants should have lower latencies than oblivious participants.

Results

Of the initial sample of 171 participants, 10 were excluded due to an absence of eye-tracking data. Most (n = 9) of these failed to successfully complete the calibration phase and thus their sessions were aborted. The final sample comprised N = 161 participants, and a total of 46,644 fixations.

Awareness of Eye Tracking

An awareness manipulation can only be successful if the eye-tracking equipment is indeed unobtrusive by default. We tested this premise immediately after completing the eye tracking phase of the study, by performing a manipulation check on participants' responses (yes/no) to whether they knew their gaze was being tracked during the study.

Of the 161 participants, 68.9% (n = 111) self-reported awareness that was consistent with the condition that they were randomly assigned to – claiming they were aware when they had been told about the eye-tracking device beforehand, or reporting non-awareness when they had not been told (see Table 5).

A Pearson's chi-squared test of independence between the awareness manipulation and their awareness self-report showed that, as desired, the two were not independent – most participants reported awareness states that were congruent with the manipulation, $\chi^2(1, N = 161) = 23.11, p < .001$. Incongruent responses were evenly distributed between the two awareness conditions: Of those who were experimentally made aware of the eye-tracking device, 30.9% reported not knowing they were being tracked; in the unaware condition, 31.2% claimed some awareness of the eye-tracking despite not being told.

The awareness manipulation was also analyzed within each guilt factor level. For benign intent participants, the pattern was similar to the overall pattern, with false positive (reporting awareness when not informed) and false negative (reporting non-awareness after being informed) rates both at 33%, as well as the expected lack of independence between the manipulation and self-report, $\chi^2(1, N = 54) = 6.00, p = .014$.

In the malintent and oblivious conditions, while the overall manipulation check failure rates were similar (29.0% and 30.8%, respectively, of participants reporting eye-tracking awareness that was the opposite of their experimental condition), their distribution was not. The false positive odds were nearly four times as high in malintent as compared to oblivious: 42.9% of those in the malintent-unaware group reported that they knew they were being tracked, whereas only 16% of the unaware oblivious did, ODDS = 3.94. This may reflect the greater stakes associated with the task – malintent participants gain strategic advantage in avoiding detecting by either being told or deducing that they are being tracked, whereas the oblivious (and benign intent) do not.

Inversely, false negatives were far more common in the oblivious condition: 44.4% of the participants who were made aware of eye-tracking reported that they were not, making false negatives four and a half times more likely to happen than for the malintent participants (14.8%), ODDS = 4.60. Oblivious participants know they plan no transgression, thus the knowledge of the presence of eye tracking is less important for them.

The high manipulation check failure rate (31.1%, n = 50 out of 161) is potentially problematic. There are many possible reasons for the size and pattern of the failure rate. Some participants (n = 9) reported, in the open-ended response to "how did you know you were being tracked", that they were tipped off after having seen the faint red glow of the eye-tracking device. Other false positives could be attributable to confusion on the participants' part. All were told that they would be performing a vision test, and all were asked to follow a dot on the screen, which they may have considered eye tracking for the purposes of the questionnaire. It is difficult to speculate on the cause of the high benign intent and malintent false positive rates. Experimenter error was improbable, as the eye-tracking software recorded the instructions that each participant was exposed to, making verification of their condition straightforward.

Due to the relatively high rate of awareness manipulation check failures, all subsequent analyses were performed twice: once with the experimentally manipulated awareness condition as a factor, and once with the awareness self-report manipulation check in its place. If the equipment is not as unobtrusive, leading several participants to become aware of it via a source other than the experimental manipulation, using their own reported eye tracking awareness would better reflect the aware-unaware difference.

When substantive differences in the pattern of the results were observed (i.e. the statistical or practical significance of any effect or comparison changed), both results were reported. When there were trivial or no differences, the analysis was reported with the experimental manipulation of awareness as the factor.

Correlations

As shown in the correlation matrix (see Table 6), many of the visual attention measures were significantly correlated with each other. The composite variable, percentage of time spent fixating on the key, was significantly correlated with its components: fixation duration, $r(158) = .61, p < .001$, and fixation count, $r(158) = .96, p < .001$, as expected. The effect size of the latter relationship was large enough to render the two variables redundant – thus, fixation time percentage is reported in favor of fixation count percentage henceforth. Fixation time percentage was also significantly negatively correlated to the distance measure, $r(158) = -.71, p < .001$, as expected; the more time spent fixating on the key, the more near-zero fixations. Encouragingly, the percentage of time spent fixating on the key was not significantly correlated with the total number of fixations, $r(159) = -.10, p = .210$, indicating that any between-condition variability in total eye-tracking data recorded would not introduce a confound in the key-percentage DV analyses below.

Means Analysis

A 2 (awareness: aware or unaware) x 3 (guilt: malintent, benign intent, or oblivious) between-subjects ANOVA was conducted on each of the above measures of visual attention, as well as pupil dilation and participants' self-report data. The following analyses address two of the studies' goals: compare intents that differ in their transgressive nature (benign vs. malintent), and to determine what moderating influence (if any) awareness of eye tracking has on those differences.

Fixation count (total). Participants averaged nearly three-hundred fixations ($M = 289.7, SD = 56.9$) over the course of the 2-minute stimulus presentation period. Total number of fixations did not differ by whether participants were aware of being tracked or not, $F(1{,}155) = 0.49, p = .486, \eta_p^2 = .003, d = 0.12$, but the guilt manipulation did influence how many times they fixated, $F(2{,}155) = 3.51, p = .032, \eta_p^2 = .043$. Participants in the malintent condition produced significantly fewer fixations than those in the oblivious condition, $M_{diff} = -27.4, p = .008, d = 0.52$, but not significantly fewer than benign individuals, $M_{diff} = -20.5, p = .080, d = 0.33$. The Guilt x Awareness interaction was not statistically significant, $F(2{,}155) = 1.62, p = .202, \eta_p^2 = .020$.

Fixation count percentage (key). Overall, a small percentage of participants' total fixations fell on the key ($M = 4.76\%, SD = 3.87$) – as one might expect for such a small object. There was no main effect of the awareness manipulation, $F(1{,}154) = 0.25, p = .615, \eta_p^2 = .002, d = 0.06$, but guilt did impact the percentage of key fixations, $F(2{,}154) = 19.06, p < .001$, $\eta_p^2 = .198$. The Awareness x Guilt interaction was not statistically significant, $F(2{,}154) = 31.05, p = .080, \eta_p^2 = .032$.

Planned comparisons of the simple effects of malintent revealed a pattern that was reversed depending on which level of awareness was present. When malintent participants were unaware they were being tracked, they fixated on the key just as frequently as the benign intent individuals, $M_{diff} = -1.05, p = .317$, 95% CI $[-2.90, 0.80], d = -0.29$, and significantly more than the oblivious, $M_{diff} = 2.40, p = .008$ 95% CI $[0.51, 4.29], d = 0.72$.

When participants were made aware that their visual attention was being monitored, those in the malintent condition avoided looking at it, and fixated on it as frequently as oblivious participants, $M_{diff} = 0.72, p = .200$ 95% CI $[-1.17, 2.61], d = 0.35$, and half as often as benign intent, $M_{diff} = -4.08, p = .002$ 95% CI $[-5.96, -2.19], d = -0.98$.

Fixation time percentage (key). Percentage of time spent fixating on the key is perhaps one of the more straightforward measures of visual attention: out of the total amount of time captured by eye-tracking, what percentage was spent fixating on the key? See Figure 3 for fixation time percentage means by condition.

We observed a significant main effect of guilt, $F(2{,}154) = 18.47, p < .001$, $\eta_p^2 = .193$, but not of awareness, $F(1{,}154) = 0.50, p = .482$, $\eta_p^2 = .003, d = 0.09$. The Guilt x Awareness interaction was statistically significant, $F(2{,}154) = 3.35, p = .038$, $\eta_p^2 = .042$.

Simple effects tests showed that when malintent participants were unaware of eye-tracking, they spent just as much time looking at the key as the benign intent participants, $M_{diff} = -1.43, p = .364, 95\%$ CI $[-4.29, 1.69], d = -0.29$, but significantly more time than the oblivious, $M_{diff} = 3.36, p = .005, 95\%$ CI $[1.32, 5.86], d = 0.73$. Similarly, unaware benign intent participants spent a significantly larger percentage of time looking at the key than oblivious participants did, $M_{diff} = 4.79, p = .001, 95\%$ CI $[1.90, 7.69], d = 1.35$.

When malintent participants were aware that their visual attention was being monitored, they spent significantly less time looking at the key when compared to the benign intent participants, $M_{diff} = -6.69, p = .001, 95\%$ CI $[-10.52, -3.28], d = -0.98$, and did not significantly differ from the percentage of time that oblivious participants spent looking at the key, $M_{diff} = 0.75, p = .225, 95\%$ CI $[-0.33, 1.93], d = 0.32$.

