


Effects of Weapons on Aggressive Thoughts, Angry Feelings, Hostile Appraisals, and Aggressive Behavior: A Meta-Analytic Review of the Weapons Effect Literature

Personality and Social Psychology Review
2018, Vol. 22(4) 347–377
© 2018 by the Society for Personality
and Social Psychology, Inc.
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1088868317725419
journals.sagepub.com/home/pspr


Arlin J. Benjamin Jr.¹ , Sven Kepes², and Brad J. Bushman³

Abstract

A landmark 1967 study showed that simply seeing a gun can increase aggression—called the “weapons effect.” Since 1967, many other studies have attempted to replicate and explain the weapons effect. This meta-analysis integrates the findings of weapons effect studies conducted from 1967 to 2017 and uses the General Aggression Model (GAM) to explain the weapons effect. It includes 151 effect-size estimates from 78 independent studies involving 7,668 participants. As predicted by the GAM, our naïve meta-analytic results indicate that the mere presence of weapons increased aggressive thoughts, hostile appraisals, and aggression, suggesting a cognitive route from weapons to aggression. Weapons did not significantly increase angry feelings. Yet, a comprehensive sensitivity analysis indicated that not all naïve mean estimates were robust to the presence of publication bias. In general, these results suggest that the published literature tends to overestimate the weapons effect for some outcomes and moderators.

Keywords

weapons effect, aggression, General Aggression Model, social priming

Guns not only permit violence, they can stimulate it as well. The finger pulls the trigger, but the trigger may also be pulling the finger.

—Leonard Berkowitz (1968, p. 22)

Obviously, using a gun can increase aggression and violence, but can just seeing a gun increase aggression? In 1967, Leonard Berkowitz and Anthony LePage conducted a randomized experiment to find out. Male college students were tested in pairs, but one of them was actually an accomplice of the experimenter who was pretending to be another participant. They evaluated each other’s performance on a task (e.g., listing ideas a used car salesperson might use to sell more cars). The “evaluations” were the number of stressful electrical shocks given, which ranged from 1 to 10. First, the accomplice evaluated the participant’s performance by using either 7 shocks (provocation condition) or 0 shocks (no provocation condition). Next, the participant “evaluated” the accomplice’s performance. The number of electrical shocks the participant chose for the accomplice was used to measure aggression. The participant was seated at a table that had a shotgun and a revolver on it, or badminton rackets and shuttles. The items on the table were described as part of

another study that another experimenter had supposedly forgotten to put away. There was also a control condition with no items on the table. The experimenter told participants to ignore the items on the table, but apparently they could not. Provoked participants who saw the guns were more aggressive than the other participants. Berkowitz and LePage called this effect the “weapons effect.” Mere exposure to weapons such as guns can increase aggression.

In later experiments, similar results were obtained when pictures of guns were used instead of actual guns (Leyens & Parke, 1975). Several field experiments tested the weapons effect outside of the lab using rifles placed in racks in the back windows of a pickup truck driven by an accomplice who refused to move when a traffic light turned green and used horn honking as the aggression measure (e.g., Turner, Layton, & Simons, 1975).

¹University of Arkansas–Fort Smith, USA

²Virginia Commonwealth University, Richmond, USA

³The Ohio State University, Columbus, USA

Corresponding Author:

Arlin J. Benjamin Jr., Department of Behavioral Sciences, University of Arkansas–Fort Smith, Fort Smith, AR 72913, USA.
Email: james.benjamin@uafs.edu

A prior meta-analysis, published in 1990, integrated the findings from weapons effect studies (Carlson, Marcus-Newhall, & Miller, 1990). Our meta-analysis provides a significant and much needed update to this important but decades old meta-analysis. This meta-analysis makes four important methodological improvements over the 1990 one. First, the sheer number of effect-sizes integrated is over 5 times more in this meta-analysis. The 1990 meta-analysis included 31 effect-size estimates, whereas this one includes 151 effect-size estimates. Second, the 1990 meta-analysis excluded unpublished studies. It is well documented that studies reporting statistically significant results are more likely to be published than are studies reporting nonsignificant results. This “prejudice against the null hypothesis” seems pervasive (Greenwald, 1975). In meta-analysis, the conditional publication of studies with significant results is called the “file drawer problem” (Rosenthal, 1979). If the null hypothesis is true, only one out of the 20 studies conducted was published and the remaining 19 studies were located in researchers’ file drawers (or garbage cans), assuming the .05 significance level is used. If publication bias is a problem, then the studies included in a meta-analysis may represent a biased subset of the total number of studies conducted on the topic. That is why we collected as many unpublished studies as possible for the present meta-analysis.

Third, we conducted a comprehensive sensitivity analysis to assess the robustness of our mean effect sizes. The robustness of published results in the social sciences, including social psychology, has been questioned (Ferguson & Heene, 2012; Fiedler, 2011; Ioannidis, 2012; Kepes & McDaniel, 2013; Open Science Collaboration, 2015; Pashler & Wagenmakers, 2012; Simmons, Nelson, & Simonsohn, 2011; Yong, 2012). Publication bias is currently one of the phenomena that has been shown to have adversely affected published meta-analytic results (e.g., Banks, Kepes, & McDaniel, 2015; Ferguson & Brannick, 2011) and, thus, distorted cumulative knowledge (Kepes & McDaniel, 2013). To assess the robustness of our results, we followed “best practice” recommendations (e.g., Greenhouse & Iyengar, 2009; Kepes, Bushman, & Anderson, 2017; Kepes, McDaniel, Brannick, & Banks, 2013) and conducted a comprehensive sensitivity analysis on the effect sizes at the study level to evaluate the robustness of the results for each individual distribution. We used seven publication bias methods, each of which is capable of estimating a “for publication bias adjusted” mean effect. Because between-studies heterogeneity is known to have adverse effects on publication bias methods (as well as the “basic” or naïve meta-analytic methods; for example, Borenstein, Hedges, Higgins, & Rothstein, 2009; Kepes & McDaniel, 2015; Terrin, Schmid, Lau, & Olkin, 2003), we performed all these analyses at the subgroup level to control for moderating effects. Fourth, we accounted for heterogeneity due to outliers by performing all analyses with and without identified outliers. Outliers can have a profound effect on meta-analytic results, and we

wanted to make sure our findings were not unduly influenced by them. Therefore, we feel strongly that our meta-analytic study is not only an update on the almost 30-year-old Carlson et al. (1990) meta-analysis, which is in desperate need of an update, but also a methodological template for future meta-analyses. To obtain more robust and accurate meta-analytic results, we suggest that future meta-analyses should follow the procedures we have outlined in this meta-analysis.

Two other very important advances of this meta-analysis over the 1990 Carson et al. meta-analysis are theoretical in nature. First, the 1990 meta-analysis did not directly compare the magnitude of the weapons effect for provoked and nonprovoked participants. Berkowitz and LePage (1967) only found a weapons effect for provoked participants. Our meta-analysis tests whether weapons have a greater effect on provoked participants than on nonprovoked participants. Second, the 1990 Carson et al. meta-analysis only tested the effects of weapons on aggressive behavior. The present one also tests the effects of weapons on aggressive thoughts, angry feelings, and hostile appraisals. This extension is important because it sheds light on *why* weapons increase aggression. For example, the most common explanation of the weapons effect is that weapons prime aggressive thoughts. Our meta-analysis directly tests this hypothesis. The theoretical foundation for this meta-analysis is the General Aggression Model (GAM; Anderson & Bushman, 2002), which is described next.

GAM

The GAM provides a useful framework for understanding the weapons effect (Figure 1). The GAM subsumes other models that have been used to explain the weapons effect, such as those based on classical conditioning, operant conditioning, and priming (e.g., Berkowitz, 1974, 1982, 1983). In the GAM, two types of input variables can influence aggression: personal and situational. Personal variables include anything the individual brings to the situation (e.g., gender, age, genetic predispositions, personality traits and other individual differences, attitudes, beliefs, values). This meta-analysis focuses on three personal variables—gender, age, and whether participants were college students or not. Some critics have argued that college students, who are often recruited from introductory psychology participant pools, are not representative of “real people” (e.g., Oakes, 1972; Sears, 1986). We tested whether the weapons effect occurs for males and females, for participants of different ages, and for student and nonstudent samples.

Situational variables include all external factors that can influence aggression (e.g., aggressive cues such as weapons, violent media exposure, provocation, frustration, alcohol, hot temperatures, crowding). This meta-analysis focuses on two situational variables—exposure to weapons and provocation. Specifically, we test whether the weapons effect occurs for provoked and nonprovoked individuals. When provoked,

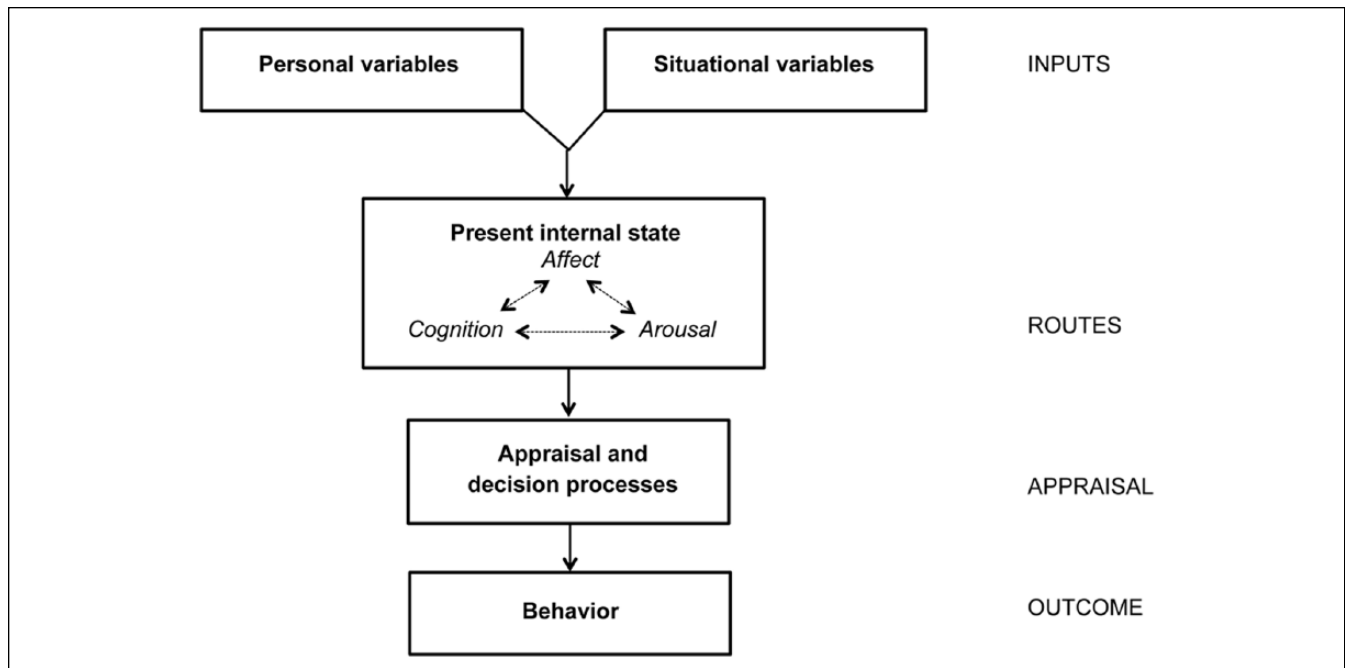


Figure 1. The General Aggression Model (Anderson & Bushman, 2002; Krahe, 2013).

individuals become physiologically aroused and ready to attack others. Thus, provoked individuals might be particularly prone to react aggressively when primed with weapons.

According to the GAM (Anderson & Bushman, 2002), personal and situational factors influence one's internal state, which can include aggressive cognition, aggressive affect, and physiological arousal. Thus, there are three possible routes to aggression—through aggressive cognition, aggressive affect, and physiological arousal. However, these routes are not mutually exclusive or even independent, as indicated by the dashed lines with double-headed arrows in Figure 1. For example, someone who has aggressive ideas might also feel angry and have elevated blood pressure. This meta-analysis examines the effects of weapons on aggressive cognition and aggressive affect. Unfortunately, only one study examined the influence of weapons on self-reported arousal (De Oca & Black, 2013). In that study, participants rated threatening items (e.g., weapons) to be more arousing than non-threatening items (e.g., trees, food, couches). We could find no studies that tested the effects of weapons on physiological arousal (e.g., heart rate, blood pressure, skin conductance).

According to the GAM, internal states can influence appraisal and decision processes. First, there is an immediate initial appraisal of whether the situation is dangerous, threatening, or warrants aggression. This initial appraisal might lead directly to an automatic or impulsive behavior, or it might lead to a reappraisal. If the initial appraisal is judged to be unsatisfactory and if the person has sufficient time and cognitive resources, reappraisal occurs (Barlett & Anderson, 2011). During reappraisal, the person considers alternative

explanations of the situation and alternative behavioral options. When the appraisal is judged to be satisfactory, or when time or resources become insufficient, the appraisal process terminates and the person engages in the behavior without reappraisal, which completes one cycle. This meta-analysis examined the influence of weapons on hostile appraisals of others. Note that in the GAM, hostile appraisals are more proximal to aggressive behavior than are internal states. Thus, hostile appraisals might have a stronger influence on aggression than internal states.

The types of appraisals and decisions people make can influence their behavior. The primary outcome variable in our meta-analysis was aggressive behavior. Most researchers define aggression as any behavior intended to harm another person who wants to avoid being harmed (Baron & Richardson, 1994).

Moderators

In addition to theoretical outcomes and moderators encompassed by the GAM, we also considered several study characteristics that might influence the magnitude of the weapons effect, including publication status (i.e., published in a peer-reviewed journal vs. unpublished), the year the study was conducted (to test whether the magnitude of the weapons effect has changed over time), and whether a between-subjects design or a within-subjects design was used. We coded several characteristics about the weapons used in studies (i.e., actual weapon vs. photo of weapon; real weapon vs. toy weapon; type of weapon).

In addition, we examined the study setting. Laboratory experiments have been criticized because they are conducted in artificial settings, with unrealistic measures, and unrepresentative samples—mainly college students (for a review see Anderson & Bushman, 1997). Moreover, participants in laboratory experiments can become suspicious about being deceived, which can contaminate the results. Field studies overcome these criticisms; yet, they are not without their own shortcomings. For instance, another important difference between laboratory experiments and field studies is control over possible confounding variables. Previous research has shown that aggression effects tend to be larger in the lab where conditions are more tightly controlled than in the field (Anderson & Bushman, 1997). In this meta-analysis, we coded whether the study was conducted in a laboratory or field setting.

Overview

The primary purpose of this meta-analysis was to examine the effects of the mere presence of weapons on aggressive thoughts, angry feelings, hostile appraisals, and aggressive behavior. Although violent media (e.g., television programs, movies, video games, Internet) also include weapons, studies that examined the effects of violent media on aggression were not included in this meta-analysis. Numerous meta-analyses have already shown that exposure to violent media can increase aggressive thoughts, angry feelings, physiological arousal, hostile appraisals, and aggressive behavior (e.g., Anderson et al., 2010; Bushman, 2016; Bushman & Huesmann, 2006; Greitemeyer & Mügge, 2014). We were interested in a more basic question: Can the mere presence of a weapon—that is not being used by one person to injure or kill another person—increase aggression? We predicted that the mere presence of weapons would increase aggression.

In an attempt to understand *why* weapons might increase aggression, we considered their effects on aggressive thoughts, angry feelings, and hostile appraisals using the GAM as a theoretical guide. Based on previous research (e.g., Anderson, Benjamin, & Bartholow, 1998), we predicted a cognitive route between exposure to weapons and aggression. Specifically, we predicted that weapons would prime or activate aggressive thoughts and increase hostile appraisals.

Because provoked individuals become physiologically aroused and ready to attack others, they might be particularly prone to react angrily and aggressively when primed with weapons. Thus, we predicted a stronger weapons effect among provoked participants than among nonprovoked participants.

We also examined several possible moderators of the weapons effect, including the gender and age of participants, whether participants were college students or not, and study characteristics (i.e., publication status, year study was conducted, between- or within-subjects design, laboratory vs.

field setting, type of weapons exposed to). However, we made no predictions about these moderators.

Method

Literature Search

To locate relevant studies, we searched the PsycINFO, PsycARTICLES, MEDLINE and SocINDEX, Google Scholar, and Dissertation Abstracts International databases from 1967 (the year the first weapons effect study was published by Berkowitz and LePage) through 2017. A thorough search was conducted to be sure that no relevant studies were excluded. We used the search terms (*gun** OR *firearm** OR *weapon**) AND (*aggress** OR *violen**). The asterisk allows terms to have all possible endings (e.g., the term *aggress** will retrieve studies that used the terms *aggress*, *aggressed*, *aggressor*, *aggressive*, and *aggression*). Thus, the article had to include the term *gun** or *firearm** or *weapon**, plus the term *aggress** or *violen**. We also searched the Social Science Citation Index for any article that cited the original weapon's effect study by Berkowitz and LePage (1967).

Five additional steps were taken to obtain any studies we might have missed. First, we searched the reference sections of relevant meta-analytic reviews (Anderson & Bushman, 1997; Bettencourt & Kernahan, 1997; Bettencourt & Miller, 1996; Bettencourt, Talley, Benjamin, & Valentine, 2006; Bushman & Anderson, 1998; Carlson et al., 1990), and narrative reviews (Berkowitz, 1971; Toch & Lizotte, 1992; Turner & Leyens, 1992; Turner, Simons, Berkowitz, & Frodi, 1977). Second, we searched the reference sections of all retrieved studies. Third, we contacted all researchers who had conducted a weapons effect study from our list of retrieved studies, and requested any published and unpublished weapons effect studies. Fourth, we searched the proceedings for eight relevant conferences for unpublished studies: (a) American Psychological Association (APA), (b) Association for Psychological Science (APS), (c) European Association of Social Psychology, (d) International Society for Research on Aggression, (e) Society of Australasian Social Psychologists, (f) Society of Experimental Social Psychology, (g) Society for Personality and Social Psychology (Division 8 of APA), and (h) Society for the Psychological Study of Social Issues. Fifth, we sent an announcement requesting unpublished and published weapons effect studies to seven listserves: (a) European Association of Social Psychology, (b) International Society for Research on Aggression, (c) Society of Australasian Social Psychologists, (d) Society of Experimental Social Psychology, (e) Society for Personality and Social Psychology (Division 8 of APA), (f) Society for the Psychological Study of Social Issues, and (g) Society for the Study of Peace, Conflict and Violence: Peace Psychology Division (Division 48 of APA).

