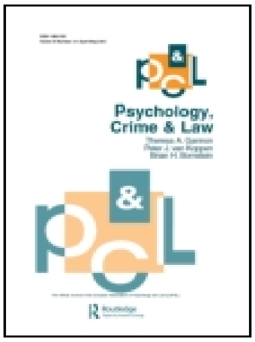

Psychology, Crime & Law

ISSN: 1068-316X (Print) 1477-2744 (Online) Journal homepage: www.tandfonline.com/journals/gpcl20

Of guns and geese: a meta-analytic review of the ‘weapon focus’ literature

Jonathan M. Fawcett, Emily J. Russell, Kristine A. Peace & John Christie

Psychology, Crime & Law
Vol. 19, No. 1, January 2013, 35–66

Of guns and geese: a meta-analytic review of the 'weapon focus' literature

Jonathan M. Fawcett[a]*, Emily J. Russell[b], Kristine A. Peace[c] and John Christie[a]

[a]*Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada;* [b]*Department of Psychology, Lakehead University, Thunder Bay, Ontario, Canada;* [c]*Department of Psychology, Grant MacEwan University, Edmonton, Alberta, Canada*

(*Received 2 April 2011; final version received 17 June 2011*)

Weapon focus is frequently cited as a factor in eyewitness testimony, and is broadly defined as a weapon-related decrease in performance on subsequent tests of memory for those elements of an event or visual scene concurrent to the weapon. This effect has been attributed to either (a) physiological or emotional arousal that narrows the attentional beam (*arousal/threat hypothesis*), or (b) the cognitive demands inherent in processing an unusual object (e.g. weapon) that is incongruent with the schema representing the visual scene (*unusual item hypothesis*). Meta-analytical techniques were applied to test these theories as well as to evaluate the prospect of weapon focus in real-world criminal investigations. Our findings indicated an effect of weapon presence overall ($g = 0.53$) that was significantly influenced by retention interval, exposure duration, and threat but unaffected by whether the event occurred in a laboratory, simulation, or real-world environment.

Keywords: weapon focus; eyewitness memory; memory; attention; cognition

Introduction

On the afternoon of 24 April 1997 a man walked into a Toronto area coffee shop and threatened to kill a hostage unless someone provided him with a small sum of money. The atmosphere was tense, but eventually the occupants met the robber's demands at which time he released his hostage and fled the scene. This may sound like a classic 'hold up', but what makes this robbery notable is the fact that the hostage was not a member of the local clientele or even a store clerk – the hostage was a very disturbed Canada Goose which the perpetrator had brought with him and dramatically threatened to choke ('Goose Held Hostage in Canada Robbery', 1997; as cited by Mitchell, Livosky, & Mather, 1998). Witnesses were stricken by the absurdity of what they had observed, apparently having devoted more time to inspecting the goose than to inspecting the robber.

The story of the 'goose robber' is an excellent illustration of how the presence of an unusual object can impair memory for the peripheral details of an event. The notion is that eyewitnesses tend to focus their attention on salient or informative objects within a visual scene, to the detriment of other important details (e.g. the identity or characteristics of the perpetrator). While the story above demonstrates

*Corresponding author. Email: jmfawcet@dal.ca

ISSN 1068-316X print/ISSN 1477-2744

http://dx.doi.org/10.1080/1068316X.2011.599325
http://www.tandfonline.com

this finding using a goose (see also Pickel, 1999), laboratory studies have historically investigated this effect using weapons such as guns and knives (e.g. Loftus, Loftus, & Messo, 1987) resulting in the term *weapon focus* (e.g. Loftus, 1979).

The potential impact of weapon focus in the real world is substantial. In the absence of a confession (e.g. Kassin & Gudjonsson, 2004), eyewitness testimony is one of the most important sources of evidence leading to a conviction (Rand Corporation, 1975; Yarmey, 2001). As such, any factor that may influence the accuracy of such testimony must be considered carefully. The purpose of this review is to evaluate the literature surrounding the weapon focus effect as well as its real-world applications. Towards this goal we present a summary of the extant literature followed by a meta-analysis considering some relevant moderator variables. But first we shall discuss the predominant theoretical explanations of weapon focus to provide a framework for the interpretation of what is to follow.

Perspectives on the weapon focus effect

Definition of weapon focus

From the perspective of an eyewitness, many details associated with a crime are important and potentially informative (e.g. the identity of the perpetrator(s), possible means of escape). However, few details are as salient as a weapon (e.g. Pickel, Ross, & Truelove, 2006). Research has suggested that the presence of a weapon captures the attention of eyewitnesses, resulting in poor recall and recognition of the perpetrator (i.e. feature accuracy) as well as diminished identification accuracy in subsequent suspect line-ups (e.g. Loftus et al., 1987). As noted above, because it was initially linked to the presence of a weapon (as opposed to an unusual object more generally), researchers have traditionally referred to this finding as the weapon focus effect (e.g. Loftus, 1979). For the purpose of this review, weapon focus shall be defined as an object-related decrease in memory performance (e.g. feature or identification accuracy) for those elements of an event or visual scene coinciding with the presence of the weapon or unusual object. This decrease must not be more parsimoniously explained by another factor (such as injury). A decrease in performance shall be considered object-related only if it can be demonstrated that the performance deficit is associated with the object's presence via the comparison of object present and object absent conditions, or comparable visibility conditions (e.g. high and low visibility; see discussion below). Although this definition is broad in scope, it is necessary to encompass the breadth of findings related to this topic. Some definitions of weapon focus assume specific mechanisms underlying this effect (e.g. concentration of gaze upon the weapon; Loftus et al., 1987) that cannot be gauged in studies conducted in the real-world or otherwise uncontrolled environments. At this time, the mechanisms underlying weapon focus remain largely unclear. However, the majority of studies on this topic have been informed by two theoretical explanations in particular: the arousal/threat hypothesis and the unusual item hypothesis.

Arousal/threat hypothesis

Past research has indicated that physiological arousal and performance are intimately related. An individual's performance may be enhanced by slightly elevating arousal (e.g. Whipple, 1915), but is often negatively affected when the individual experiences a state of over- or under-arousal (e.g. Yerkes & Dodson, 1908). Easterbrook's (1959) cue-utilization hypothesis proposed that attention and arousal interact to determine which aspects of the perceptual environment are utilized. According to this hypothesis, physiological arousal decreases the number of environmental cues that may be concurrently monitored. Therefore, in a state of high physiological arousal, only those cues that are the focus of attention are utilized (i.e. central cues). On the other hand, peripheral cues remain unmonitored, and are consequently underutilized. These findings have important implications for memory encoding and subsequent retrieval when applied to weapon focus. The arousal/threat hypothesis suggests that the presence of a weapon, by virtue of its threatening nature, induces a state of elevated physiological arousal. This results in fixation upon the source of that arousal (i.e. the weapon) as well as a decrease in the number of monitored cues and an underutilization of peripheral stimuli (Easterbrook, 1959). Thus memory for central cues (i.e. the weapon) is heightened and memory for peripheral cues (i.e. details of the perpetrator) is diminished.

Although the arousal/threat hypothesis has been substantiated by some studies directly manipulating threat or arousal (e.g. Peters, 1988), other studies have failed to find an effect of such manipulations (e.g. Maass & Kohnken, 1989; Pickel, 1998; Tooley, Brigham, Maass, & Bothwell, 1987). In fact, weapon focus has been observed using stimuli that elicit minimal arousal (e.g. Kramer, Buckhout, & Eugenio, 1990; Pickel, 1999). Inconsistent outcomes of studies predicated on arousal are limited by methods used to elicit arousal (i.e. artificial laboratory settings) and problematic operational definitions. For example, arousal could result from anxiety or fear (as might be experienced by eyewitnesses) or simply an increased state of alertness or attention (e.g. Brown, 2003; Deffenbacher, Bornstein, Penrod, & McGorty, 2004; Hope & Wright, 2007; Pickel, 2007). While previous studies have utilized physiological measures to tap into this construct (e.g. Valentine & Mesout, 2009), the source of physiological arousal may differentially influence memory recall. Due to situational variables associated with the interpretation of physiological arousal, and subsequent influences on memory, the arousal/threat hypothesis has limited support. As will be discussed in greater detail later, studies examining memory for actual crimes, where perceived threat and arousal should be greatest, have been mixed (e.g. Cooper, Kennedy, Herve, & Yuille, 2002; Pike, Brace, & Kyman, 2002; Valentine, Pickering, & Darling, 2003).

Unusual item hypothesis

In contrast to studies that have associated increased levels of arousal or threat with attentional narrowing, other research has shown that unusualness may also influence attention. For example, Antes (1974) found that objects of particular relevance or informational value are more frequently attended to and for greater durations. In particular, surprising or unexpected objects attract attention, drawing the observer's gaze more readily than objects expected given the scene's content (Loftus & Mackworth, 1978). To explain this finding, Loftus and Mackworth (1978) developed a three-stage model of early visual scene processing. First, the general qualities of the scene are analyzed to activate the schema most appropriate for the depicted context. Once a schema has been activated, preliminary processing of the scene's content is undertaken, followed by an evaluation of the degree to which each object is congruent with the activated schema. Objects exhibiting low schema congruence are further processed to resolve the incongruity, either by modifying the schema or re-interpreting the object. As a result, attentional resources are unavailable to process other aspects of the scene, leading to poor recall of those details (Loftus & Mackworth, 1978). According to the unusual item hypothesis, weapons are considered unusual in many contexts and therefore are incongruent with the schemas activated by common visual scenes. For example, guns are not commonly associated with convenience stores and therefore the attentional resources of those witnessing a convenience store robbery would be dedicated to resolving the conflict existing between the gun and the schema representing their environment. As a result, peripheral details (e.g. the perpetrator's face) would not be properly encoded.

Recently, this explanation of weapon focus has been supported by a growing number of publications demonstrating a weapon focus effect in visual scenes that contain unusual objects other than weapons (e.g. Mitchell et al., 1998; Pickel, 1998). It has also been shown that a weapon focus effect will not occur for weapons congruent with the visual scene in which they appear (e.g. a gun at a shooting range; Pickel, 1999). While these findings are explained by the unusual item hypothesis, they are incompatible with the arousal/threat view.

Evidence of the weapon focus effect

We will now review the extant literature on this topic organized by the methodological approach employed by each article. At first glance, it might appear that we should categorize each article as either a laboratory study or an investigation of an actual criminal event. However, whereas some laboratory studies have opted to use videos or slides to present the event of interest others have staged events such that participants witnessed them directly. Direct (as opposed to indirect) exposure to a criminal event could result in measurable differences with respect to arousal, perceived threat, and salience of the event and the weapon itself (e.g. Davis & Valentine, 2009; Peters, 2001). These heightened levels are not comparable to those elicited by viewing a videotape in a controlled laboratory setting and warrant separate consideration (see Deffenbacher, 2008; Reisberg & Heuer, 2007). Therefore, *Laboratory* studies were defined as any investigation conducted in a controlled environment with indirect exposure to the event of interest (often videos or slides) such that participants were explicitly aware that the studied material (e.g. the crime) was not truly occurring. *Simulation* studies were defined as any investigation conducted in a controlled environment with direct exposure to the event of interest (e.g. an enacted scene) such that participants could reasonably believe or imagine that the studied material was actually occurring. Studies of *Actual Crimes* were defined as any investigation conducted using information derived from real-world criminal events, including interviews, records from police line-ups, or related materials.