Expressed differently, awareness of eye tracking *increased* the visual attention given to the key by benign intent participants, $M_{diff} = 3.23, p = .121, 95\%$ CI $[-0.52, 7.24], d = 0.38$, but *decreased* it for malintent participants, $M_{diff} = -2.04, p = .133, 95\%$ CI $[-5.09, 0.50], d = -0.34$ – moderate effect sizes, occurring in opposite directions.

Distance. Avoidance of an object can be measured by computing the average distance between a participant’s fixation locations and the location of a given object (the key, in this case). Larger average distances would indicate fewer fixations near the object.

For each image, we first identified the centroid of the key object, then calculated the distance (in pixels) from each fixation to the center of the key. Eye-tracking awareness did not influence distance from the key, $F(1,155) = 0.73, p = .395$, $\eta_p^2 = .005$. There was also a significant main effect of guilt, $F(2,155) = 10.37, p < .001$, $\eta_p^2 = .118$, qualified by a significant Guilt x Awareness interaction, $F(2,155) = 3.63, p = .029\ , \eta_p^2 = .045$. When these analyses were performed with the awareness manipulation check responses in lieu of the awareness experimental condition as a factor, the effect size of the Guilt x Awareness interaction nearly doubled, with the nature of the interaction remaining the same, $F(2,155) = 3.63, p = .002\ , \eta_p^2 = .075$.

When participants were unaware of eye-tracking, malintent individuals' fixations were just as proximate to the key as those with benign intent, $M_{diff} = 12.18, p = .389$, 95% CI $[-18.13, 42.49], d = 0.25$, and oblivious participants, $M_{diff} = -17.39, p = .170$, 95% CI $[-48.31, 13.53], d = -0.38$.

Awareness of the eye-tracking resulted in malintent participants avoiding looking at the key, as reflected by both their higher fixation distances as compared to when they were unaware, $M_{diff} = 36.60, p = .081$, 95% CI $[1.07, 76.21], d = 0.48$, as well as more avoidance than the benign-aware participants, $M_{diff} = 70.90, p = .002$, 95% CI $[40.31, 101.47], d = 0.95$.

Fixation duration (total). The average duration for all fixations was not influenced by guilt, $F(2,155) = 0.85, p = .429\ , \eta_p^2 = .011$. Awareness did significantly affect average fixation duration, $F(1,155) = 6.01, p = .015, \eta_p^2 = .037$; the average fixation length was slightly longer for those in the aware condition than the unaware, $M_{diff} = 20.3, p = .016$, 95% CI $[3.9, 36.6], d = 0.38$. The Guilt x Awareness interaction was not significant, $F(2,155) = 1.29, p = .280, \eta_p^2 = 0.16$.

Fixation duration (key). The average length of fixations that fell on the key was influenced by guilt, $F(2,155) = 6.96, = .001\ , \eta_p^2 = .083$, but not by awareness, $F(1,155) =$

$2.81, p = .096$, $\eta_p^2 = .018$. The Guilt x Awareness interaction was not statistically significant, $F(2,155) = 1.59, p = .207$, $\eta_p^2 = .020$.

Simple effects showed that when participants were unaware of eye-tracking, there were no significant differences between malintent and benign intent, $M_{diff} = -13.87, p = .609$, 95% CI $[-67.36, 39.63], d = -0.13$, nor malintent and oblivious, $M_{diff} = 44.42, p = .110$, 95% CI $[-10.16, 98.99], d = 0.49$. When aware of the eye-tracking, malintent remained no different from oblivious, $M_{diff} = 1.98, p = .943$, 95% CI $[-52.52, 56.47], d = 0.03$, but their fixations were much shorter on average than those in the Benign group, $M_{diff} = -82.151, p = .003$, 95% CI $[-136.65, -27.65], d = -0.67$.

Latency (key). Part of an attention-avoidance pattern is determining how quickly participants looked at the key. We identified the first fixation within each image that fell on the key, and averaged the time-after-onset of these fixations for each participant. Shorter latencies (measured in ms) indicate more rapid attention.

There was a significant main effect of guilt, $F(2,152) = 5.59, p = .005$, $\eta_p^2 = .068$; malintent looked at the key sooner, but not significantly so, as compared to oblivious participants, $M_{diff} = -156.2, p = .070, d = -0.35$, malintent and benign intent participants were also equally as quick to fixate, $M_{diff} = 97.8, p = .175, d = 0.27$. Benign intent participants fixated significantly sooner than oblivious participants did, $M_{diff} = -253.9, p = .001, d = -0.74$.

Awareness of the eye-tracking equipment affected latencies as well, $F(1,152) = 4.11, p = .044$, $\eta_p^2 = .026$. Eye-tracking knowledge resulted in slightly lower latencies across the board $\left(M_{diff} = -125.8, p = .046, d = -0.30\right)$, as there was no significant Guilt x Awareness interaction, $F(2{,}152) = .086, p = .917\ , \eta_p^2 = .001$.

Pupil size. There were no effects on average pupil size attributable to guilt, $F(2{,}155) = 1.16, p = .316\ , \eta_p^2 = .015$, awareness, $F(1{,}155) = 0.61, p = .438\ , \eta_p^2 = .004$, nor was there a significant Guilt x Awareness interaction, $F(2{,}155) = 0.52, p = .595\ , \eta_p^2 = .007$. A similar lack of effects was found when considering pupil size only for the fixations that fell upon the key – no significant main effects of guilt, $F(2{,}154) = .246, p = .782\ , \eta_p^2 = .003$, or awareness were observed, $F(1{,}154) = .255, p = .615\ , \eta_p^2 = .002$, and the interaction was not statistically significant, $F(2{,}154) = .338, p = .713\ , \eta_p^2 = .004$. All relevant pairwise effect sizes were small to non-existent.

Self-Report

Motivation. There was a significant effect of guilt on participants' self-reported levels of motivation to complete their task (expressed as a percentage – higher percentages reflecting more motivation), $F(2{,}155) = 15.20, p < .001\ , \eta_p^2 = .164$. Malintent participants reported the lowest motivation $(M = 64.2, SD = 32.6)$, which was significantly lower than benign intent participants $(M = 81.2, SD = 19.3)$, $M_{diff} = 17.1, p = .002, d = 0.64$, and oblivious $(M = 88.4, SD = 13.2)$. Benign intent in turn reported significantly less motivation than the oblivious, $M_{diff} = 7.2, p = .033, d = 0.43$. The awareness manipulation did not influence motivation, $F(1{,}155) = 0.63, p = .430\ , \eta_p^2 = .004$, and the Guilt x Awareness interaction was not significant, $F(2{,}155) = 0.13, p = .877, \eta_p^2 = .002$.

Stress. The guilt manipulation influenced participants' self-reported stress levels during the study, $F(2{,}155) = 4.01, p = .020\ , \eta_p^2 = .049$. Malintent participants reported the highest levels of stress ($M = 36.1, SD = 32.2$), which were significantly greater than benign intent ($M = 20.7, SD = 30.6$), $M_{diff} = 15.4, p = .016$, 95% CI [4.0, 26.4], $d = 0.51$, and oblivious ($M = 22.9, SD = 28.6$) participants, $M_{diff} = 13.3, p = .024$, 95% CI [2.0, 24.5], $d = 0.44$. There was no main effect of awareness, $F(1,155) = 0.01, p = .928$, $\eta_p^2 = .000$, nor was there a Guilt x Awareness interaction, $F(2,155) = 0.52, p = .594$, $\eta_p^2 = .007$.

Confidence. The guilt manipulation affected participants' retrospective confidence in their ability to completing their task, $F(2,155) = 17.60, p < .001, \eta_p^2 = .185$. As the malintent participants were the only ones tasked a crime, it is unsurprising that their confidence levels were the lowest ($M = 57.3, SD = 33.8$) – significantly lower than both benign intent ($M = 83.2, SD = 26.1$), $M_{diff} = -26.0, p < .001, d = -0.87$ and oblivious participants ($M = 86.8, SD = 23.7$), $M_{diff} = -29.6, p < .001, d = -1.02$. Confidence was unaffected by awareness of the eye-tracking, $F(1,155) = 1.71, p = .193$, $\eta_p^2 = .011$, and there was no Guilt x Awareness interaction present, $F(2,155) = 0.43, p = .650$, $\eta_p^2 =$.006.

Discussion

Study 3 investigated the differences between malintent and benign intent on visual attention patterns, and how they may be dependent on an individual's awareness that eye-tracking is occurring. As shown in Studies 1 and 2, possessing malintent – in this study, concerning a mock theft – bolstered the attention given to the object relevant to that task, across several different operationalizations of visual attention. However, this effect was observed for individuals with benign intent as well: objects in a scene made salient by future plans captured additional visual attention regardless of whether the plan was a benign one or involved malintent.

In most measures of visual attention, when participants were not told that their gaze was being tracked, malintent and benign intent were largely indistinguishable from each other, and both differed considerably from naïve individuals who had no particular task or plan to interact with the object. These data do not provide support for the theory that the prohibited component of a future activity increases visual attention on its own.