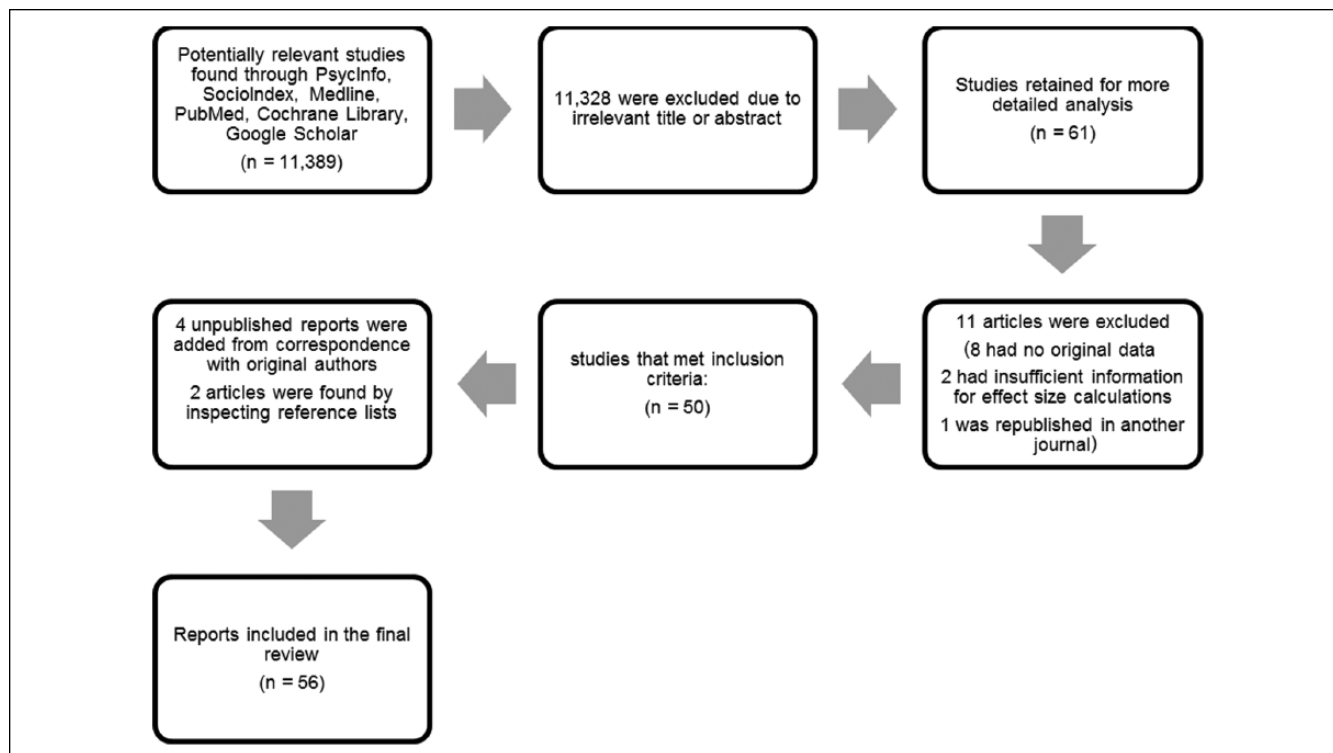


Figure 2. PRISMA flowchart of literature search and inclusion/exclusion decisions.

This thorough search yielded 11,389 articles, but not all were relevant to this meta-analysis. To determine whether articles were relevant, we read their titles, abstracts, or both. Unpublished studies, dissertations, and conference papers were also included in the database to address potential publication bias (i.e., the “file drawer problem,” Rosenthal, 1979). We found 13 unpublished research reports from master’s theses, doctoral dissertations, conference proceedings, and personal communication. These 13 reports yielded 16 studies and 40 tests of the weapons effect. A PRISMA flowchart of the literature search and study coding is shown in Figure 2.

Inclusion Criteria

Two inclusion criteria were used. First, a study needed to include a weapons condition (e.g., guns, knives, swords, hand grenades), and a no weapons (control) condition (e.g., nonviolent objects such as badminton rackets, flowers, eating utensils, nothing at all). Some studies used a between-subjects design, where participants were randomly assigned to weapons or no weapons conditions. Other studies used a within-subjects design, where participants were exposed to both the weapons and no weapons conditions in a random order. The weapons could be real weapons or toy weapons, physically present or shown in photographs. Second, a study needed to include a measure of aggressive cognition,

aggressive affect, physiological arousal, hostile appraisal, or aggressive behavior. Some studies included more than one type of outcome measure.

Two articles reported identical analyses from the same set of data (Simons & Turner, 1975, 1976). The more recent of the two articles was included in this meta-analysis because it provided a more complete set of analyses. The final sample included 56 research reports that described 78 independent studies involving 7,668 participants. We computed 151 effect-size estimates from these 56 research reports.

Moderators

Type of weapons. We coded whether the weapons were guns, knives, or a mixture of various weapons (e.g., guns, knives, swords, grenades, clubs). Most studies used either guns or knives exclusively, although some used both guns and knives (e.g., Blanchette, 2006; Sulikowski & Burke, 2014). Studies using mixtures of weapons were varied. For example, one study used guns, clubs, and swords as stimuli (Anderson et al., 1998). We coded whether participants were exposed to actual weapons or photos of weapons. We also coded whether participants were exposed to real weapons or toy weapons. Studies using mixtures of toy weapons also varied, including toy guns, daggers, bazookas, and so on (e.g., Goff, 1995; Mendoza, 1972).

Type of outcome. This meta-analysis examined four types of outcomes: (a) aggressive cognition, (b) aggressive affect, (c) hostile appraisals, and (d) aggressive behavior. Only one study investigated the effect of weapons on physiological arousal, so we could not include it as an outcome. That study found that participants rated weapons as more arousing than nonthreatening objects such as trees and food (De Oca & Black, 2013). Although some researchers have included other outcomes, there were not a sufficient number of these other outcomes to include in our meta-analysis. For example, one study found that testosterone levels increased more in men who handled a gun than in men who handled a nonviolent toy (Klinesmith, Kasser, & McAndrew, 2006). Testosterone has repeatedly been linked to aggression in research studies (Archer, 1988; Sapolsky, 1998). Next, we describe prototypical ways of measuring each outcome.

Aggressive cognition is most often measured using reaction times to aggressive and nonaggressive words (e.g., Anderson et al., 1998; Bartholow & Heinz, 2006). In other studies, participants completed word fragments by filling in missing letters to form words as quickly as possible (e.g., Bushman, in press). For example, the word fragment K I _ _ can be completed to form a nonaggressive word (e.g., KISS, KIND, KITE), or can be completed to form an aggressive word (e.g., KILL, KICK).

Aggressive affect is most often measured using mood scales. For example, participants rate how they felt at that moment ("right now") using a list of adjectives, including some that measure aggressive affect (e.g., ANGRY, FURIOUS, IRRITABLE; for example, Anderson, Anderson, & Deuser, 1996). Although weapons can also influence other emotions (e.g., anxiety, empathy), no studies included in this meta-analysis examined other emotions.

Hostile appraisals are measured in several different ways. For example, some studies have measured primary or automatic appraisals by speed of fist clenching (e.g., da Gloria, Duda, Pahlavan, & Bonnet, 1989) and by speed of identification of weapons versus neutral objects (e.g., De Oca & Black, 2013; Sulikowski & Burke, 2014). Other studies have measured secondary or controlled reappraisal by having participants indicate how disagreeable, hostile, and angry they thought a target person was (e.g., Epstein, 1980; Holbrook et al., 2014). Unfortunately, there were not enough studies to examine primary and secondary appraisals separately.

In laboratory experiments, aggression has most typically been measured by electric shocks (e.g., number, intensity, duration) given to an accomplice pretending to be another participant. Other studies have used other aversive stimuli to measure aggression, such as noise blasts (e.g., Epstein, 1980; Lindsay & Anderson, 2000), or extremely spicy hot sauce given to an accomplice who dislikes spicy (e.g., Klinesmith et al., 2006). Nonphysical measures of aggression have included negative evaluations of experimenters or accomplices (e.g., Fischer, Kelm, & Rose, 1969). In field experiments involving adults, aggression has been measured either

by the number of horn honks at an accomplice who is stalled at a traffic light (e.g., Halderman & Jackson, 1979; Turner et al., 1975) or the number of wet sponges thrown at an accomplice (e.g., Simons, Fenn, Layton, & Turner, 1976). In field experiments involving children, aggression has been measured using behaviors observed in interactions with other children, such as pushing, shoving, kicking, and hitting (e.g., Turner & Goldsmith, 1976).

Provocation. For each study, we coded whether participants were provoked or not. Like Berkowitz and LePage (1967), several researchers used electric shocks to provoke participants (e.g., Berkowitz & LePage, 1967; Frodi, 1975; Turner & Simons, 1974). Other researchers have used other unpleasant stimuli to provoke participants, such as noise blasts (e.g., Bartholow, Anderson, Carnagey, & Benjamin, 2005; da Gloria et al., 1989) or personal insults (e.g., Caprara, Renzi, Amolini, D'Imperio, & Travaglia, 1984).

Participant gender, age, and college student status. To test for potential gender differences in the weapons effects, we examined males and females separately. Most weapons effect experiments included only male participants (e.g., Berkowitz & LePage, 1967; Buss, Booker, & Buss, 1972; Klinesmith et al., 2006; Turner & Simons, 1974). However, a number of other experiments included both male and female participants (e.g., Caprara et al., 1984; da Gloria et al., 1989; Lindsay & Anderson, 2000), and one experiment included only female participants (Gallina & Fass, 2014). Because males are typically more physically aggressive and more likely to use weapons than females, weapons might have greater effects on males than on females (Caprara et al., 1984).

To test for potential age differences, we coded the average age of participants in each study. Although most weapons effect studies have used adult participants, some have used child or adolescent participants.

Because some critics have argued that college students, who are often recruited from introductory psychology participant pools, are not representative of "real people" (e.g., Oakes, 1972; Sears, 1986), we also examined college students and other participants separately. Thus, we not only tested the potential for the weapons effect to change as a function of age and whether participants were college students (typically 18 to 21 years old) or. By doing so, we hoped to gain further insight into the generalizability of the weapons effect as a function of age as well as across college student and noncollege student samples.

Study characteristics. We coded whether the study used a between- or a within-subjects design and whether the study was published in a peer-reviewed journal or not. We also coded the year on the research report to determine if the magnitude of the weapons effect has changed over time. Finally, we coded whether the study was conducted in a laboratory or field setting.

Intercoder Reliability

Two independent judges coded all studies included in the meta-analysis. There was 100% agreement on all coded characteristics.

Analysis Strategy

We used Cohen's d as the effect size estimate, which gives the number of standard deviations between the weapons and no weapons conditions. When means, standard deviations, and sample sizes were not reported, we contacted the authors and requested the missing data. Otherwise, we estimated Cohen's d from test statistics using standard formulas (Borenstein et al., 2009). Each effect size was weighted by the inverse of its variance, which is the optimal weight (Hedges & Vevea, 1998).

We used random effects (RE) meta-analytic procedures, which assume that effect sizes differ from population means by both participant-level sampling error and study-level variability (Borenstein, Hedges, & Rothstein, 2007). In contrast, fixed-effects (FE) models assume only participant-level sampling error. RE models are more conservative than FE models, but they require fewer statistical assumptions and allow for generalizations to a broader set of studies than only the ones included in the meta-analysis (Hunter & Schmidt, 2004).

For studies that reported multiple effect sizes, we used a shifting unit of analysis approach (Cooper, 1998). Each statistical test was coded as if it were independent. For example, suppose male and female participants in one study were exposed to guns versus neutral objects (control), and were provoked or not. After exposure, participants completed a measure of aggression (e.g., number of electric shocks given to an accomplice). In this study, four effect-size estimates would be coded (i.e., provoked/males; unprovoked/males; provoked/females; unprovoked/females). For the overall effect, the four effect-size estimates would be averaged so that the study provides only one effect-size estimate. For an analysis examining the effects of weapons on provoked versus unprovoked participants, the study would provide two effect-size estimates (i.e., provoked vs. unprovoked, combining males and females). For an analysis testing for gender differences in the weapons effect, the study would also provide two effect-size estimates (i.e., males vs. females, combining provoked and unprovoked conditions). The shifting unit of analysis was also used for studies that included more than one outcome measure. Shifting the unit of analysis retains as much data as possible without violating the independence assumption that underlies the validity of meta-analytic procedures.

Finally, to assess the robustness of our results to publication bias and outliers, we conducted a comprehensive sensitivity analysis (Greenhouse & Iyengar, 2009; Kepes et al., 2013) to determine the trustworthiness of our obtained results (Kepes & McDaniel, 2013).

Results

There was a significant effect of weapons on aggressive cognition, affect, appraisals, and behavior when the effects for these outcomes were combined, $k = 78$; $\bar{d}_o = 0.29$; 95% confidence interval (CI) = [0.21, 0.36]. Overall, there appears to be a weapons effect for these aggressive outcomes.

Table 1 contains the results of the initial naïve meta-analysis for the different outcomes and different moderator variables. By "naïve" we mean the meta-analytic mean effect without any "adjustment" for potential biases (Copas & Shi, 2000). In the "Sensitivity Analysis" section, we discuss the impact of publication bias and outliers on these naïve estimates. For each analysis, we display the name of the analyzed distribution, the associated number of samples (k), and individual observations (N), the mean effect size, and corresponding 95% CI. We also report Q tests for differences between classes of each moderator. However, many of these Q tests are underpowered.

Outcome Variables

As can be seen in Table 1, weapons increased aggressive thoughts, hostile appraisals, and aggressive behavior. However, weapons did not significantly increase angry feelings. The CI for angry feelings includes zero, probably due to low statistical power because it was based on only 7 samples. As Table 1 shows, the magnitude of the observed weapons effect was not the same for all outcome variables. Specifically, the magnitude of the weapons effect on hostile appraisals was the largest of all the outcomes. To increase statistical power, we combined the outcomes when examining the possible presence moderators of the weapons effect. Table 1 contains the results for the categorical moderators.

Moderator Variables

The magnitude of the observed weapons effect did not differ for provoked and unprovoked participants, and was significant for both groups. The magnitude of the weapons effect tended to be larger for lab studies than for field studies (probably due to tighter control of possible confounds), but the difference was not significant. The CI for lab studies excluded zero (i.e., the effect significantly differed from zero), but the CI for field studies did not. The magnitude of the weapons effect did not differ for between-subjects designs and within-subjects designs, and was significant for both types of designs. The magnitude of the weapons effect was larger when photos of weapons were used than when actual weapons were used. The CI for photos of weapons excluded zero, whereas the CI for actual weapons included zero. The magnitude of the weapons effect did not differ for real and toy weapons, and was significant for both groups. Although this moderator is confounded with age of participants (i.e., toy weapons are used in studies involving children whereas real

Table 1. Effect-Size Estimates and CIs for the Categories of the Moderator Variables Coded.

Variable and class	Between classes effect (Q_b)	k	N	\bar{d}_o and 95% CI
Outcome variable	$Q_b(3) = 5.69, p < .13$			
Cognition		19	3,543	0.28 [0.19, 0.37]
Affect		7	953	0.15 [-0.10, 0.40]
Appraisal		22	1,424	0.43 [0.30, 0.56]
Behavior		38	2,382	0.25 [0.07, 0.43]
Provocation level	$Q_b(1) = 0.56, p < .45$			
None/Low		57	5,550	0.27 [0.18, 0.35]
High		32	1,719	0.36 [0.14, 0.57]
Research setting	$Q_b(1) = 0.27, p < .61$			
Lab		64	6,768	0.30 [0.23, 0.38]
Field		14	900	0.22 [-0.07, 0.51]
Research design	$Q_b(1) = 1.18, p < .28$			
Between		51	5,724	0.25 [0.14, 0.36]
Within		27	1,944	0.34 [0.22, 0.45]
Photos vs. actual	$Q_b(1) = 4.46, p < .04$			
Photos of weapons		43	5,230	0.35 [0.26, 0.44]
Actual weapons		22	1,847	0.12 [-0.08, 0.31]
Real vs. toy	$Q_b(1) = 0.08, p < .78$			
Real weapons		65	7,077	0.28 [0.20, 0.36]
Toy weapons		13	591	0.32 [0.07, 0.58]
Weapon type	$Q_b(2) = 0.67, p < .72$			
Guns		60	5,498	0.28 [0.18, 0.38]
Knives		6	838	0.34 [0.16, 0.52]
Mixed		14	1,453	0.35 [0.20, 0.50]
Publication status	$Q_b(1) = 2.26, p < .14$			
Published		62	5,267	0.32 [0.23, 0.40]
Unpublished		16	2,401	0.16 [-0.02, 0.35]
Participant gender	$Q_b(1) = 1.15, p < .29$			
Males		6	355	0.70 [0.36, 1.03]
Females		6	391	0.43 [0.08, 0.79]
Participant college student status	$Q_b(1) = 0.02, p < .89$			
College student		52	4,201	0.28 [0.18, 0.38]
Nonstudent		26	3,467	0.29 [0.16, 0.42]

Note. Comparisons are for weapons versus nonweapon conditions. Positive effects indicate greater effects when exposed to weapons.

weapons are used in studies involving adults), age did not moderate the results.

The magnitude of the weapons effect did not depend on whether the weapons were guns, knives, or a mixture of various weapons (e.g., guns, knives, swords, grenades). All of the 95% CIs excluded zero.

In terms of publication status, the observed weapons effect was larger for published studies than for unpublished ones. The CI for published studies excluded zero, whereas the CI for unpublished ones included zero.

Only a limited number of studies allowed for analyzable direct comparisons of the weapons effect between male and female participants. Of the six studies that allowed for such comparisons, there was no significant gender effect, although the effect size for males tended to be larger than for females. Both CIs excluded zero; they also overlapped.

To address the question of whether the weapons effect generalized beyond college or university samples, we ran two analyses. One was a meta-regression of mean age of participants on the size of the weapons effect. The other analysis was a direct comparison of college and noncollege student samples on the magnitude of the weapons effect. The mean age of participants did not significantly influence the magnitude of the weapons effect ($b = 0.004$; 95% CI = [-0.006, 0.0143]; $z = 0.84, p < .40$). The magnitude of the weapons effect also did not differ for studies that used college student samples and studies that used nonstudent samples (Table 1). The effect sizes were nearly identical, and both CIs excluded zero.

We also tested whether the magnitude of the observed weapons effect has changed over time using publication year as a moderator variable. There was a significant positive relation between publication year and the magnitude of the weapons effect ($b = 0.005$; 95% CI = [0.0004, 0.010];

$z = 2.12$, $p < .04$), which suggests that the magnitude of the weapons effect has increased over time.

Sensitivity Analysis

As noted in the introduction, we conducted a comprehensive sensitivity analysis to assess the robustness of our results (Greenhouse & Iyengar, 2009; Kepes et al., 2017; Kepes et al., 2013). All sensitivity analyses were conducted in *R* using the metafor (Viechtbauer, 2015) and meta packages (Schwarzer, 2007) and with the recommended RE estimation model. First, we calculated the naïve observed meta-analytic mean estimate (\bar{d}_o) for each distribution as well as the associated statistics (e.g., 95%CI, 90% prediction interval [PI], heterogeneity indices). Next, we conducted a one sample removed analyses to examine the influence of each individual effect size on the obtained meta-analytic results (Borenstein et al., 2009; Kepes et al., 2013). Then, we assessed the potential for publication bias in each individual meta-analytic distribution. Following best practice recommendations (Kepes, Banks, McDaniel, & Whetzel, 2012; Kepes & McDaniel, 2015), we used a comprehensive battery of methods, including trim-and-fill (Duval, 2005), cumulative meta-analysis by precision (Kepes et al., 2012), selection models (Vevea & Woods, 2005), and PET-PEESE (precision-effect test, precision-effect estimate with standard error; Stanley & Doucouliagos, 2014) with a two-tailed significance test to determine whether the PET or PEESE estimate should be used. After these assessments, we used a multivariate battery of influence diagnostics to identify potential outliers (Viechtbauer, 2015; Viechtbauer & Cheung, 2010). We then deleted any identified outlier(s) and reran all analyses. Thus, we performed all analyses twice, once on the original distributions and once on the distributions without identified outlier(s).