Laboratory studies

In one of the earliest studies of weapon focus, Loftus et al. (1987) explored how the presence of a weapon influenced picture-viewing behavior. In two identical experiments gaze was measured while participants viewed one of two scenarios: Slides depicted a man (the target) approaching the front of a fast food restaurant where he displayed either a gun (the weapon condition) or a check (the control condition). The man then received money and left. Participants in the weapon condition fixated more frequently and for greater duration upon the critical object (i.e. the gun) than did participants in the check condition. Participants in the weapon condition also performed less accurately when administered a target present photo line-up and multiple-choice questionnaire testing details that had been presented at the same time as the weapon (Loftus et al., 1987). As a result of this research, the weapon focus effect garnered much scientific attention, and the aforementioned theoretical perspectives were applied as explanatory frameworks.

Around the same time Tooley et al. (1987) conducted a study specifically testing the arousal/threat hypothesis by manipulating attentional focus and arousal. They presented participants with a series of 24 life-sized slides, each containing a man standing in a convenience store carrying a weapon (e.g. gun) or non-weapon (e.g. can of soup) object. Each participant was told to focus his or her attention on a different aspect of the slide (i.e. face, hands, background, or free focus). As a method of manipulating arousal, half of the participants were presented with the threat of an electric shock as well as occasional bursts of white noise. Overall, the presence of a weapon relative to a non-weapon object decreased performance in a subsequent recognition task, independent of the arousal condition (Tooley et al., 1987).

Also at this time a wealth of research emerged exploring the factors influencing eyewitness testimony, including manipulations of weapon visibility (e.g. Cutler & Penrod, 1988; Cutler, Penrod, & Martens, 1987a,b; Cutler, Penrod, O'Rourke, & Martens, 1986; O'Rourke, Penrod, Cutler, & Stuve, 1989). Each study utilized approximately the same methodology involving a short video depicting an armed robbery where the robber brandished a gun (i.e. high visibility condition) or concealed a gun in his pocket (i.e. low visibility condition). Following this presentation, participants were administered a target-present or target-absent photo line-up. Of these studies, all but one (Cutler & Penrod, 1988) exhibited a small but significant effect of weapon visibility, indicating that participants in the high visibility condition showed diminished identification accuracy (Cutler et al., 1986, 1987a, b). Kramer et al. (1990, Experiment 1) observed a similar effect of weapon visibility where a bottle (the weapon) was held either at chest level (i.e. high visibility) or behind the perpetrator's back (i.e. low visibility) in a card game scenario. In that case results indicated that high weapon visibility resulted in poor recall of target features and better recall of the weapon, although no differences in identification accuracy were found on a target-present photo line-up.

Kramer et al. (1990) conducted a number of other experiments investigating the role of arousal in the weapon focus effect by attempting to elicit this effect in a stark environment with minimal arousing properties. In each of four experiments, Kramer et al. (1990, Experiments 2A–D) presented participants with six sequential slides depicting a man walking down an austere hallway, carrying a magazine or bloody meat cleaver. Two of Kramer et al.'s (1990, Experiments 2A–B) experiments indicated that weapon presence decreased memory of the visual scene, but increased memory for the weapon itself. The third and fourth studies reduced the number of slides containing the weapon or magazine. When present for only four slides, exposure to the weapon resulted in worse recall of the event, but did not affect recall of the central object (Kramer et al., 1990; Experiment 2C). However, when the weapon or magazine was present for only one slide, no significant effects were observed (Experiment 2D). Furthermore, none of the experiments (i.e. Experiments 2A–D) reported a significant effect of weapon presence on identification accuracy or self-reported arousal scores.

The finding that weapon focus could be produced without any evidence of self-reported arousal prompted some researchers to redirect their focus away from explorations of arousal to unusualness instead. The earliest such study, conducted by Shaw and Skolnick (1994), failed to produce results supporting the unusual item hypothesis. These researchers presented participants with a series of six sequential slides depicting a man or woman exiting a phone booth and walking away carrying a gun, magazine, or an unusual object (i.e. space cones, conch shell, stethoscope, or wooden stake). When tested using a target-present photo line-up and questionnaire, the effect of item type did not reach significance. Furthermore, the data exhibited a trend opposite all predictions; memory performance was greatest in the gun condition and worst in the magazine condition. The cause of this pattern is not clear, but the explanation presented by Shaw and Skolnick (1994), that the experimental conditions were not arousing or complex enough, is unlikely given the findings of Kramer et al. (1990).

Subsequent studies exploring the role of unusualness were more fruitful. Mitchell et al. (1998) presented participants with a video depicting one of four encounters between two businessmen. During each encounter one of the businessmen would reach into his briefcase and reveal nothing (i.e. his hand, followed by a handshake), a stick of celery, a gun displayed in a non-threatening manner, or a gun displayed in a threatening manner. The presence of a gun (regardless of threat) or a stick of celery decreased accuracy on a subsequent multiple-choice questionnaire (Mitchell et al., 1998; Experiment 2). Pickel (1998) independently replicated these findings using different unusual objects (e.g. a whole raw chicken) while also demonstrating that this detriment only affected details presented at the same time as the object of interest; details pertaining to people present prior to the introduction of this object or following the removal of the object were unaffected.

Subsequent studies explored the role of context. According to the unusual item hypothesis, weapon focus occurs when an object contained within a visual scene is incongruent with the schema representing that scene. As a result, attentional resources are drained as the incongruity between the object and the activated schema is resolved. This perspective would suggest that the weapon focus effect would not occur when a weapon is presented in a visual scene in which that weapon is not considered unusual. Pickel (1999) tested this hypothesis by first manipulating location (Experiment 1) and the person schema of the individual holding the gun (Experiment 2). In her first experiment, participants viewed a videotaped encounter between a woman and a man armed with a gun. Both location (sporting event vs shooting range) and level of threat (friendly vs menacing demeanor of the man) associated with the interaction were manipulated. Participants' memories were tested using a target-present photo line-up, as well as multiple-choice and cued recall questions. Recall was worse when the encounter occurred at the sporting event (incongruent condition) relative to the shooting range (congruent condition). In contrast, identification accuracy was unaffected by location or threat, and threat had no effect on recall performance (Pickel, 1999; Experiment 1). This finding was supported by Pickel's (1999) second study. Participants viewed a video depicting a friendly scene between a woman and a man dressed as either a priest or a police officer. The man carried either a gun or a cell phone. This resulted in four videos, three containing an object congruent with the visual scene (officer–phone, officer–gun, priest–phone) and one containing an object incongruent with the visual scene (priest–gun). Those in the priest–gun condition exhibited worse recall than those in the priest-phone condition; no weapon focus effect was observed for the police officer. Identification accuracy was once again unaffected by weapon presence. Similar findings have shown that the impairments caused by the presence of a weapon or unusual item is affected by the degree to which the participant associates the possession of that item with the particular characteristics of the target. Pickel (2009) observed a larger weapon focus effect when a gun was carried by a female as compared to a male (since guns are more associated with males than females; also see Shaw & Skolnick, 1999; for a similar discussion of race see Tooley et al., 1987) and a smaller or absent weapon focus effect was observed when the participants were induced to believe that the target was a dangerous individual prior to the appearance of the weapon.

One weakness common to most of the experiments supporting the unusual item hypothesis is that they have only compared the relative effects of weapons and unusual objects during a subsequent test instead of differences at the time of encoding. According to the unusual item hypothesis, visual scenes containing a weapon or usual object are more difficult to process than scenes that do not. As a result, attentional resources would be less available for other, ongoing secondary tasks. Using slides, Hope and Wright (2007) presented participants with a short scene depicting a male actor entering a small store holding a gun, feather duster, or wallet. Below each slide a box was provided in which a series of even two-digit numbers were rapidly presented at a rate of 400 ms (with 300 ms between each number). During the slide containing the object of interest an odd two-digit number was mixed into the otherwise even presentation stream and participants were required to make a speeded detection response. Participants in the control condition (wallet) recognized more features of the target individual than in either the weapon (gun) or unusual item (feather duster) conditions, which did not differ significantly. Importantly, secondary task reaction times were also slower during the weapon and unusual item conditions relative to the control condition. Slower responses for the secondary task in the presence of a weapon or unusual object (compared to a control object) supports the notion that visual scenes containing incongruent objects are harder to process.

Simulation studies

Another methodology utilized in the study of weapon focus involves experimental settings in which staged scenarios are presented to participants in order to simulate real-world experiences. In a frequently cited manuscript, Johnson and Scott (1976) detailed one of the earliest studies exploring weapon focus, and one of only a handful of simulations. Participants sat in a waiting area preceding their 'experimental session', when one of two staged scenarios occurred. In the weapon absent scenario participants heard a discussion regarding an equipment failure, following which a man entered the room carrying a grease pen, mumbled, and left. In the weapon present scenario participants heard a loud argument, including noises indicative of a fight, following which a man entered the room carrying a bloody letter opener, mumbled, and left. Participants in the weapon-present condition exhibited better memory of the carried item (i.e. the weapon) but were less accurate when selecting the perpetrator from a photo line-up.

Under more controlled conditions, Maass and Kohnken (1989) attempted to elicit weapon focus in a pseudo-medical context. They pretended to conduct a study on sports and well being, thereby making credible the possibility of an injection. Participants were seated in an office equipped with medical supplies and approached with a needle (weapon-present condition) or a pen (weapon-absent condition) and then threatened with an impending injection or not threatened at all. Participants in the weapon-present condition (compared to the control condition) exhibited greater recall of the carried item and worse performance on a suspect-absent photo line-up; neither weapon presence nor threat had any effect on recall for facial details and threat did not affect suspect identification (Maass & Kohnken, 1989). This finding was supported by work conducted by Peters (1988) who employed essentially the same procedure with the exception that each participant actually received an injection. Evidence of a weapon focus effect was evaluated by comparing memory for the nurse that administered the injection to memory for a researcher that later administered a questionnaire. A physical description and identification attempt (using a target present line-up) was less accurate for the nurse than for the researcher. In this study accuracy was negatively correlated with measurements of trait anxiety suggesting a potential role for arousal.

It was almost 20 years before Pickel et al. (2006) published the next simulation study, this time investigating whether awareness of the weapon focus effect could diminish its magnitude. These researchers seated participants in a small classroom and informed them that they were about to observe a short scene portrayed by actors. However, participants were first provided with one of two brief lectures: (a) a lecture regarding weapon focus, and how important it is to attend to perpetrator features instead of dwelling on any weapons they may carry, or (b) a lecture regarding eyewitness confidence and perceived credibility. In either case, a man interrupted the lecture by bursting into the classroom bearing a neutral object (a book) or a weapon (gun). The main finding was that participants presented with the lecture on eyewitness confidence produced fewer correct details (and more incorrect details) related to the perpetrator in the weapon condition than in the neutral object condition. However, object type had no effect on those presented with the weapon focus lecture. Pickel et al. (2006) concluded that with proper instruction, the weapon focus effect could be overcome.

Around the same time, Hulse and Memon (2006) published a similar study investigating whether the training and exposure inherent to the police force was sufficient to overcome the weapon focus effect. Officers were briefed regarding either a potential shooting (i.e. weapon-present condition) or domestic disturbance (i.e. weapon-absent condition) and then brought into a simulation room where the appropriate disturbance was displayed on a large video screen. Results indicated that officers in the weapon-present condition exhibited less complete but more accurate recall of the event (possibly reflecting a conservative reporting bias due to the severity of the crime). However, weapon presence had no effect on the officers' description or identification of the perpetrator (Hulse & Memon, 2006). Importantly, the perpetrator depicted in the weapon condition was visible both before and after the weapon providing an opportunity to encode their features unhampered by the weapon's presence (see Pickel, 1998).