Importantly, the behavior which distinguishes malintent from benign intent is the reaction of individuals to the knowledge that their eye movements are being tracked. When informed that their gaze was indeed being tracked, participants who possessed malintent acted rationally: they exercised some control over their gaze, and successfully avoided looking at the salient object, so much so that their patterns overall resembled those who were oblivious of the object or crime. No such avoidance was observed in those who did not think their act was a transgression – indeed, their reaction seemed to be a mild increase in attention captured by the relevant object.

It is in this critical difference in suspect strategy that several possible future strategies for application of this procedure lie. In this study, awareness of eye-tracking was manipulated between subjects; a within-subjects manipulation could identify the avoidance strategy individually, rather than collectively. A malintent suspect, first unaware that they are being tracked, would attend to crime-relevant objects. After being informed, they may, as observed in this study, suddenly avoid the object – the change from attention to avoidance would only be expected in those possessing the guilty knowledge. Those without (both the benign intent and oblivious) would either not change, or attend more often.

Several as-of-yet-unexplored factors could be expected to moderate the patterns shown here. Perhaps most critically is the innocent (be it benign intent or oblivious) individual's awareness of the crime. In high-profile events (e.g., a bomb left in a trash can), there may be no truly oblivious individuals – suspects or otherwise. The extent to which benign intent and malintent individuals will differ in their response to the awareness manipulation likely depends on how much information about the planned transgression the benign intent individuals have. In this study, they were only informed that a crime was in progress – they had no reason to believe that their errand overlapped with the crime, and thus had no reason to avoid looking at the key. It would be reasonable to expect them to behave similar to the malintent individuals if they knew that their visual attention patterns would make them appear suspicious.

Study 3's results may be limited by the differences in incentive between the malintent and benign intent participants. The additional monetary incentive for malintent participants could result in higher levels of motivation, and thus increased salience for the target object, resulting in differences between malintent and benign intent that are not transgression-based. Two aspects of the results suggest that this incentive-based confound did not have a severely problematic impact. First, malintent participants reported lower levels of motivation than the other conditions. Second, and more importantly, malintent participants showed quantities of visual attention that were either similar to benign intent (when unaware) or considerably lower than benign intent (when aware) – there was no malintent superiority pattern to be ambiguously attributed to either the transgressive nature of the intent or motivation. However, it is possible that if benign intent participants were equally incentivized, they would have showed increased amounts of attention, resulting in a difference between all three groups in the unaware condition, and indicating that, even when oblivious of eye-tracking equipment, the transgressive nature of the act was actually a suppressive effect. In terms of the Guided Search model, this would indicate that transgressive acts could have a *negative* weight on the visual attention given to malintent-relevant objects, even when the presence of eye tracking is not known. Given that this is a critical theoretical question, we designed Study 4 to address the incentive confound.

Study 4

Study 4 replicated the security-test paradigm from Study 3, with a focus on the unaware malintent and benign intent participants. The goals of Study 4 were to 1) generate a malintent and benign intent sample for replication and cross-validation of Study 3's results, and 2) incentivize malintent and benign intent participants equally in the task instructions, thus controlling more tightly for potential motivational differences between the two conditions.

Method

Participants & Design

Participants (*N* = 82) were randomly assigned to one of three conditions: malintent, benign intent, or oblivious. Participants were predominantly female (*n* = 55, 67%) and young (age M = 19.4, SD = 2.1). Most self-identified as Hispanic (*n* = 38, 46%); others, African-American (*n* = 12, 15%), Asian (*n* = 9, 11%), and white (*n* = 11, 13%).

Procedure

The procedure for Study 4 employed the same mock-crime paradigm from Study 3, with few changes. After providing informed consent, participants were asked to help test a new security system.

Malintent participants were asked steal $50; *Benign intent* participants were asked to fetch an envelope. Unlike in Study 3, both were informed that if they successfully completed their task, they would be entered into a lottery for $250.

Oblivious participants were asked to simply walk to a specific location and wait there.

As in Study 3, participants were intercepted before completing their task and shown the stimuli under eye-tracking. However, the cover story for eye-tracking calibration differed from Study 3. Participants were told that their identity would be recorded via an iris scan (the calibration cover story). After exposure to the stimuli, participants were asked to complete a questionnaire and debriefed.

Results

Of the 82 participants beginning the study, three were excluded due to a failure to record any eye-tracking data, resulting in a final sample of N = 79, and 23,422 total fixations.

Self-Reports

Motivation. Parity in incentives was successful – participants' self-reported motivation was high overall ($M = 79.6, SD = 25.3$), and did not significantly differ between condition, $F(2, 76) = 0.96, p = .386$.

Stress. Experimental conditions significantly differed on the amount of stress they reported experiencing during the study, $Welch's\ F(2, 48.3) = 20.44, p < .001$. Participants in the oblivious condition experienced nearly no stress ($M = 3.9, SD = 6.2$), which was significantly less than both malintent ($M = 30.9, SD = 29.0$), $M_{diff} = -27.1, p < .001, d = -1.11$ and benign intent participants ($M = 33.3, SD = 37.6$), $M_{diff} = 29.4, p < .001$, 95% CI $[17.1, 43.4]$, $d = -0.92$.

Confidence. Participants significantly differed in their retrospective confidence assessments of their ability to complete their assigned task, $Welch's\ F(2, 50.30) = 22.80, p < .001$. Oblivious participants, told only to go stand in a specific location, were highly confident

$(M = 93.9, SD = 10.2)$, significantly more so than benign intent participants $(M = 73.6, SD = 27.8)$, $M_{diff} = -20.2, p = .002, d = 0.85$. Benign intent participants were in turn significantly more confident than malintent participants $(M = 53.2, SD = 32.4)$, $M_{diff} = -20.4, p = .021, d = -0.70$.

Visual Attention Measures

Fixation time (key). The percentage of time participants spent fixating on the key significantly differed between the experimental conditions, $Welch's\ F(2,49.97) = 7.62, p = .001$. Oblivious participants spend the lowest amount of time looking at the key $(M = 2.09, SD = 1.47)$, which was significantly less than both malintent $(M = 4.56, SD = 4.70)$, $M_{diff} = 2.50, p = .026, d = 0.61$, and benign intent $(M = 5.27, SD = 4.93)$, $M_{diff} = 3.19, p = .005, 95\%$ CI $[-5.06, -1.42], d = 0.76$. As in Study 3, under conditions where participants were not aware their visual attention was being tracked, there were no significant differences in time spent gazing upon the key between the malintent and benign intent conditions, $M_{diff} = 0.72, p = .821, d = 0.14$.

Fixation duration (key). The average duration of fixations falling on the key did not significantly differ by condition, $F(2,75) = 3.04, p = .054, \eta_p^2 = .075$. Participants in the oblivious condition had the shortest average key fixation durations $(M = 268.4, SD = 90.6)$, considerably shorter than both malintent $(M = 277.8, SD = 83.2)$, $M_{diff} = 64.5, p = .029, d = 0.84$, and benign intent participants $(M = 281.1, SD = 98.8)$, $M_{diff} = 67.7, p = .022, d = 0.76$.

Distance (key). As in Study 3, there was no significant effect of guilt on the distance fixations fell from the key when participants were unaware they were being eye-tracked, $F(2{,}76) = 1.80, p = .172, \eta_p^2 = .045$.

Latency. The experimental manipulation of guilt did not affect the speed at which participants first looked at the key, $F(2{,}72) = 0.46, p = .636, \eta_p^2 = .012$.

ROC Curve Analysis

ROC curves visualize the ratio of true positives to false positives at each point along the continuum of the discrimination threshold variable (Fawcett, 2006). The more closely the curve comes to touching the upper-leftmost corner (indicating perfect discrimination), the more easily the two classifications can be distinguished from each other, as reflected in a larger area under the curve (AUC). Conversely, the more similar to a straight 45-degree line, the less useful the threshold variable is for discriminating between the two. AUCs and bootstrapped AUC confidence intervals were generated via the pROC package in R 2.51.3 (Robin et al., 2011).

Here, time percentage fixating on the key was used as the discriminating variable, as it showed the largest effect sizes in the previous studies. In the ROC visualizations (see Figures 4 and 5), the discrimination variable is expressed as a Cohen's *d*, calculated as the comparison of each Study 4 time percentage to Study 3's oblivious control group. Showing the curve's performance as a function of effect sizes (rather than time percentage, which may differ highly from object to object) provides a more generalized interpretive reference.

As expected from the lack of significant mean differences in percentage of time fixating on the key, the ROC curve comparing classification between malintent and benign intent is very poor (see Figure 4), $AUC = .45, 95\%$ CI $[.32, .60]$. The curve not only fully crosses the .50 mark (the level expected from chance classification), but is actually slightly below the diagonal, reflecting the tendency for benign intent participants to spend slightly more time fixating on the key.

The ROC curve comparing the malintent and oblivious conditions, as seen in Figure 5, indicates good, but not great, classification between malintent and oblivious, $AUC = .66, 95\%\ \text{CI}\ [.50, .82]$. This ROC curve is asymmetrical – it bulges at the low false positive end; at d = 2, none of the oblivious participants is falsely classified as having malintent, and 33% of the malintent participants are correctly labeled as having malintent. For a phenomenon like malintent with such low prevalence, minimizing false positives becomes especially important.

