Psychiatry, Psychology and Law, 2018
VOL. 25, NO. 4, 503–521
https://doi.org/10.1080/13218719.2018.1440468

Impact of Gruesome Photographic Evidence on Legal Decisions: A Meta-Analysis

Rebecca Hofstein Grady[a], Lauren Reiser[a], Robert J. Garcia[a], Christian Koeu[a] and Nicholas Scurich[b]*

[a]*Department of Psychology and Social Behavior, University of California, Irvine, USA;* [b]*Department of Criminology, Law, and Society, University of California, Irvine, USA*

Gruesome crime scene and autopsy photographs are admissible evidence under the Federal Rules of Evidence (FRE) if their probative value substantially outweighs their prejudicial impact. Despite important methodological differences and mixed results from past studies, recommendations from the psychological literature have been made about the prejudicial impact of gruesome photographs perhaps prematurely. This meta-analysis investigates whether there is sufficient empirical evidence that presenting gruesome photographs in a trial affects legal decisions. The analysis of 23 studies and 4868 participants shows a small but statistically significant effect of gruesome photographs increasing guilty/liable verdicts or punishments, Hedge's $g = 0.143$, 95% CI: [0.055, 0.232]. However, this effect is significantly, $Q(1) = 8.086$, $p = .004$, and substantially moderated by an important methodological distinction: the effect is much larger when studies compare gruesome photographs with no photographs ($g = 0.450$) than when they are compared with neutral photographs ($g = 0.077$). These results suggest that gruesome photographs do increase affirmative verdicts, both through a small effect of gruesome content as well as a larger additive of having visual material. These findings help shed light on the mixed empirical results and suggest that important additional research is needed.

Keywords: evidence; gruesome photographs; juror verdicts; meta-analysis.

Gruesome crime scene and autopsy photographs are regularly introduced in both criminal and civil trials (Bandes & Salerno, 2014; Feigenson, 2010; Risinger, 1998). Admitting gruesome photographs (meaning in this context gory, bloody, or violent photographs of injury or death to a human body) as evidence in a legal case is a complex judicial decision. The Federal Rules of Evidence (FRE) hold that evidence is presumptively admissible if it is relevant, meaning it has 'any tendency to make a fact more or less probable than it would be without the evidence; and the fact is of consequence in determining the action' FRE, 2009, rule 401). However, FRE 403 states that even relevant evidence may be excluded if 'its probative value is substantially outweighed by a danger of… unfair prejudice' (FRE, 2009). Unfair prejudice, in this case, refers to either 'the injection of undue emotionalism into the proceeding arousing hostility, anger, or sympathy' or 'the likelihood that the jury will misuse the evidence in some way or give it undue weight' (Muller & Kirkpatrick, 2003, p. 176). Thus, in deciding whether or not to admit gruesome photographic evidence, judges must determine how probative the evidence is, speculate

*Correspondence: Nicholas Scurich, 4213 Social & Behavioral Sciences Gateway, University of California, Irvine, Irvine, 92697-7085, CA, USA. Phone: 949-824-4046. Email: nscurich@uci.edu

about the possible unfair prejudicial effects of such evidence, and then determine whether or not the latter 'substantially outweighs' the former. Of course, this balancing test is inherently a legal determination, not a scientific one. But the possible effect that gruesome photographic evidence exerts on jurors is a decidedly empirical question, and one that has been addressed in numerous empirical investigations spanning nearly four decades.

These investigations tend to use a similar paradigm, though the materials may vary widely in other ways. In a typical study, research participants view a hypothetical trial scenario and must then come to a decision about the guilt of the defendant. Some participants will see a gruesome photograph as part of the trial materials, while other participants will not. Usually, research studies deliberately choose a non-probative photograph as a comparison so that any difference in the verdict can be attributed to reactions to the gruesome nature of the photograph and not any additional information provided by it. For example, if the nature of the injury in a civil case is not in question, and the only dispute is whether or not the defendant was negligent in causing it, showing a graphic image of the injury will not generally help to determine the blameworthiness of the accused.

The theory behind the research is that the gruesome photographs, either instead of or in addition to any probative value, are emotionally evocative and increase emotionality biases in legal decision-making (Feigenson & Park, 2006). Jurors may use their emotional reaction to the image as an informational cue when deciding blameworthiness or guilt (Albarracin & Kumkale, 2003; Clore & Parrott, 1991), or it may change how they process other information due to the effects of discrete negative emotions like anger and sadness (Semmler & Brewer, 2002), or it may lead them to interpret later evidence congruent with their emotional state (Forgas & Bower, 1987; Forgas & Moylan, 1987). These mechanisms are what constitute the potentially prejudicial nature of gruesome photographs, whereby the emotions elicited may bias the jurors so much that it outweighs any possible information the images may provide in helping them decide on a case.

Based on the studies that have been conducted in this area, some have called for increased attention in the courtroom to the prejudicial impact of gruesome photographs based on the emotional disturbance they may cause in jurors that can bias their decisions (e.g., Epstein & Mannes, 2016). However, the results of studies have been mixed, and key methodological differences are often overlooked. This meta-analysis synthesizes the available empirical evidence to test the hypothesis that gruesome photographs increase juror convictions compared to a case with the same evidence but without the gruesome photographs.

Gruesome Evidence in the Courts

One example of a murder trial involving questions around gruesome photographic evidence is State v. Bocharski (2001). Six highly gruesome photographs were admitted at Bocharski's trial, despite objections from the defense that some of the photographs had no probative value since the types of injuries were not disputed, and the photographs did not show that Bocharski's missing knife caused the injuries. The Arizona State Supreme Court determined that the trial judge erred in allowing two of the photographs to be admitted, finding that they were too prejudicial and should not have been presented as evidence. The Court noted that two jurors had visible, visceral reactions to the photographs, and one juror nearly hyperventilated in the courtroom. However, the Court went on to hold that the other evidence in the case was sufficient to support a finding of guilt, and thus the Court ultimately did not reverse Bocharski's conviction.

Although Bocharski's case is exceptional in some respects – e.g., it appears to be rare that a juror would hyperventilate from viewing evidence – it is not unlike the many cases which, on appeal, claim that the prejudicial impact of the photographic evidence

substantially outweighed their probative value. Moreover, like Bocharski's case, most of these claims are unsuccessful (Modin, 2006), in part because trial judges are given 'broad discretion' in applying FRE 403 (Muller & Kirkpatrick, 2003). More recently this issue has emerged in terrorism cases, where questions arise about the emotional impact of gruesome images of beheadings or extremist actions likely to cause strong reactions in jurors (Goodman-Delahunty, 2017).

Precedent for this balancing act comes in part from Old Chief v. United States (1997), a case which did not involve gruesome photographs but did involve the presentation of evidence (details from a prior felony) alleged to be prejudicial. Old Chief was convicted at trial, but the United States Supreme Court reversed the conviction, holding that the prejudicial impact of the details surrounding his prior conviction substantially outweighed any probative value. The Court went on to craft a lengthy opinion in which it reaffirmed the bedrock principle that probative value must be balanced against potential prejudicial impact, leading to this case often being discussed by legal scholars in connection with gruesome photographic evidence (e.g., Bandes & Salerno, 2014; Feigenson, 2010).

Current procedures require judges to speculate about the possible prejudicial effects of gruesome photographic evidence on jurors and legal verdicts. While many judges will make a judgment on this balancing test based on their own knowledge and experience (Tanford, 1989), those that do look to the scientific literature may not find a clear answer on the prejudicial nature of gruesome photographs.

Mixed Empirical Research

Research examining the effect of gruesome photographic evidence on verdicts has borne mixed results (Devine, 2012). This may be due, at least in part, to the many differences in methodologies that researchers have used to test the question, as well as the different legal contexts in which researchers have tested the effect.