Actual crimes

The degree to which laboratory or simulation studies are capable of emulating real crimes has been brought into question by many researchers (e.g. Cutshall & Yuille, 1989; Tollestrup, Turtle, & Yuille, 1994). It is possible that participants in traditional research studies engage different strategies and behavioral patterns than do eyewitnesses to criminal events. As a result, the weapon focus effect may manifest differently in the 'real world' (Cutshall & Yuille, 1989). Certainly, the personal relevance of watching a video is limited in its comparison to witnessing a crime such as an armed robbery (e.g. Tollestrup et al., 1994; van der Kolk, 1996). In some cases the only witness is the victim, which is a very different experience from that of a passive observer in the safety of a laboratory (Cutshall & Yuille, 1989). It is for this reason that data from actual criminal events can be helpful in terms of understanding the applicability of the weapon focus effect to the real world. Our categorization of actual crimes included field studies where witnesses were interviewed following a criminal event, as well as archival research based on reviews of historical case file information collected by police officers. In general, neither field nor archival studies have reported an effect of weapon presence on suspect identification or description accuracy.

Tollestrup et al. (1994) conducted one of the first archival analyses to explore weapon presence as a factor (see also Kuehn, 1974). Their sample included robbery and fraud cases committed in Vancouver, British Columbia between 1987 and 1989. Prior to exploring the impact of weapon presence on eyewitness accounts, Tollestrup et al. (1994) categorized each crime into one of three evidentiary categories. These categories ranged from negligible evidence to a full confession by the suspect. However, it is unclear from their report whether these evidentiary categories had any bearing upon their analysis of weapon presence. Eyewitness accounts were more complete if a weapon was involved. Weapon presence did not affect feature accuracy when compared against a description of the individual eventually charged with the crime and only marginally decreased suspect identification.

Almost a decade later Behrman and Davey (2001) conducted the next archival analysis of the weapon focus effect using police records pertaining to armed robberies and other felonies investigated by the Sacramento Police Department between 1987 and 1998. Prior to evaluating the impact of weapon presence on eyewitness accounts, these researchers first categorized each crime on the basis of extrinsic evidence of the suspect's guilt (see Tollestrup et al., 1994). The effect of weapon presence was evaluated using photographic line-ups as well as live field show-ups. Weapon presence did not significantly impact identification accuracy although the effect was in the appropriate direction for both line-up procedures once evidentiary category was collapsed.

Around the same time Pike et al. (2002) published a report for the British Home Office describing data collected from 14 separate police forces. The effect of weapon presence on identification accuracy was evaluated but did not significantly impair performance. Like Tollestrup et al. (1994) and Behrman and Davey (2001) the analysis of weapon presence was conducted largely in the context of robberies or related crimes (e.g. burglary); however, unlike those other reports Pike et al.'s (2002) sample was comprised almost entirely of live as opposed to photo line-ups. Soon after a similar governmental report was published by Mecklenburg (2006) who – instead of resorting to archival data – had police officers complete a brief form summarizing the circumstances surrounding each identification attempt soon after that attempt was made. Once again no effect of weapon presence was observed on identification accuracy. This was also true of a study conducted by Valentine et al. (2003) employing a very similar questionnaire approach within the Greater London area between January and September 2000: Weapon presence had no impact on the probability of selecting the police suspect in a live line-up, and even *decreased* the probability of selecting a foil in their exploratory analysis (this effect failed to reach significance in their global model).

Wagstaff, MacVeigh, Scott, Brunas-Wagstaff, and Cole (2003) adopted a different approach. Instead of evaluating the impact of weapon presence on identification accuracy they coded police interviews taken from witnesses or victims of robberies, assaults and rapes investigated by two separate police forces in Britain (see also Tollestrup et al., 1994). These interviews were compared against a police description of the primary suspect at the time of their arrest. Again no evidence was found of any effect of weapon presence on feature accuracy.

Finally, Cooper et al. (2002) conducted a field investigation of weapon focus for memories of sexual assault in prostitutes. During an interview, participants were instructed to recall a memory of sexual assault in as much detail as possible. Features of the narrative, such as the number of unique details recalled and weapon presence, were coded by the researchers. Their results did not yield any significant differences in the number of details recalled by prostitutes who were sexually assaulted in weapon-present versus weapon-absent conditions. The lack of a weapon focus effect may have resulted from the nature of the crime (i.e. sexual assault). The proximity of the perpetrator to the victim during the assault may have removed the weapon from their field of vision, mitigating its effects (see Pickel, 2007) or if the assault occurred in the context of a sexual transaction gone awry, it is also possible that the victim may have observed (or even known) their transgressors prior to the production of a weapon. These issues as well as other potential confounds (e.g. a population potentially more accustomed to the presence of weapons) make it difficult to draw strong conclusions from Cooper et al.'s (2002) findings.

Summary

The state of laboratory and simulation research dealing with weapon focus is promising. Given its relevance to the criminal justice system, empirical studies need to address why weapon presence appears to be clearly detrimental to feature accuracy (e.g. recall), but only sporadically affects identification accuracy. In response to this question, both Steblay (1992) and Pickel (2007) have suggested that the recall of descriptive details is simply a more sensitive measure than suspect identification. In addition, the mechanisms underlying weapon focus remain unclear, although recent findings that unusualness and context play a role are suggestive of the unusual item hypothesis (e.g. Mitchell et al., 1998; Pickel, 1998, 1999; Pickel, French, & Betts, 2003). Importantly, the precise role of arousal requires the use of objective as opposed to subjective methods of quantifying these constructs (e.g. Hulse & Memon, 2006; Peters, 1988).

The current state of real-world investigations dealing with weapon focus is much less promising. Several archival and field studies were reviewed (including two major governmental reports) none of which produced compelling evidence of the memory impairments observed in the laboratory. These studies differ from their laboratory and simulated counterparts in several ways. Measurements of feature accuracy must be interpreted cautiously due to complications with knowing the 'ground truth' of an event. This is problematic without external verification, especially because traumatic memories are known to exhibit greater detail, vividness, and consistency than non-traumatic memories (e.g. Peace & Porter, 2004; Porter & Peace, 2007).

Many researchers that evaluate memory in applied and forensic contexts have expressed concern over the lack of ground truth when providing accuracy estimates (e.g. Kuehn, 1974; Malpass, 2006a; Steblay, 2010; Wells, 2008). While measures of completeness (rather than accuracy) are one way to address this variable (e.g. Cooper et al., 2002), recent advancements in technology may assist in providing a degree of external evidence that can be used to substantiate that events occurred as recollected. For example, recordings of public locations captured on high-quality closed circuit television (CCTV) may provide ground truth and provide a further avenue of investigation for the weapon focus effect and/or eyewitness identifications (e.g. Davis & Valentine, 2009). Another method of establishing ground truth in field or archival studies is to categorize cases according to the availability and extensiveness of external corroborative evidence (see Malpass, 2006b; Ross & Malpass, 2008). This approach has been employed within the line-up identification literature and divides cases via the absence of extrinsic evidence relative to the presence of either minimal or substantive incriminating evidence (see Behrman & Davey, 2001; Tollestrup et al., 1994). Future research employing techniques to establish ground truth of events used in applied forensic research should follow the recommendations outlined above.

Despite these concerns, field studies serve a crucial role in generalizing laboratory findings as well as identifying important research questions otherwise neglected in the literature. Archival research on the other hand may provide a rich source of data, but is inherently limited by those who initially recorded details of the crime. Conclusions from both field and archival studies are often obscured by the lack of control over variables such as crime duration, witness stress, intoxication, familiarity of setting, level of questioning, and so forth. Laboratory studies are useful to tease apart the relative influence of such factors. In addition to informing laboratory research of possible variables requiring further attention, real-world investigations may also be informed by laboratory research. For example, knowing that a weapon or unusual object is likely to attract and maintain the gaze of an observer (e.g. Loftus et al., 1987), field researchers may be wise to ask witnesses to retrospectively describe their focus of attention during the event in question (e.g. on the perpetrator, the weapon, or an escape route). Self-reported attentional focus may then be used to predict the accuracy of the witness's account in the context of a field investigation and perhaps ultimately in a court of law.

Meta-analytical support of the weapon focus effect

The absence of a weapon focus effect in studies of actual criminal events may be interpreted multiple ways. It is possible that weapons make criminals feel less vulnerable and therefore act less cautiously, spend more time at the crime scene, or venture closer to eyewitnesses thus diminishing the weapon focus effect (Pickel, 1998; also see Pickel, 2007). Alternately, it could be that actual crimes are too complex, with too many other influential factors for the presence of a weapon to have a significant impact on eyewitness memory (Steblay, 1992). Another alternative is that weapon presence *does* affect eyewitness memory – only this difference is statistically obscured by the complexities alluded to above. To explore this possibility we will next conduct a meta-analysis to evaluate how systematic differences between studies could account for fluctuations in the magnitude of the weapon focus effect.

Method

Literature search

We conducted a search of the online resources *Google, Google Scholar, PsycINFO, PsychARTICLES*, and *JSTOR* using various combinations of the keywords: *weapon, focus, gun, knife, attention, memory, review, meta-analysis, eyewitness, forensic, archival, robbery*, and *crime*. This search was conducted periodically until September 2010 and supplemented by (a) articles referenced in the obtained sources, and (b) articles collected by contacting the primary authors of the obtained sources. Only articles containing comparisons fitting the definition of weapon focus provided above were considered for inclusion. Of the 34 articles identified fitting our description, two were excluded on the basis of insufficient information to estimate an appropriate effect size for a viable measure of feature or identification accuracy (Kuehn, 1974; Shaw & Skolnick, 1994). Another article was excluded on the basis that it used a special population familiar with weapons and responding to events of a violent criminal nature (police officers: Hulse & Memon, 2006). Two further studies were excluded from our primary analysis (but included in a later sub-analysis) because they measured completeness instead of feature or identification accuracy (see below; Cooper et al., 2002; Cutshall & Yuille, 1989). Comprehensive notes regarding effect size calculations are available from the first author upon request. From the remaining 28 studies we calculated 67 effect sizes – although 20 of these effect sizes were dependent upon each other (i.e. drawn from the same sample) and therefore collapsed for our overall analysis. This resulted in the 47 effect sizes summarized as a forest plot in Figure 1. Articles contributing one or more effect sizes to any of the reported analyses are indicated in our reference section by an asterisk (*).

Coding guidelines

Prior to conducting our analysis we coded each comparison in terms of six theoretically motivated moderator variables using the definitions provided below. Comparisons using weapons and those using unusual objects (or manipulating context such as to manipulate the unusualness of an object) were kept separate from each other (see below). The first author initially coded each study – however, whenever a coding decision was not readily apparent given the provided definition, two additional coders (co-authors) were consulted.

