The Dud Effect: Adding Highly Dissimilar Fillers Increases Confidence in Lineup Identifications

Steve D. Charman · Gary L. Wells ·
Stephen W. Joy

Abstract Recent research in decision-making has demonstrated the "dud-alternative effect"—the tendency to become more confident that a chosen response option is correct if it is surrounded by implausible response options (Windschitl & Chambers, J Exp Psychol 30:198–215, 2004). This finding may be applicable to a lineup task: The presence of duds (i.e., highly dissimilar fillers) may increase a witness's confidence that an identified (non-dud) lineup member is the criminal. Four studies ($N = 665$) demonstrate that the mere presence of highly dissimilar fillers inflates witnesses' confidence in a mistaken identification (Studies 1–4), provides evidence that this confidence inflation is due to the duds inflating the perceived similarity of the other lineup members to the criminal (Studies 2, 3), and delineates some conditions under which the effect holds (Studies 3, 4). The addition of highly dissimilar lineup members, far from being inert, as is often implicitly assumed, can bias witnesses' confidence reports.

Keywords Eyewitness · Confidence · Lineup bias · Lineup construction · Fillers

There is a common quip among eyewitness scientists when they assess a lineup in which some of the fillers are not plausible lineup members because they fail to fit the general description of the perpetrator, namely that those lineup members might as well have been dogs, flagpoles, refrigerators. In other words, they count for nothing. So, for instance, if the perpetrator was described by the eyewitness as a tall young white male with short dark hair and four of the five fillers were old, or had blond hair, or were short, then the lineup functionally had four fewer fillers than the nominal size of five fillers. In other words, it is the same as a two-person lineup (the suspect plus one filler) because the other four fillers were "duds" in the sense that they were not plausible choices and therefore would be ignored by the witness. In this paper, however, we argue that this conventional wisdom might not be true—that a six-person lineup comprised of four duds might very well be worse than a two-person lineup.

We agree that the addition of duds to a lineup is likely to be ignored in terms of their ability to draw choices. However, we do not necessarily agree that they are ignored by the witness and, hence, render the lineup the functional equivalent to a two-person lineup. The key to our reasoning rests with the issue of confidence. The confidence with which eyewitnesses make lineup identifications has been a major focus among eyewitness researchers. At first, this focus was largely aimed at determining the extent to which witnesses' confidence was correlated with their accuracy and whether jurors could properly use this confidence information to determine the accuracy of witnesses (Cutler & Penrod, 1989; Cutler, Penrod, & Dexter, 1990; Cutler, Penrod, & Stuve, 1988; Lindsay, Wells, & O'Connor, 1989; Sporer, Penrod, Read, & Cutler, 1995; Wells, Ferguson, & Lindsay, 1981; Wells, Lindsay, et al., 1979; Wells & Murray, 1984; Whitley & Greenberg, 1986). After determining that confidence was, under many circumstances, only modestly related to accuracy, and that jurors

S. D. Charman (✉) · S. W. Joy
Department of Psychology, DM 256, Florida International
University, Miami, FL 33199, USA
e-mail: charmans@fiu.edu

G. L. Wells
Department of Psychology, Iowa State University,
Ames, IA, USA

tend to overestimate just how strongly they are related, interest naturally shifted towards determining why this relationship was not as strong as most people expect. To that end, researchers recently started focusing on factors that inflate the identification confidence of inaccurate witnesses, such as repeated post-event questioning (Shaw & McClure, 1996), hearing a co-witness's identification (Charman, Carlucci, Vallano, & Hyman Gregory, 2010; Skagerberg, 2007), and receiving post-identification feedback (Douglass & Steblay, 2006; Wells & Bradfield, 1998). The current studies investigate an additional factor that may inflate inaccurate witnesses' post-identification confidence: The presence of lineup fillers that are highly dissimilar to the perpetrator.

Our conceptualization of the problem and our predictions borrow heavily from the judgment and decision-making literature, which has a long tradition of examining choice under conditions of uncertainty. Particularly relevant to the eyewitness identification certainty problem is theoretical and empirical work on the process by which people assess their certainty following a selection of one response option among an array of alternatives. Support theory, a popular account of how people make these certainty assessments, assumes that these judgments are largely normative;[1] the judged probability that one's decision was correct is assumed to depend purely on support for the chosen option (known as the focal hypothesis) relative to support for the alternative options (known collectively as the residual hypothesis; Tversky & Koehler, 1994). Adding alternative options should therefore increase support for the residual hypothesis as long as they have non-zero plausibility, thus decreasing the judged probability that the focal hypothesis is correct.

This assumption, however, has recently come under fire. In a series of studies, Windschitl and Chambers (2004) showed that adding very weak alternative options (known as *duds*) can actually *increase* the perceived likelihood that the focal hypothesis is correct, contrary to support theory's predictions. For example, participants' subjective, gut-level feelings about their chances of winning a lottery were greater when they held 39 lottery tickets and three others held 31, 2, and 3 lottery tickets than when they held 39 lottery tickets and one other held 31 lottery tickets, despite the fact that they were objectively less likely to win. However, this confidence-inflating effect (known as the *dud-alternative effect*) was limited to the addition of very weak alternatives; when the others held 31, 4, and 8 lottery tickets, the effect disappeared.

The favored interpretation of the dud-alternative effect is that it is the result of a contrast effect. Specifically, people base their subjective likelihood judgments on the result of a series of pair-wise comparisons between the focal option (i.e., the chosen option) and each of the individual alternatives. The presence of weak alternatives increases the number of highly favorable pair-wise comparisons, inflating the perceived support for the focal outcome, and thus inflating subjective likelihood judgments.

The necessary conditions for the dud-alternative effect to appear, then, are the following: (1) A likelihood judgment is being made; (2) a focal outcome exists; (3) alternatives to the focal outcome exist; (4) some of those alternatives are highly implausible response options. These are precisely the conditions that can exist among eyewitnesses making lineup identifications. Eyewitnesses must assess their confidence (i.e., make a likelihood judgment) following an identification (i.e., a focal outcome) from amongst an array of fillers (i.e., alternatives), some of whom may be highly implausible (i.e., they do not resemble the perpetrator). Consequently, their certainty judgments may be susceptible to the dud-alternative effect. Specifically, assuming that eyewitnesses assess their confidence somewhat similarly to how Windschitl and Chambers' (2004) participants made their likelihood judgments, they will make a series of pair-wise comparisons between the identified lineup member and each of the non-identified lineup members. The existence of duds (i.e., highly dissimilar lineup members) will increase the number of favorable pair-wise comparisons the witnesses make, thus inflating their confidence. The pertinent question therefore is whether the addition of highly dissimilar lineup members to a preexisting lineup, despite never being identified themselves, can inflate the confidence with which witnesses identify a different lineup member.

This question bears some superficial resemblance to a previously addressed issue in the eyewitness field, namely, how the composition of a lineup affects the likelihood of a witness identifying a particular lineup member and the confidence with which that identification is made. For example, constructing a lineup composed of fillers that do not resemble the suspect, instead of fillers that do resemble the suspect, leads to a higher rate of suspect identification and higher confidence in that identification, even if the suspect is not the perpetrator (Lindsay & Wells, 1980; Wells, Rydell, & Seelau, 1993). Despite their superficial similarity, however, these past studies have examined a fundamentally different issue than the one examined in the current paper. Specifically, these past studies have addressed whether the *replacement* of good fillers with poor fillers increases false identifications and confidence, whereas the current studies examine whether the mere

---

[1] Consistent with the judgment and decision making literature, we use the term normative to refer to processes or judgments that are considered to fit a rational model or align with formal proofs, such as probability theory.

addition of poor fillers can inflate confidence. The prediction that the mere addition of duds will inflate confidence in the identification of a non-dud is much less obvious because it contradicts the common sense, normative assumption that it is the *removal* rather than the *addition* of response alternatives that should raise confidence in the correctness of one's decision (c.f., support theory).