For all methods, we used the recommended methodological and statistical options. For example, we implemented trim-and-fill with the recommended FE model and the L_0 estimator (Duval, 2005; Kepes et al., 2012), and also used the RE trim-and-fill model with the same estimator to assess the robustness of the obtained results from the recommended FE model (Moreno et al., 2009). In addition to the regular cumulative meta-analysis by precision (Kepes et al., 2012), we report the cumulative meta-analytic mean of the five most precise effect sizes (for a similar approach, see Stanley, Jarrell, & Doucouliagos, 2010). To implement the selection models, we used a priori models (e.g., Hedges & Vevea, 2005) with p value cut-points to model moderate and severe instances of publication bias as recommended (Vevea & Woods, 2005). Finally, we note that all methods become less stable with small sample sizes (i.e., small distributions), partly due to second-order sampling error and low statistical power (Kepes et al., 2012; Schmidt & Hunter, 2015; Sterne et al., 2011). That is why most publication bias assessment methods should not be used with distributions containing less than 10 effect sizes, including funnel plot- and regression-based methods (Kepes

et al., 2012; Sterne et al., 2011). Therefore, we urge caution when interpreting results from distributions with less than 10 effect sizes.

Once we ran all analyses, rather than relying on a single mean estimate, we examined the range of results (i.e., the mean effect size estimates, taking publication bias and outliers into consideration) to triangulate the location of the “true” mean effect size estimate (Kepes et al., 2012). This approach is recommended to advance the methodological rigor of our sciences (Kepes et al., 2017; Orlitzky, 2012) and is aligned with customer-centric reporting of scientific evidence (Aguinis et al., 2010) as well as evidence-based practice (Kepes, Bennett, & McDaniel, 2014; Kepes & McDaniel, 2013). The results of these analyses are displayed in Table 2 (the bottom panel displays the results without the identified outliers). Columns 1 to 3 report the name of the analyzed distribution as well as the associated number of samples (k) and individual observations (N). Columns 4 to 6 display the naïve meta-analytic results, including the naïve observed mean (\bar{d}_o) and the associated 95% confidence (95% CI) and 90% prediction intervals (90% PI). Columns 7 to 9 display distinct assessments of heterogeneity, Cochran’s Q statistic, I^2 , and tau (τ). Column 10 shows the results of our one-sample removed analysis (OSR; minimum, maximum, and median \bar{d}_o). Columns 11 to 18 display the results from the trim-and-fill analyses; for the recommended FE and the RE models. For each model, we report the side of the funnel plot on which the imputed effect sizes are located (FPS), the number of the imputed effect sizes (ik), the trim-and-fill adjusted mean effect size fixed and random effects estimates ($t\&f_{FE} \bar{d}_o$ or $t\&f_{RE} \bar{d}_o$, respectively) as well as the associated 90%CI. Column 19 reports the cumulative mean for the five most precise samples ($pr_5 \bar{d}_o$); columns 20 and 21, the results from the moderate ($sm_m \bar{d}_o$) and severe selection ($sm_s \bar{d}_o$) models; and column 22, the result of the PET-PEESE ($pp \bar{d}_o$) analysis. Finally, although not discussed due to space considerations, we have included the forest plots that display the contour-enhanced funnel plots with trim-and-fill imputations (using the recommended FE model with the L_0 estimator) as well as the cumulative meta-analyses by precision in the appendix (for interpretation guidelines, see Kepes et al., 2012).

Publication bias. Publication bias seems to have affected many of the naïve meta-analytic mean estimates to a noticeable degree (i.e., by more than 20%; Kepes & McDaniel, 2015). For instance, for studies with a within-subjects design ($k = 27$), the naïve observed meta-analytic mean estimate ($\bar{d}_o = 0.34$) seems to be overestimated by between 0.06 ($sm_m \bar{d}_o = 0.28$; $\Delta = 0.06$ or 18%) and 0.23 ($pp \bar{d}_o = 0.11$; $\Delta_s = 0.23$ or 68%). However, the estimate from the PET-PEESE model ($pp \bar{d}_o$) could be an outlier because it does not converge well with the results of the other methods. Yet, even excluding this estimate from consideration, the degree of overestimation, on average,

Table 2. Meta-Analytic and Publication Bias Results.

Distribution	Publication bias analyses														PET-PEESE										
	Meta-analysis							Trim-and-fill																	
	FE trim-and-fill							RE trim-and-fill																	
	k	N	\bar{d}_o	95% CI	90% PI	Q	I^2	τ	OSR	\bar{d}_o	FPS	ik	$\frac{t\&f_{FE}}{d_o}$	95% CI		t&f _{FE}	95% CI	pr _s	\bar{d}_o	sm _m	\bar{d}_o	sm _s	\bar{d}_o	pp	\bar{d}_o
Original distributions																									
Outcome variable																									
Cognition	19	3,543	.28	[.19, .37]	[.02, .53]	48.09	62.57	.15	.26, .29; .28	L	6	.21	[.12, .30]	L	6	.21	[.12, .30]	.23	.25	.21	.23	.25	.21	.10	
Affect	7	953	.15	[-.10, .40]	[-.34, .65]	19.49	69.21	.27	.07, .24; .14	L	2	.01	[-.25, .27]	L	0	.15	[-.10, .40]	.08	.08	.01	.08	.08	.01	NA	
Appraisal	22	1,424	.43	[.30, .56]	[.01, .85]	76.26	72.46	.25	.40, .47; .43	L	6	.30	[.16, .44]	L	0	.43	[.30, .56]	.46	.38	.29	.46	.38	.29	.26	
Behavior	38	2,382	.25	[-.07, .43]	[-.58, 1.08]	199.49	81.45	.50	.21, .28; .26	L	6	.08	[-.11, .28]	L	0	.25	[.07, .43]	.06	.11	.11	.06	.11	NA	NA	
Provocation level ^a																									
None/low	57	5,550	.27	[.18, .35]	[-.16, .70]	232.18	75.88	.26	.25, .28; .27	L	0	.27	[.18, .35]	L	0	.27	[.18, .35]	.27	.17	.27	.27	.17	NA	.28	
High	32	1,719	.36	[.14, .57]	[-.53, 1.25]	147.50	78.98	.53	.31, .39; .36	L	7	.10	[-.13, .34]	L	0	.36	[.14, .57]	.04	.23	.36	.04	.23	NA	NA	
Setting																									
Lab	64	6,768	.30	[.23, .38]	[-.11, .71]	238.23	73.56	.25	.29, .32; .30	L	7	.24	[.16, .32]	L	10	.21	[.13, .30]	.27	.23	.21	.27	.23	NA	.26	
Field	14	900	.22	[-.07, .51]	[-.61, 1.06]	68.48	81.02	.48	.15, .29; .23	L	5	-.13	[-.45, .18]	L	0	.22	[-.07, .51]	-.23	.10	.22	-.23	.10	NA	NA	
Design																									
Between-subjects design	51	5,684	.25	[.14, .36]	[-.27, .77]	172.27	70.98	.31	.23, .27; .25	L	0	.25	[.14, .36]	L	4	.20	[.09, .31]	.35	.16	.20	.35	.16	NA	.27	
Within-subjects design	27	1,984	.34	[.22, .45]	[-.10, .77]	147.68	82.39	.26	.30, .36; .34	L	7	.21	[.08, .33]	L	7	.20	[.08, .33]	.27	.28	.20	.27	.28	NA	.11	
Photos vs. actual																									
Photos of weapons	43	5,230	.35	[.26, .44]	[-.04, .74]	170.83	75.41	.23	.33, .37; .35	L	10	.24	[.15, .33]	L	10	.24	[.15, .33]	.27	.30	.24	.27	.30	NA	.17	
Actual weapons	22	1,847	.12	[-.08, .31]	[-.55, .79]	88.28	76.21	.39	.07, .15; .12	L	0	.12	[-.08, .31]	L	0	.12	[-.08, .31]	-.01	.01	.12	-.01	.01	NA	.12	
Real vs. toy																									
Real weapons	65	7,077	.28	[.20, .36]	[-.16, .73]	271.92	76.46	.27	.27, .30; .28	L	5	.24	[.15, .32]	L	7	.22	[.13, .30]	.27	.20	.24	.27	.20	NA	.26	
Toy weapons	13	591	.32	[.07, .58]	[-.36, 1.01]	47.43	74.70	.39	.25, .39; .32	L	4	.12	[-.14, .38]	L	3	.18	[-.07, .43]	.02	.22	.12	.02	.22	NA	NA	
Weapon type																									
Guns	60	5,498	.28	[.18, .38]	[-.28, .84]	287.42	79.47	.34	.26, .30; .28	L	7	.20	[.09, .31]	L	7	.20	[.09, .31]	.14	.18	.20	.14	.18	NA	.23	
Knives	6	838	.34	[.16, .52]	[.00, .68]	16.51	69.72	.18	.28, .39; .33	R	1	.39	[.20, .59]	L	0	.34	[.16, .52]	.39	.33	.39	.39	.33	.29	.63	
Mixed	14	1,453	.35	[.20, .50]	[-.04, .74]	49.24	73.60	.23	.31, .39; .35	L	4	.23	[.07, .38]	L	4	.22	[.06, .38]	.28	.30	.22	.28	.30	.18	.15	
Publication status																									
Published	62	5,267	.32	[.23, .40]	[-.12, .76]	235.34	74.08	.26	.30, .33; .32	L	8	.24	[.16, .33]	L	11	.21	[.12, .30]	.27	.24	.24	.27	.24	NA	.26	
Unpublished	16	2,401	.16	[-.02, .35]	[-.39, .71]	75.38	80.10	.32	1.2, .21; .16	L	0	.16	[-.02, .35]	L	0	.16	[-.02, .35]	.19	.07	.16	.19	.07	NA	.14	
Participant gender																									
Male	6	355	.70	[.36, 1.03]	[.09, 1.31]	16.14	69.03	.33	.60, .79; .69	L	1	.63	[.32, .95]	L	0	.70	[.36, 1.03]	.77	.66	.70	.77	.66	.61	.41	
Female	6	391	.43	[.08, .78]	[-.24, 1.10]	21.10	76.30	.37	.32, .55; .42	L	0	.43	[.08, .78]	L	0	.43	[.08, .78]	.36	.37	.43	.36	.37	NA	NA	
Participant college student status																									
College student	52	4,201	.28	[.18, .38]	[-.21, .77]	222.92	77.12	.29	.26, .30; .28	L	2	.26	[.16, .35]	L	5	.22	[.11, .32]	.26	.19	.22	.26	.19	NA	.27	
Nonstudent	26	3,467	.29	[.16, .42]	[-.16, .74]	96.50	74.09	.27	.27, .32; .29	L	4	.21	[.08, .35]	L	4	.21	[.08, .35]	.26	.22	.21	.26	.22	NA	.14	

(continued)

Table 2. (continued)

		Publication bias analyses													PET-PEESE														
		Meta-analysis						Trim-and-fill			FE trim-and-fill			RE trim-and-fill			CMA	Selection models											
Distribution	k	N	\bar{d}_0	95% CI	90% PI	Q	I^2	τ	OSR	\bar{d}_0	FPS	ik	$\frac{t\&f_{FE}}{d_0}$	95% CI	$\frac{t\&f_{FE}}{d_0}$	95% CI	$\frac{t\&f_{FE}}{d_0}$	95% CI	pr_5	\bar{d}_0	sm_m	\bar{d}_0	sm_s	\bar{d}_0	pp	\bar{d}_0			
Distributions without identified outliers																													
Outcome variable																													
Cognition	19				No outlier(s) identified.																								
Affect	7				No outlier(s) identified.																								
Appraisal	22				No outlier(s) identified.																								
Behavior	37	2,346	.21	[.04, .37]	[-.54, .96]	166.01	78.31	.45	.18, .23; .21	L	5	.08	[-.09, .26]		0	.21	[.04, .37]			.06		.08	NA	NA	NA	NA	NA	NA	
Provocation level ^a																													
None/low	57				No outlier(s) identified.																								
High	32				No outlier(s) identified.																								
Setting	63	6,698	.29	[.21, .36]	[-.09, .67]	208.44	70.25	.23	.28, .30; .29	L	8	.22	[.14, .30]		9	.22	[.14, .29]			.27		.22	NA	NA	NA	NA	NA	.24	
Lab	14				No outlier(s) identified.																								
Field	14				No outlier(s) identified.																								
Design																													
Between-subjects design	51				No outlier(s) identified.																								
Within-subjects design	26	1,914	.30	[.20, .41]	[-.08, .69]	116.30	78.50	.23	.28, .33; .30	L	6	.21	[.10, .32]		6	.20	[.09, .32]			.27		.26	NA	NA	NA	NA	NA	.10	
Photos vs. actual																													
Photos of weapons	42	5,160	.33	[.25, .41]	[-.02, .68]	142.13	71.15	.21	.31, .34; .33	L	9	.24	[.15, .32]		9	.24	[.15, .33]			.27		.28	.16	.16	.16	.16	.16	.15	
Actual weapons																													
Real vs. toy	22				No outlier(s) identified.																								
Real weapons	65				No outlier(s) identified.																								
Toy weapons	13				No outlier(s) identified.																								
Weapon type																													
Guns	59	5,428	.26	[.16, .35]	[-.23, .75]	230.35	74.82	.30	.25, .27; .26	L	6	.20	[.10, .30]		7	.18	[.09, .28]			.14		.17	NA	NA	NA	NA	NA	.16	
Knives	6				No outlier(s) identified.																								
Mixed	14				No outlier(s) identified.																								
Publication status																													
Published	62				No outlier(s) identified.																								
Unpublished	16				No outlier(s) identified.																								
Participant gender																													
Male	6				No outlier(s) identified.																								
Female	6				No outlier(s) identified.																								
Participant college student status																													
College student	52				No outlier(s) identified.																								
Nonstudent	26				No outlier(s) identified.																								

Note. FE = fixed effects; RE = random effects; CMA = cumulative meta-analysis; PET-PEESE = precision-effect test-precision effect estimate with standard error; k = number of effect sizes in the analyzed distribution; N = meta-analytic sample size; \bar{d}_0 = RE weighted standardized mean difference; 90% PI = 90% prediction interval; Q = weighted sum of squared deviations from \bar{d}_0 ; I^2 = ratio of true heterogeneity to total variation; τ = between-sample standard deviation; OSR = one-sample removed, including the minimum and maximum effect size and the median \bar{d}_0 ; trim-and-fill = trim-and-fill analysis; FPS = funnel plot side (i.e., side of the funnel plot where samples were imputed; L = left, R = right); ik = number of trim-and-fill effect sizes imputed; $\frac{t\&f_{FE}}{d_0}$ = FE trim-and-fill adjusted 95% CI; $\frac{t\&f_{FE}}{d_0}$ = FE trim-and-fill adjusted 95% CI; $\frac{t\&f_{FE}}{d_0}$ = RE trim-and-fill adjusted 95% CI; $\frac{t\&f_{FE}}{d_0}$ = RE trim-and-fill adjusted 95% CI; pr_5 = RE weighted standardized mean difference of the five most precise effect sizes; sm_m , \bar{d}_0 = one-tailed moderate selection model's adjusted \bar{d}_0 ; sm_s , \bar{d}_0 = one-tailed severe selection model's adjusted \bar{d}_0 ; PET-PEESE \bar{d}_0 = PET-PEESE adjusted \bar{d}_0 ; NA = not applicable (the selection model yielded a nonsensical result due to an inflated variance estimate).

^aTwo studies, Dienstbier et al. (1998) and Page and O'Neal (1977), collapsed their data across provocation levels. Consequently, those studies were removed from the provocation-level subgroups.

appears to be around 0.10 or 29% ($\Delta = 0.13$ or 37% including PET-PEESE \bar{d}_o in the calculations), which can be considered “moderate” (Kepes et al., 2012). Thus, taking publication bias into consideration and triangulating the “true” mean effect size for within-subjects designs, it is likely to be between around 0.21 and 0.24 instead of 0.34. Interestingly, this estimate for within-subjects designs is very close to the naïve mean estimate for between-subjects designs ($\bar{d}_o = 0.25$; $k = 50$). Similarly intriguing is the finding that the publication bias adjusted means for between-subjects designs indicate that the naïve mean (0.25) could be overestimated (e.g., $sm_m \bar{d}_o = 0.16$) or underestimated (e.g., $pr_5 \bar{d}_o = 0.35$). Thus, the naïve mean for between-subjects designs is clearly not robust. On average, however, the adjusted mean is essentially identical to the naïve mean (both are 0.25). Assuming that this average is our best estimate, the design, within-subjects or between-subjects, may not have a meaningful influence of the magnitude of the “true” underlying mean estimate. Yet, we note that this assumption may be questionable given the conflicting results from the publication bias analyses for between-subjects designs.

The results for other distributions are similar. For example, the naïve mean for the “photos of weapons” distribution ($k = 43$) was estimated to be 0.35. Yet, all publication bias assessment methods indicate that this is likely an overestimate by between 0.05 ($sm_m \bar{d}_o = 0.30$; $\Delta = 0.05$ or 14%) and 0.18 ($pp \bar{d}_o = 0.17$; $\Delta = 0.18$ or 51%). On average, the degree of overestimation seems to be around 0.11 (or 30%), which is clearly noticeable and can be considered “moderate” (Kepes et al., 2012). Thus, the “true” underlying observed mean for the effect for “photos of weapons” may be around 0.24, the average of all the adjusted mean estimates. The naïve mean for “actual weapons” ($\bar{d}_o = 0.12$; $k = 22$) seems to be overestimated as well. All publication bias assessment methods yield either equivalent or severely lower magnitude mean estimates, with the difference varying between 0.00 ($t\&f_{FE} \bar{d}_o = 0.12$; $\Delta = 0.00$ or 0%) and 0.13 ($pr_5 \bar{d}_o = -0.01$; $\Delta = 0.13$ or 108%). On average, the overestimation seems to be around 0.05 (or 40%). Yet, these publication bias methods yield estimates identical to the naïve mean, suggesting that the “true” underlying mean would be between 0.07 and 0.12.

Taken together, some distributions seem to be adversely affected by publication bias, leading to over- as well as underestimates of the likely “true” underlying mean effect size. More importantly, however, the obtained results indicate that the potential distorting effect of publication bias varies greatly, sometimes depending on the method used to estimate the influence of this bias. Thus, for some distributions, the exact degree of miss-estimation due to publication bias is likely to be unknown at the moment. As an example, for the outcome variable “behavior” ($k = 38$; $\bar{d}_o = 0.25$), one publication bias assessment method (RE trim-and-fill) suggests that the naïve mean estimate of 0.25 is “free” of publication bias and thus robust. Yet, the other methods indicate that the naïve mean could be “severely” overestimated (i.e., overestimated by at least 40%; Kepes & McDaniel, 2015; pr_5

$\bar{d}_o = 0.06$, $\Delta = 0.19$ or 76%; $sm_m \bar{d}_o = 0.11$, $\Delta = 0.14$ or 56%). Thus, similar to the naïve mean for the between-subjects designs distribution, the naïve mean for aggressive behavior is unlikely to be robust and, the degree of robustness is currently unknown because the methods do not converge on a narrow range around the potentially “true” underlying mean effect.