Author(s)	Year	Experiment	Type	Retention Interval	Threat	Exposure Duration	Standardized Mean Difference [95% CI]
Tollestrup et al.	1994	Exp. 1	Actual Crime	Long	High	Long	1.20 [0.11 , 2.28]
Behrman and Davey	2001	Photo Lineups	Actual Crime	Long	High	Long	0.08 [−0.30 , 0.47]
Behrman and Davey	2001	Field Lineups	Actual Crime	Long	High	Long	0.32 [−0.40 , 1.03]
Pike et al.	2002	Exp. 1	Actual Crime	Long	High	Long	0.20 [0.04 , 0.36]
Valentine et al.	2003	Exp. 1	Actual Crime	Long	High	Long	−0.03 [−0.36 , 0.29]
Wagstaff et al.	2003	Exp. 1	Actual Crime	Long	High	Long	0.23 [−0.25 , 0.71]
Mecklenburg	2006	Exp. 1	Actual Crime	Long	High	Long	0.09 [−0.14 , 0.32]
Cutler et al.	1986	Exp. 1	Laboratory	Short	High	Long	0.00 [−0.28 , 0.28]
Cutler et al.	1986	Exp. 2	Laboratory	Long	High	Long	0.28 [−0.02 , 0.58]
Loftus et al.	1987	Exp. 1	Laboratory	Short	High	Short	0.78 [−0.08 , 1.64]
Loftus et al.	1987	Exp. 2	Laboratory	Short	High	Short	0.58 [0.04 , 1.13]
Tooley et al.	1987	Exp. 1	Laboratory	Immediate	Low	Intermediate	0.25 [−0.16 , 0.65]
Cutler et al.	1987a	Exp. 1	Laboratory	Long	High	Long	0.56 [0.16 , 0.96]
Cutler et al.	1987b	Exp. 1	Laboratory	Long	High	Long	0.04 [−0.25 , 0.33]
Cutler and Penrod	1988	Exp. 1	Laboratory	Long	High	Long	−0.17 [−0.57 , 0.24]
O'Rourke et al.	1989	Exp. 1	Laboratory	Long	High	Long	0.48 [0.03 , 0.93]
Kramer et al.	1990	Exp. 1	Laboratory	Immediate	High	Intermediate	0.83 [−0.12 , 1.78]
Kramer et al.	1990	Exp. 2	Laboratory	Immediate	Low	Intermediate	0.69 [0.11 , 1.27]
Kramer et al.	1990	Exp. 3	Laboratory	Immediate	Low	Intermediate	0.79 [−0.06 , 1.63]
Kramer et al.	1990	Exp. 4	Laboratory	Immediate	Low	Intermediate	0.76 [0.09 , 1.42]
Kramer et al.	1990	Exp. 5	Laboratory	Immediate	Low	Short	0.11 [−0.58 , 0.80]
Mitchell et al.	1998	Exp. 1	Laboratory	Immediate	Low	Short	−0.05 [−0.67 , 0.57]
Mitchell et al.	1998	Exp. 1	Laboratory	Immediate	High	Short	0.49 [−0.13 , 1.11]
Mitchell et al.	1998	Exp. 2	Laboratory	Immediate	Low	Short	0.31 [−0.13 , 0.75]
Mitchell et al.	1998	Exp. 2	Laboratory	Immediate	High	Short	0.55 [0.10 , 0.99]
Pickel	1998	Exp. 1	Laboratory	Short	High	Intermediate	0.45 [−0.02 , 0.93]
Pickel	1998	Exp. 2	Laboratory	Short	High	Intermediate	0.62 [0.16 , 1.08]
Pickel	1999	Exp. 1	Laboratory	Short	Low	Intermediate	0.51 [0.14 , 0.88]
Pickel	1999	Exp. 1	Laboratory	Short	High	Intermediate	0.29 [−0.06 , 0.65]
Pickel	1999	Exp. 2	Laboratory	Short	Low	Intermediate	0.23 [−0.19 , 0.65]
Shaw and Skolnick	1999	Exp. 1	Laboratory	Short	High	Intermediate	0.14 [−0.24 , 0.52]
Pickel et al.	2003	Exp. 1	Laboratory	Short	High	Intermediate	0.34 [−0.09 , 0.77]
Pickel et al.	2003	Exp. 2	Laboratory	Short	High	Long	0.24 [−0.26 , 0.74]
Hope and Wright	2007	Exp. 1	Laboratory	Short	High	Short	1.10 [0.33 , 1.88]
Pickel et al.	2008	Exp. 1	Laboratory	Immediate	High	Intermediate	1.00 [0.74 , 1.26]
Davies et al.	2008	Exp. 1	Laboratory	Short	Low	Intermediate	0.99 [0.33 , 1.64]
Pickel	2009	Exp. 1	Laboratory	Immediate	High	Intermediate	1.78 [1.37 , 2.19]
Pickel	2009	Exp. 2	Laboratory	Immediate	High	Intermediate	1.38 [1.04 , 1.72]
Pickel	2009	Exp. 3	Laboratory	Immediate	High	Intermediate	1.08 [0.82 , 1.35]
Saunders	2009	Exp. 1	Laboratory	Immediate	High	Short	0.87 [0.58 , 1.16]
Saunders	2009	Exp. 2	Laboratory	Immediate	High	Short	0.80 [0.53 , 1.06]
Johnson and Scott	1976	Exp. 1	Simulation	Long	High	Short	0.62 [−0.22 , 1.46]
Peters	1988	Exp. 1	Simulation	Long	High	Intermediate	0.45 [0.11 , 0.79]
Maass and Kohnken	1989	Exp. 1	Simulation	Short	Low	Intermediate	0.84 [0.07 , 1.61]
Maass and Kohnken	1989	Exp. 1	Simulation	Short	High	Intermediate	0.36 [−0.41 , 1.13]
Pickel et al.	2006	Exp. 1	Simulation	Immediate	High	Intermediate	1.26 [0.78 , 1.74]
Pickel et al.	2006	Exp. 2	Simulation	Immediate	High	Intermediate	1.28 [0.70 , 1.85]
Actual Crimes							**0.15 [0.04 , 0.26]**
Laboratory							**0.55 [0.40 , 0.70]**
Simulation							**0.82 [0.46 , 1.17]**
OVERALL							**0.53 [0.40 , 0.66]**

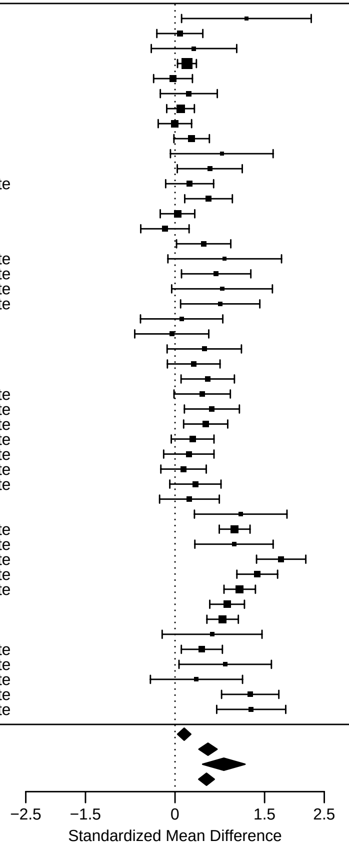

Figure 1. Aggregate effect sizes and confidence intervals as a function of experiment arranged into a forest plot. Polygons are provided depicting the estimated effect size as calculated by separate random-effects models applied to each level of study type, as well as to the overall data.

Study type

Each effect size was initially coded to reflect the nature of the research from which it arose using the definitions provided above.

Dependent variable

Each effect size was next coded to reflect the nature of the task from which it arose. Feature accuracy was defined as any measure testing participants for the details of the event such that their performance was rewarded for producing correct responses and penalized for producing incorrect responses. Common tasks falling within this category included multiple-choice and fill-in-the-blank questionnaires as well as recall – so long as errors were included in the measure (e.g. by subtracting the number of incorrect details from the number of correct details). Identification accuracy was defined more narrowly as the application of a simulated or real forensic line-up. In this regard it did not matter whether the line-up was conducted in-person or using photographs or whether the target/suspect was present or absent. Finally, Completeness was defined as any measure testing participants for the details of the event such that errors were *not* taken into account. Tasks falling within this category were invariably recall tasks or interviews wherein the provided details were counted but not scored for accuracy.

Retention interval

The time elapsed between the offset of the studied material and the onset of the associated memory task was initially coded as a continuous variable but was then converted into an ordinal scale. This conversion was conducted to permit the inclusion of studies for which the precise delay was not known but for which an approximation could be made. In a small number of instances multiple dependent effect sizes having different retention intervals were combined for the purpose of analysis. In these instances the average interval was computed and used for the purpose of determining the ordinal value. Retention interval was coded as (a) *immediate* for an interval of less than 10 min, (b) *short* for an interval of more than 10 min but less than 24 h, or (c) *long* for an interval of 24 h or more. Thus, our handling of this variable is comparable to a logarithmic transformation of time grouped into three levels. Therefore, our analysis treats retention interval as a continuous variable.

Threat

Comparisons were coded as high in threat if there was real, perceived, potential, or implied bodily harm to the participant, actors, or witnesses involved in or (in the case of laboratory stimuli) portraying the studied material. Otherwise comparisons were coded as low in threat. The presence of a weapon was not considered to warrant potential bodily harm so long as it was not presented in a menacing or criminal manner (Pickel, 1998) or if assurances were given that the weapon was not to be used for the purpose of violence (Maass & Kohnken, 1989).

Exposure duration

The duration for which the participant was exposed to the critical object was initially coded as a continuous variable but was then converted into an ordinal scale. Much like retention interval this conversion was conducted to permit the inclusion of studies for which the precise duration of exposure was not known but for which an approximate estimate could be made. Exposure duration was coded as (a) *short* if exposure lasted 10 s or less, (b) *intermediate* if exposure lasted between 10 and 60 s, or (c) *long* if exposure lasted for more than 60 s. As with our treatment of the ordinal delay variable above, our coding of exposure duration is likewise comparable to a logarithmic transformation of time and was therefore treated as a continuous variable. Because exposure duration was not reported for actual criminal events we instead assumed that the weapon had been presented for at least 60 s in these crimes. This assumption is supported by reports indicating the average duration of an armed robbery (upon which much of the archival and field research is based) to be somewhere between one and three min (e.g. Gray, 1971; Mastrobuoni, 2010; Mastrobuoni, 2011).

Object type

As described in detail above, a relatively small but growing number of studies have investigated the contribution of unusualness to the weapon focus effect. Scenarios were assigned to the *weapon* condition if they contained an object commonly associated with injury (e.g. handgun) or pain (e.g. syringe). Scenarios were assigned to the *unusual* condition if they instead contained an object that was unexpected given the context in which it was presented. Many of the experiments having manipulated unusualness also included a more traditional weapon condition. In most cases both the unusual and weapon conditions were dependent upon the same control condition and therefore sample. As a result both novel and weapon comparisons could not be included in the same analysis forcing us to conduct a separate analysis for this variable.

Effect size calculation and analysis

Effect sizes for feature accuracy and completeness were calculated as standardized mean differences using the *escalc* function from the *metafor* package (Viechtbauer, 2010) within R version 2.12.1 (R Development Core Team, 2010). This function employs the procedure recommended by Hedges (1982) with a correction for positive bias (see Hedges & Olkin, 1985). Due to its dichotomous nature, a measure of effect size for identification accuracy was needed that was comparable to the standardized mean difference. A probit transform was chosen (see Glass, McGaw, & Smith, 1981) with the sampling variance estimated using the procedure recommended by Sánchez-Meca, Marín-Martínez, and Chacón-Moscoso (2003, see Equation (21)). While some meta-analysts have opted to segregate continuous and dichotomous dependent variables (e.g. Hazell, O'Connell, Heathcote, Robertson, & Henry, 1995) recent theorists have recognized the benefits of aggregating these data into a single analysis (see Chinn, 2000; Lipsey & Wilson, 2000; Sánchez-Meca et al., 2003; Whitehead, Bailey, & Elbourne, 1999). Most prominently, aggregation maximizes statistical power and increases the range of potential moderators producing more robust predictions. For these reasons we have opted to adopt this approach. However, whereas both feature and identification accuracy may be viewed as addressing the fidelity of an eyewitness account, completeness instead speaks only to the vividness of this account. Accuracy and vividness are separable and often unrelated in memory research (e.g. Koriat, Goldsmith, & Pansky, 2000; Porter & Peace, 2007); given our interest in how weapon presence impacts eyewitness accuracy we have therefore aggregated only measures of feature and identification accuracy. Completeness was analyzed in a separate sub-analysis. Unless otherwise noted all effect sizes have been calculated such that a positive value represents greater performance in the *control* as opposed to the *weapon* condition. Therefore, higher (positive) effect sizes represent a larger weapon focus effect.