An appreciation for this distinction is important for methodological reasons. Without it, it might be thought that testing whether duds inflate confidence would require a comparison between witnesses who view a lineup composed of six people—none of whom are duds—and witnesses who view a lineup composed of six people that contains a certain number of duds. But this would be erroneous for at least two reasons. First, because the dud effect refers to the *addition* of duds, not the *replacement* of non-duds with duds, a proper test of the dud-alternative effect requires comparing witnesses' confidence following an identification decision from a lineup containing non-duds to witnesses' confidence following an identification decision from a lineup containing the same non-duds *plus* some number of duds.

Second, the dud effect is interesting partly because it predicts that duds will inflate witnesses' confidence despite their inability to actually draw identifications. If, however, we compared two six-person lineups—one with duds and one without—we would create differences in choosing across conditions. Witnesses in the no dud condition would distribute their choices among all six lineup members whereas witnesses in the dud condition would identify only one of the non-duds. Consequently, the witnesses' reported confidence would be in reference to different lineup members across dud conditions, and we would lose our ability to make meaningful comparisons across conditions. However, by only manipulating the addition of duds—lineup members that almost no one would ever identify as the criminal—we can examine the effect of duds on confidence without also changing the distribution of witnesses' choices.

The dud effect has important implications. Both researchers and police officers creating a lineup may reason that because lineups of larger nominal size are typically perceived as providing better protection to an innocent suspect, adding a few additional lineup fillers should improve the lineup, or at worst—if the additional fillers completely fail to resemble the perpetrator—should fail to have any impact on the quality of the lineup. In other words, it would seem intuitive that the duds would just be ignored. This intuition has at least two important consequences. First, adding additional fillers in an attempt to increase the quality of a lineup may, under some circumstances, backfire and, ironically, decrease the quality of

that lineup by resulting in an inflation of a witness's confidence in a mistaken identification.[2]

Second, the failure to account for dud-induced confidence inflation may lead researchers to underestimate the bias produced by poor lineups. This would occur because current conceptualizations of lineup fairness treat highly dissimilar lineup members as irrelevant. For example, one of the standard measures among lineup bias eyewitness researchers is the concept of functional size, which is determined by having a large number of mock witnesses—people who never witnessed the actual perpetrator—attempt an identification from a lineup based solely on the description given to the police by the witness (Wells, Leippe, et al., 1979). The functional size is calculated by dividing the total number of mock witness by the number of mock witnesses who identify the suspect (e.g., if 50% of mock witnesses identify the suspect, the lineup has a functional size of 2). Because duds by definition do not resemble the perpetrator (and thus should not resemble the description given to police by the witness), mock witnesses should tend to avoid identifying them, and consequently, the functional size of a lineup without duds should be equal to the functional size of the same lineup with duds. In other words, the addition of duds to a lineup would be considered an inert operation. This same result occurs when measuring the fairness of the lineup using effective size (Malpass, 1981) rather than functional size; lineup members who draw no choices in a mock witness test are "functionally" or "effectively" considered to be nonexistent, just as if they were not in the lineup at all. Showing that the presence of duds may actually bias the lineup (if not by an increase in choosing, then by an increase in confidence in a mistaken identification) would demonstrate a shortcoming with using functional size and effective size as measures of lineup fairness.

It is important to note that neither functional size nor effective size were designed to account for anything other than how the composition of a lineup can bias choice distributions. Nothing in the above analysis indicates that these metrics themselves are wrong in this respect, and we are certainly not suggesting that the existence of a dud effect would undermine their use. The sole point being made is that lineup composition may bias not only witnesses' *choices*, but also their *confidence* in those choices, which, because our current measures of lineup bias do not model confidence, is a type of bias that may occur unacknowledged. The observation of a dud effect would simply indicate that these metrics, although still necessary, are

---

[2] Of course, the inclusion of duds could also theoretically increase witness confidence in a correct identification as well; however, the general point to be made is that dud-induced confidence inflation is a *distortion* of a witness's 'true' confidence, since it derives from a source considered not relevant to the judgment.

insufficient by themselves to measure all of the ways in which the lineup can be biased.

Study 1

Experiment 1 provides a simple test of the dud-alternative effect as it applies to lineups. Participants viewed a mock crime and were shown either a two-person target-absent lineup (composed of relatively good fillers), or a six-person target-absent lineup (composed of the same two relatively good fillers in addition to four extremely poor fillers). We expected that although no one would identify a dud, their mere existence in the lineup would increase the confidence with which participants identified a non-dud.

Method

Participants

One hundred ten undergraduate students from a large Midwestern university participated in this study in exchange for extra credit.

Materials

Participants viewed a 45-s mock crime of a man planting a bomb on the roof of a building, a video that has been successfully used in other eyewitness experiments. Participants viewed either a two-person "non-dud" lineup or a six-person "dud" lineup. The dud lineup was composed of the same two non-duds plus four duds. Lineups were pretested by showing participants ($N = 34$) a looped clip from the mock crime that clearly showed the criminal's face beside each individual lineup member in turn. Participants rated the similarity of the criminal to each lineup member on a scale from 1 (very dissimilar) to 7 (very similar). Results from this pre-testing indicated that similarity between the perpetrator and each of the non-duds was moderate (Ms = 5.3 and 4.3, 95% CIs [4.9, 5.7] and [3.8, 4.7], respectively), and that the similarity between the perpetrator and each of the duds was extremely low (Ms = 2.1, 1.2, 1.6, and 1.9, 95% CIs [1.7, 2.5], [1.1, 1.4], [1.3, 2.0], [1.4, 2.3], respectively). The average non-dud similarity was significantly greater than the average dud similarity, $t(33) = 17.01$, $p < .001$, $d = 5.92$. Appendix A displays the dud lineup along with the actual perpetrator.

Procedure

The experiment was administered over a computer. After viewing the mock crime, participants were informed that they were witnesses and would be attempting a lineup identification. Participants were randomly assigned to view either a dud lineup or a non-dud lineup. Their respective lineup was displayed on screen and participants indicated their identification via computer. They were not given a "not there" option, so participants were forced to identify a lineup member to continue. Participants were then asked to indicate their confidence in their choice on a scale from 0 to 10.

Results and Discussion

95% Confidence intervals are reported in square brackets for all means. Two participants who viewed a dud lineup identified a dud. Because they did not allow a test of whether the presence of duds would increase confidence in a non-dud, they were excluded from analyses. All results were thus based on the remaining 108 participants. Consistent with predictions of the dud-alternative effect, the mean confidence among participants who viewed a dud lineup ($M = 7.9$ [7.2, 8.5]) was significantly higher than the mean confidence of participants who viewed a non-dud lineup ($M = 6.6$ [5.8, 7.3]), $t(106) = 2.79$, $p = .006$, $d = .54$. The presence of duds did not significantly change the distribution of identifications of the two non-duds, χ^2 ($df = 1$, $n = 108$) = 1.80, $p = .18$, $\varphi = .13$.

Study 2

Study 1 demonstrated the existence of the dud-alternative effect within a lineup context. Why does the dud-alternative effect occur? In their original conceptualization of the effect, Windschitl and Chambers (2004) provided evidence that it was due to a series of pair-wise comparisons. Specifically, when forming likelihood judgments, people compare the focal point to each of the alternatives, and thus the presence of duds increases the number of favorable comparisons made, increasing the perceived support for the focal point. Certainly, a similar process may be occurring within the lineup context as well: Confidence in one's identification may be ascertained by comparing the focal point (i.e., the identified lineup member) individually to each of the alternatives (i.e., the other lineup members). The existence of highly dissimilar lineup fillers increases the number of favorable comparisons, thus increasing overall perceived support in the identified lineup member. We call this the *pair-wise comparison* hypothesis.