As previously discussed, a typical paradigm involves giving mock jurors a shortened version of a criminal trial, either by reading vignettes, listening to an audio recording of a mock trial, or watching a video of a simulated trial. These studies usually use students or online samples and a brief set of materials (depicting a real trial). As part of the evidence in the mock trial, some participants are shown one or more gruesome photographs, such as that of a mangled corpse, a bloody injury, or a graphic autopsy. The rest of the participants do not see these photographs; instead, they either see no photographs at all or some neutral photographs (e.g., a cleaned body or bloodless crime scene photograph). The researchers then compare the dependent measure, generally a verdict or sentencing decision, between the groups to see if those who see the gruesome photograph give harsher decisions (convict more often, longer sentences) than those who do not. Some studies also look at potential mediators of harsher judgments, such as emotional arousal or lowered conviction threshold. One key distinction in methodology is whether the gruesome-photograph condition is compared to a neutral-photograph condition or a no-photograph condition, the results of which have important theoretical distinctions.

Comparing Gruesome Photographs to Neutral Photographs or the Absence of Photographs

An important manipulation in the study of how gruesome photographs affect juror verdicts has to do with whether the control group sees non-gruesome photographs in their materials or no photographs at all; this manipulation affects whether or not the study can rule out the possibility that the mere presence of any visual material, compared to the prejudicial nature of a gruesome photograph in particular, is what influences verdicts. For example, in one of the earliest experiments that addressed the impact of gruesome

photographs, participants read a case summary of a negligence trial and saw either no photographs or four gruesome photographs of the victim's injured hand (Oliver & Griffit, 1976). The participants were not asked for a verdict, but they were asked to rate the likelihood of guilt and to decide what proportion of the sought damages should be awarded. The group who saw the four gruesome photographs showed increased punitiveness in the proportion of damages awarded, even without a statistically significant increase in the likely guilt of the plaintiff. While this showed an important effect that researchers continue to cite, the fact that the control group did not see any photograph leaves open a potential confound: the researchers cannot parse out the effect of seeing any photograph from that of seeing a gruesome photograph.

There is literature from other areas that support this possibility, such as in studies which show that images of a brain can increase the rating of a neuroscience article's credibility even more than bar graphs and topographical maps (McCabe & Castel, 2008). Some have dubbed this a 'truthiness' effect, defined as a tendency for judgments of veracity to increase with the presentation of relevant but non-probative information, such as photographs that do not help establish the truth of a statement (Newman, Garry, Bernstein, Kantner, & Lindsay, 2012). Newman et al. (2012) gave participants names of unfamiliar celebrities, along with a statement of either 'This famous person is alive' or 'This famous person is dead', and participants had to respond true or false to each claim. When the statement also included a photograph of the celebrity, it was rated as true more often, no matter whether the statement suggested the celebrity was alive or dead. Similar results were shown for trivia items paired with non-probative photographs. The results suggest that the photograph makes people more likely to accept a claim or statement, even when the photograph does not add any helpful information in deciding whether or not the claim is true. The truthiness effect may encompass the presentation of gruesome photographs at trial; even if photographs seem to lack probative value, they may ultimately play a role in swaying mock jurors toward believing the plaintiff's or defendant's story. For example, in the experiment above, they may have been more persuaded by the plaintiff's claim for the amount of money needed based on the photographs, even though the accident and medical treatment needed was the same. When jurors are exposed to unfamiliar evidence, as frequently happens during trials, photographic evidence (gruesome or otherwise) may increase jurors' perception of the veracity of the evidence.

In one experiment that had both neutral-photograph and no-photograph conditions, Bright and Goodman-Delahunty (2006) manipulated the nature of visual and verbal evidence. The mock jurors either saw five gruesome photogtaphs (e.g., the deceased victim with deep wounds to the neck), five neutral photographs (e.g., a partially damaged door, house exterior), or no photographs as part of a murder trial. Additionally, the participants read a 20-page trial summary with either gruesome or neutral testimony about the degree of injury caused to the victim. The results indicate that the verbal description of the injuries did not have a significant effect on the verdicts, but the type of visual evidence did influence the verdicts. The mock jurors exposed to either the gruesome or neutral photographs were significantly more likely to convict the defendant compared to the participants who saw no photographs. There was not a statistically different conviction rate between the mock jurors who saw the gruesome photographs and those who saw the neutral photographs; thus, the inclusion of a photograph – regardless of its nature – was found to be sufficient to increase guilty verdicts.

Both of these studies suggest the possibility that the gruesome nature of a photograph may be less important than the presence of any photograph at all, but this is not generally mentioned in discussions of the gruesome photograph effect (e.g., Bandes & Salerno,

2014; Devine, 2012; Simon, 2011), and the results are varied across studies within each type of comparison. Although a few studies have included both control conditions to assess the separate effects, most studies have chosen to use either neutral photographs or no photographs. In both types of comparison group – neutral photographs and no photographs – there have been some studies that show a significant effect of gruesome photographs and some that do not show a significant effect.

Overview of the Present Study

As these studies illustrate, the effect of gruesome photographs on juror decisions is inconsistent, potentially due to important theoretical differences in the choice of control group. This meta-analysis tests the hypothesis that gruesome photographs increase juror verdicts and assesses whether or not the choice of control group significantly moderates any effect. If it is solely the gruesome nature of gruesome photographs that causes harsher jury decisions then the effect size increases in juror verdicts will not differ based on control groups. If the gruesome content is not especially prejudicial relative to the effect based on the mere presence of any photograph then there will only be a significant effect overall for studies that compare a gruesome-photograph condition to a no-photograph condition, but not in studies that compare to a neutral-photograph condition. It is also possible that both contribute independent effects, in which case both comparison types will show a significant effect but the no-photograph condition will be significantly higher.

Method

Search Strategy

This meta-analysis includes both published and unpublished English-language studies of the effect of gruesome photographic evidence on both criminal and civil trials. The inclusion criteria stipulate that studies must have at least one experimental condition that includes a gruesome photograph of a person, along with a control condition, which can either include neutral photographs or feature no photographic evidence. Gruesome photographs are operationally defined as those explicitly designated by the authors as victim or autopsy photographs that are '(highly) gruesome', 'gory', or 'highly violent'. For example, gruesome photographs include victims with bloody stab wounds, mangled limbs, or severe burns. Neutral photographs are defined as any other photographs or visual materials presented to participants that are not highly gruesome, graphic, or bloody, which range from inanimate objects (e.g., photographs of a home or a car) to those with victim photographs featuring no blood or gore. The conditions have to feature the same evidence outside of the photograph; i.e., the different photographs cannot be confounded with different injury types or different evidence levels.

The studies also need to include a legal judgment by mock jurors. The primary dependent variable (DV) in most studies is the verdict, either as a guilty/not guilty verdict in criminal cases or liable/not liable verdict in civil cases. In three studies identified that met the inclusion criteria, punishment measures (e.g., preferred length of the incarceration for the defendant in criminal cases; monetary damage awards in civil cases) are the primary DV. Verdict data are not available in these studies, and in at least one study are not the primary decision because the trial scenario implies guilt.

Several steps were taken to locate studies dealing with gruesome photographs in the courtroom. The search began with the use of PsycINFO, ProQuest Criminal Justice, Social Science Research Network, and Google Scholar search engines to search for papers with one of the words 'gruesome', 'gory', 'violent', and 'graphic', one of the words 'photo*', 'picture', and 'evidence', and one of the words 'jur*', 'verdict', and 'trial' (note

that asterisks indicate wildcards).[1] When papers were found that match the topic, including non-empirical papers (e.g., Bandes & Salerno, 2014; Bornstein & Nemeth, 1999; Feigenson, 2010), the reference section of each paper was examined in order to find other related papers. Additionally, the authors of relevant empirical and review papers were contacted by email to ask for details of any unpublished work they were aware of, recommendations for other researchers to contact, and any necessary statistical clarification. Posts were also created on the Society for Personality and Social Psychology Open Forum to request details of any pertinent unpublished work. Lastly, all American Psychology-Law Society programs posted online for previous years (2008 to 2015, except for 2012, which was unavailable) were searched for any posters or talks that are relevant to the present analysis, and the authors were emailed to gather the relevant information. The searches were completed in September 2016.