Discussion

Study 4 compared malintent, benign intent, and oblivious participants' visual attention patterns under conditions in which they were unaware their gaze was being tracked. The results were in line with the pattern shown in Study 3: on every measure of visual attention, no significant differences were found between malintent and benign intent participants. This study provided no evidence to support the theory that the transgressive nature of an intention increases the visual saliency of intent-relevant objects.

While Studies 3 and 4 found no differences between malintent and benign intent participants when they were unaware of the presence of eye-tracking, it is possible that this is a product of the low stakes for malintent participants. The malintent for these participants was not personally relevant to them, nor was the reward for successful execution particularly large. It will be critical, during further exploration of this paradigm, to investigate the incentive and motivational effects on the malintent-benign intent (non-) difference. One can imagine a scenario in which malintent individuals are extremely motivated to complete a high-stakes task, resulting in a visual salience boost over a low-stakes, mundane task planned by benign intent individuals.

General Discussion

These studies represent the first attempts at detecting malintent using visual attention measures. Our goals were to identify whether malintent increased attention given to plans-specific objects, in which of the multiple measures of this increase manifested, compare malintent to benign intent, and determine whether any of the prior effects were influenced by awareness of the presence of eye-tracking equipment. We constructed predictions for each of these goals, based on existing theory from two domains: 1) deception detection – the Guilty Knowledge Test and orienting response, and 2) basic visual attention research – the Guided Search model. Both GS and orienting response predict that objects relevant to one's malintent should receive a more rapid and increased attentional response due to their higher importance. We made several specific predictions, some of which received support in the data collected here, while others did not.

The most important prediction was that malintent individuals would attend more to malintent-relevant objects than someone without malintent. The data from each of the four studies was consistent with this prediction. Additionally, studies 1 and 2 showed that the malintent effect is object-specific; visual attention was concentrated on the more salient object, while attention for the other objects was unchanged or diminished.

We made no specific prediction regarding effect sizes in general, nor did we predict which of the many measures of visual attention recorded here would yield the largest effect. With the large total number of participants present in these studies (N = 476), it would have been a trivial effort to find a few statistically significant effects of malintent. However, in each study, effect sizes of the comparisons between malintent and oblivious participants were consistently in the range that Cohen (1988) deems “large” (most ds > 1). The largest effects were generally found in the time percentage measure, which has the advantage of being perhaps the most straightforward to interpret – of the time spent fixating on a given image, how much of it was targeted at the object.

Generally, larger effects were seen for fixation count measures than fixation duration, indicating that the participants viewed the images as a sort of implicit extended visual search. In an ordinary visual search, presentation ceases once the target is found – but the images here were presented for fixed durations (2-3s). If they are performing a visual search, then one can envision a process resembling the following: 1) malintent participants first locate and attend to the target object, 2) as they have no need to continue attending, they then direct their attention to the next high-salience object (thus the only modest increases in fixation durations for the target), and finally 3) the target object refreshes its top-down salience, thus increasing the probability that it is attended to, and restarting the visual search. The fixation durations observed in these studies are consistent with the amount of time needed to intentionally shift attention from one object to another – about 200-500 milliseconds (Theeuwes, Godijn, & Pratt, 2004; Wolfe, Alvarez, & Horowitz, 2000).

Our prediction that these large malintent effects would be present when compared to benign intent individuals was not supported by the data. We anticipated that the visual salience of an object would receive a boost from both the intention to interact with it as well as the transgressive nature of that interaction. The critical comparison, between malintent and benign intent participants, showed no evidence of the second predicted saliency boost – the gaze patterns were similar across all measures of visual attention. One could argue then that there was no *mal*intent effect observed at all. This is particularly relevant for our extension of Guided Search; here, top-down influences were much the same for the two tasks, even though they had very different meanings. This suggests that in a visual search context, such as the implicit one here, the search's qualities (i.e., the visual attention measures recorded) are influenced by the narrow short-term goals of the search, the "what am I looking for?" rather than the "why am I looking for it?"

Malintent and benign intent individuals did not behave identically in all contexts, however; their response to knowing their gaze was being tracked by investigators was strong, in opposite directions. Our last goal was to determine whether this eye-tracking awareness moderated the malintent (and benign intent) effect. We observed the predicted pattern: that those with malintent behaved rationally, by avoiding looking the object, and successfully suppressing their gaze such that they resembled someone who had no intentions towards the object. Those with benign intent had no reason to avoid; there was no transgression for them. The malintent participants' behavior here is consistent with the self-presentational perspective on deception (DePaulo, 1992); though there is no statement, malintent participants know that they want to conceal their plan, and thus they exercise top-down control over their gaze to avoid looking at the potentially incriminating object.

In sum, we have adopted and extended orienting response and Guided Search in several important ways. First, these results are consistent with the theory that the orienting response extends to oculomotor behavior; here, the increased saliency of the object produced more rapid attention after stimulus onset. However, this version of the orienting response is subject to motivational influences, and can be suppressed by a motivated viewer. Second, Guided Search seems to be a good model for describing how individuals peruse visual stimuli after setting out to perform a task which involved interacting with an object in those stimuli. The top-down influence of the intention did increase visual saliency, and thus attracted more frequent attention, even after the object had been located in the scene.

These are the first studies known to explore detecting malintent using eye tracking. As with any research in its infancy, the applications warrant discussion only with a generous helping of caution – a great deal of research is yet to be conducted before a protocol inspired by these studies could be ethically deployed in the field. That said, eye-tracking procedures may have the potential as a valuable investigative tool for law enforcement. The use of eye tracking as an investigatory tool to detect individuals with guilty knowledge, or malintent, offers several potential advantages, pending considerable future research. First, it allows investigators to leverage information about a crime they have already gathered through traditional techniques by comparing their knowledge of a (past or potential) crime scene with the difference between a suspect and known innocent's visual attention patterns when shown those photographs. Second, many conventional suspect-interrogation strategies (e.g. the polygraph) require a nominally cooperative suspect. A suspect who refuses to speak may be willing to look at photographs shown to them, especially if they are unaware that their visual attention is being tracked. Third, it allows investigators not only to identify potential suspects, but also to possibly clear persons of interest – the exoneration of the innocent and the incrimination of the guilty both being important. Given the low false positive rate of Study 1, we are particularly optimistic about this goal. Finally, creating a baseline of innocence is simple and not resource-demanding. All that is required is a handful of known innocents to view the stimuli and create a distribution of normal, naïve responses against which to compare suspect data.

While existing malintent-detection programs (SPOT) focus on a security screening context, eye-tracking procedures similar to these would likely only be appropriate in a secondary screening context – when investigators have enough information to allow for the generation of meaningful stimuli and the identification of a set of suspects. There are two potential strategies for untangling malintent and benign intent in an investigatory context. First, a suspect's statement may be used as a prior probability for the two non-guilty options. If they make no mention of plans that involve components of the suspected malintent, benign intent is less likely, and thus investigators can opt to conceal the presence of eye tracking, and compare the suspect's patterns a known innocent group.

The second strategy involves a within-subjects awareness manipulation – presenting the stimuli first with the suspect unaware of the presence of eye tracking, then repeating the procedure after revealing that their gaze is being tracked. If the pattern found in Study 3 holds, suspects with malintent should show a large drop in their visual attention measures when going from unaware to aware, a drop not seen for innocent and benign intent suspects.

Limitations and Future Directions

These four studies are just the beginning of testing eye tracking's effectiveness in detecting malintent – several issues demand further investigation. The issue of information contamination could completely alter the differential response that malintent and benign intent participants have to eye-tracking awareness, whether the contamination comes from an interrogation, or press reports of the plot. In this study, benign intent participants did not know that the planned theft involved interaction with a key; if they had, they may well have altered their gaze as well, in an attempt to not appear guilty.

The stimuli used in these studies contained mundane, office-congruent objects. Incongruent objects attract greater visual attention (Rayner, 1998) – would this wash out any differences between malintent and innocent suspects? Or would malintent offer an additional boost on top of the baseline? The same issue exists if such a procedure were to be used for past crimes – crime scene photographs including weapons, victims' bodies, or other obviously crime-relevant objects; congruent or not, a high baseline level of attention would be expected.

At the outset, we defined malintent as a plan to commit a specific transgressive task. It may be that transgression is not necessary to observe data such as these – and that a desire to conceal one's intent is sufficient to produce both the increased salience for task-relevant objects and the avoidance pattern seen here. For example, imagine a man planning a surprise party for his friend – a plan he certainly wishes to remain concealed from his friend, but that few would classify as transgressive. Future research would be necessary to differentiate what, if any, difference exists between the visual attention patterns shown by those with malintent and concealed intent.

The legal implications of eye tracking raise many questions. Does recording an individual's eye movements constitute a search, invoking the individual's 4th amendment protections against unreasonable search? Would investigators be required to show probable cause before asking a suspect to view photographs? Furthermore, is the suspect's visual attention pattern a form of statement, and thus subject to 5th amendment protection? This is a more straightforward issue with interrogations and polygraphs, where a suspect both knows a procedure is occurring and makes some verbal or written statement. However, no such utterance is necessary when viewing photographs. Just as a great deal of research is required before any field implementation is justified, a discussion amongst legal scholars should precede any practical use of these procedures.