But a lineup task is different from the types of general knowledge tasks involved in the original Windschitl and Chambers (2004) study. For instance, unlike general knowledge tasks, it involves both perceptual and memorial

components (e.g., witnesses presumably assess the perceived similarity between the various lineup members and their memory for the criminal). These differences introduce new mechanisms through which duds may artificially inflate witnesses' confidence. There are at least three alternatives that should be considered in addition to the pair-wise comparison hypothesis.

First, the effect of lineup duds might operate at a perceptual level. For example, the mere presence of highly dissimilar lineup members may, by way of perceptual contrast, make the non-duds appear even more similar to the perpetrator. If confidence in one's identification is determined in part by the degree to which a lineup member matches one's memory for the perpetrator, this increased perception of similarity should then translate into increased confidence in one's identification. We call this possibility the *perceptual contrast* hypothesis.

Second, the effect of lineup duds may operate via a self-perception route (Bem, 1967). Specifically, if witnesses find that they can easily reject the duds as being the perpetrator, they may use the ease of these eliminations to infer that their memory for the event is strong (e.g., see Shaw, 1996, for evidence that witnesses use ease of judgments to infer their confidence). To the extent that witnesses' confidence judgments reflect their meta-cognitive beliefs, this inflated perception of one's memory should translate into higher confidence. We call this the *memory inference* hypothesis.

Third, the effect of duds may operate via a cognitive dissonance route (Festinger & Carlsmith, 1959). If post-identification confidence inflation is viewed as a way to reduce the cognitive dissonance that might be created if one were to entertain the possibility of having mistakenly identified an innocent person (see Charman et al., 2010), then the addition of more lineup members might amplify these concerns. After all, the more options there are to choose from, the more likely that one's specific identification was wrong. Consequently, the confidence inflation resulting from duds may simply be a response to the greater pressure to reduce dissonance when presented with a larger lineup (see Harmon-Jones & Harmon-Jones 2002; Liberman & Förster, 2006). We call this the *dissonance-reduction hypothesis*.

These hypotheses can be easily tested. Table 1 lists these hypotheses and some of the different predictions they make. For example, both the pair-wise comparison hypothesis and the dissonance-reduction hypothesis predict that dud-induced inflated confidence should be specific for the identified lineup member, in the former case because only the identified lineup member is used as the focal point for multiple comparisons, and in the latter case because only the confidence inflation of the identified lineup member can reduce a witness's dissonance. Conversely, the perceptual contrast hypothesis predicts that dud-induced confidence inflation should occur for all non-duds, because the dud-induced confidence inflation is a result of a general context effect on all non-duds. (It is unclear whether the memory inference hypothesis predicts a specific or general effect of duds on confidence; inflated beliefs about one's memory may inflate witnesses' beliefs that any non-dud is the perpetrator, or may only inflate witnesses' beliefs that their specific identification is correct.)

Also, two of the hypotheses—perceptual contrast and cognitive dissonance—predict that dud-induced confidence inflation should be accompanied by an inflation in the perceived similarity of the non-duds. According to the former, the effect of duds on confidence is mediated by an inflation of perceived similarity; according to the latter, an inflation in perceived similarity is a way to reduce one's dissonance. However, the memory inference hypothesis does not predict a dud-induced inflation in perceived similarity; inferences about one's memory should only affect meta-cognitive beliefs, such as confidence, not perception. (It is unclear whether the pair-wise comparison hypothesis predicts an inflation of perceived similarity; multiple favorable comparisons may directly influence confidence or the effect may be mediated by an inflation in the perceived similarity of the identified lineup member.)

One of the purposes of the remaining three studies therefore is to conduct tests that differentiate between these accounts. Study 2 is the first of such tests. Specifically, it examines (1) whether the dud-induced confidence inflation is specific to the identified non-dud (as predicted by the pair-wise comparison and dissonance-reduction hypotheses) or not (as predicted by the perceptual contrast hypothesis); (2) whether duds inflate perceived similarity

Table 1 Predictions made by the various hypotheses concerning the cause of the dud-alternative effect

	Increase in self-perceived memory?	Increase in perceived similarity?	Are effects general or specific to one lineup member?	Effect dependent on having made an identification?
Pair-wise comparison	No	No	Specific	Yes
Perceptual contrast	No	Yes	General	No
Memory inference	Yes	No	Unsure	Unsure
Dissonance reduction	No	Yes	Specific	Yes

of the non-duds (as predicted by the perceptual contrast and dissonance-reduction hypotheses) or not (as predicted by the memory inference hypothesis), and if any such effect is specific to the identified non-dud or not; and (3) whether duds increase witnesses' beliefs about the quality of their memory (as predicted by the memory inference hypothesis) or not (as predicted by the pair-wise comparison, perceptual contrast, and dissonance-reduction hypotheses).

In addition, participants in Study 2 were given the option of not identifying anyone, which accomplishes three things: (1) It allow us to test the generality of the effect by examining whether the dud effect occurs when witnesses are given the option of not making an identification; (2) it allows an analysis of whether witnesses who do not make an identification still show evidence of the dud-alternative effect (as predicted by the perceptual contrast hypothesis) or not (as predicted by the pair-wise comparison and dissonance-reduction hypotheses); and (3) it allows an analysis of whether the presence of duds affects the likelihood that a witness will identify a non-dud.

Method

Participants

Two hundred three university undergraduates from a Midwestern university participated in exchange for research credit.

Materials

The same mock crime and the same lineups were used as in Study 1.

Procedure

The experiment was administered over computer. Participants viewed the mock crime and were then randomly assigned to a lineup condition (dud versus no dud). Participants were given unbiased instructions that the perpetrator may or may not be in the lineup and were explicitly given an option of responding 'not there.' Following their identification decision, witnesses provided their confidence in their decision, and then rated the similarity of each of the non-duds to their memory for the criminal (and witnesses in the dud condition also rated the similarity of the duds to their memory for the criminal), which they did on a 1 (very dissimilar) to 7 (very similar) scale. All participants then rated the quality of their memory for the criminal on a 1 (very poor memory) to 7 (very strong memory) scale.

Results

Unsurprisingly, participants who viewed a dud lineup perceived the duds as being highly dissimilar to their memory for the perpetrator—the mean similarity ratings for each of the duds among these participants ($n = 103$) were 1.2 [1.1, 1.3], 1.2 [1.1, 1.3], 1.1 [1.0, 1.2], and 1.3 [1.2, 1.4]. No participant identified a dud, and thus all 203 participants were included in the analyses.

Effect of Duds on Choosing

119 participants (58.6%) responded that the criminal was not in the lineup. Participants were significantly more likely to make an identification of a non-dud from a dud lineup (50.4%) than from a no dud lineup (32.0%), χ^2 ($df = 1$, $n = 203$) $= 7.15$, $p < .01$, $\varphi = .19$. In other words, the presence of duds increased the likelihood of making an identification.

Effect of Duds on Confidence

Overall, participants who viewed a dud lineup were significantly more confident in their response ($M = 7.9$ [7.5, 8.2]) than those participants who viewed a non-dud lineup ($M = 7.2$ [6.8, 7.6]), $t(201) = 2.31$, $p = .02$, $d = .33$. Participants were then split according to whether they made an identification or not. Among those participants who made an identification, the confidence of those who viewed a dud lineup ($M = 7.5$ [6.9, 8.0]) was marginally significantly higher than the confidence of those who viewed a non-dud lineup ($M = 6.7$ [6.0, 7.4]), $t(82) = 1.87$, $p = .065$, $d = .41$. Interestingly, among participants who did not make an identification (i.e., who rejected the lineup), the confidence of those who viewed a dud lineup ($M = 8.3$ [7.7, 8.8]) was significantly higher than the confidence of those who viewed a non-dud lineup ($M = 7.4$ [6.9, 8.0]), $t(117) = 2.10$, $p = .04$, $d = .39$. In other words, participants who viewed a dud lineup had greater confidence in their decision regardless of whether they decided to identify someone or not (although the effect was only marginally significant for choosers).