To gain a clearer understanding of the lack of stability among our publication bias assessment measures, we ran some preliminary supplementary analyses for the “behavior” outcome. Recall that Carlson et al. (1990) argued that weapons were most likely to influence aggression under conditions of high provocation and when participants were not aware of the hypotheses or that they were being evaluated. When provoked participants were aware of the hypotheses or that they were being evaluated, the naïve average effect size is -0.21 (-0.82 , 0.39 ; $k = 4$). When provoked participants were unaware of the hypotheses or that they were being evaluated, the average effect size is 0.41 (0.15 , 0.67 ; $k = 27$). These findings shed light on why the behavioral effects might be so heterogeneous. Yet, we urge caution when interpreting these results because of the very small number of samples in one of the distributions ($k = 4$).

Outliers. Although publication bias seems to have affected the majority of naïve observed mean effect size estimates, our results indicate that outliers tend to have no noticeable effect on the naïve meta-analytic results. Out of the 23 analyzed original distributions, the comprehensive multivariate battery of influence diagnostics identified outliers in only five ($5/23 = 22\%$). However, some of those five distributions were noticeably affected by the identified outliers. For instance, the influence diagnostics identified one outlier in the “within-subjects” distribution, reducing the number of effect sizes from 27 to 26. In turn, the observed mean was estimated to be 0.30 , a decrease of 0.04 or 12% (before outlier removal: $\bar{d}_o = 0.34$). Furthermore, all but two (i.e., sm_s and PET-PEESE) of the publication bias assessment methods yielded identical magnitude adjusted mean estimates, and the estimates from these two methods were very close to the estimate before outlier removal (i.e., 0.26 vs. 0.28 and 0.10 vs. 0.11). These results indicates that the “true” underlying observed mean may be around 0.21 once publication bias and outliers are taken into consideration.

In sum, publication bias and/or outliers seem to have adversely affected some of the observed naïve meta-analytic mean estimates. For some distributions (e.g., within-subjects design), after taking publication bias and, if necessary, outliers into consideration, our results tended to be fairly robust and aligned with the interpretation that the weapons effect is “real” and of similar magnitude to other important effects in the social sciences (Bosco, Aguinis, Singh, Field, & Pierce, 2015; Richard, Bond, & Stokes-Zoota, 2003). For other distributions (e.g., between-subjects design), our results did not

converge, indicating that the naïve meta-analytic mean is not robust and that the magnitude of the “true” underlying effect is currently unknown. Finally, we note that we did not correct the effect size data for measurement error in the outcome variable or other sources of potential error (Kepes et al., 2013; Schmidt & Hunter, 2015). Such psychometric corrections would likely have increased the magnitude of the obtained effect size estimates.

Discussion

The results of the basic or naïve meta-analysis show that merely seeing a weapon can increase aggressive thoughts, hostile appraisals, and aggressive behavior. Our findings extend previous reviews of the weapons effect literature (e.g., Carlson et al., 1990). In particular, the obtained results not only provide additional evidence that the mere presence of weapons can potentially increase aggressive behavior, but more importantly, provide insights into why such an increase might occur. Based on the GAM, there are three possible routes to aggression—a cognitive route, an affective route, and an arousal route. The weapons effect appears to use the cognitive route. Our naïve results indicate that merely seeing a weapon can prime or activate aggressive thoughts in one’s memory. This might partially explain why seeing weapons can increase aggressive behavior. People who have aggressive thoughts active in their minds might be more likely to act on those thoughts and behave in an aggressive manner. Our findings also show that the mere presence of weapons can also increase hostile appraisals. That is, the mere presence of weapons can cause people to believe other people are aggressive and will respond in an aggressive manner in ambiguous situations. This hostile perception of others should increase the likelihood of aggression.

The naïve meta-analysis showed that the weapons effect is quite robust. It occurred inside and outside the lab, for many different kinds of weapons (e.g., guns, knives, spears, swords, hand grenades), for real and toy weapons, for males and females, for college students and nonstudents, and for people of all ages, regardless of whether they were provoked. For some distributions, the weapons effect was also robust to the influence of publication bias and/or outliers. Yet, for other distributions, the weapons effect was not robust to these phenomena. The results from the sensitivity analyses showed that publication bias had a small-to-moderate impact on the cognitive and appraisal outcomes. Given the difficulty in triangulating around a likely “true” effect size for affective and behavioral outcomes, for instance, we recommend interpreting their mean estimates with considerable caution.

Among studies examining behavioral outcomes, there was tremendous variability in terms of research designs, independent and dependent variables, and experimental contexts. Several behavioral studies were designed specifically to reduce the weapons effect by increasing participants’ awareness of the hypothesis or level of evaluation

apprehension (see, for example, Simons & Turner, 1976; Turner & Simons, 1974), which might account for the smaller effect size for aggression. Preliminary tests suggest that when participants were made aware of the hypothesis or were made apprehensive of evaluation, the effect of weapons on aggressive behavior was suppressed, even though participants were highly provoked. However, there is some reason to expect that when participants are unaware of the hypothesis and are highly provoked, the effect of weapons on aggressive behavior is moderate in size. That said, we urge caution in interpreting these supplementary results, especially given that the distribution for behavioral outcomes in which participants were made aware of the hypothesis or made apprehensive of evaluation was very small ($k = 4$).

One counterintuitive finding in our analyses concerned the comparison of real weapons and images of weapons. Specifically, the magnitude of the effect for images of weapons was larger than for real weapons. Although there is no particular theoretical reason for why there should be a difference between real weapons and images of weapons, the difference was significant in the naïve meta-analysis, and the difference remained significant after taking publication bias and potential outliers into consideration. Perhaps participants were more suspicious when they saw real weapons.

Consistent with the dynamics that lead to publication bias (see Kepes et al., 2012), our results suggest that the weapons effect was noticeably larger in published studies than in unpublished studies (Table 2). It is worth noting that the weapons effect was significant in published studies, but not in unpublished studies (Table 1). However, the 95% CIs overlapped for the naïve estimates, and also for both trim-and-fill estimates (Table 2), indicating that the two mean estimates are not statistically significantly different from each other. After taking publication bias into consideration, the magnitude of the mean effect from published studies was noticeably reduced, leading to the differences between published and unpublished studies to be markedly smaller in magnitude but still noticeable. Thus, our results clearly indicate that the literature on the weapons effect, like so many other literatures in the social sciences (e.g., Banks et al., 2015), has been affected by publication bias. Our contour-enhanced funnel plots (see the appendix) shows this graphically; they tend to suggest that studies from the left side of several distributions have been suppressed from the publicly available literature.

It is also worth noting that the weapons effect tended to increase slightly over time. This particular finding is important given that many of the studies included in the meta-analysis, especially the earlier studies, used small samples. This is worth further comment for at least two reasons. First, initial reports based on smaller underpowered samples are particularly vulnerable to being nonreplicable in subsequent studies (Ioannidis, 2005; Trikalinos & Ioannidis, 2005). In other words, initial findings may be inflated, and a decline in effect size may be observed subsequently. Second, many of the

early failures to replicate the weapons effect were themselves based on relatively small samples. More recent studies have been based on larger samples than earlier studies, and arguably the reported effects have been more stable. These assertions are supported by supplemental analyses we conducted. As can be seen from the cumulative meta-analyses by year of publication in the Supplemental Materials, after the earliest studies many effects are relatively stable across time, providing credence for the assertion that many effects have not declined over time. Our supplemental analyses also suggest that, on average, more recent studies have tended to use relatively large samples when compared with earlier published samples, which are likely to get published regardless of the obtained effect size magnitude (because large sample studies tend to obtain statistically significant results regardless of the magnitude of the effect). When examining the forest plots depicting the cumulative meta-analyses by year for drift, we once again urge caution when the distribution is relatively small, especially when it contains less than 10 effect sizes.

In summary, after taking publication bias and outliers into account, we can conclude that the weapons effect is real and noticeable for some outcome variables (e.g., aggressive cognition and hostile appraisals) and some moderators (e.g., photos of weapons) but we cannot necessarily conclude the same about other outcome variables (e.g., aggressive affect and aggressive behavior) and other moderators (e.g., actual weapons).

Theoretical Implications

Although the obtained findings are consistent with the GAM, it is not the only model that can explain these findings. The findings from this meta-analysis are also consistent with other models that have been used to explain the weapons effect, such as those based on classical conditioning, operant conditioning, and priming (e.g., Berkowitz, 1974, 1982, 1983). For example, Berkowitz (1974) used a stimulus–response learning model to explain the weapons effect. According to Berkowitz (1974), weapons become associated with aggression through their frequent pairing with aggression in the mass media and in everyday life. Once that association is made, the mere presence of weapons can elicit aggressive responses when individuals are exposed to them. Other scholars have argued that operant conditioning served as an alternative explanation for both successful and unsuccessful instances of the weapons effect (Ellis, Weinir, & Miller, 1971). According to this perspective, participants not only process the presence of the weapon but the likelihood of reinforcement or punishment when behaving aggressively toward another person. Researchers were already discussing the potential for weapons to serve as cognitive cues that primed aggression as early as the 1970s (e.g., Turner et al., 1977). However, cognitive priming theories specific to aggression did not fully develop until the late 1980s and early 1990s. One example is cognitive neo-association

theory (Berkowitz, 1990), which proposes that aggressive thoughts are linked together in memory, forming an associative network. Once an aggressive thought is processed or stimulated, activation spreads through the network and primes associated thoughts as well. Thus, seeing a weapon can prime or activate other aggressive thoughts in memory. Having aggressive thoughts accessible in memory can increase the likelihood of aggressive behavior.

Most recently, the situated inference model (SIM; Loersch & Payne, 2011) has been used to explain the weapons effect (Engelhardt & Bartholow, 2013). According to this model, exposure to a weapon will lead to an increased accessibility of aggressive thoughts that may lead to aggression if individuals attribute any arousal activated by the weapon to internal processes rather than the weapon itself. This model has been used to explain some of the early failures to replicate the weapons effect, such as when individuals have been made aware of experimental hypotheses. If individuals are aware that weapons can influence their behavior, they presumably make external attributions for any arousal that occurs from exposure to weapons, thus leading to an inhibition of aggression. On the contrary, if individuals are unaware that weapons can influence their behavior, they presumably make internal attributions for any arousal that occurs from exposure to weapons, thus leading to a facilitation of aggression. This theory could not be tested in this meta-analysis, because not enough studies have measured physiological responses to seeing weapons. It is also worth noting that the GAM subsumes many of these other theories, which is why we used it as the theoretical foundation for our meta-analysis.

More generally, our obtained findings shed light on the controversy regarding social priming effects (Benjamin & Bushman, 2016). As scholars have observed (e.g., Molden, 2014), social psychologists have accepted as a given that mere exposure to any of a number of social stimuli will facilitate changes in thoughts, attitudes, and behavioral outcomes, often outside of conscious awareness or control. However, over the past several years, there have been a number of noteworthy failures to replicate highly cited social priming experiments (e.g., Pashler, Coburn, & Harris, 2012; Shanks et al., 2013), leading some scholars to express serious doubts about social priming research as it is currently conducted (Kahneman, 2012). In the case of social cues that should facilitate aggressive outcomes, one would want to consider if the initial observed effects are consistently replicated over time. Based on our findings from the available published and unpublished literature, it appears that weapons function as social primes insofar as they facilitate aggressive thoughts, hostile appraisals, and aggressive behavior. However, the extent to which weapons increase aggressive behavior is less clear. Although priming aggressive thoughts and hostile appraisals should theoretically increase the likelihood of aggression, other factors may play a role in either facilitating or inhibiting aggression. In addition, aggression is further

“down-stream” than cognition and hostile appraisals, and like other social primes may be sensitive to contextual factors that may facilitate or inhibit aggressive behavioral outcomes. These findings tend to be consistent with research on other social stimuli that have been demonstrated to facilitate aggression, such as violent video games (e.g., Anderson et al., 2010; Kepes et al., 2017). Of course, this meta-analysis cannot conclusively settle the social priming debate, but social priming does appear to potentially occur with weapons. (For other examples of social priming, see the recent special issue on the topic published in *Current Opinion in Psychology*; Strack & Schwarz, 2016).

Limitations and Future Research

Like all meta-analyses, this one has limitations. Relatively few studies measured aggressive affective outcomes (i.e., anger), and although the overall estimated effect size was not negligible, it was not significant. Of those studies, most only examined the effect of weapons on anger in nonprovoking situations. It is conceivable that the mere presence of weapons on anger is stronger under provoking circumstances than nonprovoking circumstances. Future research should include measures of anger for both provoked and nonprovoked individuals. Thus, the findings regarding the influence of weapons on anger are inconclusive.

Furthermore, almost no research has examined the effects of weapons on physiological arousal. The one experiment we are aware of (De Oca & Black, 2013) is suggestive, but was conducted using a small sample for the purpose of selecting stimulus materials for a subsequent experiment. Given that participants in experiments who failed to replicate a weapons effect on behavioral outcomes may have misattributed arousal under highly provoking conditions to either the weapon or to some other feature of the experimental context other than the provoking stimulus, as with other scholars (Engelhardt & Bartholow, 2013), it strikes us as imperative that arousal is measured as a means of testing this possibility in a more prospective manner.

Although recent research on the weapons effect has included both male and female participants, very few researchers have directly tested for possible differences in the size of the weapons effect between males and females. That may be due to researchers failing to report null findings when conducting initial tests of gender as a potential moderator, or due to the tendency for contemporary participant pools to be disproportionately female, making such comparisons difficult to test. The few studies available in which gender was a moderator suggest that although the weapons effect might be larger for males than females, the difference was nonsignificant. Thus, we cannot definitively say whether there are gender differences in the weapons effect.

With very few exceptions, researchers have left unexamined the potential moderating role of individual differences in aggressive personality traits (e.g., Caprara et al., 1984) or

life experience with weapons (e.g., Bartholow et al., 2005; Korb, 2016). Thus, we were unable to include such individual differences in this meta-analysis. It is also worth noting that the vast majority of the weapons effect research has been conducted using samples from the United States and Europe. Only one study was conducted in the Middle Eastern sample (Mahjoob, Leyens, & Yzerbyt, 1992), and only two studies were conducted on an Asian sample (Guo, Egan, & Zhang, 2016; Zhang, Tian, Zhang, & Rodkin, 2016). Additional research on the weapons effect utilizing participants from a wider variety of cultures would allow for greater confidence in the generalizability of the weapons effect, much like cross-cultural research has done for other risk factors for aggression, such as violent video games (e.g., Anderson et al., 2010).

In addition, we recommend that future behavioral studies include a provocation manipulation, because earlier researchers found that weapons are most likely to influence aggression under conditions of high provocation (e.g., Carlson et al., 1990). Most behavioral studies in this meta-analysis were underpowered. Thus, we recommend future behavioral studies include larger sample sizes. We concur with Kahneman's (2012) advice to researchers investigating social primes: the onus is on researchers interested in exploring the extent to which weapons influence behavioral outcomes to improve the soundness of their research designs, including determining that their designs are sufficiently powered prior to conducting their research. Failing to do so will merely continue to invite skepticism of the priming effect of weapons on aggressive behavior.

As we noted earlier, our analyses could not take into account factors such as measurement error. Very little attention in the experimental literature has been paid to the reliability of the outcome variables (for a notable exception, see Anderson et al., 1996). Future researchers should report how reliable their measures of, for instance, aggressive thoughts, affect, appraisal, and behavior are.

With regard to our sensitivity analysis, we also need to acknowledge some limitations. First, some of our methods (e.g., contour-enhanced funnel plot, trim-and-fill) tend to attribute the degree of asymmetry in the funnel plot to publication bias (Duval, 2005; Kepes et al., 2012; Moreno et al., 2009). Asymmetry in a funnel plot asymmetry can also be caused by other heterogeneous influences, such as moderators and outliers. To account for such factors, we formed subgroups based on the moderators we identified in our review. In addition, we accounted for outliers by using a thorough battery of multivariate influence diagnostics to identify potential outliers (Viechtbauer, 2015; Viechtbauer & Cheung, 2010), deleting them from the distribution, and then re-analyzing the entire distribution without the identified outlier(s). Consequently, our approach minimized the influence of heterogeneous effects on the reported results. Still, to alleviate additional concerns regarding such influences, we used the contour-enhanced funnel plot (see Appendix) to differentiate between

publication bias and other causes of funnel plot asymmetry, especially the small sample bias (Kepes et al., 2012; Sterne et al., 2011).

We also need to address the performance of the severe selection model, which was recommended by Vevea and Woods (2005) but, in our study, often did not yield a credible adjusted mean estimate. This lack of a credible result is likely because of inflated variance estimates associated with the respective mean estimate (Kepes et al., 2012; Vevea & Woods, 2005). In some prior research, the occurrence of such instances was noticeably smaller (e.g., Kepes et al., 2012; Kepes & McDaniel, 2015). Yet, some other research (e.g., Harrison, Banks, Pollack, O'Boyle, & Short, 2017; O'Boyle, Rutherford, & Banks, 2014) omitted this method to assess publication bias, likely partly because of a high frequency of noncredible results. We thus call for research to investigate the situations under which the selection model does not yield a credible adjusted mean estimate.

Perhaps most important, the reason for asymmetry and thus the bias in the naïve observed mean estimates, publication bias or otherwise, may not be of primary concern. What may matter more is that we know whether something is biased (or not) as well as the approximate degree of the bias on the obtained naïve estimate; in our case the weapons effect. In addition, some of our publication bias methods (e.g., selection models) are less affected by heterogeneous influences and should thus be relatively robust regardless of whether such influences are present or not (Vevea & Woods, 2005). Although we acknowledge that not all methods always converged, many of the methods tended to provide highly confirmatory results for some distributions. Put differently, by using several publication bias assessment methods, we estimated "multiple reference points" to triangulate (Jick, 1979) the location of the "true" underlying observed mean effect. In the instances where the methods converged on a relatively narrow range of possible locations for the "true" observed mean, we can have confidence in our obtained results and the associated conclusions (Kepes & McDaniel, 2013). In the instances where the methods did not converge, we call for more research to better determine the "true" magnitude of the weapons effect.

Finally, we wish to reiterate an important caveat we mentioned earlier. Just as it is for other statistical methods, sample size is of utmost importance when determining the confidence one can have in results obtained from meta-analytic or publication bias methods. Prior research has recommended a minimum of 10 effect sizes in a meta-analytic distribution when conducting publication bias analyses (Kepes et al., 2012; Sterne et al., 2011). Very few, but some (e.g., 4/23 [17%, before outlier removal]), of the distributions we analyzed contained less than the recommended minimum of 10 effect sizes; yet we still applied all methods to them. We did this primarily for transparency reasons. Also, if one feels comfortable reporting a naïve meta-analytic mean and related results (e.g., 95% CI), one should also feel comfortable reporting the results of publication bias and

related methods (Kepes, Banks, & Oh, 2014). However, we once again urge caution when interpreting the results of the naïve meta-analysis and publication bias analyses with distributions containing fewer than 10 effect sizes.

Recommendations for Future Meta-Analyses

We also have recommendations for future meta-analyses. Our results showed that both outliers and publication bias can have distorting effects on naïve meta-analytic results. Hence, suggestions regarding the irrelevance of outlier and publication bias analyses (Aguinis et al., 2011; Dalton, Aguinis, Dalton, Bosco, & Pierce, 2012) seem to be an urban myth. Furthermore, we found that publication bias had much stronger adverse effect on our obtained results than outliers did. This is what previous research found as well (Kepes & McDaniel, 2015). Yet, it may not be that way in other literature areas. Furthermore, although the adverse effects of publication bias were much stronger and more widespread than the effects of outliers, the latter did have a noticeable distorting effect on some meta-analytic results.