Sample

Out of the initial 18 published empirical papers identified as potentially meeting criteria based on title and abstract, 10 were retained. The other 8 were excluded because they did not include gruesome photographs ($n = 4$), the content of the photograph was not manipulated between groups (i.e., all participants saw the same gruesome photographs so there was no control group; $n = 2$), the data needed for analysis was not available in the paper or from the author ($n = 1$), or the change in the photograph was confounded with a change in case facts/probative evidence ($n = 1$).

Searching for unpublished work yielded 13 additional studies that met inclusion criteria: 4 dissertations (one of which contains 2 relevant studies), 4 conference presentations, and 4 under-review or unpublished experiments. This search also located 6 unpublished studies that were not included because they did not vary the gruesome nature of the photographs ($n = 1$), the gruesome photographs were not part of trial evidence ($n = 2$), or there was not enough information available in the paper and from the author to complete the analyses ($n = 3$).

The final number of studies included in the meta-analysis is thus 23, consisting of 10 published papers (9 from peer-reviewed journals and 1 from the non-peer-reviewed journal *The Jury Expert*), 5 from dissertations or theses, and 8 unpublished experiments, with a total of 4868 participants.

Coding and Analysis Plan

Some experiments featured a control group that saw neutral photographs, while other studies had a control group that did not see any photographs. This is the primary moderator for the present research, given the theoretically distinct nature of these two comparisons, as previously discussed. If the comparison group (i.e., the group that did not see gruesome photographs) saw any photographs or visuals in their trial materials, they were coded as neutral-photograph conditions. No-photograph conditions were coded when the participants did not see any visual materials at all in their trial scenario, only text. Details of each study, including the type of case (civil or criminal), type of crime, the sample population, the study modality (in-person, online, or via mail), the location in which the study was conducted, and the current publishing status of the experiment, were independently coded by two authors, with the first author resolving any disputes. The reliability analyses show high agreement between the two raters: the intraclass correlation coefficient (ICC) for the numeric effect sizes is .916, and the percentage agreement for the categorical moderator codings is 94.9%. Table 1 shows all the included studies and methodological details.

The final sample size is 23 studies incorporating 4868 participants. Of these studies, 3 used both comparison types and thus have

Table 1. Characteristics of the studies included in the meta-analysis.

Study	Comparison group(s)	Trial type	Crime[a]	Sample	Modality	Location	Status
Ahola et al. (2010)	No photo	Criminal	Homicide	Undergraduates	In-person	Stockholm	Published
Bright (2008)	Neutral	Criminal	Homicide	Undergraduates	Online	New South Wales	Dissertation
Bright and Goodman-Delahunty (2006)	Both	Criminal	Homicide	Undergraduates	In-person	New South Wales	Published
Bright and Goodman-Delahunty (2011)	Neutral	Civil	Negligence	Undergraduates	In-person	New South Wales	Published
Burd (2016)	Neutral	Criminal	Homicide	MTurk	Online	Mturk	Dissertation
Cush and Goodman-Delahunty (2006)	Neutral	Criminal	Homicide	Schoolteachers	Mailed surveys	Sydney	Published
Douglas et al. (1997)	Neutral	Criminal	Homicide	Undergraduates	In-person	British Columbia	Published
Edelman (2009)	Neutral	Civil	Negligence	Community members	Online	California, New York, Illinois, Texas, Florida	Published (non-peer-reviewed journal)
Edwards and Mottarella (2012)	Both	Criminal	Homicide	Undergraduates	Online	University of Central Florida	Published
Finkelstein and Bastounis (2010)	No photo	Criminal	Assault causing death without intent	Trainees and undergraduates	In-person	Paris	Published
Griffin et al. (2015)	Neutral	Criminal	Homicide	Undergraduates	In-person	University of Alabama	Unpublished
Koeu et al. (2015)	Neutral	Criminal	Homicide and robbery	Undergraduates	Online	University of California, Irvine	Presentation
Matsuo and Itoh (2015)	Both	Criminal	Homicide	Undergraduates	In-person	Japan	Published
Modin (2006)	Neutral	Criminal	Homicide	Community members	In-person	Denver, CO	Dissertation
Nemeth (2002), Study A	Neutral	Criminal	Homicide	Undergraduates	In-person	University of Nebraska–Lincoln	Dissertation
Nemeth (2002), Study B	Neutral	Criminal	Homicide	Undergraduates	In-person	University of Nebraska–Lincoln	Dissertation
Oliver and Griffitt (1976)	No photo	Civil	Negligence	Undergraduates	In-person	Kansas State University	Published
Peter-Hagene et al. (2015)	Neutral	Criminal	Homicide	MTurk	Online	MTurk	Unpublished
Salerno and Peter-Hagene (2013a)	Neutral	Criminal	Homicide	Undergraduates	In-person	University of Illinois at Urbana–Champagne	Presentation
Salerno (2015a)	Neutral	Criminal	Homicide	MTurk	Online	MTurk	Unpublished
Salerno (2015b), Study 1	No photo	Criminal	Homicide	MTurk	Online	MTurk	Unpublished
Salerno (2015b), Study 2	Neutral	Criminal	Homicide	MTurk	Online	MTurk	Unpublished
Salerno et al. (2015)	Neutral	Criminal	Homicide	Undergraduates	In-Person	Arizona State University	Unpublished

[a]Published papers may have included additional crimes in other conditions; the ones listed here are those that were used in this analysis.

both comparisons in the analysis.[2] This yielded a total of $n = 26$ comparisons for analysis: $n = 19$ for comparisons of the gruesome-photograph and neutral-photograph conditions, and $n = 7$ for comparisons of the gruesome-photograph and no-photograph conditions. Given that meta-analyses often need as few as 3 to 5 studies to find a robust effect that closely approximates the effect size that a much larger meta-analysis would find (Herbison, Hay-Smith, & Gillespie, 2011), this sample is sufficiently large to test the research question at hand.

The software package Comprehensive Meta-Analysis (CMA; Version 3.3.070; Biostat, 2014) was used to combine effect sizes from different data types and draw statistical conclusions. Because of its capacity for broad usability across outcome types in meta-analysis and its correction for studies with a small sample, Hedge's g is used as the primary effect size measure, although odds ratios (ORs) are reported as well. Random effects models are used in order to make inferences for a broader population than the population with which the studies were conducted. All significance tests are evaluated at $\alpha = .05$.

Results

The overall analysis of 26 comparisons shows a small but significant effect of gruesome photographs in increasing juror convictions, Hedge's $g = 0.143$, 95% CI: [0.055, 0.232], $OR = 1.298$. Although only 7 of the 26 comparisons on their own show a significant difference between conditions at $p < .05$, on average – across all studies – the participants who saw gruesome photographs were more likely to convict the defendant (or find for the plaintiff in a civil case, or provide harsher punishments) than those who saw only neutral photographs or who did not see any photographs. The individual Hedge's g effect sizes range from -0.209 to 1.074 (see Table 2). A one-study-removed analysis shows that there is no single study driving the effect, as the effect is still significant with any one study removed.