Once effect sizes were computed and aggregated, a random-effects model was fitted to the overall data to estimate the impact of weapon presence on memory performance. Mixed-effects models were then fitted to the data to test each of the moderator variables. Prior to developing our model each moderator variable was centered by subtracting the mean of that variable. Both the random- and mixed-effects models were generated using the *rma* function from the *metafor* package. While developing our model we employed a maximum likelihood estimator of the heterogeneity of variance within our model; however, once our model was completed we switched to a restricted maximum-likelihood estimator instead, which is what we report below. Restricted maximum-likelihood estimates are less biased than maximum-likelihood estimates – however, the latter was required to permit comparison between models as we added and removed moderators (see Viechtbauer, 2010).

Results

Overall effect of weapon presence on memory performance

Results indicated a moderate effect of weapon presence within the aggregate data, $g = 0.53$, $CI_{95\%} = [0.40, 0.66]$, although our model also indicated a significant degree of heterogeneity across measures, $Q(46) = 223.48$, $p < 0.01$. These data correspond to the bottom polygon provided in Figure 1. Having established evidence of a weapon focus effect, we next calculated a fail-safe N to estimate the number of additional findings averaging to null ($g = 0$) needed for the effect of weapon presence to become non-significant (see Rosenthal, 1978). The outcome of this analysis suggested that 4552 additional comparisons unsupportive of this effect would need to be included in our analysis to change the outcome from significant to non-significant. This is over 100 times the number of known articles on this topic. We next considered our moderators – the model is presented in Table 1.

Retention interval was found to account for a significant portion of the heterogeneity within the current data, $b = -0.27$, $CI_{95\%} = [-0.45, -0.09]$. As retention interval increased effect size decreased resulting in a large effect when the test followed soon after the studied material, $g = 0.91$, a moderate to large effect after a short delay, $g = 0.66$, and a small to moderate effect after a long delay, $g = 0.38$ (see Figure 2a). This pattern could result from either a gradual increase in performance within the weapon condition or a gradual decrease in performance

Table 1. The final model for our meta-analysis of the aggregate effect sizes. Moderators were centered prior to analysis (see text for details). Reported values have been rounded to two decimal places.

Variable	*b*	SE	*Z*	CI	
Intercept	0.67	0.07	9.07	0.52,	0.81
Threat	0.39	0.13	3.05	0.14,	0.64
Retention interval	−0.27	0.09	−3.01	−0.45,	−0.09
Exposure duration	−0.08	0.10	−0.78	−0.27,	0.12
Exposure duration^2	−0.29	0.11	−0.29	−0.50,	−0.08

$Q_{\text{Moderators}}(4) = 46.68$, $p < 0.01$; $Q_{\text{Error}}(42) = 84.26$, $p < 0.01$.

within the control condition as time elapses. This issue is difficult to disentangle within the context of the current data – too few experiments systematically varied the time between study and test to evaluate performance in the control and experimental conditions. Comparison of raw performance measures between studies is complicated because these measures differ in their respective ranges even within the same dependent variable.

The main effect of *Threat* also reached significance, $b = 0.39$, $CI_{95\%} = [0.14, 0.64]$, with larger effect sizes in the context of threatening, $g = 0.76$, as opposed to non-threatening scenarios, $g = 0.37$ (see Figure 2b). This finding offers provisional support for the role of threat as predicted by the arousal/threat hypothesis described above. There was a significant relation between *exposure duration* and effect size: The first-order polynomial failed to reach significance, $b = -0.08$, $CI_{95\%} = [-0.27, 0.12]$, but was qualified by a significant effect of the second-order polynomial, $b = -0.29$, $CI_{95\%} = [-0.50, -0.08]$. Effect sizes were largest at the intermediate exposure duration, $g = 0.67$, with smaller effect sizes when the weapon was presented for an especially short period of time, $g = 0.42$, or long period of time, $g = 0.34$ (see Figure 2c). This pattern suggests that the effect builds with exposure until it reaches some maximal value at which point the impact of weapon presence begins to dissipate. The precise nature of the processes behind this relationship is unclear at this time. It could be that a weapon requires a certain period of exposure to ensure that the weapon is perceived and captures attention. With prolonged exposure the shock of the weapon could fade and the viewer could then begin directing their attention away from the weapon to get a better analysis of their surroundings. This relationship remains a mystery at this time, but we will discuss it further below.

Our full model is presented in Table 1 and although it accounts for 75% of the heterogeneity within our results a significant amount of unexplained variance remains, $Q(42) = 84.26$, $p < 0.01$. Figure 3 depicts a funnel plot characterizing this variance. This plot presents the standard error associated with each effect size (in descending order) as a function of the residual value calculated within our full model. It is expected that, as the standard error goes down (higher in the plot), residuals will become smaller because the sample estimate will tend toward the true value. Presuming that our residual values are approximately normally distributed this should result in a pattern similar to a funnel in shape. Overall our residual values conform nicely to the expected shape with only two studies falling more than two standard errors from the predicted values: one such study fell above the predicted value (Pickel, 2009) and another below the predicted value (Shaw & Skolnick, 1999).

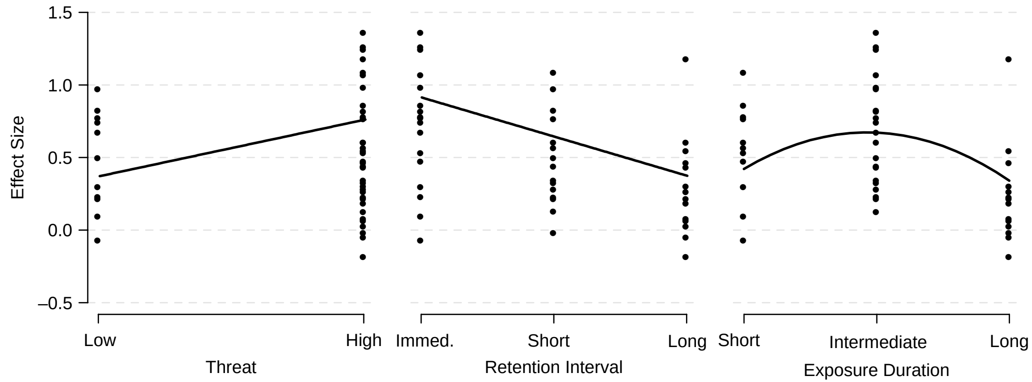

Figure 2. Aggregate effect sizes plotted as a function of (a) threat, (b) retention interval, and (c) exposure duration. Lines represent the predicted values derived from our meta-analytical model (see Table 1).

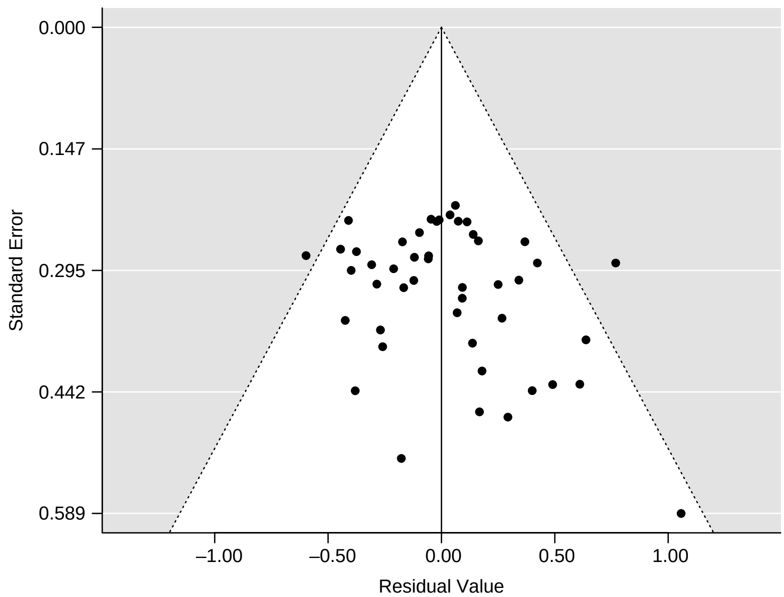

Figure 3. Funnel plot depicting the relationship between standard error and residuals in our model. While we did have significant heterogeneity in the full model it can be seen that all but two of the studies (Pickel, 2009; Shaw & Skolnick, 1999) are less than two standard errors from the predicted value.

Figure 3 further demonstrates our residuals to be roughly symmetrical around 0. This assuages any concern that our analysis might have inadvertently excluded a large body of unpublished articles that had failed to obtain significance. This form of publication bias occurs within a given literature when it is difficult to publish non-significant findings and results in an overestimation of the mean effect by eliminating studies having observed a small or even reversed difference. When present this bias should manifest within the funnel plot as an asymmetry favoring positive outcomes, most notably in the lower regions of the plot. This does not appear to be the case and indeed due to the controversial nature of weapon focus many studies have been published reporting non-significant results.

Although initially included in our model, the effect of *study type* accounted for less than 2% of the heterogeneity observed within our data once other moderators were included; overall this effect failed to reach significance and was removed from our model. Nonetheless, due to the relevance of this variable to public policy and the applicability of weapon focus to actual criminal events we calculated separate effect sizes for each category of study. The polygons provided at the bottom of Figure 1 demonstrate this effect to be significant in each. This outcome is surprising given the summary of real-world studies provided above – it has been general knowledge within this literature that whereas weapon focus was readily observed in the laboratory there have been no reports to date of a negative effect of weapon presence on actual eyewitness performance. However, glancing at Figure 1 demonstrates that in almost every reported case the effect has been present – it has simply been too small to evaluate in the context of a single experiment.

An astute reader may argue that study type does appear to impact the magnitude of weapon focus and that this impact is demonstrated by comparing the mean effects for each type presented in Figure 1. The effect in studies of actual crimes ($g = 0.15$) is clearly lower than that representing laboratory ($g = 0.55$) or simulation studies ($g = 0.82$), and the confidence intervals do not overlap. An analysis with only this variable included as a moderator shows a significant effect of study type explaining 12% of the heterogeneity within those results. The effect of study type vanishes as soon as any of the moderators presented above are added to the model. This is because they are better predictors of the variability within our data. One must keep in mind that the polygons presented in Figure 1 were calculated in the absence of any moderator variables only to demonstrate the aggregate effect of weapon presence within those studies. Given our model, one would expect lower effect sizes in studies of actual criminal events on the basis that these studies are characterized by both long weapon exposure durations and retention intervals (see above). Importantly, the mean effect size observed for laboratory studies with similar properties ($g = 0.22$) is comparable to those observed in the real-world studies ($g = 0.15$) included in our model. Further, the effect was numerically largest for simulation studies – which apply the control inherent in laboratory work to real events experienced in a naturalistic setting. These findings suggest that the current difficulty associated with identifying a strong weapon focus effect in the real world is in large part due to the nature of criminal events: Witnesses are often exposed to the perpetrator and the weapon for a protracted period and are not asked to identify the suspect for some time after the event itself.

Comparing dependent variables

Having completed our primary analysis of the weapon focus effect, separate analyses were conducted to estimate its magnitude within each of the dependent categories. Separating the data in this manner reduced the overall range of each moderator within a given subset preventing us from conducting a full moderator analysis as undertaken above. Therefore each of the analyses described below represent separate random-effects models.