Current malintent-detection processes are focused on security screening, but primary security screening suffers from a base rate problem. Malintent in a security setting is a rare occurrence, especially for terrorism-related crimes. Hartsfield-Jackson Atlanta International Airport alone processed almost 90 million passengers in 2010. FAST claims a classification accuracy rate in the mid-70s (Weinberger, 2010). If every one of those passengers were subjected to FAST, and 1% (a very high estimate) had some form of malintent, 22 million innocent passengers would be incorrectly flagged as potentially possessing malintent by FAST – an impractically massive number to subject to further screening. The procedures we introduce here are not appropriate for screening contexts, primarily because some knowledge of the crime scene (in order to generate stimuli) is necessary.

It is our hope that these procedures and data help inspire a host of malintent-detection and eye-tracking research, as there are numerous issues of theoretical and practical importance waiting to be explored.

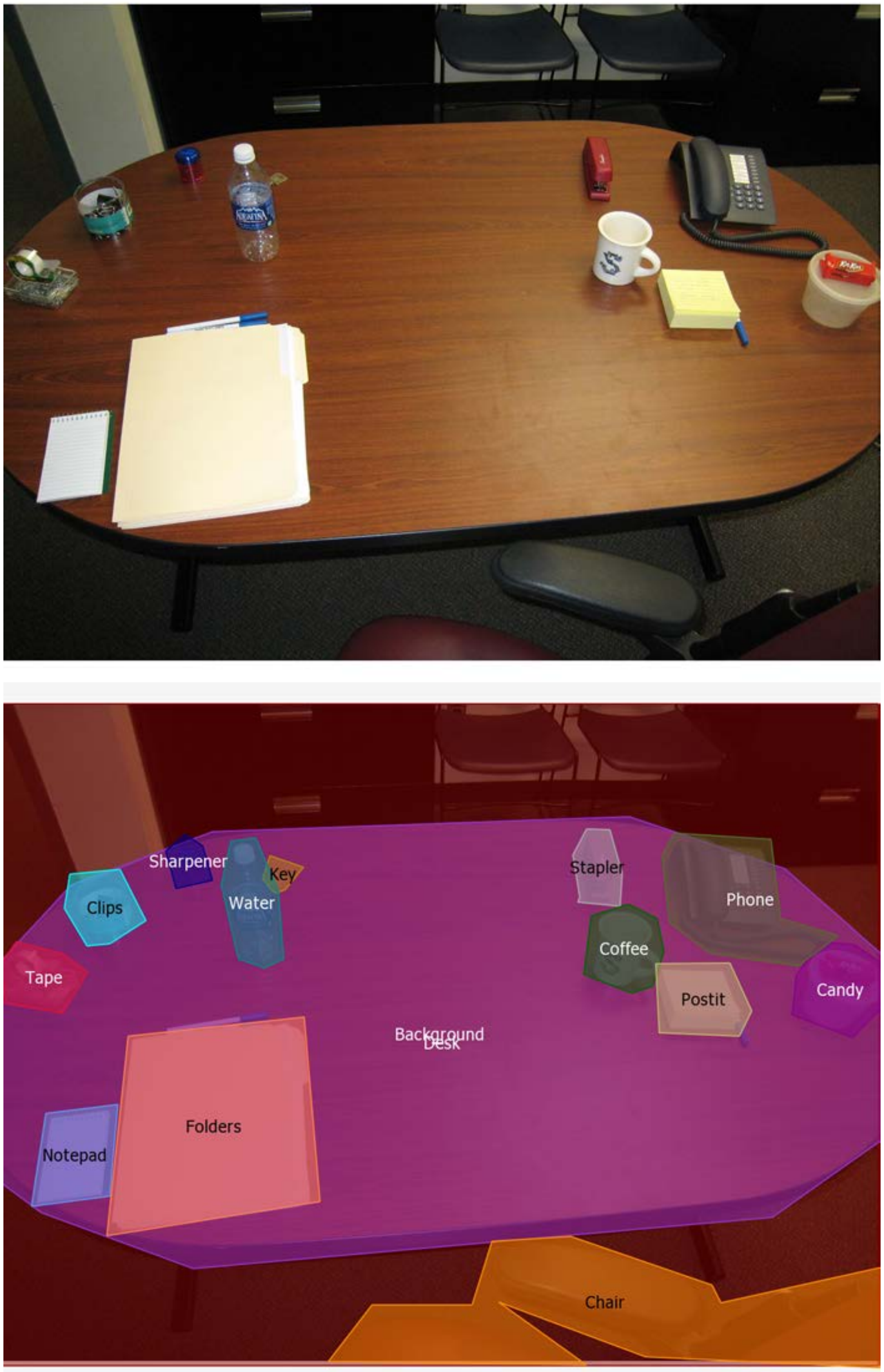

Figure 1. One image of the desk, before and after coding.

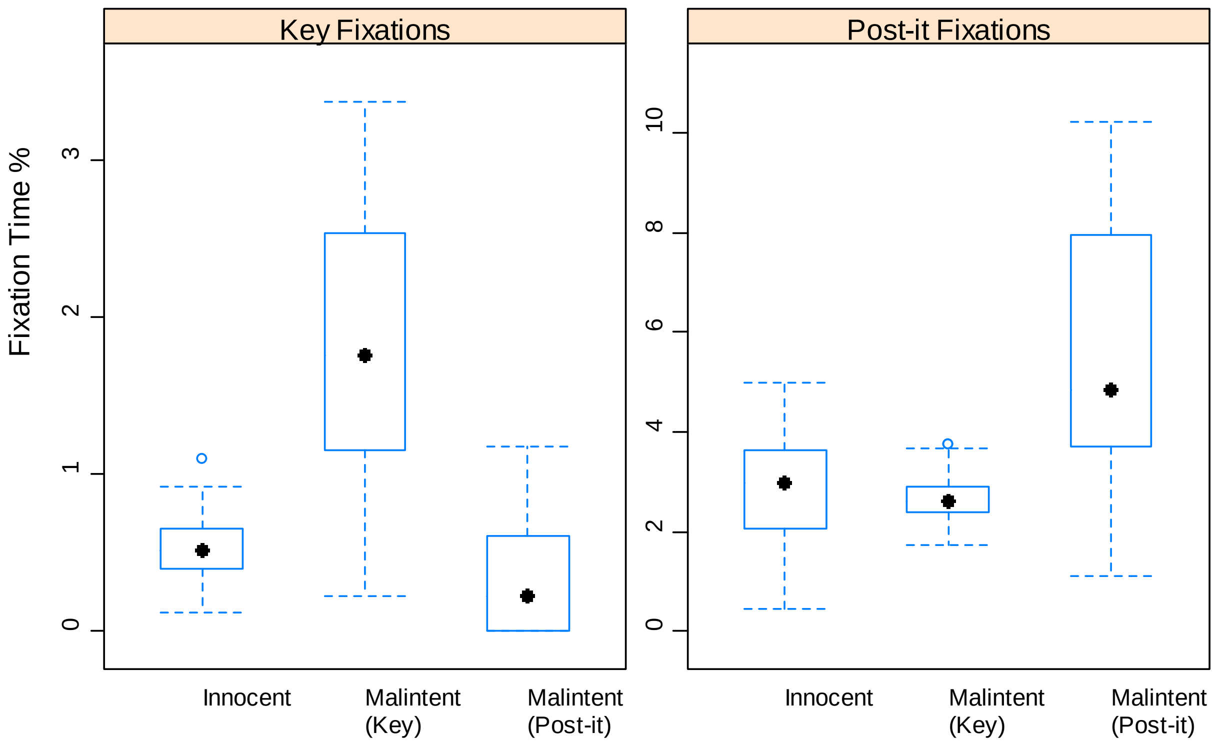

Figure 2. Box and whisker plot for percentage of total fixation time spent on the two target objects, by condition. (Study 1)