Comparison Group Moderator

The primary moderator is the choice of comparison group, either neutral photographs or no photographs. The effect of gruesome photographs is significantly larger, $Q(1) = 8.086$, $p = .004$, when gruesome photographs are compared to a no-photograph condition, $g = 0.450$, 95% CI: [0.202, 0.697], $OR = 2.284$, than to a neutral-photograph condition, $g = 0.077$, 95% CI: [0.003, 0.150], $OR = 1.150$. Figure 1 presents a forest plot of the studies grouped by comparison type. Both comparison types are still statistically significant on their own, although a one-study-removed analysis shows that the comparison with the neutral-photograph condition is much less robust; there are five separate studies where if any one of them is removed, this renders the overall comparison with the neutral-photograph condition no longer significant at $p = .05$. No single study however changes the overall significance of the comparison with the no-photograph condition.

Other Moderator Analyses

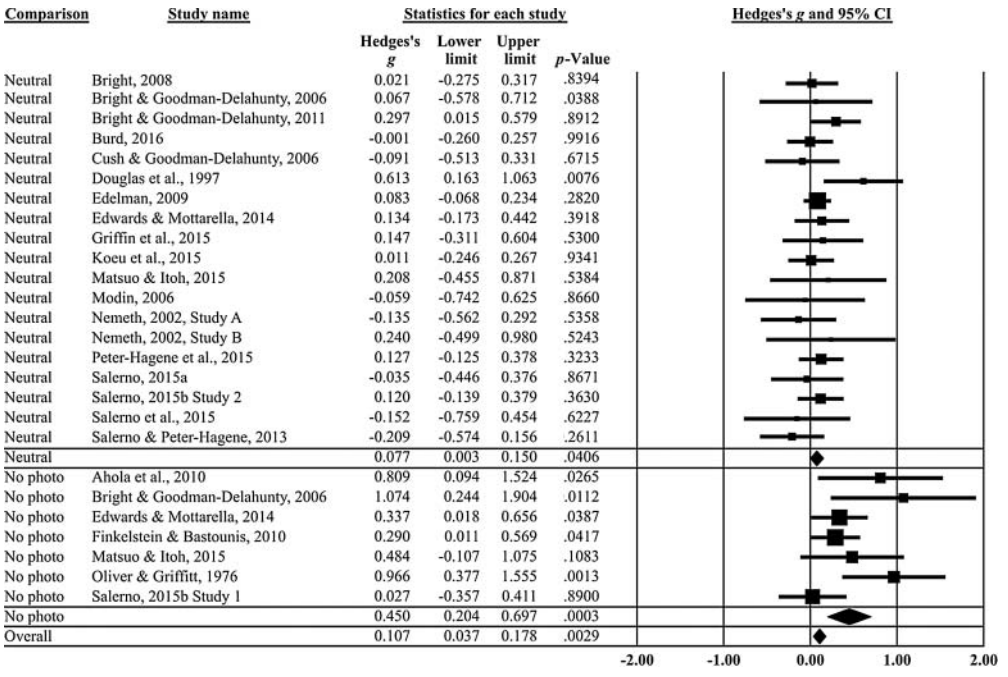
Five other pre-planned moderator analyses were tested based on status or methodological differences in studies: case type (civil or criminal), study modality (in-person or remote), DV type (verdict or punishment), sample population (all-student sample or other adult samples), and publication status in a peer-reviewed journal (yes or no). The moderators are all selected for theoretical or methodological reasons. The case type is included because there are important procedural differences between civil and criminal cases such as the standard of evidence, meaning that the type of case may differentially affect mock jurors' decisions (Devine, Clayton, Dunford, Seying, & Pryce, 2001). The modality of the study has an impact on the participant experience, for example that

Table 2. Effect sizes for included experiments.

			Hedge's g[a]		Odds ratio[b]	
Study	DV type	Total n[c]	Neutral	No photo	Neutral	No photo
Ahola et al. (2010)	Years sentenced	31		0.809*		4.515*
Bright (2008)	Verdict	240	0.021		1.038	
Bright and Goodman-Delahunty (2006)	Verdict	102	0.067	1.074*	1.131	7.235*
Bright and Goodman-Delahunty (2011)	Verdict	240	0.297*		1.716*	
Burd (2016)	Verdict	292	−0.001		0.997	
Cush and Goodman-Delahunty (2006)	Verdict	108	−0.091		0.846	
Douglas et al. (1997)	Verdict	120	0.613*		3.064*	
Edelman (2009)	Verdict	940	0.083		1.162	
Edwards and Mottarella (2012)	Verdict	532	0.134	0.337*	1.277	1.845*
Finkelstein and Bastounis (2010)	Years sentenced	198		0.290*		1.695*
Griffin et al. (2015)	Verdict	99	0.147		1.307	
Koeu et al. (2015)	Verdict	284	0.011		1.020	
Matsuo and Itoh (2015)	Verdict	112	0.208	0.484	1.466	2.436
Modin (2006)	Verdict	52	−0.059		0.897	
Nemeth (2002), Study A	Verdict	130	−0.135		0.782	
Nemeth (2002), Study B	Verdict	41	0.240		1.560	
Oliver and Griffitt (1976)	Damages awarded	48		0.966*		5.937*
Peter-Hagene et al. (2015)	Verdict	419	0.127		1.259	
Salerno (2015a)	Verdict	147	−0.035		0.938	
Salerno (2015b), Study 1	Verdict	193		0.027		1.051
Salerno (2015b), Study 2	Verdict	329	0.119		1.244	
Salerno and Peter-Hagene (2013a)	Verdict	143	−0.209		0.756	
Salerno et al. (2015)	Verdict	68	−0.152		0.756	

Note: [a]Hedge's g represents the bias-corrected standardized mean difference between the gruesome-photograph condition and the condition specified in the listed dependent variable (DV), where a positive value indicates more guilty verdicts in the gruesome-photograph condition than the comparison condition; [b]Odds ratios (ORs) represent the odds of a guilty verdict (or a conversion to such from a sentencing measure) in the gruesome-photograph condition relative to the comparison condition, where a values greater than one indicates an increased likelihood of punitive judgment in the gruesome-photograph condition; [c]The sample size listed here may be smaller than the full sample reported in published papers if the study included other experimental conditions that were not part of the present meta-analysis; *Effect significant at $p < .05$.

in-person studies may be more immersive or realistic than studies done online. The DV type is important to test because the guilt and sentencing phases of a trial are distinct, and it is possible that gruesome photographs might have more of an effect on one than the other. The sample population is included in order to test whether there are differences between college undergraduates and a more diverse online sample (Sears, 1986) and to evaluate whether reactivity to emotional materials differs by participant age (since online and other adult samples will be older than student samples; Charles, Mather, & Carstensen, 2003; Mather et al., 2004). Finally, the publication status is included to test for biases in publication (discussed further below). Because some categories include only a few studies in each cell (the lowest are case type with three civil trials and DV type with three non-verdict measures; the rest have at least eight per cell), some of these moderators may be underpowered or confounded and thus do not provide definitive conclusions (Hoyle, 1999; Schmidt & Hunter, 2014). Additional research studies would be needed to provide