Feature accuracy

Thirty-three of the calculated effect sizes used measures of feature accuracy ranging from cued-recall questions or scored interviews to object identification. This analysis demonstrated a sizable effect of weapon presence, $g = 0.75$, $CI_{95\%} = [0.60, 0.89]$, although as before our model also indicated a significant degree of heterogeneity across measures, $Q(32) = 116.12$, $p < 0.01$. We computed a separate fail-safe N for this analysis that suggested 4803 additional comparisons unsupportive of this effect would need to be included to change the outcome from significant to non-significant.

Identification accuracy

Thirty-three of the calculated effect sizes used measures of identification accuracy. These measures included a variety of line-up procedures (e.g. photos, etc.) and both suspect present and absent arrays. To ease the comparison between feature and identification accuracy we opted to maintain the probit-transformed response data instead of using a more traditional metric such as log-odds ratio – both analyses lead to precisely the same conclusion and fail-safe *N*. Identification accuracy reveals only a small to moderate effect lower in magnitude than that observed for feature accuracy, $g = 0.22$, $CI_{95\%} = [0.13, 0.32]$. Unlike our previous models, heterogeneity across measures did not vary significantly, $Q(32) = 44.61$, $p > 0.05$. The finding of a smaller effect for identification accuracy is not surprising and has been found by an earlier meta-analysis of this literature (see Steblay, 1992). The fail-safe *N* for this analysis suggests that 431 additional comparisons unsupportive of this effect would need to be included to change the outcome from significant to non-significant.

Completeness

Nine of the calculated effect sizes used measures of completeness. These measures typically included interviews wherein the number of details produced were counted but not scored for accuracy. In several of the laboratory studies we are also able to calculate completeness by adding the number of correct details recalled to the number of incorrect details recalled instead of subtracting them (Johnson & Scott, 1976; Pickel, 2009; Pickel et al., 2006). These values are included in this analysis. There was no effect of weapon presence on the number of details produced, $g = -0.04$, $CI_{95\%} = [-0.35, 0.28]$, although there was significant heterogeneity amongst the included effect sizes, $Q(8) = 42.47$; $p < 0.05$. We have insufficient data to properly explore any moderator variables, but striking is that all real world and simulation studies ($n = 4$) using completeness tended towards more details produced in the weapon condition ($g = -0.55$). The laboratory studies ($n = 5$) tended towards fewer details produced in the weapon condition ($g = 0.33$). These trends may suggest that completeness is sensitive to the nature of the experiment: in laboratory studies there is little pressure to produce especially detailed accounts (especially in a typical recall task) whereas in a criminal investigation more details may be encouraged, particularly when a weapon is involved. Alternately, it is possible that completeness is sensitive to one or more of the moderator variables that tend to vary between real world and laboratory investigations. While research concerning memory for stimuli that induce negative arousal has been conducted (e.g. Christianson & Hübinette, 1993; Kensinger, Garoff-Eaton, & Schacter, 2006), there are insufficient studies exploring how *weapon* presence influences the *vividness* of a criminal event to make any definitive determination at this time. Moreover, these analyses should be considered cautiously due to our relatively small sample.

The role of threat

The role of threat has already been investigated between studies in the analysis presented above. A more direct approach with regards to disentangling the contribution of threat would be to calculate an effect size comparing a weapon condition in which no one was threatened or injured to a weapon condition in which someone *was* threatened or injured. Four experiments specifically manipulated threat independently of weapon presence, permitting just such an analysis. In using the threatening and non-threatening scenarios as the respective control and experimental conditions we ensured that as in the above analyses an effect of threat would manifest as a positive value. This analysis revealed a non-significant trend where the weapon focus effect was larger in the context of a threatening scenario, $g = 0.14$, $CI_{95\%} = [-0.21, 0.48]$. The heterogeneity amongst these scores failed to reach significance, $Q(3) = 5.00$; $p > 0.05$. Further studies are required explicitly manipulating this variable to develop a better notion as to its contributions.

The role of unusualness

The contribution of unusualness to the weapon focus effect was investigated using two separate approaches. First, we fit a random-effects model to those studies manipulating unusualness to estimate the relation between the presence of an unusual object and memory performance. Six comparisons were included in this model, producing a moderate effect, $g = 0.40$, $CI_{95\%} = [0.14, 0.65]$, without significant heterogeneity amongst the included measures, $Q(5) = 8.17$, $p > 0.05$. This effect is slightly smaller than that observed in the overall analysis of weapon presence presented above. Second, we approached this issue in the same manner we approached the role of threat in the preceding section. We calculated a new measure of effect size comparing the unusual object and weapon conditions. This comparison specifically addressed the question: Is there something special about a weapon that negatively impacts performance even when controlling for unusualness? If this were the case, we would expect a positive effect representing lower performance in the weapon condition. The six comparisons included in our original analysis were recalculated in this manner and joined with three additional comparisons that had been excluded from that analysis because their manipulation of unusualness involved a weapon (e.g. a priest with a gun vs a police officer with a gun; Pickel, 1999). This analysis revealed a non-significant tendency for performance to be *higher* in the weapon as opposed to unusual object condition, $g = -0.25$, $CI_{95\%} = [-0.63, 0.12]$. This time the heterogeneity in the model was significant, $Q(8) = 52.88$, $p < 0.05$. This non-significant effect combined with the direction of the trend provides provisional support for the position that unusualness is capable of producing an effect of similar magnitude to that of a weapon.

General discussion

The current paper sought to evaluate the extant literature regarding the impact of weapon presence on eyewitness memory. Our narrative review was largely in line with modern conceptualizations of this effect: weapon presence has consistently demonstrated a negative effect on both feature accuracy and identification accuracy under controlled conditions (see Steblay, 1992; Pickel, 2007). It has only once demonstrated even a marginally negative effect on actual suspect identification (Tollestrup et al., 1994) and has even been shown to decrease false alarms in another preliminary analysis (Valentine et al., 2003). If we were to stop there we might have concluded that weapon focus was limited to laboratory and simulation experiments as opposed to actual criminal events (see Cooper et al., 2002). This position has gained some support over the years as both field and archival studies have failed to produce a reliable effect of weapon presence (e.g. Cooper et al., 2002; Valentine et al., 2003). Our current evaluation of this literature instead suggests that this effect has been there all along, obscured but not eliminated by the complexities inherent to real-world crime.

One of most shocking findings presented above is the disconnection between the circumstances at play during the witnessing of a crime and the circumstances under which current laboratory research is being conducted. With some exceptions laboratory and simulation studies expose their participants to a weapon for only a very brief period and then test their memory shortly thereafter. Insofar as retention interval is concerned this arrangement is desirable for pragmatic reasons. In the context of a laboratory experiment a protracted retention interval would increase the danger of attrition within the research sample. Retention intervals greater than several months are not even feasible for researchers reliant upon undergraduates who are prone to change courses on roughly the same timeline. The importance of additional studies emulating the progression of interrogation observed in the real world is driven home by the fact that the marginal effect of weapon presence observed by Tollestrup et al. (1994) only reached marginal significance when retention interval was accounted for as a covariate.

Of course, as noted earlier, the precise nature of the relationship between retention interval and weapon focus remains unclear. The decrease observed in this effect could result from either an increase in performance within the weapon condition or a decrease in performance within the control condition. The current data cannot easily address this issue, but Tollestrup et al. (1994) did observe a substantial decrease in overall positive suspect identifications as retention interval increased (see also Valentine et al., 2003). It seems reasonable to predict that the pattern depicted in Figure 2a arises from this very trend. As time passes perhaps performance decreases in both weapon and control conditions, but more rapidly in the control condition as they begin to equate with those exposed to a weapon. This is a hypothesis highly tractable to experimentation – and even those conducting field or archival research would be wise to follow Tollestrup et al.'s (1994) example in controlling for this variable.

More surprising is the lack of any systematic exploration of exposure duration within the experimental weapon focus literature. Kramer et al. (1990) are perhaps the only researchers to explicitly manipulate this variable, but they did so by reducing exposure duration within a narrow time window as opposed to testing the effects of shorter vs longer exposures *per se*. Cutler et al. (1987b) manipulated weapon exposure as well, but confounded it with exposure to the perpetrator. Tooley et al. (1987) employed a high exposure duration overall using a series of discrete visual scenes containing weapons but this was both indirect and likely unintentional. This exposure was not continuous and the role of exposure duration was not explicitly assessed. Figure 2c suggests that – as with most effects in psychology – the impact of weapon presence is bounded by under- and over-exposure (e.g. Yerkes & Dodson, 1908). Both of the theoretical frameworks detailed above predict this conclusion. The unusual item hypothesis dictates that attentional resources are drained only so long as the incongruity between the weapon and the schema representing the scene remains unresolved. Presumably this conflict would first build and later resolve over time, permitting attention to process other details of the visual scene. A similar argument could be made for the arousal/threat hypothesis. As time passes, the mind may become habituated to the presence of the weapon, and therefore less physiologically aroused by it, allowing greater utilization of peripheral cues (Easterbrook, 1959). Laboratory exposure times have often been limited to a measure of seconds whereas some real-world crimes may have a much longer duration (e.g. Gray, 1971).

Our analysis of exposure duration was relatively coarse due to the quality of the information currently available on this topic. Neither field nor archival studies commonly report the duration for which eyewitnesses are exposed to a weapon (when present) or the precise duration of the crime itself. These are important details especially when considered in combination. Another factor that has remained largely unexplored (again with the exception of Kramer et al., 1990) is the relationship between exposure to the weapon and exposure to the *perpetrator*. Virtually all laboratory investigations of the weapon focus effect confound these variables – with the weapon appearing and disappearing in synchrony with the individual holding the weapon. While one might imagine that this is a realistic portrait of actual criminal activity there are surely crimes involving *some* pre- or post-exposure to the perpetrator. In fact laboratory studies would do well to explore both pre- and post-exposure to the perpetrator once the weapon has been obscured, as there may be differences in the relative benefits. On the one hand, eyewitnesses may not benefit as much from pre-exposure as they may not yet realize the importance of the perpetrator; on the other hand, eyewitnesses may not benefit as much from post-exposure as they may experience some lingering effects of the weapon even following its removal. Notably the arousal/threat hypothesis might especially predict the latter as the physiological reaction to the weapon would require time to fade. This is in contrast to the unusual item hypothesis that would have a harder time explaining residual processing of the weapon in relation to the activated schema once the weapon had been removed. Overall, some real-world crimes *can* last for a prolonged duration and the duration of the crime does not necessarily equate with the duration of weapon exposure. Therefore, knowledge of the specific time course on which the weapon focus effect builds and dissipates would greatly benefit our understanding of its role in eyewitness testimony.

Theoretical implications

In conducting our meta-analysis we have attempted to address both of the traditional theoretical frameworks used in explaining the weapon focus effect. Neither theory has been definitively supported. We observed significantly larger effect sizes in the context of threatening as opposed to non-threatening scenarios when threat was compared between studies (see Figure 2b). We also observed a trend towards the same general pattern when we computed effect sizes comparing threatening and non-threatening events within studies. While this certainly suggests that arousal or threat might play a role in weapon focus, we also identified a sizable effect of unusualness on aggregate effect sizes when conducted between studies and a non-significant trend towards a larger impact of unusual items (as opposed to traditional weapons) when conducted within studies. It would appear that these accounts are in a stalemate, suggesting a common underlying mechanism (e.g. positive and negative arousal) or that both arousal and unusualness impact performance (suggesting that the weapon focus effect is an emergent property of this interaction).