Tests of Pair-Wise Comparisons, Perceptual Contrast, Memory Inference, and Cognitive Dissonance Hypotheses

In contrast to the memory inference hypothesis, there was no significant effect of the presence of duds on participants' perceived memory for the criminal either overall ($Ms = 4.7$ [4.5, 4.9] and 4.6 [4.4, 4.8] for duds and non-duds, respectively), $t(201) = .81$, $p = .42$, $d = .11$, or looking only at choosers ($Ms = 4.8$ [4.5, 5.1] and 4.6 [4.3, 5.0] for duds and non-duds, respectively), $t(82) = .73$, $p = .47$,

$d = .16$ or non-choosers (Ms = 4.6 [4.3, 4.9] and 4.6 [4.3, 4.8] for duds and non-duds, respectively), $t(117) = .23$, $p = .82$, $d = .04$. However, consistent with the perceptual contrast, dissonance-reduction, and pair-wise comparison hypotheses, the presence of duds significantly increased the average perceived similarity of the non-duds to the witnesses' memory of the criminal, Ms = 4.5 [4.3, 4.7] and 3.7 [3.5, 3.9], $t(201) = 5.98$, $p < .001$, $d = .84$.

The perceptual contrast hypothesis predicts a general increase in perceived similarity—duds should increase the similarity of all non-duds. In contrast, both the dissonance-reduction and pair-wise comparison hypotheses predict a selective increase in perceived similarity—duds should only increase the similarity of the identified non-dud. Data support the perceptual contrast account: Duds increased perceived similarity of both the identified non-dud (Ms = 5.9 [5.7, 6.2] and 5.3 [5.0, 5.6]), $t(82) = 3.35$, $p = .001$, $d = .74$, as well as of the non-identified non-dud (Ms = 3.5 [3.1, 4.0] and 2.3 [1.8, 2.8]), $t(82) = 3.53$, $p = .001$, $d = .78$. The increase in similarity for the identified non-dud was not significantly different from the increase in similarity for the non-identified non-dud, $F(1, 82) = 2.38$, $p = .13$, Cohen's $f = .17$. In fact, the non-significant trend was in the opposite direction.

Additional evidence for the perceptual contrast hypothesis of a general increase in perceived similarity, and against the dissonance-reduction and pair-wise comparison hypotheses of a selective increase in perceived similarity, can be derived from an analysis of the non-choosers' data. Specifically, despite the fact that non-choosers by definition made no identification, the presence of duds nonetheless increased the average perceived similarity of the non-duds to the non-choosers' memory of the criminal, Ms = 4.3 [4.0, 4.6] and 3.6 [3.4, 3.9], $t(117) = 3.55$, $p = .001$, $d = .66$.

Discussion

Again, the presence of duds increased the confidence with which witnesses made their decisions (although this effect was only marginally significant among choosers). They did so despite participants being able to choose to reject the lineup. Furthermore, results from this study supported a perceptual contrast account for the observed dud-alternative effect: The presence of duds increased the similarity of the non-duds. Assuming that confidence in one's identification is at least partly a function of the similarity of the identified lineup member to one's memory of the criminal, this increased perceived similarity would lead to an increased confidence in one's identification.

In addition, results contradicted predictions made by both the memory inference hypothesis and the dissonance-

reduction hypothesis. Specifically, the presence of duds did not increase witnesses' self-perceptions of their memory, in contrast to memory inference predictions. And dud-induced confidence inflation occurred not only for the identified lineup member, but also for non-identified lineup members, in contrast to dissonance-reduction predictions.

Interestingly, this study also indicated that the presence of duds increased the likelihood of identifying a non-dud. Although this last finding was not directly predicted by the dud-alternative effect, it is easy to reconcile it with the previous findings: The presence of duds increased the similarity of the non-duds via a perceptual contrast effect, which made participants both more likely to identify a non-dud as being the perpetrator as well as increased their confidence with which they made that identification.

Study 3

Consistent with predictions made by the dud-alternative effect, Study 1 showed that the presence of highly dissimilar lineup members inflated witnesses' confidence in their mistaken identification of a non-dud, and Study 2 showed that this is consistent with a perceptual contrast account. Study 3 attempts to extend the dud-alternative finding in five ways. First, neither of the first two studies manipulated the instructions given to witnesses. Study 3 thus included an instruction manipulation; half of the participants were forced to make an identification (as in Study 1), and half were given the explicit option of indicating that the perpetrator was not in the lineup (as in Study 2).

Second, to assess whether the dud-alternative effect is limited to confidence decisions or whether it also affected witnesses' perceptions of other details about the crime, participants also responded to a number of other questions, concerning factors such as the view they had of the criminal, the amount of attention they paid the criminal, etc. This not only allows us to examine the generalizability of the dud-alternative effect, but it also allows for some secondary analyses. In particular, studies that have shown confidence inflation resulting from post-identification feedback have also shown commensurate inflation of responses to these other crime-related variables (Charman et al., 2010; Douglass & Steblay, 2006; Neuschatz et al., 2005; Wells & Bradfield, 1998, 1999; Wells, Olson, & Charman, 2003). An interesting question then is whether witnesses' responses to these other crime-related variables are simply based on their confidence judgments, or whether they are made independently of their confidence judgments. If, for instance, duds are shown to inflate confidence, but not witnesses' reports of how good a view they had of the criminal, how much attention they paid to the criminal's face, etc., then this would suggest that these other variables are assessed independently of confidence.

Third, to determine whether the dud-induced confidence inflation is limited to the identified lineup member, we also collected witnesses' assessments of their confidence that the non-identified non-dud was the perpetrator. This allows us to accomplish two things: First, it allows a test between the perceptual contrast hypothesis and Windschitl and Chambers' (2004) pair-wise comparison hypothesis—the former predicts that duds should inflate the confidence of the non-identified lineup member whereas the latter does not. (In fact, the perceptual contrast hypothesis may even predict that the confidence inflation should be greater for the non-identified lineup member since that person should tend to be perceived as less similar to the perpetrator than the identified lineup member, and thus less susceptible to ceiling effects.) Second, it allows a more direct test of whether the effect of duds on confidence is mediated by increases in perceived similarity. Although this mediation analysis could have been done on data obtained from Study 2, the results would be less informative since it could only have been run on witnesses who chose to make an identification (since witnesses who chose not to make an identification only provided confidence that their non-identification was correct, and did not provide confidence for the likelihood that a given lineup member was the criminal). Because the decision to make an identification or not was a self-selected variable, any mediation analyses based on that self-selected group would be misleading and may not generalize beyond witnesses who made an identification. However, because all participants in Study 3 provided their confidence that each non-dud was the criminal, regardless of their actual identification decision, a proper mediation analysis could be conducted.

Fourth, to assess whether dud-induced confidence inflation occurs even when witnesses are instructed prior to a lineup that their confidence in their identification will be assessed, we included such an instruction in our methodology. Some police departments do just this before showing witnesses a lineup and, in fact, it is a standard part of the instruction recommended by the U.S. Department of Justice (Technical Working Group, 1999). It is conceivable that directing witnesses' attention to their own internal states would serve to buffer them against any confidence-inflation potentially induced by duds. Thus, showing dud-induced confidence inflation under this condition would extend the generalizability of the dud-alternative effect.