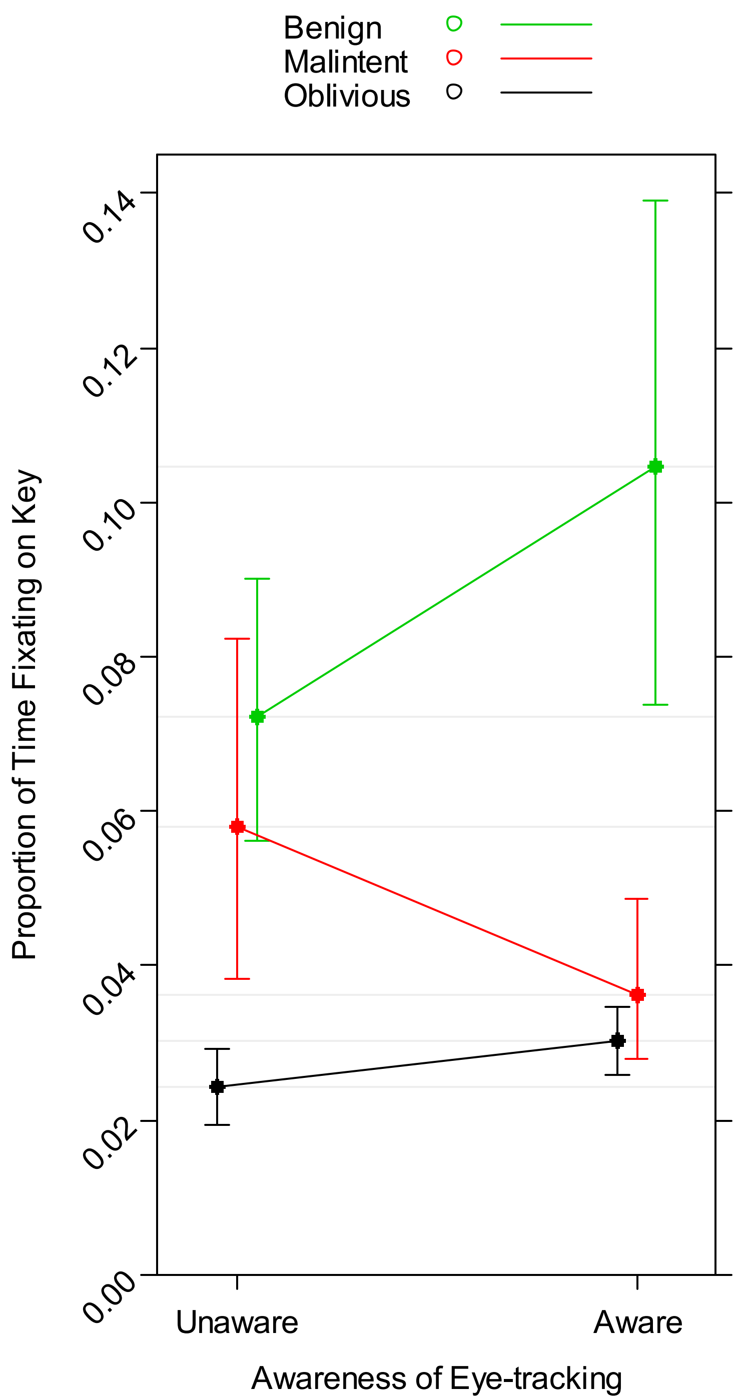

Figure 3. Proportion of total time spent fixating on the key, by group. Error bars represent bootstrapped 95% confidence intervals. (Study 3)

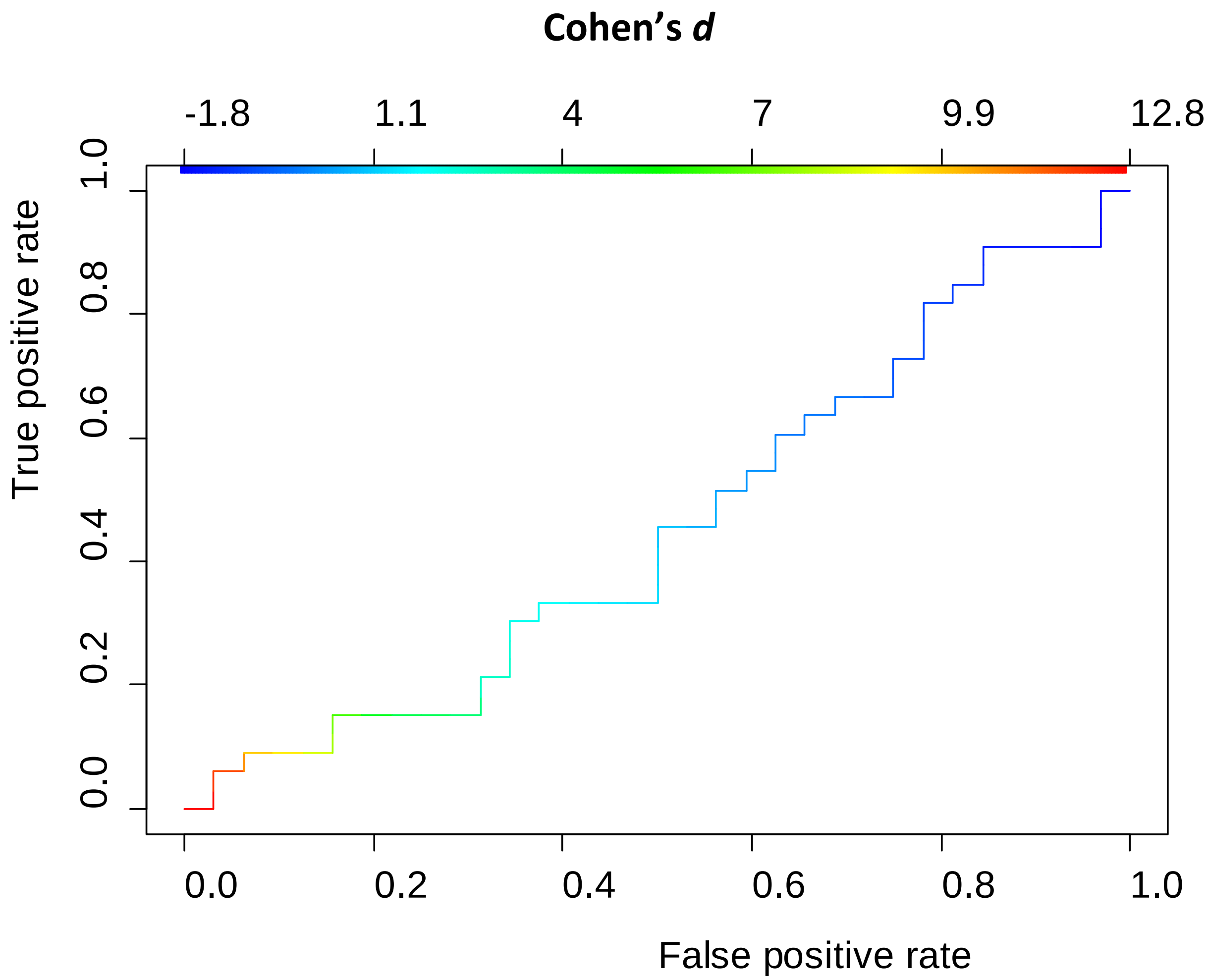

Figure 4. ROC plot comparing classification between malintent and benign intent conditions. The discriminant variable, Cohen's *d* (displayed in the chromatic scale along the top), was calculated comparing each participant's percentage of time fixating on key as to Study 3's oblivious condition. (Study 4)