Taken together, aligned with prior research from other literature areas in the social sciences (e.g., Banks et al., 2015; Kepes, Banks, & Oh, 2014), our results point to the quite obvious conclusion that the published literature on the weapons effect has been affected by publication bias. We do not know whether other areas in social psychology are affected as well and, if so, the degree to which they are. Therefore, we suggest that future meta-analytic studies assess the robustness of their results by using the comprehensive battery of publication bias methods recommended in the literature and used in this article (Kepes et al., 2012; Kepes et al., 2017; Kepes & McDaniel, 2015). That way, one can have much greater confidence in the robustness of meta-analytic results. Given the current "crisis of confidence" in many of the psychological sciences (Pashler & Wagenmakers, 2012), such a comprehensive assessment may be more important now than ever to ensure that our results and the associated conclusions are trustworthy (Kepes, Bennett, & McDaniel, 2014; Kepes & McDaniel, 2013).

Conclusion

In conclusion, weapons do appear to increase aggressive thoughts and hostile appraisals, although their effect on aggressive behavior is currently less clear. The effect of weapons on aggressive behavior will become clearer when researchers conduct high power studies with a provocation manipulation. That said, overall, the magnitude of the weapons effect may even be increasing over time although that may be due to the fact that much of this research has focused on cognitive and appraisal outcomes since the 1990s. The National Rifle Association correctly notes, "Guns don't kill people; people kill people." But guns are not just neutral stimuli either. As Professor Len Berkowitz noted, although the finger pulls the trigger of a gun, "the trigger may also be

pulling the finger.” When Carlson et al. (1990) published their findings, that statement appeared accurate. As we have shown, the extent to which weapons influence aggressive behavior is still debatable and worthy of further research.

Appendix

Interpretation for Contour-Enhanced Funnel Plots

Funnel plots display the precision (i.e., inverse of the standard error) of an effect size (e.g., a correlation or a standardized mean difference) against its magnitude. Precision is typically shown on the vertical axis (Y-axis) and the effect size magnitude on the horizontal axis (X-axis). Samples with large magnitude effect sizes are plotted on the right side of the X-axis; samples with small magnitude effect sizes on the left side. Because effect sizes from larger samples tend to be more precise (i.e., they are less affected by sampling error) than effect sizes from smaller samples, they tend to cluster at the top of the funnel plot around the population mean. By contrast, smaller samples are usually spread throughout the bottom of the funnel plot because these samples tend to contain more sampling error, resulting in their effect sizes deviating to a greater degree from the population mean.

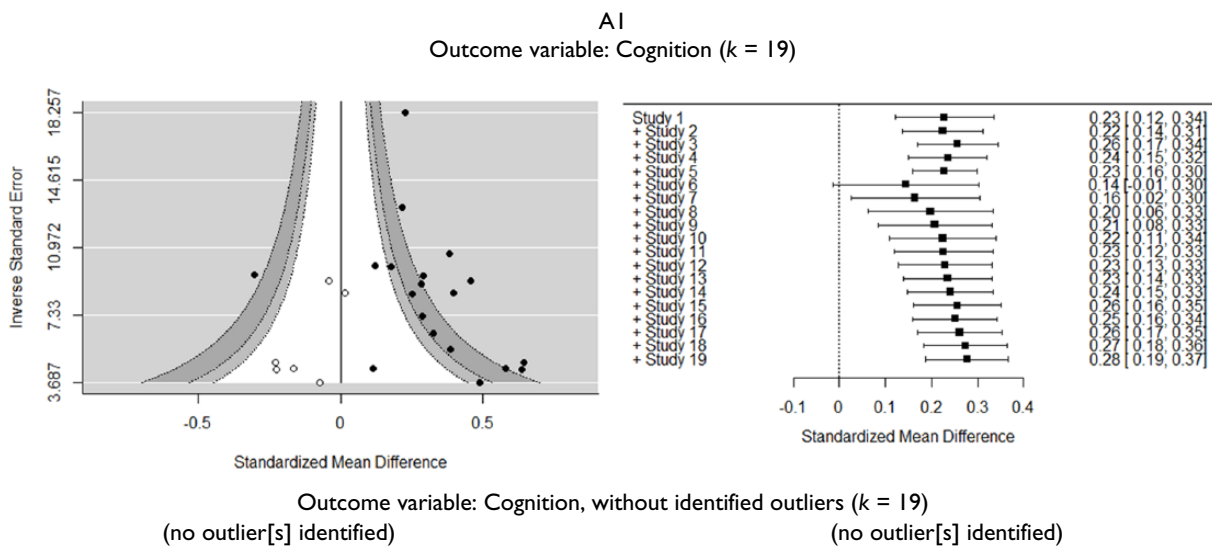
If the distribution is symmetrical, publication bias is likely to be absent. By contrast, if small-sample studies with insignificant effect sizes are not included in the funnel plot, but large-sample studies with statistically significant effect sizes are, the distribution will be asymmetrical. Causes for asymmetry include publication bias and the small sample bias. The contour-enhanced funnel plot helps to distinguish publication bias from the small sample bias and related

causes of funnel plot asymmetry by incorporating contour lines that correspond to values of statistical significance (e.g., $p < .05$, and $p < .1$). If the funnel plot distribution is asymmetric and the “missing” effect sizes are located in areas of statistical insignificance (e.g., $p > .1$), credence is provided to the possibility that funnel plot asymmetry is due to publication bias. By contrast, if the distribution is asymmetric because of missing samples in areas of statistical significance (i.e., the light gray area in the contour-enhanced funnel plot represented by $p < .05$), the small sample bias could be present (for an alternative view, see Kepes, Banks, & Oh, 2014). For more detailed information, see Kepes et al. (2012); Kepes and McDaniel (2015); Peters, Sutton, Jones, Abrams, and Rushton (2008); and Sterne et al. (2011).

Interpretation for the Forest Plots Displaying the Cumulative Meta-Analysis by Precision

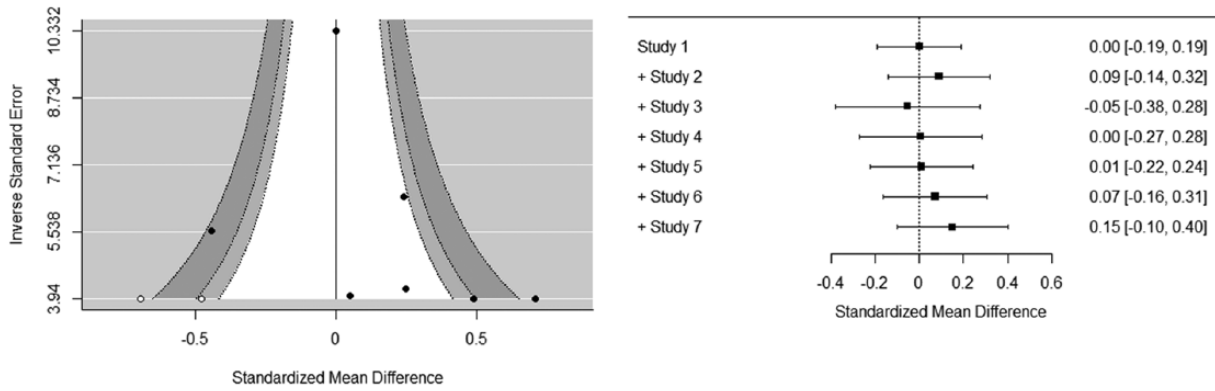
To obtain the plots, the averages study effect sizes were sorted from most precise to least precise and entered into the meta-analysis one at a time in an iterative manner. The lines around the plotted means are the 95% confidence intervals (CIs) for the meta-analytic means. The numbers on the right of each forest plot represent the weighted meta-analytic mean estimate following each iteration and its corresponding CI. A drift from smaller to larger cumulative meta-analytic means is consistent with an inference of statistically insignificant correlations from smaller sample size studies being suppressed (i.e., publication bias). For more detailed information, see Kepes et al. (2012); see also, for example, Kepes et al. (2014) and Kepes and McDaniel (2015).

Funnel Plots (With FE Trim and Fill Imputations) and Forest Plots (Displaying the Cumulative Meta-Analysis by Precision)



A2

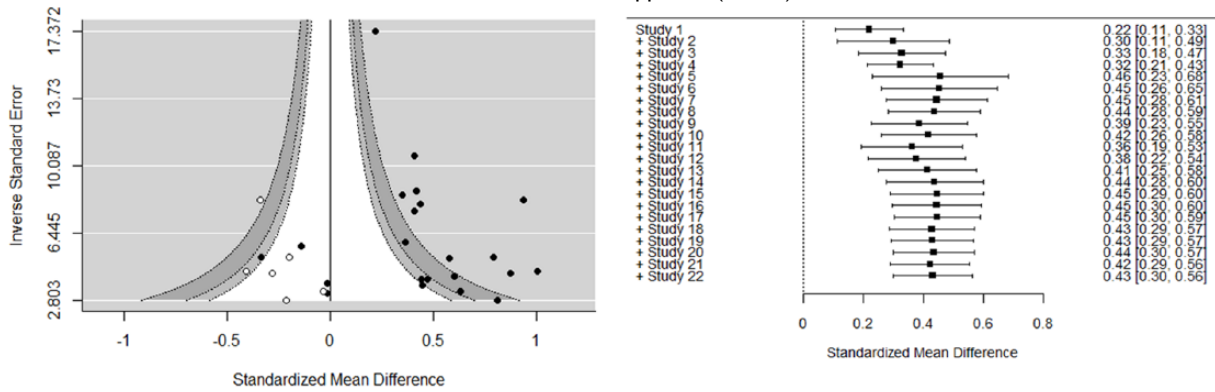
Outcome variable: Affect (k = 7)



Outcome variable: Affect, without identified outliers (k = 7)
 (no outlier[s] identified) (no outlier[s] identified)

A3

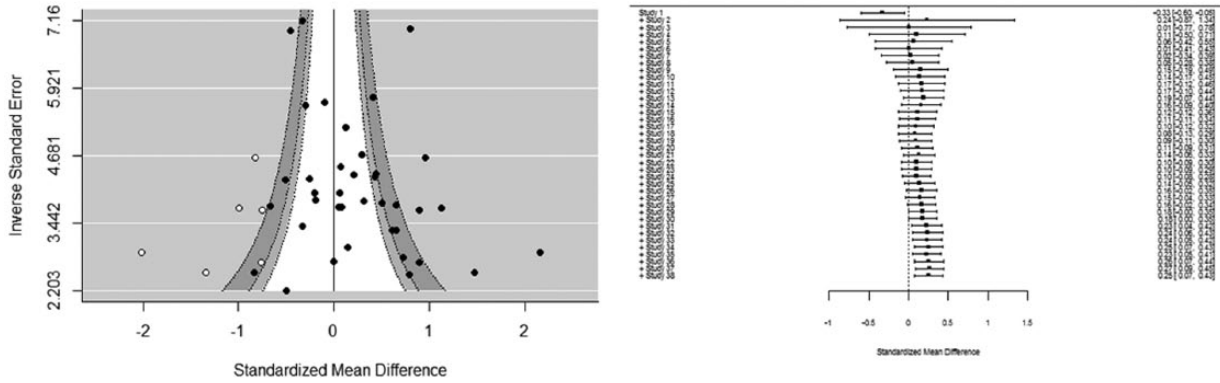
Outcome variable: Appraisal (k = 22)



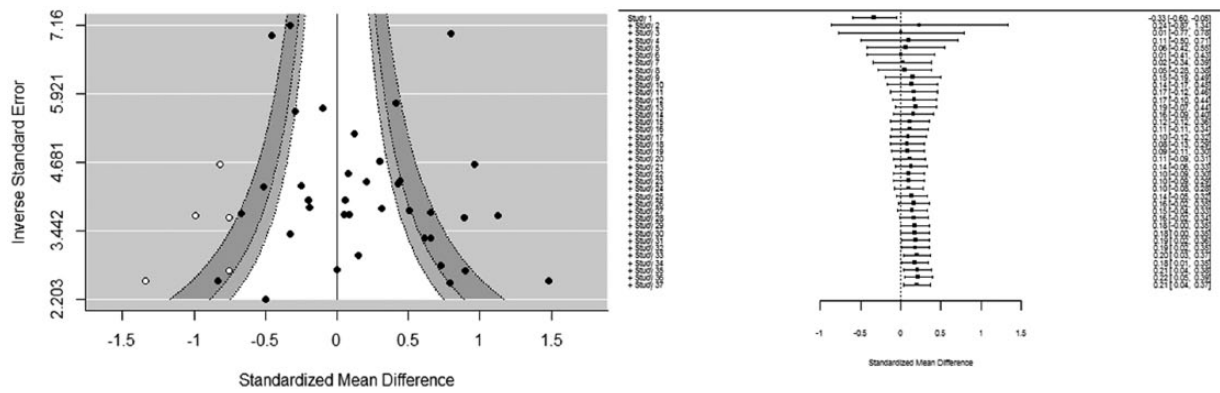
Outcome variable: Appraisal, without identified outliers (k = 22)
 (no outlier[s] identified) (no outlier[s] identified)

A4

Outcome variable: Behavior (k = 38)

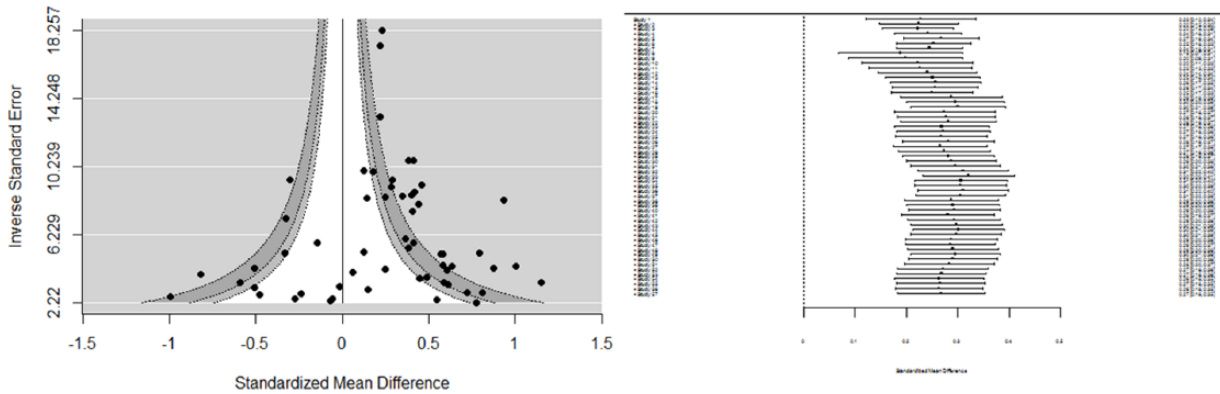


Outcome variable: Behavior, without identified outliers (k = 37)



A5

Provocation level: None/low (k = 57)

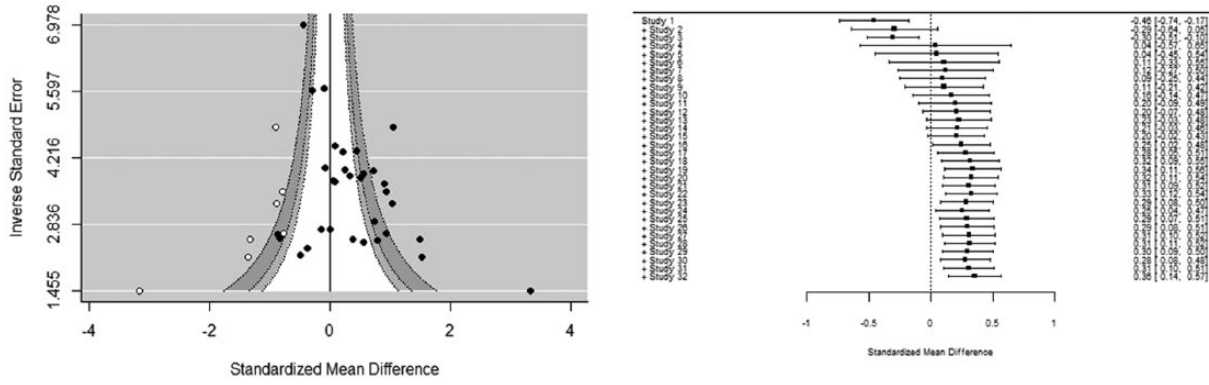


Provocation level: None/low, without identified outliers (k = 57)
(no outlier[s] identified)

(no outlier[s] identified)

A6

Provocation level: High (k = 32)

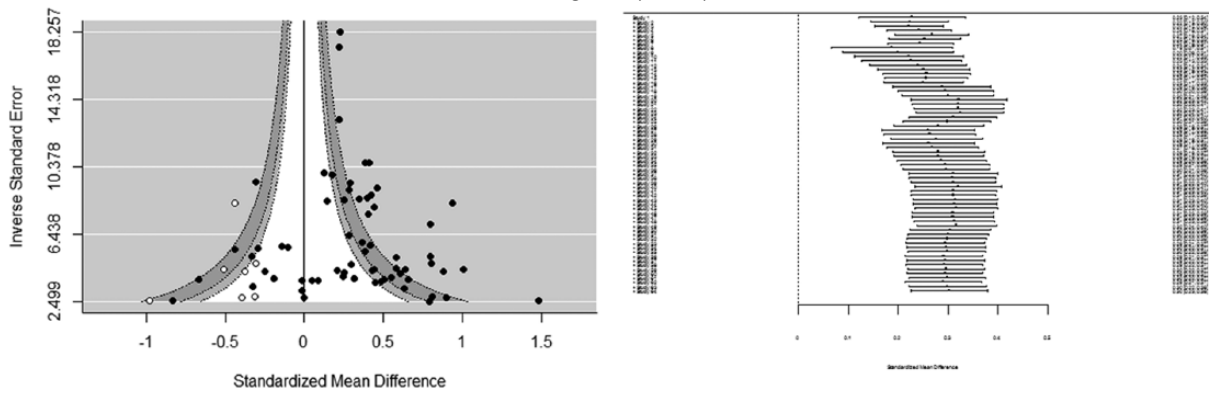


Provocation level: High, without identified outliers (k = 32)
(no outlier[s] identified)

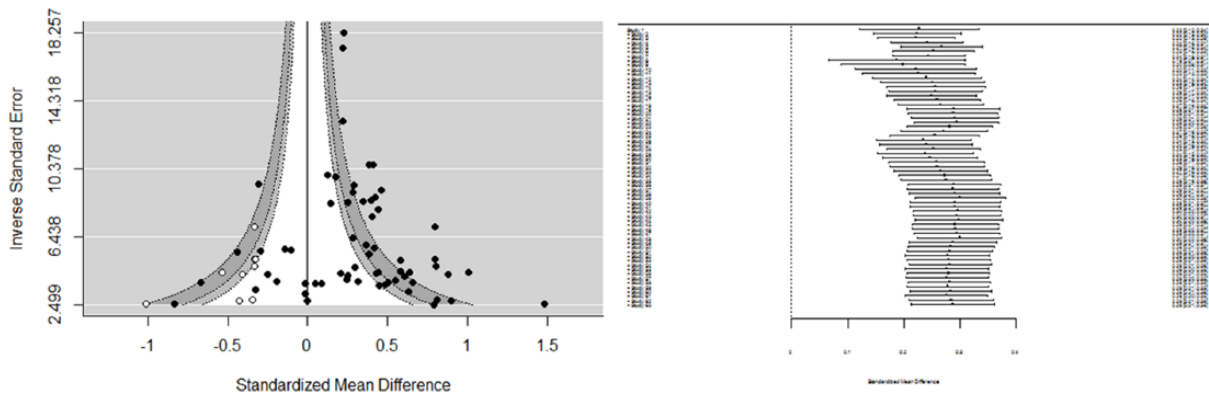
(no outlier[s] identified)