Comparison	Study name	Hedges's g	Lower limit	Upper limit	p-Value
Neutral	Bright, 2008	0.021	-0.275	0.317	.8394
Neutral	Bright & Goodman-Delahunty, 2006	0.067	-0.578	0.712	.0388
Neutral	Bright & Goodman-Delahunty, 2011	0.297	0.015	0.579	.8912
Neutral	Burd, 2016	-0.001	-0.260	0.257	.9916
Neutral	Cush & Goodman-Delahunty, 2006	-0.091	-0.513	0.331	.6715
Neutral	Douglas et al., 1997	0.613	0.163	1.063	.0076
Neutral	Edelman, 2009	0.083	-0.068	0.234	.2820
Neutral	Edwards & Mottarella, 2014	0.134	-0.173	0.442	.3918
Neutral	Griffin et al., 2015	0.147	-0.311	0.604	.5300
Neutral	Koeu et al., 2015	0.011	-0.246	0.267	.9341
Neutral	Matsuo & Itoh, 2015	0.208	-0.455	0.871	.5384
Neutral	Modin, 2006	-0.059	-0.742	0.625	.8660
Neutral	Nemeth, 2002, Study A	-0.135	-0.562	0.292	.5358
Neutral	Nemeth, 2002, Study B	0.240	-0.499	0.980	.5243
Neutral	Peter-Hagene et al., 2015	0.127	-0.125	0.378	.3233
Neutral	Salerno, 2015a	-0.035	-0.446	0.376	.8671
Neutral	Salerno, 2015b Study 2	0.120	-0.139	0.379	.3630
Neutral	Salerno et al., 2015	-0.152	-0.759	0.454	.6227
Neutral	Salerno & Peter-Hagene, 2013	-0.209	-0.574	0.156	.2611
Neutral		0.077	0.003	0.150	.0406
No photo	Ahola et al., 2010	0.809	0.094	1.524	.0265
No photo	Bright & Goodman-Delahunty, 2006	1.074	0.244	1.904	.0112
No photo	Edwards & Mottarella, 2014	0.337	0.018	0.656	.0387
No photo	Finkelstein & Bastounis, 2010	0.290	0.011	0.569	.0417
No photo	Matsuo & Itoh, 2015	0.484	-0.107	1.075	.1083
No photo	Oliver & Griffitt, 1976	0.966	0.377	1.555	.0013
No photo	Salerno, 2015b Study 1	0.027	-0.357	0.411	.8900
No photo		0.450	0.204	0.697	.0003
Overall		0.107	0.037	0.178	.0029

Figure 1. Forest plot of Hedge's *g* individual effect sizes from included studies, grouped by comparison type. Note: Positive values indicate gruesome photographs yielded increased convictions or punishments compared to the comparison group.

stronger support for the lack of any subgroup or moderator effect.

Of these moderators, two were found to have a significant effect at $\alpha = .05$: DV type, $Q(1) = 4.408$, $p = .036$, and peer-review status, $Q(1) = 11.528$, $p = .001$ (see Table 3 for a list of all moderated effect sizes). The verdict measures show a significantly smaller effect than the punishment measures, although both subgroups still show a statistically significant effect of gruesome photographs. There are no significant differences (all ps $> .05$) based on sample population, study modality, and case type. The studies published in peer-review journals found significantly higher effects than those that were not, and the effects sizes for the latter are overall not significantly different from zero ($p = .336$).

Publication Bias

As with any meta-analysis, publication bias is always a concern, and this is the main source of potential bias that was identified. A significant difference was found between published and unpublished studies, meaning that publication bias may be affecting the work that is available for analysis. To counter this, multiple avenues were used to find unpublished work, resulting in over half of the sample comprising of unpublished studies, which shows the effectiveness of this search strategy and bolsters confidence in the results with the inclusion of so much unpublished work. Additionally, out of the 13 comparisons from published studies, only 7 are statistically significant at $p < .05$ on their own. Thus, a non-significant finding does not seem to be a barrier to publication. This is likely because these papers offer other important contributions, such as the effect the photographs have on other outcomes, the interaction of the photographs with other manipulations, and implications for the emotional processes behind the potential effect. It is also possible that the unpublished studies are of lower

Table 3. Other methodological moderator analyses for the effect of gruesome photographs.

Moderator	k[a]	Hedge's g	Odds ratio
Published in peer-reviewed journal[***]			
Yes	12	0.351	1.899
No	14	0.040	1.075
Case type			
Civil	3	0.347	1.883
Criminal	23	0.118	1.239
Modality			
In-person	15	0.266	1.629
Online/By mail[b]	11	0.077	1.151
Dependent variable[*]			
Verdict	23	0.102	1.204
Punishment	3	0.617	3.128
Sample population			
Students	18	0.223	1.503
Other adults	8	0.061	1.118

Note: [*]Effect significant at $p < .05$; [***]Effect significant at $p < .001$; [a]k indicates the number of studies in each subgroup; [b]Modality comparison is still non-significant if the single mailed study is not included. Moderators marked in italic are statistically significant using the Q-test for heterogeneity.

quality – however, the majority of these are authored by the same scholars as the ones who behind the published studies, and the majority were being written for publication at the time they were obtained for this meta-analysis.

Rosenthal's fail-safe N[3] was also calculated for the analysis (Rosenthal, 1979), which gives the number of studies with a completely null effect that would need to be added to make the main effect no longer significant at $p < .05$. This yielded a fail-safe N of 148, meaning there would have to be 148 additional included studies with a null effect to render the overall effect of gruesome photographs non-significant. The small size of the field and the extent of the search make it unlikely that that many null unpublished studies might be sitting in a file-drawer. However, this value differs substantially if testing the comparison subgroups separately; for the studies which used a no-photograph condition, the fail-safe N is 64, while for the studies with the neutral-photograph condition it is only 4. The small number for the latter subgroup indicates a stronger possibility that there could be other unpublished studies – which our analysis shows have smaller effect sizes – that would nullify the result. This is especially worrisome given that 2 studies were located during the search phase that reported a non-significant effect but could not be included because the data were not available in the correct format.

Discussion

A meta-analytic approach was utilized to investigate existing studies involving the effect of gruesome photographs on mock juror verdicts. The results indicate that the presence of gruesome photographs increases the number affirmative jury decisions. This is also significantly and substantially moderated by the comparison group used in the research study, supporting the hypothesis that the gruesome nature of the photographs and the mere presence of any photograph both serve to increase the rate of guilty judgments.

Although the overall result reaches statistical significance, the effect when limited to studies comparing gruesome to neutral photographs is small ($g = 0.077$) and not particularly robust. In legal decisions alone (not considering the impact of gruesome photographs on other measures, such as anger at the defendant), the gruesome content of photographs appears to have only a small effect on increasing juror convictions when compared to having only neutral photographs, based on this literature. While the overall analysis is not driven by any one study, removing any one of five different studies makes the effect of the studies which use the neutral-photograph condition no longer significant, meaning that this result warrants additional research before firm conclusions can be drawn

On the other hand, the moderator analysis shows that the effect of gruesome photographs when compared to the no-photograph condition is much higher (Hedge's $g = 0.450$) and more robust (i.e., removing any one study would not change the level of the statistical significance). This suggests that some, though not all, of the effect of gruesome photographs is likely driven by the mere presence of any photograph.

However, it is also possible that this effect is also partially driven by a methodological confound; a significantly higher effect was found for the three studies that used punishment measures compared to the verdict measures, and all three punishment measures happen to be in the studies which use the no-photograph condition. Had the search been limited to only verdicts instead of legal judgments (which was decided before seeing any data), the primary moderator would no longer have been statistically significant using the random effects model that was chosen a priori, $Q(1) = 2.677$, $p = .102$, although the effect is still larger for the no-photograph condition ($g = 0.308$) than the neutral-photograph condition ($g = 0.077$, unchanged). This is likely due to the lower statistical power of the test being reduced because of the smaller sample; the moderator would have been significant with a fixed effects analysis, which has more power with smaller samples.

One reason why the effect is significantly moderated by DV type (although the fact that there are only three punishment measures makes this result preliminary) is that the punishment measures have more possible variability in them, since they are scale measures instead of dichotomous. If gruesome photographs raise the harshness of judgments to a small degree, it may be enough to cause a change in a continuous measure but not enough to change a verdict from not guilty to guilty. There is also the possibility that ambiguity in a case may be necessary, or at least helpful, in finding an effect of gruesome photographs; if the other evidence in the case is very clearly indicative of either guilt or innocence then the small biasing effect of gruesome photographs may not have an impact on dichotomous verdicts. In the more ambiguous cases, or in sentencing phases with more variability, however, the gruesome photographs may be more prejudicial in swaying undecided jurors to a greater degree towards harsh judgments.