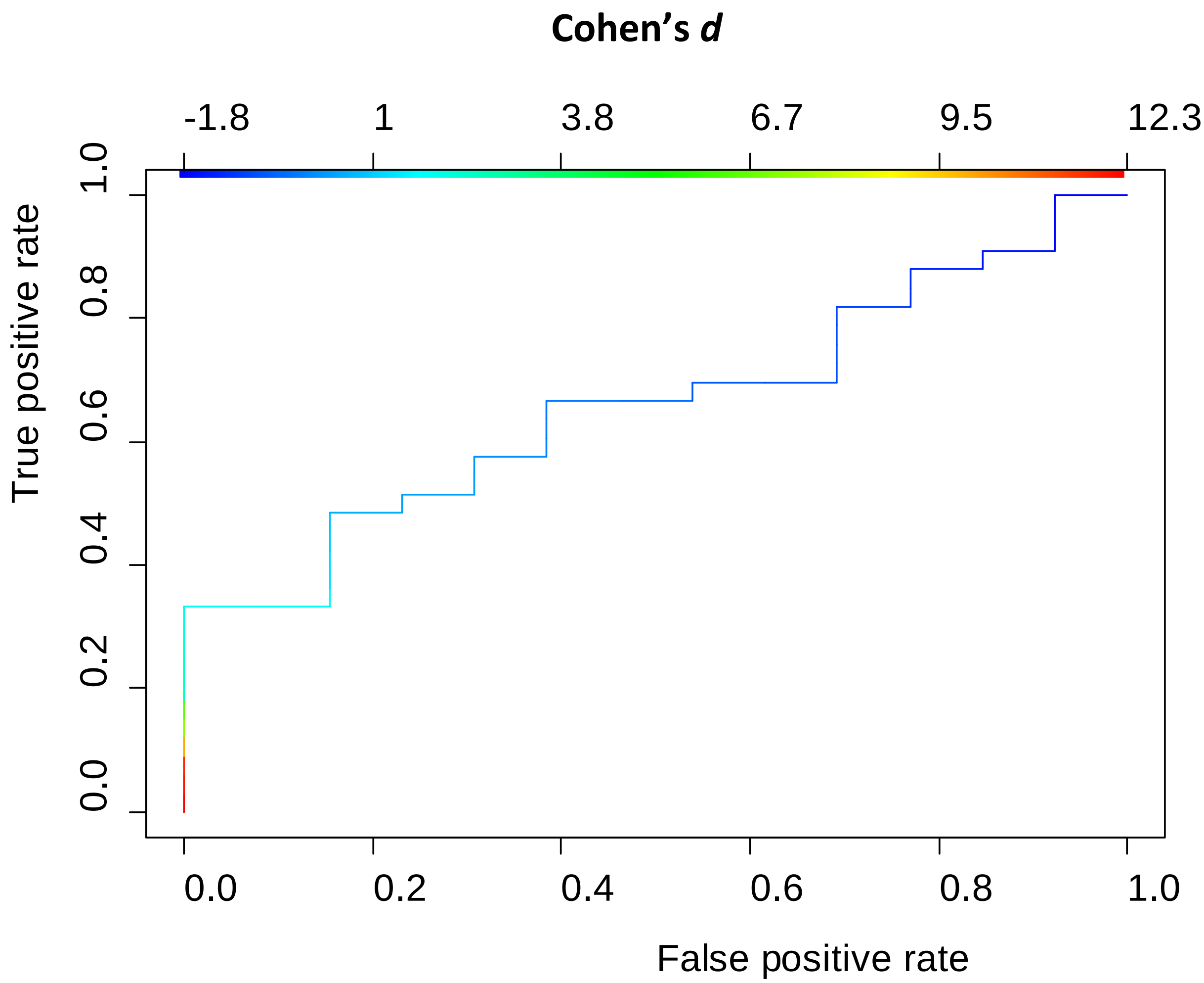

Figure 5. ROC plot comparing classification between malintent and oblivious conditions. The discriminant variable, Cohen's *d* (displayed in the chromatic scale along the top), was calculated comparing each participant's percentage of time fixating on key as to Study 3's oblivious condition. (Study 4)

Table 1

(Study 1) Visual Attention Measures by Target Object and Condition

		Object Poisoners		Innocent			
Object	Visual Attention Measure	*M*	*SD*	*M*	*SD*	Cohen's *d*	*p*
Key	Fixation Count (%)	1.75	1.27	0.63	0.28	1.12	.001
	Fixation Time (%)	2.04	1.65	0.55	0.26	1.44	.001
	Fixation Duration M (ms)	277.3	67.7	227.4	100.6	0.71	.093
	Fixation Duration SD	139.7	58.6	70.6	56.1	1.31	.001
Postit	Fixation Count (%)	5.48	3.79	2.79	1.14	1.09	.009
	Fixation Time (%)	6.37	5.11	2.89	1.25	1.07	.011
	Fixation Duration M (ms)	288.6	114.4	220.5	69.6	0.32	.336
	Fixation Duration SD	176.2	91.7	121.2	55.7	0.76	.046

Note: Cohen's *d* and *p* reflect within-row means comparisons

Table 2

(Study 1) Varimax-Rotated Component Loadings for Principal Component Analysis

Object	Measure	Component 1	Component 2
Key	Fixation Duration M	.039	.726
	Fixation Duration SD	−.061	.790
	Fixation Count	−.131	.892
	Fixation Time	−.162	.903
Postit	Fixation Duration M	.898	.066
	Fixation Duration SD	.865	−.168
	Fixation Count	.902	−.086
	Fixation Time	.933	−.125

Table 3

(Study 2) Object Selection Frequencies and Discriminant Analysis Classification Accuracy

Target Chosen by Participant	Discriminant Analysis Classification										
	Folder	Key	Kitkat	Mug	Innocent	Phone	Postit	Sharpener	Stapler	Water	Total
Folder	2 (50%)	1	0	0	1	0	0	0	0	0	4
Key	0	7 (47%)	0	1	6	1	0	0	0	0	15
Kitkat	0	1	0 (0%)	0	3	1	1	0	0	1	7
Mug	0	0	1	6 (38%)	5	3	1	0	0	0	16
Innocent	0	2	2	0	23 (70%)	5	1	0	0	0	33
Phone	2	0	0	0	14	23 (55%)	0	0	0	3	42
Post-it	0	0	0	1	2	3	0 (0%)	0	0	0	6
Sharpener	0	1	0	0	1	0	0	0 (0%)	0	0	2
Stapler	0	1	1	0	3	0	0	0	3 (38%)	0	8
Water	0	0	1	1	2	2	1	0	1	8 (50%)	16
Total											72/149 (48%)

Note: Correct classification percentages along diagonal.

Table 4

(Study 2) Percentage of Time Spent Fixating on the Top 4 Chosen Objects, by Condition

	Object Poisoners		Innocent			
Object	*M*	*SD*	*M*	*SD*	Cohen's *d*	*p*
Phone	18.2	11.8	10.5	2.6	0.88	< .001
Key	6.2	4.1	2.1	1.5	1.69	.002
Mug	18.3	13.3	5.6	2.6	1.68	.002
Water Bottle	15.9	9.8	8.1	3.5	1.29	.007

Note: Each row's means significantly differ at $p < .05$

Table 5

(Study 3) Percentage of Participants Reporting Awareness of Eye-Tracking

	Awareness Manipulation		
	Unaware	Aware	Total
Oblivious	16.0% (4/25)	55.6% (15/27)	36.5% (19/52)
Benign Intent	33.3% (9/27)	66.7% (18/27)	50.5% (27/54)
Malintent	42.9% (12/28)	85.2% (23/27)	63.3% (35/55)
Total	31.2% (25/80)	69.1% (56/81)	50.3% (81/161)

Table 6

(Study 3) Correlation Matrix for Eight Measures of Visual Attention

	Fixation Count		Fixation Time		Fixation Duration		Distance	Latency
	Total	% Key	Total	% Key	M_{Total}	M_{Key}		
Fixation Count (Total)	1.0							
Fixation Count (% Key)	−.083	1.0						
Fixation Time (Total)	.631*	.148	1.0					
Fixation Time (% Key)	−.100	.961*	.123	1.0				
Fixation Duration (M_{Total})	−.247*	.236*	.578*	.224*	1.0			
Fixation Duration (M_{Key})	−.228*	.479*	.307*	.609*	.572*	1.0		
Distance	−.141	−.717*	−.268*	−.710*	−.172*	−.383*	1.0	
Latency	.054	−.249*	−.114	−.244*	−.203*	−.238*	.076	1.0

*$p < .05$

References

Aikins, D. E., Martin, D. J., & Morgan, C. A. (2010). Decreased respiratory sinus arrhythmia in individuals with deceptive intent. *Psychophysiology*, *47*, 633–636.

Barrie, A. (2008). Homeland security detects terrorist threats by reading your mind. Retrieved from http://www.foxnews.com/story/0,2933,426485,00.html

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information wi**th the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology*, *88*, 131–151. doi:10.1037/0021-9010.88.1.131

Bond, C. F. J., & DePaulo, B. M. (2006). Accuracy of Deception Judgments. *Personality and Social Psychology Review*, *10*, 214–234. doi:10.1207/s15327957pspr1003_2

Bond, C. F. J., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, *134*, 477–492. doi:10.1037/0033-2909.134.4.477

Bond, G. D. (2008). Deception detection expertise. *Law and Human Behavior*, *32*, 339–351. doi:10.1007/s10979-007-9110-z

Bronner, E. (2012, November 14). Adultery, an Ancient Crime Still on Many Books. *The New York Times*. Retrieved from http://www.nytimes.com/2012/11/15/us/adultery-an-ancient-crime-still-on-many-books.html

Carver, R. P. (1978). The case against statistical significance testing. *Harvard Educational Review*, *48*, 378–399.

Carver, R. P. (1993). The case against statistical significance testing, revisited. *Journal of Experimental Education*, *61*, 287–292.

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition*, *29*, 1146–1152.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Psychology Press.

Committee to Review the Scientific Evidence on the Polygraph, National Research Council. (2003). *The Polygraph and Lie Detection*. Washington, D.C.: The National Academies Press.

Cook, A. E., Hacker, D. J., Webb, A. K., Osher, D., Kristjansson, S. D., Woltz, D. J., & Kircher, J. C. (2012). Lyin' eyes: Ocular-motor measures of reading reveal deception. *Journal of Experimental Psychology: Applied*, *18*, 301–313. doi:10.1037/a0028307

Department of Homeland Security. (2011). Human Factors Behavioral Sciences Projects. Retrieved March 20, 2011, from http://www.dhs.gov/files/programs/gc_1218480185439.shtm#9

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, *111*, 203–243. doi:10.1037/0033-2909.111.2.203

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74–118.

Dionisio, D. P., Granholm, E., Hillix, W. A., & Perrine, W. F. (2001). Differentiation of deception using pupillary responses as an index of cognitive processing. *Psychophysiology*, *38*, 205–211.

Donk, M., & Soesman, L. (2010). Salience is only briefly represented: Evidence from probe-detection performance. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 286–302. doi:10.1037/a0017605

Donk, M., & van Zoest, W. (2008). Effects of salience are short-lived. *Psychological Science*, *19*, 733–739. doi:10.1111/j.1467-9280.2008.02149.x

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*, 53–78. doi:10.1348/000712601162103

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*, 171–185.