Regarding the former position, previous research has suggested that while negative arousal (e.g. a defensive response such as fear) may degrade memory, positive arousal (e.g. an orienting response such as pleasant surprise) may exert a similar effect (e.g. Christianson, 1986). Perhaps while viewing unusual objects participants experience a slight feeling of surprise at the absurdity of the scenario (e.g. imagine viewing a video of someone trying to rob a store with a goose!). The net result could be a state of heightened physiological arousal that narrows attentional focus in a manner similar to the presentation of a weapon. If this were the case, it could explain why unusual objects impact the ability to respond to peripheral stimuli (e.g. Hope & Wright, 2007) and in general diminish memory for peripheral details (e.g. Pickel, 1998, 1999).

This view is not fundamentally different from the unusual item hypothesis detailed above. It differs only in proscribing the effects associated with the presence of a weapon or unusual object to the arousing or emotional properties of that object as opposed to its integration with the surrounding schema. Furthermore this finding is not even necessarily inconsistent with the probabilistic analysis proposed by Loftus and Mackworth (1978). Consolidating these views would require only that arousal is understood as one possible consequence of that analysis or its outcome. In the case of unusual objects one could dissociate the role of unusualness (i.e. schema incongruence) from the amusement it entails by independently controlling for these variables using the same sort of object rating procedure employed by Pickel (1998) to designate objects as high or low in unusualness or threat. Although no one has manipulated amusement in this literature, Mitchell et al. (1998) observed that the effect of unusualness remained even when accounting for self-reported emotional responses. This provides provisional support for the position that positive arousal (e.g. pleasant surprise) is not the *sole* cause of this effect. However, only by objectively measuring arousal (e.g. heart rate) at the time of encoding and relating this variable to subsequent memory performance may its role be clearly determined.

Future directions

Despite the growing number of publications exploring the effect of weapons and other unusual objects on eyewitness memory, there are still many avenues left unexplored. Below we will briefly discuss several research questions that we believe would further our knowledge and application of this effect.

Individual differences

Little work has been done investigating how individual differences shape eyewitness reactions to the presentation of a weapon or unusual object. Both Hulse and Memon (2006) and to a lesser degree Cooper et al. (2002) studied the impact of weapon presence in a population perhaps more familiar with weapons and violent criminal events. Curiously, neither demonstrated a convincing effect of weapon presence – although this may be attributed to methodological differences as easily as sample characteristics (see review above). It would be interesting to evaluate the role of prior exposure to weapons – and perhaps the nature of that exposure – to determine whether such exposure can mitigate the effect. This work need not be done in the laboratory, as field studies could attempt to make direct comparison between crimes involving a weapon in a neighborhood where guns are common vs an area where guns would be considered unusual.

Weapon presence and weapon visibility

Several studies (e.g. Cutler et al., 1987a) have treated weapon visibility (concealed or unconcealed) as a proxy for weapon presence (present or absent). However, no studies have explored the impact of a concealed weapon relative to a weapon absent control condition. Is a concealed weapon still capable of producing some degree of interference? The answer to this question is likely dependent upon the manner and degree of concealment. For example, would the edge of a gun peeking under a perpetrators' shirt be more effective than holding a gun in one's pocket such that the gun's presence is evident despite the gun itself being occluded, and how would these compare to a condition in which no gun was evident?

Limited data outside the visual modality

The vast majority of studies investigating weapon focus have dealt exclusively with the visual modality. However, there is one exception: Pickel et al. (2003) explored the plausibility of a cross-modal weapon focus effect by presenting participants with concurrent visual and auditory information containing the presence of a weapon or non-weapon object. In each of two experiments, participants were presented with a video depicting an encounter between a man and a woman. The man approached the woman carrying a weapon or mundane object and delivered a dialogue that was easy or difficult to comprehend. Participants were then administered separate suspect-present or suspect-absent photo and voice line-ups, as well as a questionnaire separately measuring their memory of the visual and auditory information presented in the scene. Both experiments produced the same pattern of results. Participants in the weapon condition exhibited worse memory for visual details presented at the same time as the weapon (e.g. perpetrator details). Furthermore, an interaction was observed between weapon presence and comprehension difficulty indicating that participants exhibited worse memory for the semantic content of the suspect's speech when a weapon was present and the comprehension difficulty was high. Weapon presence had no effect on photo or voice line-ups, or memory for the auditory features of the man's voice (Pickel et al., 2003).

Further, the modality through which the weapon is presented may influence subsequent recall. For example, verbal (e.g. someone yelling 'He's got a gun!') or tactile (e.g. feeling a weapon touch the body) presentation of a weapon without the corresponding visual stimulus may also result in decreased feature or identification accuracy. Additional research is required to characterize the parameters under which weapon presence can affect memory for auditory (or other sensory) details.

Completeness and accuracy

The outcome of our sub-analysis exploring the impact of weapon presence on the completeness of an eyewitness account was inconclusive. However it did suggest that completeness may be differentially sensitive to one or more of our moderator variables – or to the demands associated with different methodological approaches. It would not be difficult to study completeness in the laboratory alongside feature and identification accuracy. We revealed this in our own analysis by computing completeness scores for experiments that did not directly explore this variable. This computation required only that authors separate correct and incorrect details so that they may be summated (in addition to subtracted). Our preliminary analysis suggests that completeness *may* behave differently in the laboratory relative to the real world. However we cannot truly make this determination until we have a larger sample of laboratory and real-world studies using this variable. This is an important topic because, in the real world, completeness in the absence of accuracy, risks the possibility of introducing misinformation and possibly misdirecting a criminal investigation.

Functional novelty

Studies could further elaborate the potential effect of novelty on visual scene processing by investigating the effect of functional novelty (i.e. using a familiar object in a novel way). Kramer et al.'s (1990) first experiment unintentionally alluded to this topic by using a bottle as a weapon, although this concept has not been investigated directly. Variations on existing research may help to evaluate the influence of functional novelty. For example, studies with police officers have indicated a familiarity with the presence of a gun (e.g. Hulse & Memon, 2006), however they may not expect the perpetrator to throw the gun at their target. While the unusual item hypothesis would clearly predict that functional novelty within a visual scene would impair subsequent memory performance (as the use of the object is resolved against the schema used to represent that object or even the event itself) the arousal/threat hypothesis would predict such an effect only if such novelty produced significant changes in arousal (see discussion above).

Conclusion

In the past, weapon focus was given little attention relative to other forensic topics, and has only recently grown in credibility as a viable area of inquiry (e.g. Egeth, 1993; Kassin, Tubb, Hosch, & Memon, 2001). In laboratory studies, the weapon focus effect has been found using a variety of methodologies, ranging from slides (e.g. Loftus et al., 1987) to videos (e.g. Cutler et al., 1987a) to staged crimes (e.g. Maass & Kohnken, 1989), and using items ranging from guns (e.g. Hulse & Memon, 2006) to switchblade knives (Pickel, 2009). Despite a growing body of research, there has remained a disconnection between experimental findings and real-world applications of this effect. We have attempted to bring this literature back into focus by providing both a narrative overview of what has been done as well as an empirical evaluation of the status of this effect.

Our findings bear very important implications for the development of public policy on the credibility of the weapon focus phenomenon in a court of law. Although some studies have found that mock jurors are not heavily influenced by witness testimony concerning weapon presence (e.g. Cutler, Penrod, & Dexter, 1990), their conclusions apply more to guilt judgments and *not* witness credibility. It is not well understood how judges and jurors would interpret testimony of an eyewitness who has experienced weapon focus; however, perceptions of testimonial accuracy are diminished by statement characteristics such as inconsistencies (Brewer, Potter, Fisher, Bond, & Luszcz, 1999). That being said, the practical implications of testimony being discounted because of weapon focus (and corresponding memory inaccuracies) are dramatic. Both prosecution and defense lawyers would benefit from expert evidence concerning weapon focus in relevant cases to assure witness testimony does not lead to miscarriages of justice. Given the general failure to identify any negative impact of weapon presence in the real world, some researchers have suggested that the weapon focus effect should be dismissed as irrelevant to the topic of eyewitness testimony. This argument has recently been supported by two major governmental reports issued in both the USA (Mecklenburg, 2006) and the UK (Pike et al., 2002). In each case the authors reported no negative impact of weapon presence on identification accuracy despite the fact that, upon closer inspection, *both* reports demonstrated a trend in this direction. The finding of an effect of weapon presence on actual suspect identification attempts validates the inclusion of expert testimony on this topic.

While our focus has been on resolving the tension between laboratory and real-world research, the mechanisms behind the weapon focus effect remain to be fully understood. Given the propensity of unusual objects to create an analogous effect, perhaps it should instead be referred to as *salient feature focus* or an *object saliency effect* (see also Mitchell et al., 1998; Shaw & Skolnick, 1994, 1999; although see Hope & Wright, 2007). Even if the mechanisms underlying this effect are more relevant to unusualness than arousal, this does not attenuate its real-world utility. The example presented at the beginning of this article exemplifies only one unusual situation that law enforcement officers have encountered. In fact, this was not an isolated incident: only one month after his notorious coffee shop robbery, the Goose Robber struck again – this time threatening to cudgel a baby raccoon with a large rock if pedestrians on the streets of Toronto refused to give him cash ('Beggar Threatens Raccoon to Get Cash', 1997). This example further accentuates the need to understand how the features (e.g. unusualness, arousal) of a crime may affect a witness' memory and perception of the criminal event.

Although our review and meta-analysis offer much in the way of understanding the nature of weapon focus in real-world crime, it is more a starting point than a final synopsis on the subject. Upon collating the data for analysis we repeatedly came across areas in need of further characterization (see discussion above). This is largely the reason behind our not exploring a more complicated model of analysis and why we could not explore any potential interactions. The range and variability within our moderators could simply not support such ambitions. Our challenge to experimentalists is to broaden the parameters under which they conduct their research to (at least at times) better approximate actual eyewitness testimony (e.g. Cutler et al., 1987a, b); our challenge to those working in the field or archives is to better characterize and report the details of the event (e.g. exposure duration, crime duration, etc.) for the benefit of future analyses like the one presented above.

Acknowledgements

JMF was supported by an NSERC Canada Graduate Scholarship and a Killam Predoctoral Scholarship, and would like to thank Dr. Ray Klein for comments on an earlier draft of this manuscript. KAP would also like to thank Jessica Gilbert for her research assistance. Finally, we would like to thank Dr. Kerri Pickel, Dr. Brian Bornstein and our anonymous reviewers for their feedback.

References

(Articles contributing one or more effect sizes to any of the reported analyses in the main text are indicated by an asterisk.)

Antes, J.R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*, 62–70.

Beggar threatens raccoon to get cash. (1997, May 13). *Los Angeles Times*, p. 8.

*Behrman, B.W., & Davey, S.L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behaviour*, *25*, 475–491.

Brewer, N., Potter, R., Fisher, R.P., Bond, N., & Luszcz, M.A. (1999). Beliefs and data on the relationship between consistency and accuracy of eyewitness testimony. *Applied Cognitive Psychology*, *13*, 297–313.

Brown, J.M. (2003). Eyewitness memory for arousing events: Putting things into context. *Applied Cognitive Psychology*, *17*, 93–106.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, *19*, 3127–3131.

Christianson, S.-Å. (1986). Effects of positive emotional events on memory. *Scandinavian Journal of Psychology*, *27*, 287–299.

Christianson, S.-Å., & Hübinette, B. (1993). Hands up! A study of witnesses' emotional reactions and memories associated with bank robberies. *Applied Cognitive Psychology*, *7*, 365–379.

*Cooper, B.S., Kennedy, M.A., Herve, H.F., & Yuille, J.C. (2002). Weapon focus in sexual assault memories of prostitutes. *International Journal of Law and Psychiatry*, *25*, 181–191.

*Cutler, B.L., & Penrod, S.D. (1988). Improving the reliability of eyewitness identification: Lineup construction and presentation. *Journal of Applied Psychology*, *73*, 281–290.