Future researchers addressing the effects of gruesome evidence on juror decisions should compare with a neutral-photograph condition if they want to solely focus on any differences arising from gruesome content, as opposed to just the presence of any photograph in the materials – or ideally include both comparison groups in order to study the mechanism behind the differential effect (such as in Bright & Goodman-Delahunty, 2006). They may also want to examine the moderating effects of evidence strength and judgment measure on the impact of gruesome photographs. A few plausible explanations for why the gruesome photographs may have led to verdict bias and why the effect is weakened relative to the neutral-photograph condition are discussed below.

Truthiness and Placebic Reasoning

The present findings fit well with truthiness research on how non-probative photographs (regardless of content) can make a claim seem more credible, as discussed previously. This is particularly supported in one of the included studies, which found that adding photographic evidence from the defense's side (showing the plaintiff's positive recovery long after the injury) counteracted the increased verdicts from having the plaintiff's side alone present gruesome photographs of the plaintiff's injury (Edelman, 2009; in most studies the photographs supported the prosecution's case). The truthiness effect of photographs has been demonstrated to persist 48 hours after viewing a photograph (Fenn, Newman, Pezdek, & Garry, 2013), making it plausible that photographs may have

influence for a lengthy portion of a given trial. In both the neutral-photograph and gruesome-photograph conditions, the charge against the defendant may seem more credible (as compared with the same trial without any photographs). This explains why the effect of gruesome photographs is so much stronger when compared to the absence of photographs: when comparing gruesome photographs to neutral ones, the truthiness effect impacts both conditions and thus does not add to the difference (leaving only the gruesome content itself to impart the effect), whereas in the no-photograph condition, the conviction rate is elevated by *both* the gruesome content *and* the truthiness effect of having photographic evidence.

Similarly, the effect of gruesome photographs on verdicts may reflect placebic reasoning processes in jurors. Placebic reasoning is the tendency to accept propositions or suggestions if justification is provided, regardless of the actual content and validity of the justification (Langer, Blank, & Chanowitz, 1978). For instance, because the prosecution provides a gruesome photograph of the victim, jurors may consider the presentation of such photographs as justification of the defendant's guilt, despite the potential lack of probative value. Although there may be probative value in actual photographs presented at trials, many of these studies were specifically designed by the researchers to have little or no probative value (for both the gruesome and neutral photographs) in order to isolate and assess their prejudicial nature. Truthiness and placebic reasoning help explain why the effect size for the comparison of the gruesome-photograph and no-photograph conditions would be much higher than that for the comparison of the gruesome-photograph and neutral-photograph conditions; this literature has shown that mere exposure to any photographic evidence may act as a tool to promote mock jurors' beliefs that the defendant is guilty or that a given piece of evidence is credible. In the neutral-photograph condition this effect raises the rate of guilty verdicts in both conditions and thus is cancelled out when looking at the difference, while in the no-photograph condition this effect is likely to significantly raise the rate of guilty verdicts in the gruesome-photograph condition just due to the presence of a photo (before any effect of gruesome content additionally).

Emotional Responses

While the truthiness literature sheds light on the larger effect of the comparison with the no-photograph condition, there is a still significant effect – although smaller – of gruesome content on juror decisions. Previous work on the effect of gruesome photographs has often focused on the negative emotions they can elicit in jurors, which may in turn influence verdicts. In legal decision-making, emotions are predicted to be problematic and prejudicial by yielding increased punitive judgments (Salerno & Bottoms, 2009). There are a variety of ways in which emotional states, as may be elicited by gruesome photographic evidence, can influence legal decisions, as previously mentioned.

In the studies that assess emotion in participants, a variety of scales have been used, including the Juror Negative Affect Scale, the Positive and Negative Affect Schedule, and the Profile of Mood States. Although they differ in various ways, they all ask participants to self-report the degree to which they are feeling a variety of emotional reactions, which may include anxiety, disgust, revulsion, anger, fear, hostility, and many more. It is particularly important to investigate these effects within the legal system because the often-touted Aristotelian ideal of law being free from passion is in conflict with psychological findings of emotion as an inherent part of decision-making – and various discrete emotions and affective states are known to influence judgments in specific ways. When people are in a good mood they tend to be better at inductive reasoning, whereas people in a negative mood often perform better at deductive reasoning (Forgas, 1995). Within

the family of negative emotions, discrete emotions can differentially influence decision-making (Bless & Schwarz, 1999). Anger and sadness are both negatively valenced, but anger is associated with less systematic processing, whereas sadness is associated with more substantive information processing (Tiedens & Linton, 2001). For example, feeling angry after exposure to gruesome photographs could cause a juror to react heuristically and select a verdict without thoroughly evaluating the evidence, whereas a sad juror might be more sensitive to carefully examining all of the evidence (Semmler & Brewer, 2002). Memory congruency effects have been applied to discrete emotions as well; for example, when participants are manipulated to feel angry or sad, they predict that angry or sad events, respectively, are more likely to occur (DeSteno, Petty, Wegener, & Rucker, 2000), an effect which could influence jurors when they need to make judgments about both the defendant and the strength of the evidence.

Emotional responses to gruesome photographs are best captured in the studies that compare the gruesome-photograph and neutral-photograph conditions, since, unlike in the studies which use the no-photograph condition, the effect cannot be explained by the mere presence of any photograph. Given the small and less robust effect of such comparisons in the present study, it is possible that emotional arousal may not be large enough to greatly increase dichotomous verdicts, which may be less sensitive to variations than other continuous judgments, like the blameworthiness of the defendant. It may also be that not enough jurors in these studies were emotionally aroused by the stimulus photographs. If the effect of gruesome photographs on verdicts is mediated by an increase in negative emotions such as anger (Bright & Goodman-Delahunty, 2011) or disgust (Salerno, 2015b), it may be that the studies which do not show a significant effect are the ones where the photograph failed to arouse negative emotion in participants.

Practical Implications

The small effect size observed and the significant moderation based on comparison groups makes it difficult at this time to give a strong recommendation about the use of gruesome photographs in the courtroom. While there have been calls from the psychological literature to give substantial weight to the prejudicial nature of gruesome photographs in court admissibility decisions, this meta-analysis indicates that additional research is needed in order to address the boundary conditions of when gruesome photographs are likely to cause substantial prejudicial impact (e.g., perhaps when the case is most ambiguous, or only in those who are highly emotionally affected by the photographs). Though a statistically significant effect of the presentation of gruesome photographs on increasing punitive juror decisions was found, it is possible that most of this effect is driven by the presentation of any photograph or the type of DV used, as opposed to specifically the gruesome nature alone. The effect of gruesome photographs compared to neutral photographs is not particularly robust, so while it cannot be concluded that gruesome images have no effect on trials, this effect may be limited to some types of case or people, and may not have a noticeable generalized effect. Further research in this small field is needed before blanket statements about the prejudicial nature of gruesome photographs should be made.

In any particular case, the judge will need to weigh the probative nature of the gruesome photograph – which will be unique in every instance – against the prejudicial nature, which science can more broadly speak to. Many judges rely on intuition and personal biases to decide how prejudicial a gruesome photograph is likely to be (Tanford, 1989); this meta-analysis, along with future studies that help to disentangle potential confounds in the existing literature, can hopefully provide a more scientific tool with which to inform those judgments by synthesizing the

available scientific research. While the judge will always be the gatekeeper in determining how the probative/prejudicial balance weighs out in a particular case, this work helps inform the legal system in a more comprehensive manner, as opposed to relying on a single study that may not capture the nuance of the literature as a whole.

Limitations and Future Directions

Several limitations must be kept in mind when considering the findings reported in this meta-analysis. One is the sample size of available studies, which some may consider small for a meta-analysis. Although there is clear interest in the field on the effect that emotion has on legal decision-making, there is a relatively small number of studies that look specifically at gruesome photographs and their effects on verdicts. The extensive search for unpublished studies conducted as part of this meta-analysis helped to bolster the number of studies, but the lack of a large and heterogeneous set of studies makes the moderator analyses more preliminary than a much larger sample would allow. Despite the relatively small number of studies used, the primary comparison moderator is still statistically significant, likely due to the magnitude of the difference between comparison types.