Ekman, P., & O'Sullivan, M. (2006). From Flawed Self-Assessment to Blatant Whoppers: The Utility of Voluntary and Involuntary Behavior in Detecting Deception. *Behavioral Sciences & the Law*, *24*, 673–686. doi:10.1002/bsl.729

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, *27*, 861–874.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. New York, NY US: Oxford University Press.

Granhag, P. A. (2010). On the psycho-legal study of true and false intentions: Dangerous waters and some stepping stones. *The Open Criminology Journal*, *3*, 37–43. doi:10.2174/1874917801003010037

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*, 274–279. doi:10.1111/1467-9280.00255

Griffin, Z. M., & Oppenheimer, D. M. (2006). Speakers gaze at objects while preparing intentionally inaccurate labels for them. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 943–948. doi:10.1037/0278-7393.32.4.943

Hahm, J., Ji, H. K., Jeong, J. Y., Oh, D. H., Kim, S. H., Sim, K.-B., & Lee, J.-H. (2009). Detection of concealed information: Combining a virtual mock crime with a P300-based Guilty Knowledge Test. *CyberPsychology & Behavior*, *12*, 269–275. doi:10.1089/cpb.2008.0309

Hartwig, M., & Bond, C. F. J. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, *137*, 643–659. doi:10.1037/a0023589

Henderson, J. M. (1992). Identifying objects across saccades: Effects of extrafoveal preview and flanker object context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 521–530. doi:10.1037/0278-7393.18.3.521

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*, 498–504. doi:10.1016/j.tics.2003.09.006

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 210–228. doi:10.1037/0096-1523.25.1.210

Honts, C. R., Devitt, M. K., Winbush, M., & Kircher, J. C. (1996). Mental and physical countermeasures reduce the accuracy of the concealed knowledge test. *Psychophysiology*, *33*, 84–92. doi:10.1111/j.1469-8986.1996.tb02111.x

Hu, X., Hegeman, D., Landry, E., & Rosenfeld, J. P. (2012). Increasing the number of irrelevant stimuli increases ability to detect countermeasures to the P300-based complex trial protocol for concealed information detection. *Psychophysiology*, *49*, 85–95. doi:10.1111/j.1469-8986.2011.01286.x

Isaacowitz, D. M. (2006). Motivated gaze: The view from the gazer. *Current Directions in Psychological Science*, *15*, 68–72. doi:10.1111/j.0963-7214.2006.00409.x

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

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.

Kwak, S.-M., Na, D. L., Kim, G., Kim, G. S., & Lee, J.-H. (2007). Use of eye movement to measure smokers' attentional bias to smoking-related cues. *CyberPsychology & Behavior*, *10*, 299–304. doi:10.1089/cpb.2006.9953

Lam, S. Y., Chau, A. W.-L., & Wong, T. J. (2007). Thumbnails as online product displays: How consumers process them. *Journal of Interactive Marketing*, *21*, 36–59. doi:10.1002/dir.20073

Lykken, D. T. (1974). Psychology and the lie detector industry. *American Psychologist*, *29*, 725–739. doi:10.1037/h0037441

MacLaren, V. V. (2001). A quantitative review of the Guilty Knowledge Test. *Journal of Applied Psychology*, *86*, 674–683. doi:10.1037/0021-9010.86.4.674

Meixner, J. B., & Rosenfeld, J. P. (2011). A mock terrorism application of the P300-based concealed information test. *Psychophysiology*, *48*, 149–154. doi:10.1111/j.1469-8986.2010.01050.x

Miltner, W. H. R., Krieschel, S., Hecht, H., Trippe, R., & Weiss, T. (2004). Eye Movements and Behavioral Responses to Threatening and Nonthreatening Stimuli During Visual Search in Phobic and Nonphobic Subjects. *Emotion*, *4*, 323–339. doi:10.1037/1528-3542.4.4.323

Noordraven, E., & Verschuere, B. (2013). Predicting the sensitivity of the reaction time-based Concealed Information Test. *Applied Cognitive Psychology*, *27*, 328–335.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion*, *6*, 257–268. doi:10.1037/1528-3542.6.2.257

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107–123. doi:10.1016/S0042-6989(01)00250-4

Peth, J., Kim, J. S. C., & Gamer, M. (2013). Fixations and eye-blinks allow for detecting concealed crime related memories. *International Journal of Psychophysiology*. doi:10.1016/j.ijpsycho.2013.03.003

Porter, S., & ten Brinke, L. (2008). Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science*, *19*, 508–514. doi:10.1111/j.1467-9280.2008.02116.x

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422. doi:10.1037/0033-2909.124.3.372

Rinck, M., & Becker, E. S. (2006). Spider fearful individuals attend to threat, then quickly avoid it: Evidence from eye movements. *Journal of Abnormal Psychology*, *115*, 231–238. doi:10.1037/0021-843X.115.2.231

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*, 77. doi:10.1186/1471-2105-12-77

Rosenfeld, J. P., Labkovsky, E., Winograd, M., Lui, M. A., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): A new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology*, *45*, 906–919. doi:10.1111/j.1469-8986.2008.00708.x

Schall, J. D., & Thompson, K. G. (1999). Neural selection and control of visually guided eye movements. *Annual Review of Neuroscience*, *22*, 241–259. doi:10.1146/annurev.neuro.22.1.241

Schwedes, C., & Wentura, D. (2012). The revealing glance: Eye gaze behavior to concealed information. *Memory & Cognition*, *40*, 642–651. doi:10.3758/s13421-011-0173-1

Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess “guilty knowledge”. *Journal of Applied Psychology*, *85*, 30–37. doi:10.1037/0021-9010.85.1.30

Siddle, D. A. (1991). Orienting, habituation, and resource allocation: An associative analysis. *Psychophysiology*, *28*, 245–259. doi:10.1111/j.1469-8986.1991.tb02190.x

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

Theeuwes, J., Godijn, R., & Pratt, J. (2004). A new estimation of the duration of attentional dwell time. *Psychonomic Bulletin & Review*, *11*, 60–64. doi:10.3758/BF03206461

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 478–499. doi:10.1037/0096-1523.20.3.478

United States Government Accountability Office. (2010). Efforts to validate TSA’s passenger screening behavior detection program underway, but opportunities exist to strengthen validation and address operational challenges.

US Department of Homeland Security. (2008). Privacy impact assessment for the Screening of Passengers by Observation Techniques (SPOT) Program. Retrieved from http://www.dhs.gov/xlibrary/assets/privacy/privacy_pia_tsa_spot.pdf

Verschuere, B., Ben-Shakhar, G., & Meijer, E. (Eds.). (2011). *Memory Detection: Theory and Application of the Concealed Information Test* (1st ed.). Cambridge University Press.

Verschuere, B., Crombez, G., Degrootte, T., & Rosseel, Y. (2010). Detecting concealed information with reaction times: Validity and comparison with the polygraph. *Applied Cognitive Psychology*, *24*, 991–1002. doi:10.1002/acp.1601

Verschuere, B., Crombez, G., & Koster, E. H. W. (2004). Orienting to guilty knowledge. *Cognition and Emotion*, *18*, 265–279. doi:10.1080/02699930341000095

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities (2nd ed.)*. New York, NY US: John Wiley & Sons Ltd.

Vrij, A., Granhag, P., Mann, S., & Leal, S. (2011). Lying about flying: The first experiment to detect false intent. *Psychology, Crime & Law*, *17*, 611–620. doi:10.1080/10683160903418213

Vrij, A., Leal, S., Mann, S. A., & Granhag, P. A. (2011). A comparison between lying about intentions and past activities: Verbal cues and detection accuracy. *Applied Cognitive Psychology*, *25*, 212–218. doi:10.1002/acp.1665

Warmelink, L., Vrij, A., Mann, S., Leal, S., Forrester, D., & Fisher, R. P. (2010). Thermal imaging as a lie detection tool at airports. *Law and Human Behavior*. doi:10.1007/s10979-010-9251-3

Webb, A. K., Honts, C. R., Kircher, J. C., Bernhardt, P., & Cook, A. E. (2009). Effectiveness of pupil diameter in a probable-lie comparison question test for deception. *Legal and Criminological Psychology*, *14*, 279–292. doi:10.1348/135532508X398602

Weinberger, S. (2010, May). Airport security: Intent to deceive? *Nature*, *465*, 412–415.

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238. doi:10.3758/BF03200774

Wolfe, J. M. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. D. Gray (Ed.), *Integrated models of cognitive systems.* (pp. 99–119). New York, NY US: Oxford University Press.

Wolfe, J. M., Alvarez, G. A., & Horowitz, T. S. (2000). Attention is fast but volition is slow. *Nature, 406*. doi:10.1038/35021132

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 419–433. doi:10.1037/0096-1523.15.3.419

Wolfe, J. M., & Gancarz, G. (1997). Guided Search 3.0: A model of visual search catches up with Jay Enoch 40 years later. In V. Lakshminarayanan (Ed.), *Basic and Clinical Applications of Vision Science* (pp. 189–192). Dordrecht, Netherlands: Kluwer Academic Publishers.

Yantis, S. (2005). How visual salience wins the battle for awareness. *Nature Neuroscience*, *8*, 975–977. doi:10.1038/nn0805-975

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.