A7

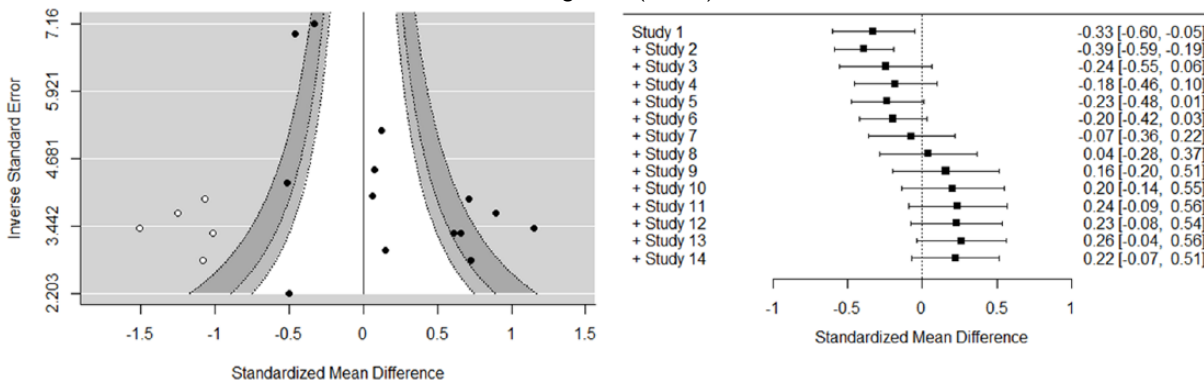
Setting: Lab (k = 64)



Setting: Lab, without identified outliers (k = 63)



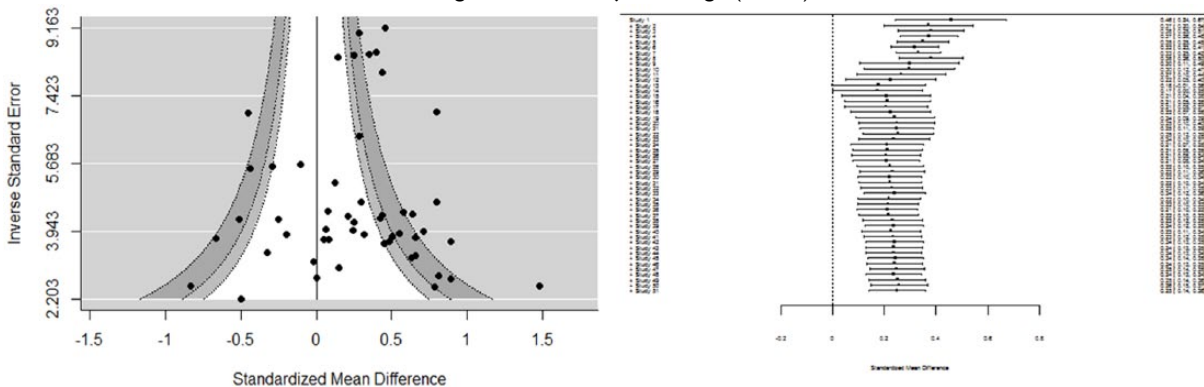
A8
Setting: Field (k = 14)



Setting: Field, without identified outliers (k = 14)
(no outlier[s] identified) (no outlier[s] identified)

A9

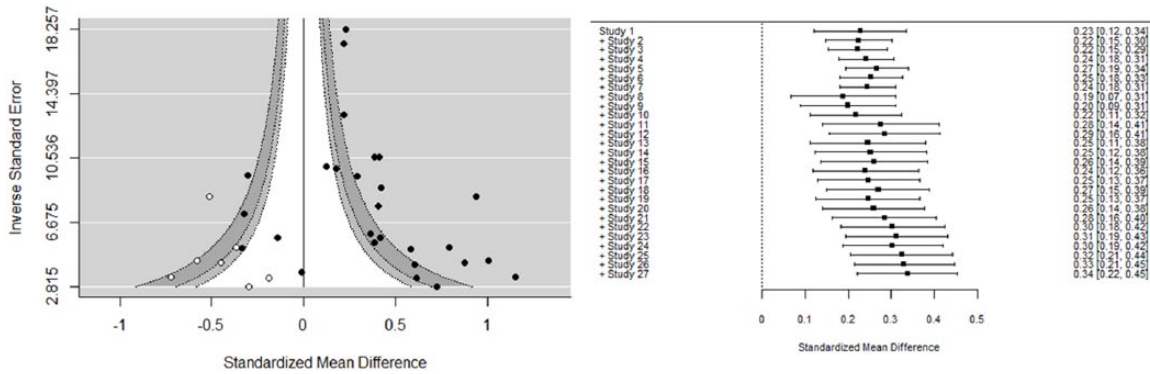
Design: Between-subjects design (k = 51)



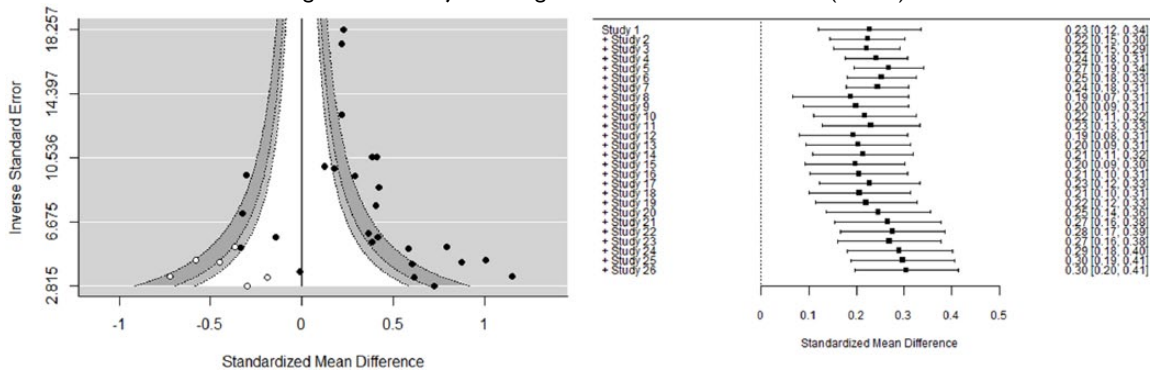
Design: Between-subjects design, without identified outliers (k = 51)
(no outlier[s] identified) (no outlier[s] identified)

A10

Design: Within-subjects design (k = 27)

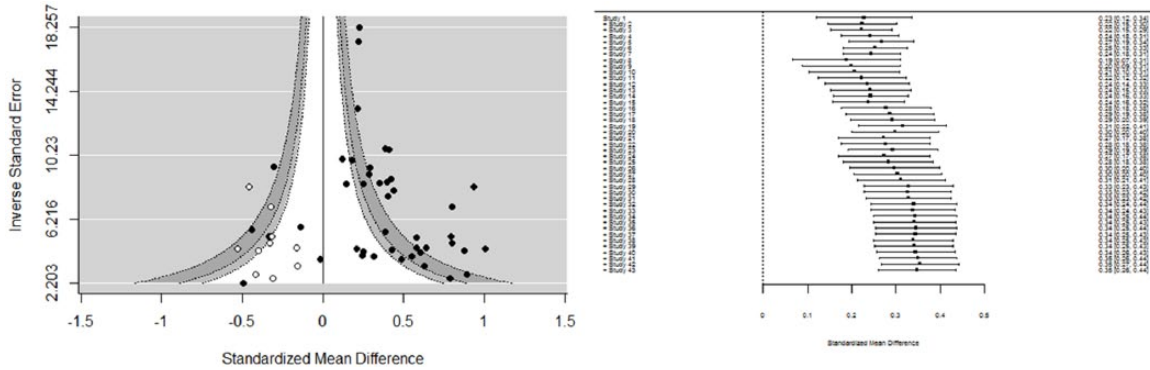


Design: Within-subjects design, without identified outliers (k = 26)

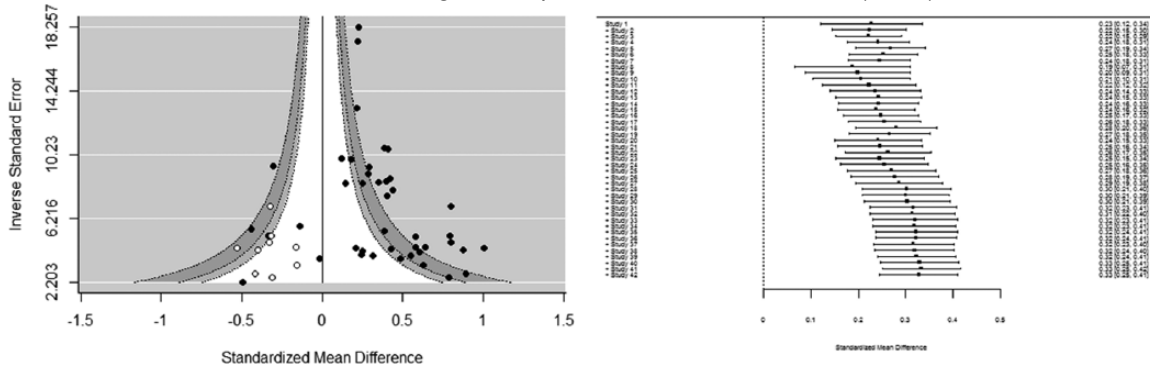


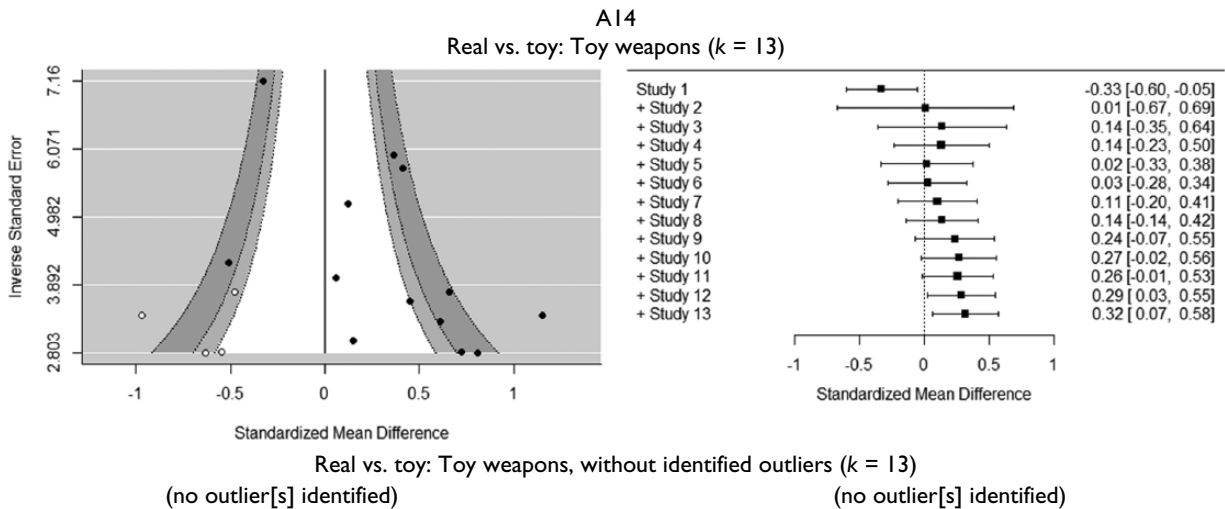
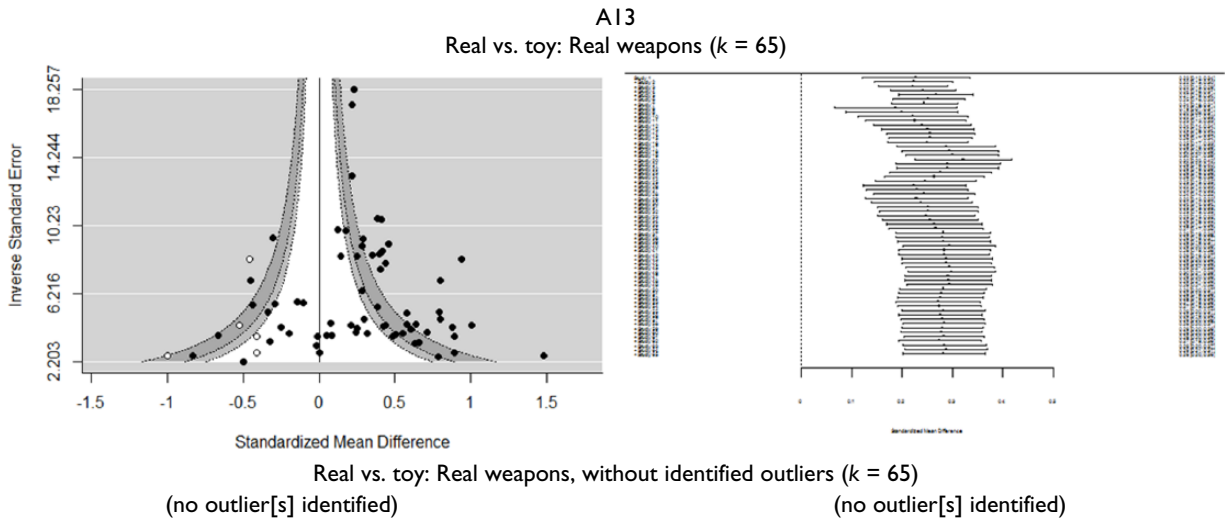
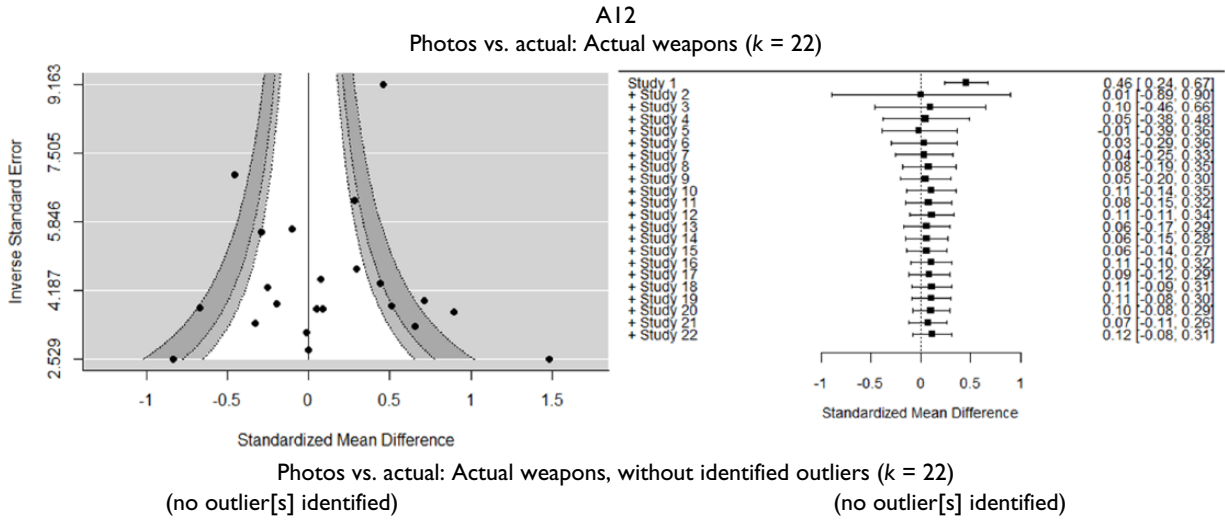
All

Photos vs. actual: Photos of weapons (k = 43)



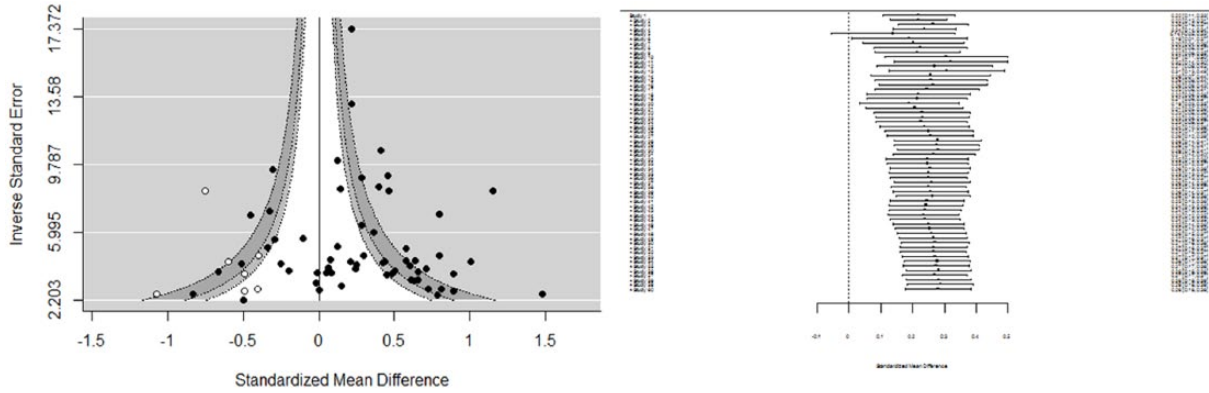
Photos vs. actual: Images of weapons, without identified outliers (k = 42)



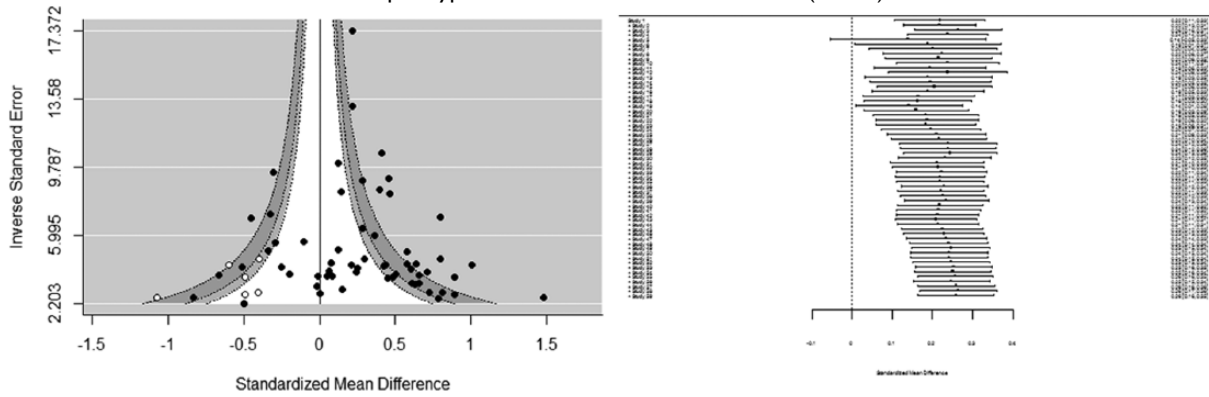


A15

Weapon type: Guns (k = 60)