Fifth, in order to minimize the possibility that our previous studies' findings were the result of the specific crime that was viewed, the criminal that was seen, the lineups that were used, or the population from which the sample was drawn, Study 3 used a different mock crime with a different criminal, different lineups with different duds, and a sample that was obtained from a different population.

Method

Participants

One hundred forty university undergraduates from a large Southeastern university participated in this experiment in exchange for course credit.

Materials

Participants viewed a 30-s mock crime of two people—one male and one female—stealing a car. Participants viewed either a two-person non-dud lineup (comprised of two fillers who resembled the male perpetrator) or a six-person dud lineup (comprised of the same two fillers plus an additional four duds). Lineups were pre-tested by having participants ($N = 34$) view a looped clip from the mock crime in which the male criminal's face was clearly seen beside each of the individual lineup members in turn. Participants indicated the similarity between the perpetrator and each of the lineup members on a scale from 1 (very dissimilar) to 7 (very similar). Results from this pre-testing indicated that the perceived similarity of the non-duds to the criminal was moderate (Ms = 3.4 [2.9, 3.9] and 4.8 [4.4, 5.2]), and that the similarity between the perpetrator and each of the duds was extremely low (Ms = 1.0 [1.0, 1.1], 1.1 [1.0, 1.2], 1.0 [1.0, 1.1], and 2.4 [1.8, 2.9]). The average non-dud similarity was significantly greater than the average dud similarity, $t(33) = 17.15$, $p < .001$, $d = 5.97$. Appendix B displays the dud lineup along with the actual perpetrator.

Participants responded to a memory questionnaire. The exact questions that were asked are displayed in Table 2.

Procedure

Participants were randomly assigned to a lineup condition (dud vs no dud) and to a lineup instruction condition (biased instructions vs non-biased instructions). Participants viewed the mock crime individually on a computer and then performed a 5-min filler task (which consisted of rating a series of unrelated stories for consistency). After this filler task, the experimenter informed the participants that they were a witness to a car theft and that they would be seeing a lineup comprised of males. Participants in the biased instructions condition were told to "tell me the number of the individual that you think stole the car." Participants in the non-biased instructions condition were told that the actual car thief may or may not be in the lineup and to "tell me the number of the individual that you think stole the car or say 'not there' if you do not think he is in the lineup." In order to make their confidence processes salient, all participants were told that they would be asked

Table 2 Dependent measures questionnaire

Dependent measures question	Scale
At the time you identified the person in the lineup, how **certain** were you that the person you identified from the lineup was the person you saw in the video?	0% (not at all certain) to 100% (totally certain)
How good a **view** did you get of the person in the video?	1 (very poor) to 10 (very good)
How much **attention** were you paying to the person's face while viewing the video?	1 (none) to 10 (my total attention)
For how long would you estimate the man's **face** was in view during the video?	1 (very little time) to 10 (quite a bit of time)
How well were you able to make out specific **features** of the person's face from the video?	1 (not at all) to 10 (very well)
How **far** away was the man in the video?	1 (not far) to 10 (very far)
To what extent to you feel that you had a good **basis** (enough information) to make an identification?	1 (no basis at all) to 10 (a very good basis)
How **easy** or difficult was it for you to figure out which person in the lineup was the person you saw in the video?	1 (extremely easy) to 10 (extremely difficult)
After you were first shown the lineup, how **long** do you estimate it took you to make an identification?	1 (I needed almost no time to pick him out) to 10 (I had to look at the photos for a long time to pick him out)
On the basis of your memory of the person you saw in the video, how willing would you have been to **testify** in court that the person you identified was the person in the video?	1 (not at all willing) to 10 (totally willing)
Generally, how good is your recognition **memory** for the faces of strangers you have encountered on only one prior occasion?	1 (very poor) to 10 (excellent)
How **clear** is the image you have in your memory of the person you saw in the video?	1 (not at all clear) to 10 (very clear)

about their confidence in their decision following the identification. Participants were then shown the lineup that corresponded to their condition and made their identification decision. Participants provided confidence reports (on a 1–10 scale) that each of the two non-duds was the criminal, as well as similarity ratings. Finally, participants responded to the memory questionnaire while the lineup was still in view.

Results

No participant identified a dud, and thus all analyses are based on all 140 participants.

Effect of Duds on Choosing

Among witnesses who received unbiased instructions, the presence of duds had no significant effect on choosing rates, χ^2 ($df = 1$, $n = 73$) $= .39$, $p = .53$, $\varphi = .07$.

Effects on Confidence

To assess the effect of duds and instructions on confidence, 2 (lineup: dud vs no dud) $\times$ 2 (instructions: biased vs not biased) ANOVAs were conducted on the non-dud

confidence ratings. Because witnesses provided confidence ratings for both non-duds, these confidence ratings were averaged for these initial analyses. The addition of duds significantly increased confidence in a non-dud being the criminal ($Ms = 4.7$ [4.2, 5.3] and 3.8 [3.3, 4.2]), $F(1, 136) = 9.00$, $p = .003$, Cohen's $f = .26$. Interestingly, biased instructions also significantly increased the confidence in a non-dud being the criminal ($Ms = 3.4$ [2.9, 3.9] and 5.1 [4.7, 5.6]), $F(1, 136) = 27.26$, $p < .001$, Cohen's $f = .45$. The interaction between these variables was not significant, $F(1, 136) = .02$, $p = .90$, Cohen's $f = .00$.

To assess whether the presence of duds increased confidence that a non-dud was the criminal for both choosers and non-choosers, a 2 (lineup) $\times$ 2 (choice: identification or no identification) ANOVA was conducted on the average of the confidence ratings of the two non-duds among witnesses in the nonbiased instructions condition. The presence of duds increased confidence that a non-dud was the criminal ($Ms = 2.9$ [2.4, 3.5] and 3.9 [3.1, 4.8]), $F(1, 69) = 4.41$, $p = .04$, Cohen's $f = .25$. A lack of a significant interaction indicated that the magnitude of this confidence inflation was not significantly different for choosers compared to non-choosers, $F(1, 69) = .10$, $p = .75$, Cohen's $f = .03$.

To examine whether the dud-induced confidence inflation was selective or general, a 2 (lineup) $\times$ 2 (non-dud: identified vs non-identified) mixed ANOVA (with non-dud

as a within-subjects factor) was conducted among witnesses who made an identification. Results were consistent with the perceptual contrast hypothesis and contrary to both the dissonance-reduction hypothesis and the pair-wise comparison hypothesis: Although the addition of duds significantly increased average confidence in the non-duds being the criminal (Ms = 4.6 [4.2, 5.1] and 5.6 [5.1, 6.1]), $F(1, 94) = 8.25$, $p = .005$, Cohen's $f = .30$, a lack of a significant interaction indicated that the magnitude of this increase was not significantly different for the identified non-dud compared to the non-identified non-dud, $F(1, 94) = 1.16$, $p = .28$, Cohen's $f = .11$.

Effects on Similarity

Similarity ratings for the two non-duds were averaged for these first analyses. The effects on similarity paralleled those found for confidence. The addition of duds significantly increased the perceived similarity of the non-duds to the criminal (Ms = 5.3 [4.9, 5.7] and 4.7 [4.4, 5.1]), $F(1, 136) = 5.35$, $p = .02$, Cohen's $f = .20$. Biased instructions also significantly increased these similarity scores (Ms = 5.6 [5.3, 5.9] and 4.5 [4.1, 4.9]), $F(1, 136) = 17.73$, $p < .001$, Cohen's $f = .36$. The interaction between these two variables was not significant, $F(1, 136) = .01$, $p = .93$, Cohen's $f = .00$.