The methodology used in the constituent studies must also be considered. In particular, these studies were all conducted under the tightly-controlled conditions of the laboratory. Differences between the laboratory and the real world abound, and these differences could have important implications for the observed effects. For example, consider the potential impact of consequentiality (Bornstein & McCabe, 2005); real jurors may be more affected by gruesome photographs knowing that the case is real and the potential perpetrator is sitting in front of them. Although many studies used real crime scene or autopsy photographs in their materials, the participants know they are taking part in a study and not truly judging a case. However, in other jury decision-making research with the same limitations, research has found broadly similar results between student-based studies, laboratory studies and more realistic or diverse studies (Bornstein, 1999, Bornstein et al., 2016). Another limitation often noted is the lack of deliberation in many jury decision-making studies. Two of the studies included in this meta-analysis feature mock juror deliberation (Finkelstein & Bastounis, 2010; Modin, 2006), but do not find any difference in the effect of gruesome photographs between judgments made before and after deliberation – although future studies aiming for applied recommendations may wish to conduct more studies with deliberation. Finally, the heterogeneity of stimuli between studies is important to note. The studies varied in the level of gruesomeness, the number of photographs shown, whether the photograph was of an injury or a dead body, and other factors that cannot be evaluated in every case. Although this could be considered a weakness, if some cases were not well designed to be able to demonstrate the effect it can also be considered a strength. While these variations make internal validity less strong because they are not tightly controlled, this does mirror reality in the sense that real-world court cases will differ greatly in the details of the gruesome photographic content that is admitted as evidence. Although what one author may consider gruesome may seem less emotional to another, each effect size was computed within studies such that it was always a significantly more gruesome photo compared to a less gruesome or non-gruesome photograph.

Although these methodological limitations are important to acknowledge, they arguably do not devastate the studies, individually or collectively, and offer suggestions for future work in this area. The main finding of this article is the importance of choosing a control group that is relevant to the research question. Studies that are solely interested in the effect of gruesome content may wish to present neutral photographs to other

participants, unless the goal is specifically to study the impact of adding a single gruesome photograph to a case that was previously devoid of photographic evidence. Rating the level of gruesomeness and emotional impact of photographs is also important. If photographs purported to be gruesome by researchers do not elicit negative emotions (e.g., disgust, anger) in participants – who may be desensitized to such images if they are not worse than what is regularly shown on television – then they are less likely to have an effect. For example, Koeu, Grady, Joudi, and Loftus (2015) used photographs from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2008) that are normatively rated to evoke strong negative emotions. In the experiment, while the gruesome photograph did raise negative emotions compared to the control group, participants were still generally feeling more positive than negative, which may explain the lack of an effect on verdicts. Studies may wish to utilize manipulation checks and other measures to see if their particular sample found a photograph gruesome – as unique populations will have unique reactions – as well as investigate the effects of varying levels of gruesomeness. This will also help determine if different types of gruesome photographs or cases lead to different types of emotional reaction, in order to determine when disgust, anger, and sadness are driving any impact on verdicts.

It is important to note that, despite claims that gruesome photographs are commonly introduced at trial (e.g., Douglas, Lyon, & Ogloff, 1997; Feigenson, 2010), there is no research documenting the prevalence of gruesome photographic evidence in actual court cases. The admission of gruesome photographs likely differs from state to state, county to county, and even judge to judge. Additionally, gruesome photographs may be admitted for some types of crime but not others. It is important for future researchers to investigate how frequently, when, and for which types of case gruesome photographs are admitted as evidence in real-world cases. Empirically documenting the prevalence of gruesome photographs may provide insight into how concerned the legal community should be about the potentially biasing effects of gruesome photographic evidence.

Conclusion

Past research has yielded mixed results on whether or not gruesome photographs lead to increased juror convictions. The present meta-analysis found that gruesome photographs have the effect of leading jurors towards more guilty or liable verdicts, although the effect is larger and more robust when comparing gruesome photographs to the absence of photographs than when comparing gruesome photographs to neutral ones. Future research can help determine how much of this result is driven by the gruesome content of the photographs and how much is due to the presence of any photograph. Other promising avenues include what other factors might moderate and mitigate this effect, such as the level of ambiguity in the case, the subjective gruesomeness of the photographs, and the type of outcome (verdict, sentence) being affected. While the judgment of whether or not a gruesome photograph is more probative than prejudicial needs to be balanced in every particular case, this meta-analysis brings together the relevant scientific findings in order to inform the courts and the research community as a whole.

Acknowledgements

This article does not contain any studies with human participants or animals performed by any of the authors.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Notes

1. For example, in PsycINFO the search was the following: TI,AB(gruesome OR gory OR

violent OR graphic) AND TI,AB(evidence OR photo* OR picture) AND TI,AB(trial OR verdict OR jury).

2. In order to account for this non-independence, the sample size for the shared group was divided in half as per the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins & Green, 2011). This makes it possible to account for the shared variance while still including information from both comparisons separately for analyses.
3. While some other failsafe statistics use effects sizes rather than *p* values to compute this measure of potential publication bias, there was no standard to use in relation to how big of an effect of gruesome photographs would be meaningful in such a serious issue of guilty verdicts. While this measure is imperfect, it is a rough indicator of how likely it is that publication bias would significantly change the overall result of the meta-analysis.

References

Ahola, A. S., Hellström, Å., & Christianson, S. Å. (2010). Is justice really blind? Effects of crime descriptions, defendant gender and appearance, and legal practitioner gender on sentences and defendant evaluations in a mock trial. *Psychiatry, Psychology and Law, 17*, 304–324.

Albarracin, D., & Kumkale, G. T. (2003). Affect as information in persuasion: A model of affect identification and discounting. *Journal of Personality and Social Psychology, 84*, 453–469.

Bandes, S. A., & Salerno, J. M. (2014). Emotion, proof and prejudice: The cognitive science of gruesome photos and victim impact statements. *Arizona State Law Journal, 46*, 1004–1054.

Biostat. (2014). Comprehensive Meta-Analysis [Computer Software]. Retrieved from https://www.meta-analysis.com/.

Bless, H., & Schwarz, N. (1999). Sufficient and necessary conditions in dual-process models: The case of mood and information processing. In S. Chaiken & Y. Trope (Eds.), *Dual-Process Theories in Social Psychology,* (pp. 423–440). New York, NY: Guilford Press.

Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior, 23*, 75–91.

Bornstein, B. H., Golding, J. M., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C., & Luecht, K. (2016, October 20). Mock juror sampling issues in jury simulation research: A meta-analysis. *Law and Human Behavior*. Advance online publication. doi:0.1037/lhb0000223

Bornstein, B. H., & McCabe, G. G. (2005). Jurors of the absurd: The role of consequentiality injury simulation research. *Florida State Law Review, 32*, 443–467.

Bornstein, B. H., & Nemeth, R. J. (1999). Jurors' perception of violence: A framework for inquiry. *Aggression and Violent Behavior, 4*, 77–92.

Bright, D. A. (2008). *The influence of gruesome evidence on juror emotion and decision making* (Unpublished doctoral dissertation). University of New South Wales, New South Wales.

Bright, D. A., & Goodman-Delahunty, J. (2006). Gruesome evidence and emotion: Anger, blame, and jury decision-making. *Law and Human Behavior, 30*, 183–202.

Bright, D. A., & Goodman-Delahunty, J. (2011). Mock juror decision making in a civil negligence trial: The impact of gruesome evidence, injury severity, and information processing route. *Psychiatry, Psychology and Law, 18*, 439–459.