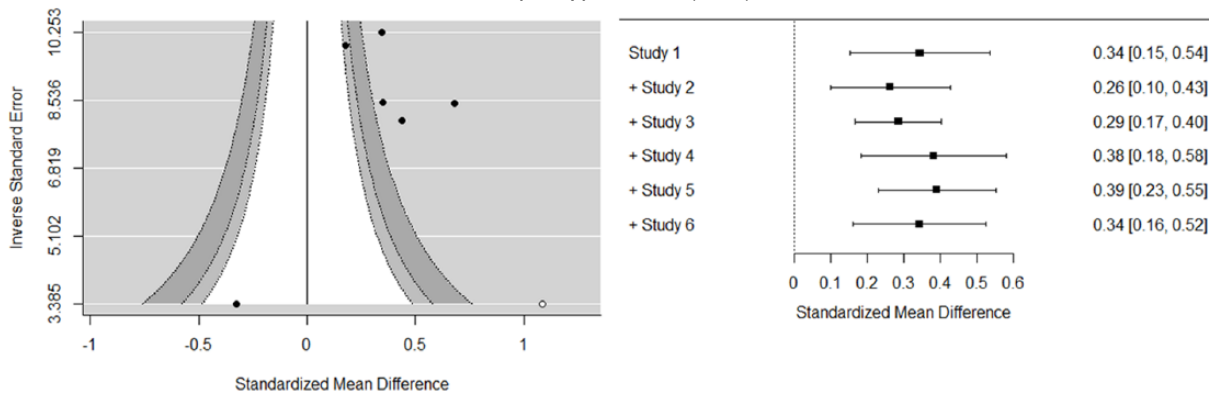


Weapon type: Guns, without identified outliers (k = 59)



A16

Weapon type: Knives (k = 6)

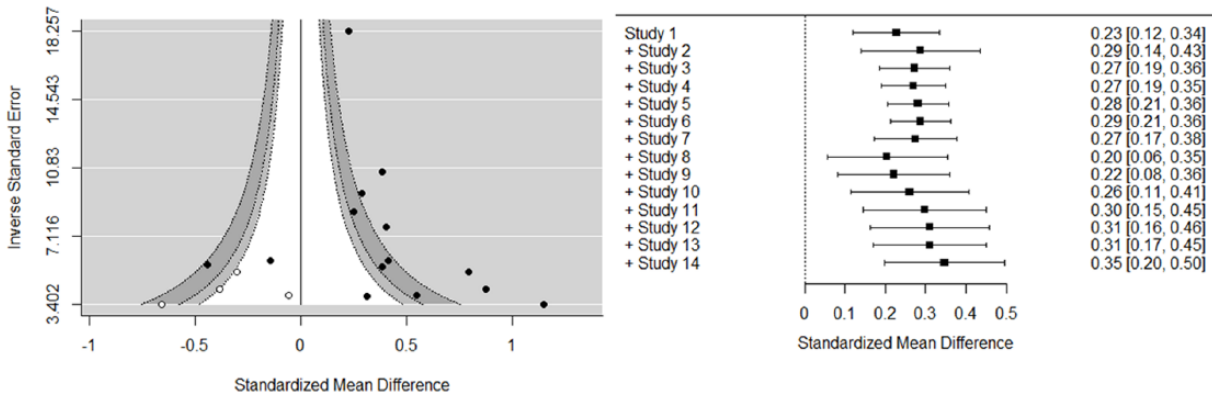


Real vs. toy: Knives, without identified outliers (k = 6)

(no outlier[s] identified)

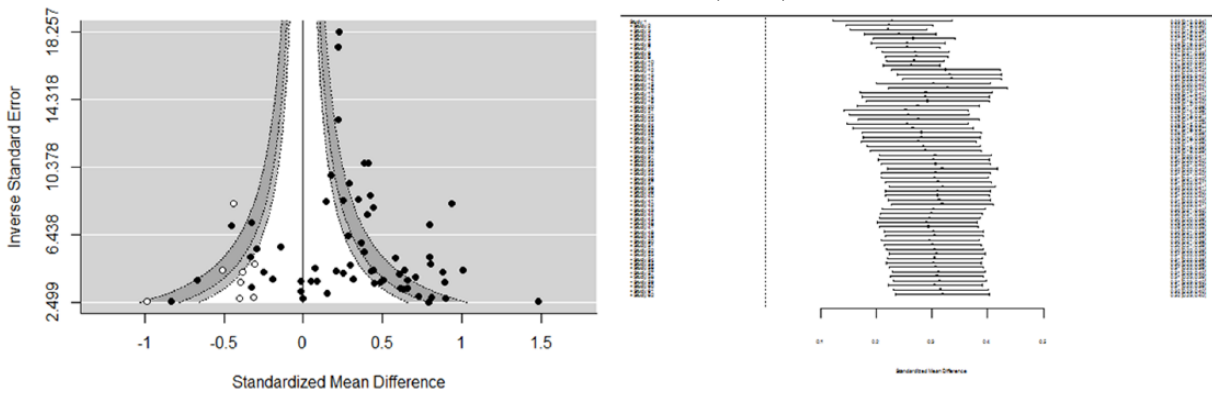
(no outlier[s] identified)

A17
Weapon type: Mixed (k = 14)



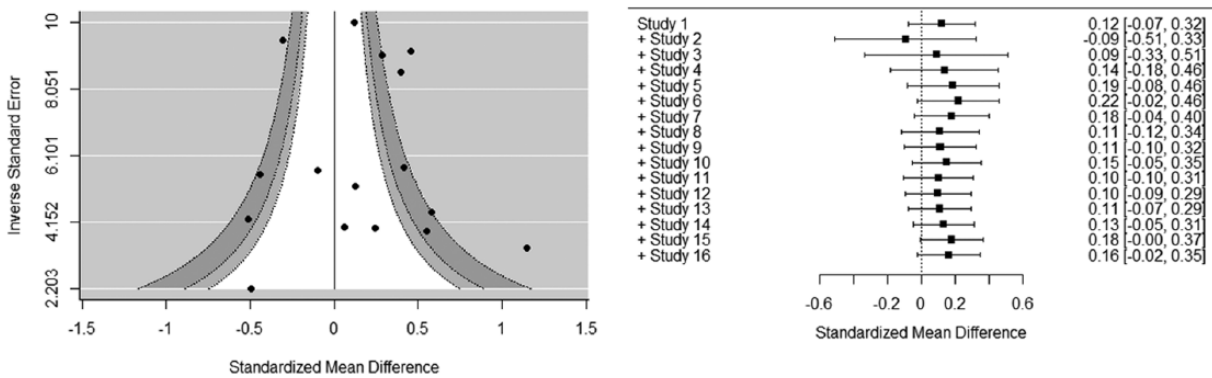
Weapon type: Mixed, without identified outliers (k = 14)
(no outlier[s] identified) (no outlier[s] identified)

A18
Publication status: Published (k = 62)



Publication status: Published, without identified outliers (k = 62)
(no outlier[s] identified) (no outlier[s] identified)

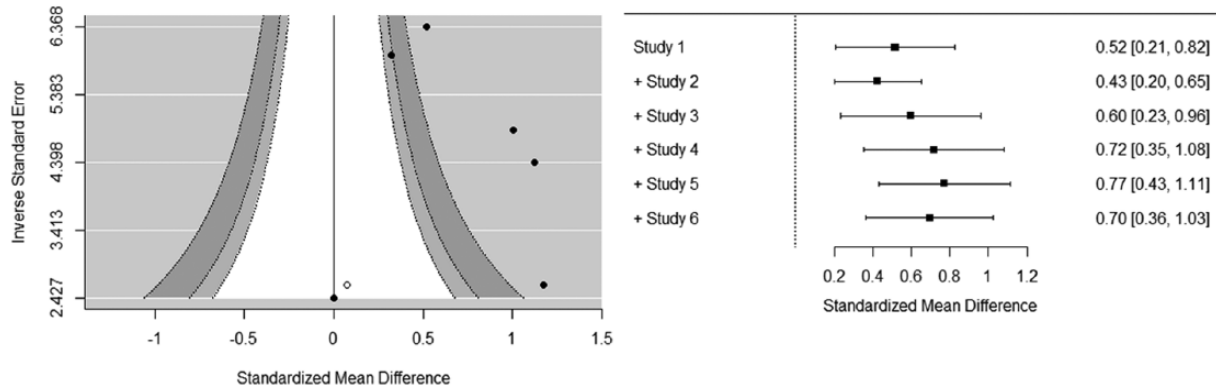
A19
Publication status: Unpublished (k = 16)



Publication status: Unpublished, without identified outliers (k = 16)
(no outlier[s] identified) (no outlier[s] identified)

A20

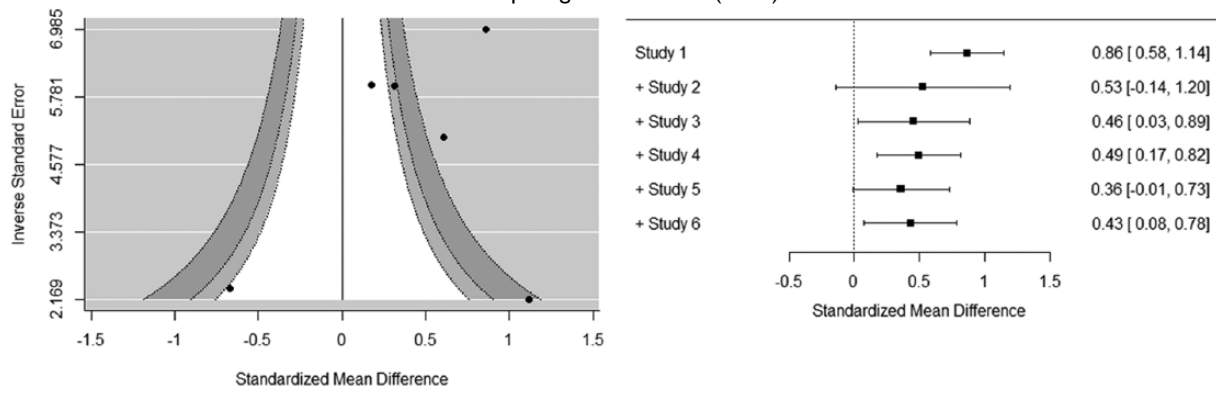
Participant gender: Male (k = 6)



Participant gender: Male, without identified outliers (k = 6)
 (no outlier[s] identified) (no outlier[s] identified)

A21

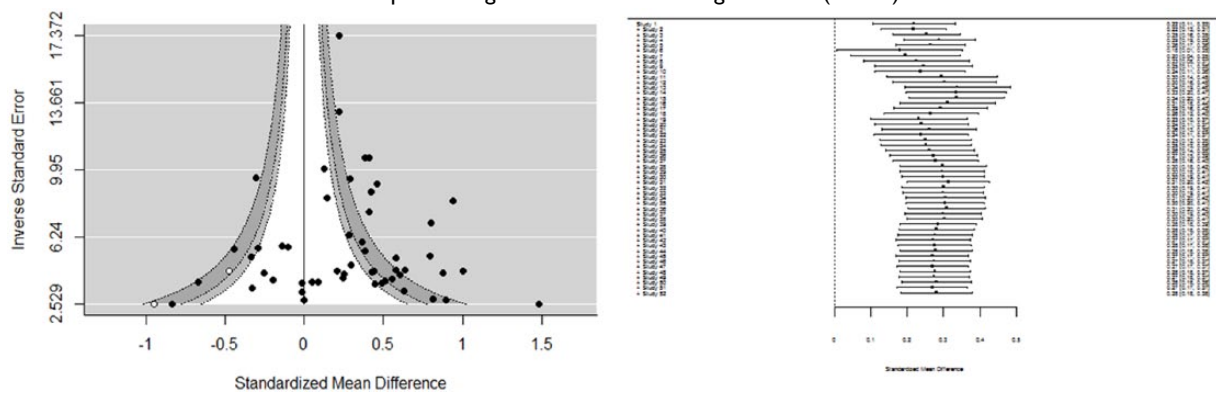
Participant gender: Female (k = 6)



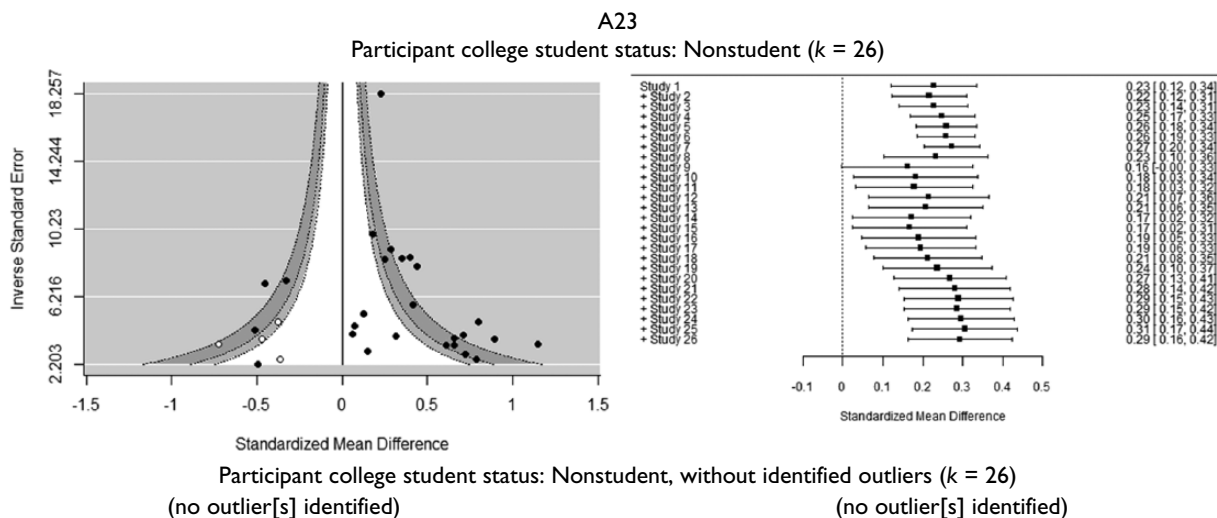
Participant gender: Female, without identified outliers (k = 6)
 (no outlier[s] identified) (no outlier[s] identified)

A22

Participant college student status: College student (k = 52)



Participant college student status: College student, without identified outliers (k = 52)
 (no outlier[s] identified) (no outlier[s] identified)



Authors' Note

The data for the meta-analysis can be obtained from the first author. This article is significantly revised from the version that appeared originally on OnlineFirst due to an error in the database upon which the original analyses were based. We corrected the error and reanalyzed the data. The results of our re-analyses are presented in the revisions to both Table 1 and Table 2. There were three major changes. First, we found it prudent to revise our discussion of the behavioral outcomes and advise greater caution in interpreting the obtained results. We offered one possible explanation for the heterogeneity in behavioral outcomes. Specifically, behavioral effects were negative when participants knew the researchers were testing the hypothesis that weapons can increase aggression, whereas behavioral effects were positive when participants were unaware of the hypothesis being tested. We deleted the subsection in the Discussion section regarding practical implications, given the inconclusive nature of the findings regarding behavioral outcomes. Second, the magnitude of the weapons effect appears significantly larger for weapon images than for real weapons, even when taking the impact of publication bias and outliers into account. Third, the robustness of the obtained results changed noticeably. Using the revised database, the results of our sensitivity analyses indicated that several naïve meta-analytic means were noticeably affected by publication bias. Consequently, not all naïve mean estimates are necessarily robust and caution should be exercised when interpreting some of the naïve mean estimates.

We would like to thank Dr. Joseph Hilgard and Dr. Michael Borenstein for their helpful comments and suggestions.

Acknowledgments

The authors would like to thank Meagan Crosby for her assistance in coding studies. The authors would also like to thank Dr. Joseph Hilgard and Dr. Michael Borenstein for their helpful comments and suggestions.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Arlin J. Benjamin Jr.  <https://orcid.org/0000-0002-5802-115X>

Supplemental Material

Supplemental material is available online with this article.

References

- References marked with an asterisk indicate studies included in the meta-analysis.
- Aguinis, H., Werner, S., Abbott, J. L., Angert, C., Park, J. H., & Kohlhausen, D. (2010). Customer-centric science: Reporting significant research results with rigor, relevance, and practical impact in mind. *Organizational Research Methods, 13*, 515-539. doi:10.1177/1094428109333339
- Aguinis, H., Pierce, C. A., Bosco, F. A., Dalton, D. R., & Dalton, C. M. (2011). Debunking myths and urban legends about meta-analysis. *Organizational Research Methods, 14*, 306-331. doi:10.1177/1094428110375720
- *Anderson, C. A., Anderson, K. B., & Deuser, W. E. (1996). Examining an affective aggression framework: Weapon and temperature effects on aggressive thoughts, affect, and attitudes. *Personality and Social Psychology Bulletin, 22*, 366-376. doi:10.1177/0146167296224004
- *Anderson, C. A., Benjamin, A. J., Jr., & Bartholow, B. D. (1998). Does the gun pull the trigger? Automatic priming effects of weapon pictures and weapon names. *Psychological Science, 9*, 308-314. doi:10.1111/1467-9280.00061
- Anderson, C. A., & Bushman, B. J. (1997). External validity of "trivial" experiments: The case of laboratory aggression. *Review of General Psychology, 1*, 19-41. doi:10.1037//1089-2680.1.1.19

- Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53*, 27-51. doi:10.1146/annurev.psych.53.100901.135231
- Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., . . . Barlett, C. P. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries: A meta-analytic review. *Psychological Bulletin, 136*, 151-173. doi:10.1037/a0018251
- Archer, J. (1988). *The behavioral biology of aggression*. Cambridge, UK: Cambridge University Press.
- Banks, G. C., Kepes, S., & McDaniel, M. A. (2015). Publication bias: Understanding the myths concerning threats to the advancement of science. In C. E. Lance & R. J. Vandenberg (Eds.), *More statistical and methodological myths and urban legends* (pp. 36-64). New York, NY: Routledge.
- Barlett, C. P., & Anderson, C. A. (2011). Reappraising the situation and its impact on aggressive behavior. *Personality and Social Psychology Bulletin, 37*, 1564-1573. doi:10.1177/0146167211423671
- Baron, R. A., & Richardson, D. (1994). *Human aggression*. New York, NY: Plenum Press.
- *Bartholow, B. D., Anderson, C. A., Carnagey, N. L., & Benjamin, A. J., Jr. (2005). Interactive effects of life experience and situational cues on aggression: The weapons priming effect in hunters and nonhunters. *Journal of Experimental Social Psychology, 41*, 48-60. doi:10.1016/j.jesp.2004.05.005
- *Bartholow, B. D., & Heinz, A. (2006). Alcohol and aggression without consumption: Alcohol cues, aggressive thoughts, and hostile perception bias. *Psychological Science, 17*, 30-37. doi:10.1111/j.1467-9280.2005.01661.x
- Benjamin, A. J., Jr., & Bushman, B. J. (2016). The weapons priming effect. *Current Opinion in Psychology, 12*, 45-48.
- *Benjamin, A. J., Jr., Crosby, M., & Bushman, B. J. (2015). *Guns prohibited images prime aggressive thoughts*. Unpublished manuscript.
- Berkowitz, L. (1968). Impulse, aggression, and the gun. *Psychology Today, 2*, 19-22.
- Berkowitz, L. (1971). The "weapons effect," demand characteristics, and the myth of the compliant subject. *Journal of Personality and Social Psychology, 20*, 332-338. doi:10.1037/h0031804
- Berkowitz, L. (1974). Some determinants of impulsive aggression: Role of mediated associations with reinforcements for aggression. *Psychological Review, 81*, 165-176. doi:10.1037/h0036094
- Berkowitz, L. (1982). Aversive conditions as stimuli to aggression. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 15, pp. 249-285). Orlando, FL: Academic Press.
- Berkowitz, L. (1983). The experience of anger as a parallel process in the display of impulsive, "angry" aggression. In R. Geen & E. I. Donnerstein (Eds.), *Aggression: Theoretical and empirical reviews* (Vol. 1, pp. 103-133). Orlando, FL: Academic Press.
- Berkowitz, L. (1990). On the formation and regulation of anger and aggression: A cognitive-neoassociationistic analysis. *American Psychologist, 45*(4), 494-503. doi:10.1037/0003-066X.45.4.49
- *Berkowitz, L., & LePage, A. (1967). Weapons as aggression-eliciting stimuli. *Journal of Personality and Social Psychology, 7*, 202-207. doi:10.1037/h0025008
- Bettencourt, B. A., & Kernahan, C. (1997). A meta-analysis of aggression in the presence of violent cues: Effects of gender differences and aversive provocation. *Aggressive Behavior, 21*, 447-456. doi:10.1002/(SICI)1098-2337(1997)23:6<447::AID-AB4>3.0.CO;2-D
- Bettencourt, B. A., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin, 119*, 422-447. doi:10.1037/0033-2909.119.3.422
- Bettencourt, B. A., Talley, A., Benjamin, A. J., Jr., & Valentine, J. (2006). Personality and aggressive behavior under provoking and neutral conditions: A meta-analytic review. *Psychological Bulletin, 132*, 751-777. doi:10.1037/0033-2909.132.5.751
- Blanchette, I. (2006). Snakes, spiders, guns, and syringes: How specific are evolutionary constraints on the detection of threatening stimuli? *The Quarterly Journal of Experimental Psychology, 59*(8), 1484-1504. doi:10.1080/02724980543000204
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, UK: John Wiley.
- Borenstein, M., Hedges, L. V., & Rothstein, H. R. (2007). *Meta analysis: Fixed effect versus random effects*. Retrieved from <http://www.Meta-Analysis.com>
- Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology, 100*, 431-449. doi:10.1037/a0038047
- Bushman, B. J. (2016). Violent media exposure and hostile appraisals: A meta-analytic review. *Aggressive Behavior, 42*, 605-613. doi:10.1002/ab.21655
- *Bushman, B. J. (in press). Guns automatically prime aggressive thoughts, regardless of whether a "good guy" or "bad guy" holds the gun. *Social Psychological and Personality Science*. Advance online publication. doi:10.1177/1948550617722202
- Bushman, B. J., & Anderson, C. A. (1998). Methodology in the study of aggression: Integrating experimental and nonexperimental findings. In R. G. Geen & E. Donnerstein (Eds.), *Human aggression: Theories, research, and implications for policy* (pp. 23-48). New York, NY: Academic Press.
- Bushman, B. J., & Huesmann, L. R. (2006). Short-term and long-term effects of violent media on aggression in children and adults. *Archives of Pediatrics & Adolescent Medicine, 160*, 348-352. doi:10.1001/archpedi.160.4.348
- *Bushman, B. J., Kerwin, T., Whitlock, T., & Weisenberger, J. M. (2017). The weapons effect on wheels: Motorists drive more aggressively when there is a gun in the vehicle. *Journal of Experimental Social Psychology, 73*, 82-85. doi:10.1016/j.jesp.2017.06.007
- *Buss, A. H., Booker, A., & Buss, E. (1972). Firing a weapon and aggression. *Journal of Personality and Social Psychology, 22*, 296-302. doi:10.1037/h0032869
- *Caprara, G. V., Renzi, P., Amolini, P., D'Imperio, G., & Travaglia, G. (1984). The eliciting cue value of aggressive slides reconsidered in a personological perspective: The weapons effect and irritability. *European Journal of Social Psychology, 14*, 313-322. doi:10.1002/ejsp.2420140306
- Carlson, M., Marcus-Newhall, A., & Miller, N. (1990). Effects of situational aggression cues: A quantitative review. *Journal of Personality and Social Psychology, 58*, 622-633. doi:10.1037/0022-3514.58.4.622

- Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: SAGE.
- Copas, J., & Shi, J. Q. (2000). Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics*, *1*, 247-262. doi:10.1093/biostatistics/1.3.247
- *da Gloria, J., Duda, D., Pahlavan, F., & Bonnet, P. (1989). "Weapons effect" revisited: Motor effects of the reception of aversive stimulation and exposure to pictures of firearms. *Aggressive Behavior*, *15*, 265-271. doi:10.1002/ab.2480150401
- Dalton, D. R., Aguinis, H., Dalton, C. M., Bosco, F. A., & Pierce, C. A. (2012). Revisiting the file drawer problem in meta-analysis: An assessment of published and non-published correlation matrices. *Personnel Psychology*, *65*, 221-249. doi:10.1111/j.1744-6570.2012.01243.x
- *De Oca, B. M., & Black, A. A. (2013). Bullets versus burgers: Is it threat or relevance that captures attention? *The American Journal of Psychology*, *126*, 287-300. doi:10.5406/amerjpsyc.126.3.0287
- Dienstbier, R. A., Roesch, S. C., Mizumoto, A., Hemenover, S. H., Lott, R. C., & Carlo, G. (1998). Effects of weapons on guilt judgments and sentencing recommendations for criminals. *Basic and Applied Social Psychology*, *20*(2), 93-102. doi:10.1207/s15324834basps2002_1
- Duval, S. J. (2005). The "trim-and-fill" method. In H. R. Rothstein, A. Sutton, & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 127-144). West Sussex, UK: John Wiley.
- *Ellis, D. P., Weinir, P., & Miller, I. I. (1971). Does the trigger pull the finger? An experimental test of weapons as aggression-eliciting stimuli. *Sociometry*, *34*, 453-465.
- Engelhardt, C. R., & Bartholow, B. D. (2013). Effects of situational cues on aggressive behavior. *Social and Personality Psychology Compass*, *7*, 762-774. doi:10.1111/spc3.12067
- *Epstein, J. F. (1980). *Firearms experience and the weapons effect* (Doctoral dissertation). Tulane University, New Orleans, LA.
- Ferguson, C. J., & Brannick, M. T. (2011). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods*, *17*, 120-128. doi:10.1037/a0024445
- Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, *7*, 555-561. doi:10.1177/1745691612459059
- Fiedler, K. (2011). Voodoo correlations are everywhere—Not only in neuroscience. *Perspectives on Psychological Science*, *6*, 163-171. doi:10.1177/1745691611400237
- *Fischer, D. G., Kelm, H., & Rose, A. (1969). Knives as aggression-eliciting stimuli. *Psychological Reports*, *24*, 755-760. doi:10.2466/pr0.1969.24.3.755
- Frodi, A. (1975). The effect of exposure to weapons on aggressive behavior from a cross-cultural perspective. *International Journal of Psychology*, *10*(4), 283-292. doi:10.1080/00207597508247340
- *Gallina, M. F., & Fass, W. (2014). The weapons effect in college females. *Violence and Gender*, *1*, 165-169. doi:10.1089/vio.2014.0020
- *Goff, K. E. (1995). *The relation to violent and nonviolent toys to play behavior in preschoolers* (Unpublished doctoral dissertation). Iowa State University, Ames.
- Greenhouse, J. B., & Iyengar, S. (2009). Sensitivity analysis and diagnostics. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed.). (pp. 417-433). New York, NY: Russell Sage Foundation.
- Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, *82*, 1-20.
- Greitemeyer, T., & Mügge, D. O. (2014). Video games do affect social outcomes: A meta-analytic review of the effects of violent and prosocial video game play. *Personality and Social Psychology Bulletin*, *40*, 578-589. doi:10.1177/0146167213520459
- *Guo, X., Egan, V., & Zhang, J. (2016). Sense of control and adolescents' aggression: The role of aggressive cues. *Psych Journal*, *5*, 263-274. doi:10.1002/pchj.151
- *Halderman, B. L., & Jackson, T. T. (1979). Naturalistic study of aggression: Aggressive stimuli and horn honking: A replication. *Psychological Reports*, *45*, 880-882. doi:10.2466/pr0.1979.45.3.880
- Harrison, J. S., Banks, G. C., Pollack, J. M., O'Boyle, E. H., & Short, J. (2017). Publication bias in strategic management research. *Journal of Management*, *43*, 400-425. doi:10.1177/0149206314535438
- Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, *3*, 486-504.
- Hedges, L. V., & Vevea, J. L. (2005). Selection methods approaches. In H. R. Rothstein, A. Sutton, & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment, and adjustments* (pp. 145-174). West Sussex, UK: John Wiley.
- *Hemenway, D., Vrinotis, M., & Miller, M. (2006). Is an armed society a polite society? Guns and road rage. *Accident Analysis and Prevention*, *38*, 687-695. doi:10.1016/j.aap.2005.12.014
- *Holbrook, C., Galperin, A., Fessler, D. M. T., Johnson, K. L., Bryant, G. A., & Haselton, M. G. (2014). If looks could kill: Anger attributions are intensified by affordances for doing harm. *Emotion*, *14*, 455-461. doi:10.1037/a0035826
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis* (2nd ed.). Thousand Oaks, CA: SAGE.
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124. doi:10.1371/journal.pmed.0020124
- Ioannidis, J. P. A. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science*, *7*, 645-654. doi:10.1177/1745691612464056
- Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, *24*, 602-611. doi:10.2307/2392366
- Kahneman, D. (2012, September 26). A proposal about how to deal with priming effects. *Nature*. Retrieved from https://www.nature.com/polopoly_fs/7.6716.1349271308!/supinfoFile/Kahneman%20Letter.pdf
- Kepes, S., Banks, G. C., & Oh, I.-S. (2014). Avoiding bias in publication bias research: The value of "null" findings. *Journal of Business and Psychology*, *29*, 183-203. doi:10.1007/s10869-012-9279-0
- Kepes, S., Bennett, A. A., & McDaniel, M. A. (2014). Evidence-based management and the trustworthiness of our cumulative scientific knowledge: Implications for teaching, research, and

- practice. *Academy of Management Learning & Education*, 13, 446-466. doi:10.5465/amle.2013.0193
- Kepes, S., Bushman, B. J., & Anderson, C. A. (2017). Violent video game effects remain a societal concern: Comment on Hilgard, Engelhardt, and Rouder (2017). *Psychological Bulletin*, 143, 775-782.
- Kepes, S., & McDaniel, M. A. (2013). How trustworthy is the scientific literature in industrial and organizational psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 252-268. doi:10.1111/iops.12045
- Kepes, S., & McDaniel, M. A. (2015). The validity of conscientiousness is overestimated in the prediction of job performance. *PLoS ONE*, 10, e0141468. doi:10.1371/journal.pone.0141468
- Kepes, S., McDaniel, M. A., Brannick, M. T., & Banks, G. C. (2013). Meta-analytic reviews in the organizational sciences: Two meta-analytic schools on the way to MARS (the Meta-analytic Reporting Standards). *Journal of Business and Psychology*, 28, 123-143. doi:10.1007/s10869-013-9300-2
- *Klinesmith, J., Kasser, T., & McAndrew, F. T. (2006). Guns, Testosterone, and aggression: An experimental test of a mediational hypothesis. *Psychological Science*, 17, 568-571. doi:10.1111/j.1467-9280.2006.01745.x
- *Korb, A. (2016). *Life experience as a moderator of the weapons priming effect* (Unpublished master's thesis). Texas A&M University Corpus Christi.
- Krahé, B. (2013). *The social psychology of aggression* (2nd ed.). New York, NY: Psychology Press.
- *Leyens, J., & Parke, R. D. (1975). Aggressive slides can induce a weapons effect. *European Journal of Social Psychology*, 5, 229-236. doi:10.1002/ejsp.2420050207
- *Lindsay, J. J., & Anderson, C. A. (2000). From antecedent conditions to violent actions: A general affective aggression model. *Personality and Social Psychology Bulletin*, 26, 533-547. doi:10.1177/0146167200267002
- Loersch, C., & Payne, B. K. (2011). The situated inference model: An integrative account of the effects of primes on perception, behavior, and motivation. *Perspectives on Psychological Science*, 6, 234-252. doi:10.1177/1745691611406921
- *Mahjoob, A., Leyens, J., & Yzerbyt, V. (1992). *The weapons effect among children in an armed-conflict environment*. Unpublished manuscript.
- *Mendoza, A. (1972). *The effects of exposure to toys conducive to violence* (Unpublished doctoral dissertation). University of Miami, FL.
- Molden, D. C. (2014). Understanding priming effects in social psychology: What is "social priming" and how does it occur? *Social Cognition*, 32(Suppl.), 1-11. doi:10.1521/soco.2014.32.suppl.1
- Moreno, S. G., Sutton, A., Ades, A. E., Stanley, T. D., Abrams, K. R., Peters, J. L., & Cooper, N. J. (2009). Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Medical Research Methodology*, 9, Article 2.
- Oakes, W. (1972). External validity and the use of real people as subjects. *American Psychologist*, 27, 959-962.
- O'Boyle, E. H., Rutherford, M. W., & Banks, G. C. (2014). Publication bias in entrepreneurship research: An examination of dominant relations to performance. *Journal of Business Venturing*, 29, 773-784. doi:10.1016/j.jbusvent.2013.10.001
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349, Article 6251. doi:10.1126/science.aac4716
- Orlitzky, M. (2012). How can significance tests be deinstitutionalized? *Organizational Research Methods*, 15, 199-228. doi:10.1177/1094428111428356
- Page, D., & O'Neal, E. (1977). 'Weapons effect' without demand characteristics. *Psychological Reports*, 41(1), 29-30. doi:10.2466/pr0.1977.41.1.29
- Pashler, H., Coburn, N., & Harris, C. R. (2012). Replicate effects on Social and food judgments. *PLoS ONE*, 7(8), e42510. doi:10.1371/journal.pone.0042510
- Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science. *Perspectives on Psychological Science*, 7, 528-530. doi:10.1177/1745691612465253
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61, 991-996. doi:10.1016/j.jclinepi.2007.11.010
- Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7, 331-363. doi:10.1037/1089-2680.7.4.331
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638-641. doi:10.1037/0033-2909.86.3.638
- Sapolsky, R. M. (1998). *The trouble with testosterone: And other essays on the biology of the human predicament* (p. 150). New York, NY: Scribner.
- Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Newbury Park, CA: SAGE.
- Schwarzer, G. (2007). Meta: An R package for meta-analysis. *The Newsletter of the R Project*, 7, 40-45.
- Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51, 515-530.
- Shanks, D. R., Newell, B. R., Lee, E. H., Balakrishnan, D., Ekelund, L., Cenac, Z., . . . Moore, C. (2013). Priming intelligent behavior: An elusive phenomenon. *PLoS ONE*, 8(4), e56515. doi:10.1371/journal.pone.0056515
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359-1366. doi:10.1177/0956797611417632
- *Simons, L. S., Fenn, M. R., Layton, J. F., & Turner, C. W. (1976). Aggressive behavior in a game at the amusement park. In J. Koch (Ed.), *Altruismus und aggression: Das fieldexperiment in der sozialpsychologie 1* (Altruism and aggression: The field experiment in social psychology 1) (pp. 141-148). Weinheim, Germany: Beltz Verlag.
- *Simons, L. S., & Turner, C. W. (1975). A further investigation of the weapons effect. *Personality and Social Psychology Bulletin*, 1, 188-190. doi:10.1177/014616727400100164
- *Simons, L. S., & Turner, C. W. (1976). Evaluation apprehension, hypothesis awareness, and the weapons effect. *Aggressive*

- Behavior*, 2, 77-87. doi:10.1002/1098-2337(1976)2:1<77::AID-AB2480020108>3.0.CO;2-A
- Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, 5, 60-78. doi:10.1002/jrsm.1095
- Stanley, T. D., Jarrell, S. B., & Doucouliagos, H. (2010). Could it be better to discard 90% of the data? A statistical paradox. *The American Statistician*, 64, 70-77. doi:10.1198/tast.2009.08205
- Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., . . . Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal*, 343, 302-307. doi:10.1136/bmj.d4002
- Strack, F., & Schwarz, N. (2016). Editorial overview: Social priming—Information accessibility and its consequences. *Current Opinion in Psychology*, 12, iV-Vii. doi:10.1016/j.copsyc.2016.11.001
- *Subra, B., Muller, D., Bègue, L., Bushman, B. J., & Delmas, F. (2010). Automatic effects of alcohol and weapon cues on aggressive thoughts and behaviors. *Personality and Social Psychology Bulletin*, 36, 1052-1057. doi:10.1177/0146167210374725
- *Sulikowski, D., & Burke, D. (2014). Threat is in the sex of the beholder: Men find weapons faster than do women. *Evolutionary Psychology*, 12, 888-906. doi:10.1177/147470491401200505
- Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, 22, 2113-2126. doi:10.1002/sim.1461
- Toch, H., & Lizotte, A. J. (1992). Advocating gun control. In P. Suedfeld & P. E. Tetlock (Eds.), *Psychology and social policy* (pp. 223-249). New York, NY: Hemisphere Publishing.
- Trikalinos, T. A., & Ioannidis, J. P. A. (2005). Assessing the evolution of effect sizes over time. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta analysis: Prevention, assessment and adjustments* (pp. 241-259). West Sussex, UK: John Wiley.
- *Turner, C. W., & Goldsmith, D. (1976). Effects of toy guns and airplanes on children's antisocial free play behavior. *Journal of Experimental Child Psychology*, 21, 305-315. doi:10.1016/0022-0965(76)90044-8
- *Turner, C. W., Layton, J. F., & Simons, L. S. (1975). Naturalistic studies of aggressive behavior: Aggressive stimuli, victim visibility, and horn honking. *Journal of Personality and Social Psychology*, 31, 1098-1107. doi:10.1037/h0076960
- Turner, C. W., & Leyens, J. (1992). The weapons effect revisited: The effects of firearms on aggressive behavior. In P. Suedfeld & P. E. Tetlock (Eds.), *Psychology and social policy* (pp. 201-222). New York, NY: Hemisphere Publishing.
- *Turner, C. W., & Simons, L. S. (1974). Effects of subject sophistication and evaluation apprehension on aggressive responses to weapons. *Journal of Personality and Social Psychology*, 30, 341-348. doi:10.1037/h0036894
- Turner, C. W., Simons, L. S., Berkowitz, L., & Frodi, A. (1977). The stimulating and inhibiting effects of weapons on aggressive behavior. *Aggressive Behavior*, 3, 355-378. doi:10.1002/1098-2337(1977)3:4<355::AID-AB2480030405>3.0.CO;2-G
- Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods*, 10, 428-443. doi:10.1037/1082-989X.10.4.428
- Viechtbauer, W. (2015). *Meta-analysis package for R: Package "metafor."* (R Package Version 1.9-5). Retrieved from <http://www.metafor-project.org/doku.php>
- Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1, 112-125. doi:10.1002/jrsm.11
- Yong, E. (2012). Replication studies: Bad copy. *Nature*, 485, 298-300. doi:10.1038/485298a
- *Zhang, Q., Tian, J., Zhang, D., & Rodkin, P. (2016). Exposure to weapon pictures and subsequent aggression during adolescence. *Personality and Individual Differences*, 90, 113-118. doi:10.1016/j.paid.2015.09.017