To examine whether this non-dud similarity inflation was selective or general, a 2 (lineup) × 2 (non-dud: identified vs non-identified) mixed ANOVA (with non-dud as a within-subjects factor) was conducted among witnesses who made an identification. Results were again consistent with the perceptual contrast hypothesis and contrary to both the dissonance-reduction hypothesis and the pair-wise comparison hypothesis: Although the addition of duds significantly increased average similarity of the non-duds to the criminal (Ms = 5.2 [4.9, 5.6] and 6.0 [5.6, 6.4]), $F(1, 94) = 8.17$, $p = .005$, Cohen's $f = .29$, this increase was significantly greater for the non-identified non-dud (Ms = 3.2 [2.7, 3.7] and 4.6 [4.0, 5.2]) compared to the identified non-dud (Ms = 7.3 [6.8, 7.7] and 7.4 [7.0, 7.9]), $F_{\text{interaction}}(1, 94) = 6.07$, $p = .02$, Cohen's $f = .25$.

An examination of non-choosers revealed that the addition of duds did not significantly increase perceived similarity (Ms = 3.7 [3.0, 4.3] and 3.8 [3.0, 4.5] for non-dud lineup and dud lineup, respectively), $t(42) = .18$, $p = .86$, $d = .06$, nor confidence that a non-dud was the criminal (Ms = 2.0 [1.4, 2.7] and 2.8 [1.7, 3.8] for non-dud lineup and dud lineup, respectively), $t(42) = 1.29$, $p = .20$, $d = .40$. Although this latter result seems at first glance to contradict the earlier omnibus analysis that showed that the effect of duds on confidence was not moderated by whether the witnesses identified someone or not from the lineup, note that this analysis has much less

power to detect a difference (since it is based on data from only 44 non-choosers) compared to the omnibus analysis (which is based on data from 140 participants), and the lack of a significant effect of duds among non-choosers is very likely attributable to a lack of power to find the effect.

Effects on Other Testimony-Relevant Variables

A composite measure was created for all remaining testimony-relevant variables by averaging across them (with appropriate variables reverse-scored) so that larger numbers represent a better witnessing experience. A 2 (lineup) × 2 (instructions) ANOVA on this composite measure revealed no significant effect for lineup, $F(1, 136) = .03$, $p = .87$, Cohen's $f = .00$, or instructions, $F(1, 136) = 2.56$, $p = .11$, Cohen's $f = .14$, and no significant interaction between these variables, $F(1, 136) = .06$, $p = .80$, Cohen's $f = .00$.

Test of Mediation

According to the perceptual contrast hypothesis, the addition of duds inflates the perceived similarity of the non-duds to the witness's memory of the criminal, which in turn inflates the confidence that the non-dud is the criminal. To test for this mediational account, Baron and Kenny's (1986) method of mediation testing was conducted. All conditions for mediation were met. The presence of duds significantly predicted average confidence that the non-duds were the criminal, $\beta = .22$, $t = 2.70$, $p = .008$. The presence of duds significantly predicted the average similarity score for the non-duds, $\beta = .18$, $t = 2.14$, $p = .03$. Average non-dud similarity scores predicted average non-dud confidence scores when controlling for the presence of duds, $\beta = .69$, $t = 11.42$, $p < .001$. Finally, the relationship between the presence of duds and average non-dud confidence was reduced to non-significance when controlling for average non-dud similarity, $\beta = .10$, $t = 1.65$, $p = .10$. Sobel's (1982) test, which tests the significance of the mediated path, was significant, $z = 2.10$, $p = .04$.

Discussion

The results of Study 3 again show that duds can increase the confidence with which witnesses believe a non-dud to be the criminal, and demonstrate the generality of the dud-alternative effect. This effect occurred despite using a different mock crime, a different lineup with different lineup members, a sample drawn from a different population, and different pre-lineup instructions (specifically, that witnesses would need to report their identification confidence) than Studies 1 and 2. The effect occurred regardless

of whether the pre-lineup instructions were biased or not, and was not moderated by whether witnesses identified a lineup member or not. This generality underscores the potential danger of including duds in a lineup.

In addition, the results again supported a perceptual contrast account of the dud-alternative effect. The dud-induced confidence inflation occurred for both the identified as well as the non-identified non-dud. Further, the addition of duds inflated the perceived similarity of both of the two non-duds. These findings are inconsistent with the cognitive dissonance or pair-wise comparison hypotheses (which both predict a selective inflation in confidence and similarity for only the identified lineup member), but are consistent with a perceptual contrast account. A mediational analysis also indicated the effect of duds on confidence is mediated by an increase in perceived similarity of the non-duds.

Despite inflating confidence, the addition of duds failed to significantly inflate other testimony-relevant measures, such as how much attention the witnesses paid to the criminal, how good a view they had of the criminal, etc., suggesting that confidence may be assessed differently from these other testimony-relevant variables. In retrospect, this lack of an effect is not surprising, given some of our earlier findings. Whereas confidence is likely based at least in part on perceived similarity of the lineup member to one's memory of the criminal, responses to these other variables are probably largely driven by witnesses' meta-cognitive beliefs, such as their inferences about the quality of the memory (i.e., "I have a good memory, therefore I must have paid a lot of attention, had a good view, etc."). But, as Study 2 demonstrated, the effect of the presence of duds is to inflate perceived similarity, and not to inflate witnesses' meta-cognitive beliefs. This would explain why manipulations that influence one's perceived memory (such as post-identification feedback) tend to influence confidence and other measures similarly (Douglass & Steblay, 2006) whereas manipulations that influence perceived similarity (such as the presence of duds) influence only confidence and not these other measures (i.e., the current study).

Study 4

Although Study 3 largely supports the perceptual contrast account for the dud-alternative effect, we found some ambiguous data. Specifically, although omnibus analyses indicated that the dud effect was not moderated by whether witnesses identified a lineup member or not, simple main effects failed to indicate a significant effect of duds on confidence among non-choosers. Although this null effect was likely due to insufficient power, we cannot completely

rule out the possibility that the effect of duds on confidence does depend on whether the witness made a choice or not. One of the purposes of Study 4 is to examine this possibility further.

The only way we have so far examined whether the dud effect exists among non-choosers was to examine witnesses who decided not to make an identification. But instead of allowing witnesses to decide themselves whether to make an identification or not, it would be highly beneficial for us to directly manipulate whether witnesses make an identification decision or not. Thus, we could examine whether the dud effect is dependent on having staked oneself to a choice. If the perceptual contrast account of the dud-alternative effect is correct, then whether witnesses make an identification decision prior to providing confidence and similarity ratings is irrelevant—the similarity of the non-duds should be inflated regardless. However, both the pair-wise comparison hypothesis and the cognitive dissonance hypothesis predict that the dud effect is dependent on having staked oneself to an identification decision. In the former case, it matters because the identification provides the focal point against which the comparisons are made. In the latter case, it matters because a commitment, such as an identification, is necessary to induce dissonance-reduction strategies (Brehm & Cohen, 1962). Thus, to determine whether the dud effect is dependent on having made an identification, Study 4 includes a condition in which witnesses assess their confidence that the lineup members are the criminal without having previously made any identification decision.