Burd, K. A. (2016). *Mitigating the biasing impact of emotions and morality: Cognitive training, juror emotionality, and juror decision-making* (Unpublished dissertation). Cornell University, Ithaca, New York.

Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General, 132*, 310.

Clore, G. L., & Parrott, W. G. (1991). Moods and their vicissitudes: Thoughts and feelings as information. In J. Forgas (Ed.), *Emotion and social judgments* (pp. 107–123). New York: Pergamon Press,

Cush, R. K., & Delahunty, J. G. (2006). The influence of limiting instructions on processing and judgments of emotionally evocative evidence. *Psychiatry, Psychology and Law, 13*, 110–123.

DeSteno, D., Petty, R. E., Wegener, D. T., & Rucker, D. D. (2000). Beyond valence in the perception of likelihood: The role of emotion specificity. *Journal of Personality and Social Psychology, 78*, 397–417.

Devine, D. J. (2012). *Jury decision making: The state of the science*. New York: New York University Press.

Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law, 7*, 622. doi:10.1037//1076-8971.7.3.622

Douglas, K. S., Lyon, D. R., & Ogloff, J. R. (1997). The impact of graphic photographic evidence on mock jurors' decisions in a murder trial: Probative or prejudicial? *Law and Human Behavior, 21*, 485–501.

Edelman, B. (2009). The impact of graphic injury photographs on liability verdicts and noneconomic damage awards. *The Jury Expert, 21*, 1–4

Edwards, E. R., & Mottarella, K. E. (2014). Preserving the right to a fair trial: An examination of prejudicial value of visual and auditory evidence. *North American Journal of Psychology, 16*, 397–414.

Epstein, J., & Mannes, S. (2016, March 17). "Gruesome" evidence, science, and Rule 403. *The Judicial Edge*. Retrieved from http://www.judges.org/gruesome-evidence-science-and-rule-403/

Federal Rules of Evidence. 28a U.S.C. §§ 401–403 (2012).

Feigenson, N. (2010). Visual evidence. *Psychonomic Bulletin & Review, 17*, 149–154.

Feigenson, N., & Park, J. (2006). Emotions and attributions of legal responsibility and blame: A research review. *Law and Human Behavior, 30*, 143–161.

Fenn, E., Newman, E. J., Pezdek, K., & Garry, M. (2013). The effect of nonprobative photographs on truthiness persists over time. *Acta Psychologica, 144a*, 207–211.

Finkelstein, R., & Bastounis, M. (2010). The effect of the deliberation process and jurors' prior legal knowledge on the sentence: The role of psychological expertise and crime scene photo. *Behavioral Sciences & the Law, 28*, 426–441.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117*, 39–66. doi:10.1037/0033-2909.117.1.39

Forgas, J. P., & Bower, G. H. (1987). Mood effects on person-perception judgments. *Journal of Personality and Social Psychology, 53*, 53–60.

Forgas, J. P., & Moylan, S. (1987). After the movies: Transient mood and social judgments. *Personality and Social Psychology Bulletin, 13*, 467–477. doi:10.1177/0146167287134004

Goodman-Delahunty, J. (2017). Assessing unfair prejudice from extremist images in terrorism trials. In D. Tait & J. Goodman-Delahunty (Eds.), *Juries, science and popular culture in the age of terror*. London, England: Palgrave Macmillan.

Griffin, D. A., Neal, T. M. S., Parrot, C. T., & Griffin, M. P. (2015). *Let emotions be the judge: Graphic evidence and need for affect in legal decision-making*. Manuscript in preparation.

Herbison, P., Hay-Smith, J., & Gillespie, W. J. (2011). Meta-analyses of small numbers of trials often agree with longer-term results. *Journal of Clinical Epidemiology, 64*, 145–153.

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions Version 5.1.0*. Retrieved from www.cochrane-handbook.org.

Hoyle, R. H. (1999). *Statistical strategies for small sample research*. Thousand Oaks, CA: Sage Publications.

Koeu, C., Grady, R. H., Joudi, D., & Loftus, E. F. (2015, May). *Are gruesome photographs more prejudicial than probative*? Poster presented at the annual meeting of the 22nd annual Undergraduate Research Symposium, Irvine, CA.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL.

Langer, E. J., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology, 36*, 635–642.

Mather, M., Canli, T., English, T., Whitfield, S., Wais, P., Ochsner, K., … Carstensen, L. L. (2004). Amygdala Responses to Emotionally Valenced Stimuli in Older and Younger Adults. *Psychological Science, 15*, 259–263.

Matsuo, K., & Itoh, Y. (2015). Effects of emotional testimony and gruesome photographs on mock jurors' decisions and negative emotions. *Psychiatry, Psychology and Law*. Advanced Online Publication. doi:10.1080/13218719.2015.1032954

McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107*. 343–352. doi:10.1016/j.cognition.2007.07.017

Modin, K. (2006). *The impact of photographic crime scene evidence on juror verdicts, jurors' feelings of confidence, beliefs about impartiality, and emotional reactions* (Unpublished doctoral dissertation). Alliant International University, San Diego.

Muller, C. B., & Kirkpatrick, L. C. (2003). *Evidence*. New York; Aspen.

Nemeth, R. J. (2002). *The impact of gruesome evidence on mock juror decision making: The role of evidence characteristics and emotional response* (Unpublished doctoral dissertation). Louisiana State University, Baton Rouge.

Newman, E. J., Garry, M., Bernstein, D. M., Kantner, J., & Lindsay, D. S. (2012). Nonprobative photographs (or words) inflate truthiness.

Psychonomic Bulletin and Review, 19, 969–974. doi:10.3758/s13423-012-0292-0

Old Chief v. United States, 519 U.S. 172. (1997).

Oliver, E., & Griffitt, W. (1976). Emotional arousal and 'objective' judgment. *Bulletin of the Psychonomic Society, 8*, 399–400.

Peter-Hagene, C. L., Salerno, J. M., Greenaway, K., & Kalokerinos, E. (2015). [Study on gruesome photographs, results shared by second author]. Unpublished raw data.

Risinger, D. M. (1998). "A syllogism is not a story": Overproof of undisputed facts in pursuit of what? United States v. Old Chief John Henry Wigmore, Johnny Lynn Old Chief, and "legitimate moral force" – Keeping the courtroom safe for heartstrings and gore. *Hastings Law Journal, 49*, 403–1511.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638–64. doi:10.1037/0033-2909.86.3.638

Salerno, J. M. (2015a). [Pilot test of gruesome photographs, results shared by first author]. Unpublished raw data.

Salerno, J. M. (2015b). *Seeing red: Disgust drives the differential effect of color versus black and white gruesome photographs on legal judgments*. Manuscript submitted for publication.

Salerno, J. M., Baker, D., & Malik, S. (2015). [Study on gruesome photographs, results shared by first author]. Unpublished raw data.

Salerno, J. M., & Bottoms, B. L. (2009). Emotional evidence and jurors' judgments: The promise of neuroscience for informing psychology and law. *Behavioral Sciences and the Law, 27*, 273–296.

Salerno, J. M., & Peter-Hagene, C. L. (2013, March). *Incidental – but not integral – emotional evidence moderates the effect of defendant race on guilt judgments*. Talk presented at the annual meeting of the American Psychology-Law Society, Portland, OR.

Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Thousand Oaks, CA: Sage Publications.

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.

Semmler, C., & Brewer, N. (2002). Effects of mood and emotion on juror processing and judgments. *Behavioral Sciences & the Law*, *20*, 423–436. doi:10.1002/bsl.502

Simon, D. (2011). The limited diagnosticity of criminal trials. *Vanderbilt Law Review, 64*, 143–223.

State v. Bocharski, 200 Ariz. 50. (2001).

Tanford, J. A. (1989). A political choice approach to limiting prejudicial evidence. *Indiana Law Journal, 64*, 831–872.

Tiedens, L., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology, 81*, 973–988.