Cutler, B.L., Penrod, S.D., & Dexter, H.R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior*, *14*, 185–191.

*Cutler, B.L., Penrod, S.D., & Martens, T.K. (1987a). The reliability of eyewitness identification: The role of system and estimator variables. *Law and Human Behavior*, *11*, 223–258.

*Cutler, B.L., Penrod, S.D., & Martens, T.K. (1987b). Improving the reliability of eyewitness identification: Putting context into context. *Journal of Applied Psychology*, *72*, 629–637.

*Cutler, B.L., Penrod, S.D., O'Rourke, T.E., & Martens, T.K. (1986). Unconfounding the effects of contextual cues on eyewitness identification accuracy. *Social Behavior*, *1*, 113–134.

*Cutshall, J.L., & Yuille, J.C. (1989). Field studies of eyewitness memory of actual crimes. In D.C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 97–124). New York, NY: Springer.

*Davies, G.M., Smith, S., & Blincoe, C. (2008). A 'weapon focus' effect in children. *Psychology, Crime & Law*, *14*, 19–28.

Davis, J.P., & Valentine, T. (2009). CCTV on trial: Matching video images with the defendant in the dock. *Applied Cognitive Psychology*, *23*, 482–505.

Deffenbacher, K.A. (2008). Estimating the impact of estimator variables on eyewitness identification: A fruitful marriage of practical problem solving and psychological theorizing. *Applied Cognitive Psychology*, *22*, 815–826.

Deffenbacher, K.A., Bornstein, B.H., Penrod. S. D, & McGorty, E.K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior*, *28*, 687–706.

Easterbrook, J.A. (1959). The effect of emotion on the utilization and organization of behavior. *Psychological Review*, *66*, 183–201.

Egeth, H.E. (1993). What do we not know about eyewitness identification? *American Psychologist*, *48*, 577–580.

Goose held hostage in Canada robbery. (1997, April 26). *Toledo Blade*, p. 2.

Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Gray, P. (1971). Robbery and assault of bus drivers. *Operations Research*, *19*, 257–269.

Hazell, P., O'Connell, D.O., Heathcote, D., Robertson, J., & Henry, D. (1995). Efficacy of tricyclic drugs in treating child and adolescent depression: A meta-analysis. *British Medical Journal*, *310*, 897–901.

Hedges, L.V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, *92*, 490–499.
Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego, CA: Academic Press.
*Hope, L., & Wright, D. (2007). Beyond unusual? Examining the role of attention in the weapon focus effect. *Applied Cognitive Psychology*, *21*, 951–961.
Hulse, L.M., & Memon, A. (2006). Fatal impact? The effects of emotional arousal and weapon presence on police officers' memories for a simulated crime. *Legal and Criminological Psychology*, *11*, 313–325.
*Johnson, C., & Scott, B. (1976). *Eyewitness testimony and suspect identification as a function of arousal, sex of witness and scheduling of interrogation*. Paper presented at the American Psychological Association Annual Meeting, Washington, DC.
Kassin, S.M., & Gudjonsson, G.H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, *5*, 33–67.
Kassin, S.M., Tubb, V.A., Hosch, H.M., & Memon, A. (2001). On the 'general acceptance' of eyewitness testimony research: A new survey of the experts. *American Psychologist*, *56*, 405–416.
Kensinger, E.A., Garoff-Eaton, R.J., & Schacter, D.L. (2006). Memory for specific visual details can be enhanced by negative arousing content. *Journal of Memory and Language*, *54*, 99–112.
Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, *51*, 481–537.
*Kramer, T.H., Buckhout, R., & Eugenio, P. (1990). Weapon focus, arousal and eyewitness memory: Attention must be paid. *Law and Human Behaviour*, *14*, 167–184.
Kuehn, L.L. (1974). Looking down a gun barrel: Person perception and violent crime. *Perceptual and Motor Skills*, *39*, 1159–1164.
Lipsey, M.W., & Wilson, D.B. (2000). *Practical meta-analysis.* Thousand Oaks, CA: Sage.
Loftus, E.F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.
*Loftus, E.F., Loftus, G.R., & Messo, J. (1987). Some facts about 'weapon focus'. *Law and Human Behaviour*, *1*, 55–62.
Loftus, G.R., & Mackworth, N.H. (1978). Cognitive determinates of fixation location during picture viewing. *Journal of Experimental Psychology*, *4*, 565–572.
*Maass, A., & Kohnken, G. (1989). Eyewitness identification: Simulating the 'weapon effect'. *Law and Human Behaviour*, *13*, 397–408.
Malpass, R.S. (2006a). A policy evaluation of simultaneous and sequential lineups. *Psychology, Public Policy and Law*, *12*, 394–418.
Malpass, R.S. (2006b). Notes on the Illinois pilot program on sequential double-blind identification procedures. *Public Interest Law Reporter*, *11*(2), 5–47.
Mastrobuoni, G. (2010). Everybody stay cool, this is a robbery. In N. Garoupa (Chair), *Crime, labor, health*. Panel conducted at the meeting of the American Law and Economics Association, Princeton University.
Mastrobuoni, G. (2011). *Optimal criminal behavior and the disutility of jail: Theory and evidence on bank robberies.* Unpublished manuscript. Retrieved from http://www.carloalberto.org/people/mastrobuoni/doc/ValueFreedom4.pdf
*Mecklenburg, S.H. (2006). *Report to the legislature of the State of Illinois: The Illinois Pilot Program on double-blind, sequential lineup procedures.* Springfield, IL: Illinois State Police.
*Mitchell, K.J., Livosky, M., & Mather, M. (1998). The weapon focus effect revisited: The role of novelty. *Legal and Criminological Psychology*, *3*, 287–303.
*O'Rourke, T.E., Penrod, S.D., Cutler, B.L., & Stuve, T.E. (1989). The external validity of eyewitness identification research: Generalizing across subject populations. *Law and Human Behaviour*, *13*, 385–395.
Peace, K.A., & Porter, S. (2004). A longitudinal investigation of the reliability of memories for trauma and other emotional experiences. *Applied Cognitive Psychology*, *18*, 1143–1159.
*Peters, D.P. (1988). Eyewitness memory and arousal in a natural setting. In M. Gruneberg, P. Morris, & R. Skyes (Eds.), *Practical aspects of memory: Current research and issues: Vol. 1. Memory in everyday life* (pp. 89–94). Chichester: Wiley.
Peters, J.D. (2001). Witnessing. *Media, Culture, and Society*, *23*, 707–723.

*Pike, G., Brace, N., & Kyman, S. (2002, March). *The visual identification of suspects: Procedures and practices.* (Briefing Note 2/02). London: Policing and Reducing Crime Unit Home Office Research, Development and Statistics Directorate. Retrieved from http://rds.homeoffice.gov.uk/rds/pdfs2/brf202.pdf

*Pickel, K.L. (1998). Unusualness and threat as possible causes of 'weapon focus'. *Memory*, *6*, 277–295.

*Pickel, K.L. (1999). The influence of context on the 'weapon focus' effect. *Law and Human Behaviour*, *23*, 299–311.

Pickel, K.L. (2007). Remembering and identifying menacing perpetrators: Exposure to violence and the weapon focus effect. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *The handbook of eyewitness psychology, Vol II: Memory for people* (pp. 339–360). Mahwah, NJ: Lawrence Erlbaum.

*Pickel, K.L. (2009). The weapon focus effect on memory for female versus male perpetrators. *Memory*, *17*, 664–678.

*Pickel, K.L., French, T., & Betts, J. (2003). A cross-modal weapon focus effect: The influence of a weapon's presence on memory for auditory information. *Memory*, *11*, 277–292.

*Pickel, K.L., Narter, D.B., Jameson, M.M., & Lenhardt, T.T. (2008). The weapon focus effect in child eyewitnesses. *Psychology, Crime & Law*, *14*, 61–72.

*Pickel, K.L., Ross, S.J., & Truelove, R.S. (2006). Do weapons automatically capture attention? *Applied Cognitive Psychology*, *20*, 871–893.

Porter, S., & Peace, K.A. (2007). The scars of memory: A prospective, longitudinal investigation of the consistency of traumatic and positive emotional memories in adulthood. *Psychological Science*, *18*, 435–441.

R Development Core Team. (2010). *R: A language and environment for statistical computing*. Vienna: R Development Core Team.

Rand Corporation. (1975, October). *The criminal investigation process* (Vol. 1–3; Report No. R-1776-DOJ, R-1777-DOJ). Santa Monica, CA: The Rand Corporation.

Reisberg, D., & Heuer, F. (2007). The influence of emotion on memory for forensic settings. In M.P. Toglia, J.D. Read, D.F. Ross, & R.C.L. Lindsay (Eds.), *The handbook of eyewitness psychology: Memory for events* (pp. 81–116). Mahwah, NJ: Erlbaum.

Rosenthal, R. (1978). Combining results of independent studies. *Psychological Bulletin*, *85*, 185–193.

Ross, S.J., & Malpass, R.S. (2008). Moving forward: Response to 'Studying eyewitness investigations in the field'. *Law and Human Behavior*, *32*, 16–21.

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, *8*, 448–467.

*Saunders, J. (2009). Memory impairment in the weapon focus effect. *Memory & Cognition*, *37*, 326–335.

Shaw, J.I., & Skolnick, P. (1994). Sex differences, weapon focus and eyewitness reliability. *Journal of Social Psychology*, *134*, 413–420.

*Shaw, J.I., & Skolnick, P. (1999). Weapon focus and gender differences in eyewitness accuracy: Arousal versus salience. *Journal of Applied Social Psychology*, *29*, 2328–2341.

Steblay, N.M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, *16*, 413–424.

Steblay, N.M. (2010). What we know now: The Evanston Illinois field lineups. *Law and Human Behavior*, *35*, 1–12.

*Tollestrup, P.A., Turtle, J.W., & Yuille, J.C. (1994). Actual victims and witnesses to robbery and fraud: An archival analysis. In D. Ross, D. Read, & C. Ceci (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 144–162). New York, NY: Press Syndicate of the University of Cambridge.

*Tooley, V., Brigham, J.C., Maass, A., & Bothwell, R.K. (1987). Facial recognition: Weapon effect and attentional focus. *Journal of Applied Social Psychology*, *17*, 845–859.

Valentine, T., & Mesout, J. (2009). Eyewitness identification under stress in the London Dungeon. *Applied Cognitive Psychology*, *23*, 151–161.

*Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, *17*, 969–993.

van der Kolk, B.A. (1996). Trauma and memory. In B. van der Kolk, A. McFarlane, & L. Weisaeth (Eds.), *Traumatic stress* (pp. 279–302). New York, NY: Guildford Press.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48.

*Wagstaff, G.F., MacVeigh, J., Scott, L., Brunas-Wagstaff, J., & Cole, J. (2003). Can laboratory findings on eyewitness testimony be generalized to the real world? An archival analysis of the influence of violence, weapon presence, and age on eyewitness accuracy. *Journal of Psychology*, *137*, 17–28.

Whipple, G. (1915). Psychology of testimony. *Psychological Bulletin*, *12*, 221–224.

Whitehead, A., Bailey, A., & Elbourne, D. (1999). Combining summaries of binary outcomes with those of continuous outcomes in a meta-analysis. *Journal of Biopharmaceutical Statistics*, *9*, 1–16.

Yarmey, A.D. (2001). Expert testimony: Does eyewitness memory research have probative value for the courts? *Canadian Psychology*, *42*, 92–100.

Yerkes, R.M., & Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, *18*, 459–482.

Wells, G.L. (2008). Field experiments on eyewitness identification: Towards a better understanding of pitfalls and prospects. *Law and Human Behavior*, *32*(1), 6–10.