Study 4 also manipulated the level of dissimilarity of the duds in order to ascertain the boundary conditions of the dud effect. This is important for at least three reasons. First, although (as previously explained in the introduction) a proper examination of the dud effect is accomplished by manipulating the addition of duds, this technically leaves the possibility that the dud effect is the result not of duds per se, but of lineup size. For instance, the cognitive dissonance hypothesis predicts that a greater array size, by increasing the number of alternatives people have to choose from, makes the task more difficult and thus increases dissonance concerns (Liberman & Förster, 2006), which would increase the tendency to inflate the similarity and confidence of the chosen lineup member. If true, this effect would occur independently of the existence of duds per se. In fact, the cognitive dissonance hypothesis predicts that the effect should get *stronger* as the similarity of the fillers to the criminal increases, because that makes the lineup task even more difficult. A purely perceptual contrast account, on the other hand, predicts that the confidence-inflating effects are critically dependent on the existence of duds per se, since it is the presence of duds that creates the context that produces the perceptual contrast, and thus the effect should disappear as the similarity

of the fillers to the criminal increases. Including lineups of the same nominal size that vary only with respect to the levels of dissimilarity of the 'duds' thus allows us to test this possibility.

Second, it is important to know how dissimilar the duds have to be in order for the dud-alternative effect to occur. Will it occur only when the alternatives are extremely dissimilar, or will it also occur for slightly dissimilar alternatives? Having some sort of measure of just how dissimilar duds have to be to inflate confidence provides more information to an expert witness who must decide whether a dud-alternative effect may have been operating in a given lineup.

Third, it is possible that the dud effect will actually disappear if the duds get too dissimilar. This could happen if the duds are so dissimilar that they are no longer perceived as providing an appropriate context from which to evaluate one's confidence in a lineup identification. To test this, we included duds that were even more dissimilar to the criminal than we have used in the previous three studies.

Finally, to increase generalizability even more, we used a different mock crime with a different suspect and different lineups with different fillers from the previous three studies. Furthermore, whereas duds in the previous three studies were selected by the experimenters, duds in Study 4 were 'selected' in pre-testing by participants who rated the similarity of the criminal to a series of individual faces. Our selection of fillers in the various lineups were chosen based on these participants' a priori similarity ratings.

Method

Design

A 5 (filler dissimilarity: high dissimilarity, moderately high dissimilarity, moderate dissimilarity, low dissimilarity, no duds) × 3 (identification requirement: forced identification, elective identification, and prohibited identification) between-subjects design was used.

Participants

Two hundred twelve university undergraduates from a large Southeastern university participated in this experiment in exchange for course credit.

Materials

Mock Crime Video. A 30-s mock crime video was created that featured a young man and a young woman unlocking a car with a slim-jim and driving off quickly. The video included approximately 8 s where the young man's face was prominently featured.

Lineups. We required five lineups for our comparisons—one that includes just two non-duds (control condition) and four that include the same two non-duds plus four 'duds.' These latter four lineups differed from each other in terms of the degree of dissimilarity of the duds. Because the term 'duds' refers specifically to highly dissimilar lineup members, in the present study we refer to those four individuals whose similarity we manipulated as "fillers," and the two highly similar individuals who remain constant across all lineups as "false targets." In order to create multiple lineups with varying levels of filler dissimilarity, we showed undergraduate students ($n = 91$) 51 photographs of people's faces one at a time in random order beside a looped 8-s video clip of the male perpetrator. Participants rated the similarity of each photograph to the perpetrator on a 1 (extremely dissimilar) to 9 (extremely similar) scale. Each photograph was then rank ordered according to the average similarity rating from 1 (most dissimilar) to 51 (most similar). The two photographs rated highest in similarity (Ms = 4.52 [4.0, 5.1] and 4.41 [3.9, 5.0]) were selected to serve as the two false targets for each lineup. The next four most highly rated photographs (rank order numbers 46-49; $M = 3.72$ [3.4, 4.1]) were selected to be used as the low dissimilarity fillers, rank order numbers 32-35 ($M = 2.52$ [2.2, 2.8]) were selected to be used as the moderate dissimiliarity fillers, rank order numbers 17-20 ($M = 1.91$ [1.7, 2.1]) were selected to be used as the moderately high dissimilarity fillers, and the four lowest rated photographs (rank order numbers 1–4; $M = 1.23$ [1.1, 1.4]) were selected to be used as the high dissimilarity fillers.

Pair-wise comparisons confirmed significant differences between the means of each level of dissimilarity (high dissimilarity vs moderately high dissimilarity $t(90) = 5.94$, $p < .001$, $d = 1.25$; moderately high dissimilarity vs moderate dissimilarity $t(90) = 4.97$, $p < .001$, $d = 1.05$; moderate dissimilarity vs low dissimilarity $t(90) = 7.04$, $p < .001$, $d = 1.48$). All lineups were created from the two false targets plus the four appropriate fillers. Appendix C displays all dud lineups and the actual perpetrator.

Measures. All similarity and confidence ratings were assessed on 1 (low confidence/low similarity) to 10 (high confidence/high similarity) scales.

Procedure

Participants completed this study online. They viewed the mock crime and then responded to questions concerning personal and demographic information (as a short filler task) and were randomly assigned to all conditions. In the *forced identification* conditions, participants were given written instructions via the computer that they were

to make an identification from the lineup (i.e., they could not indicate that the criminal was not in the lineup). In the *elective identification* conditions, participants were given written instructions that they were to either identify a lineup member or say that the perpetrator was not in the lineup. In the *prohibited identification* conditions, participants were given no instructions regarding making an identification, and were never asked for any identification decision. Participants then viewed their assigned lineup (and either made an identification or not, according to instruction condition). All participants, regardless of instruction type, then indicated their confidence that each lineup member was the perpetrator, and indicated how similar each lineup member was to their memory of the perpetrator. Participants were then debriefed and excused.

Results

Table 3 displays mean confidence and similarity ratings of false target lineup members as a function of filler dissimilarity and identification requirement condition.

The Dud Effect as a Function of Dud Dissimilarity

The first set of results examines whether the dud effect depends on the degree of dissimilarity of the fillers and thus collapses across identification requirement conditions.

Effect of Duds on Confidence. To test whether the presence of duds increased confidence in the non-duds, we compared the mean false target confidence ratings in each of the four filler-present lineups to mean false target confidence ratings in the control condition. In order to account for multiple comparisons, a Bonferroni adjustment was made, which set our critical alpha level to .05/4 = .012.

The presence of fillers significantly inflated average confidence that a false target was the criminal only in the highly dissimilar filler condition, $t(77) = 3.17$, $p < .01$, $d = 0.72$, but in none of the other conditions (moderately high dissimilarity: $t(79) = 1.51$, $p = .13$, $d = 0.34$; moderate dissimilarity: $t(78) = 1.04$, $p = .30$, $d = 0.24$: low dissimilarity, $t(96) = 0.05$, $p = .96$, $d = 0.01$).

Effect of Duds on Similarity. To test whether the presence of duds increased perceived similarity, we

Table 3 Mean confidence scores and similarity ratings (SDs in parentheses and 95% CIs in square brackets) of the average non-dud lineup members as a function of filler dissimilarity and identification requirement condition in Study 4

Identification requirement	Filler dissimilarity to target				
	No fillers	High dissimilarity	Moderately high dissimilarity	Moderate dissimilarity	Low dissimilarity
Confidence					
Forced identification	5.2 (1.3)	6.8 (1.8)	5.4 (2.5)	3.4 (2.3)	4.6 (2.3)
	[4.5, 5.9]	[5.6, 7.9]	[4.1, 6.8]	[1.5, 5.4]	[3.3, 5.9]
Elective identification	3.0 (1.7)	5.2 (2.4)	5.4 (2.5)	4.6 (2.9)	3.8 (2.7)
	[1.9, 4.1]	[3.8, 6.6]	[3.2, 7.5]	[2.5, 6.7]	[2.5, 5.2]
Prohibited identification	5.3 (1.3)	6.1 (2.2)	5.2 (2.9)	4.1 (2.8)	5.1 (2.4)
	[4.5, 6.0]	[4.7, 7.6]	[3.5, 6.8]	[2.8, 5.3]	[4.1, 6.2]
Overall	4.6 (1.7)	6.0 (2.2)	5.3 (2.6)	4.1 (2.7)	4.6 (2.5)
	[4.0, 5.1]	[5.3, 6.7]	[4.5, 6.2]	[3.2, 4.9]	[3.9, 5.2]
Similarity					
Forced identification	6.6 (2.2)	8.0 (1.8)	6.3 (2.7)	5.4 (2.5)	3.8 (2.3)
	[5.4, 7.7]	[6.9, 9.2]	[4.8, 7.7]	[3.4, 7.5]	[2.5, 5.1]
Elective identification	4.4 (1.5)	7.0 (2.4)	5.7 (2.5)	4.3 (2.7)	4.2 (2.1)
	[3.5, 5.3]	[5.6, 8.3]	[3.6, 7.8]	[2.3, 6.3]	[3.1, 5.2]
Prohibited identification	5.7 (1.3)	7.7 (2.0)	6.1 (2.8)	4.5 (2.8)	5.2 (2.7)
	[4.9, 6.4]	[6.4, 9.0]	[4.5, 7.6]	[3.2, 5.8]	[4.0, 6.3]
Overall	5.7 (1.9)	7.5 (2.1)	6.1 (2.6)	4.7 (2.7)	4.5 (2.4)
	[5.1, 6.3]	[6.8, 8.2]	[5.2, 6.9]	[3.8, 5.5]	[3.8, 5.1]

compared the mean false target similarity ratings in each of the four filler-present lineups to mean false target similarity ratings in the control condition. In order to account for multiple comparisons, a Bonferroni adjustment was made, which set our critical alpha level to $.05/4 = .012$.

The presence of duds significantly inflated similarity scores only in the highly dissimilar filler condition, $t(77) = 4.15$, $p < .01$, $d = 0.95$, but not for the moderately high dissimilarity condition, $t(79) = .807$, $p = .42$, $d = 0.18$, or the moderate dissimilarity condition, $t(78) = 1.96$, $p = .05$. $d = 0.44$. Interestingly, the presence of low dissimilarity fillers lowered the similarity scores of the false targets, $t(96) = 2.59$, $p = .01$, $d = 0.53$.

Effect of Duds on Choosing. To determine whether the presence of duds increased choosing rates, we must examine only participants in the elective identification instruction conditions because that is the only condition where participants were allowed to choose whether or not to make an identification. Because we only found the dud effect among participants who viewed a lineup with highly dissimilar fillers, we only examined that condition. The presence of highly dissimilar fillers did not significantly affect the likelihood that a witness would make an identification χ^2 $(df = 1, n = 26) = .52, p = .47, \varphi = .14$.

The Dud Effect as a Function of Identification Requirement Condition

This set of analyses examine whether the magnitude of the dud effect is dependent on having previously committed oneself to a response option. Because the dud effect was only found in the high dissimilarity filler lineup condition, all comparisons involve only this highly dissimilar filler condition and the control condition.

If committing oneself to a decision drives the dud effect, it should disappear in the condition in which participants did not make an identification. A 3 (identification requirement) × 2 (filler dissimilarity: highly dissimilar vs control) ANOVA failed to produce a significant interaction on both the average non-dud confidence, $F(2,73) = .85$, $p = .43$, Cohen's $f = .15$, and the average non-dud similarity, $F(2,73) = .56$, $p = .58$, Cohen's $f = .12$, suggesting that the dud effect was present regardless of whether witnesses were forced to make an identification, were unable to make an identification, or chose themselves whether to make an identification or not.

Further evidence supporting the idea that committing oneself to a decision has no bearing on the dud effect can be derived by looking just at the choosers' data. Specifically, if committing oneself to a decision is necessary to produce the dud effect, then the dud effect should be exhibited for the

Table 4 Mean confidence and similarity ratings (SDs in parentheses and 95% CIs in square brackets) of non-dud lineup members as a function of filler dissimilarity condition and non-dud identification status among witnesses who made an identification in Study 4

Filler dissimilarity condition	Non-dud identification status	
	Identified	Not identified
Confidence		
High dissimilarity fillers ($n = 16$)	8.4 (1.6)	4.7 (3.0)
	[7.5, 9.2]	[3.1, 6.3]
No fillers ($n = 18$)	7.5 (2.2)	2.7 (1.7)
	[6.4, 8.6]	[1.9, 3.5]
Similarity		
High dissimilarity fillers ($n = 16$)	8.9 (1.2)	6.3 (3.0)
	[8.3, 9.6]	[4.7, 7.9]
No fillers ($n = 18$)	7.8 (2.4)	5.0 (2.7)
	[6.6, 9.0]	[3.6, 6.4]

identified false target, but not for the non-identified false target. To examine this, 2 (filler dissimilarity: control vs highly dissimilar) × 2 (non-dud: identified vs not identified) mixed ANOVAs (with non-dud as a within-subjects factor) were conducted on all witnesses who made an identification (i.e., all witnesses in the forced identification condition, and witnesses who chose to make an identification in the elective identification condition). Mean scores are displayed in Table 4. The presence of highly dissimilar fillers inflated both confidence, $F(1, 32) = 75.13, p < .001$, Cohen's $f = 1.53$, and similarity ratings, $F(1, 32) = 36.49$, $p < .001$, Cohen's $f = 1.07$. More importantly for present purposes, the interaction was non-significant for confidence ratings, $F(1, 32) = 1.36$, $p = .25$, Cohen's $f = .21$ indicating that duds inflated confidence ratings for all false targets regardless of whether the false target had been identified or not. Similarly, the interaction was also non-significant for similarity ratings, $F(1, 32) = .03$, $p = .87$, Cohen's $f = .03$, indicating that duds inflated similarity ratings for all false targets, regardless of whether the false target had been identified or not.

Test of Mediation

To test whether the effect of duds on confidence was mediated by inflated perceptions of similarity, Baron and Kenny's (1986) method of mediation testing was conducted. Because the dud effect was only found when comparing the control condition to the highly dissimilar filler condition, only these conditions were used in the analysis. All conditions for mediation were met. The presence of duds significantly predicted average confidence that the false targets were the criminal, $\beta = .70$,

$t = 3.17$, $p = .002$. The presence of duds significantly predicted the average similarity score for the false targets, $\beta = .93$, $t = 4.15$, $p < .001$. Average similarity score for the false targets predicted average false target confidence score, even when controlling for the presence of duds, $\beta = .58$, $t = 6.33$, $p < .001$. Finally, the relationship between the presence of duds and average false target confidence was reduced to non-significance when controlling for average similarity score for the false targets, $\beta = .16$, $t = .81$, $p = .42$. Sobel's (1982) test, which tests the significance of the mediated path, was significant, $z = 3.47$, $p < .001$.

Discussion

There are three main findings from Study 4: (1) the dud effect occurred only when there was a high level of dissimilarity between the duds and the criminal, and disappeared otherwise; (2) the dud effect was mediated by increases in perceived similarity of the non-duds to the criminal; (3) the dud effect did not seem to depend on having previously made an identification. All of these findings are consistent with the perceptual contrast hypothesis: The presence of highly dissimilar fillers makes others in the lineup appear more similar to the criminal, increasing witnesses' confidence that those individuals are the criminal, and because they do so at a purely perceptual level, the effect does not depend on having previously made an identification. In contrast, these findings contradict predictions made by other possible theoretical accounts of the dud effect. For example, the cognitive dissonance hypothesis predicts that the dud effect should have gotten stronger the more similar the fillers were to the criminal, and should only appear for lineup members who had been previously identified. Consequently, the results of Study 4 are consistent with the previous studies: Duds inflate confidence that a non-dud is the perpetrator by means of increasing the perceived similarity of that non-dud to the criminal.

Not only is it true that the perceived similarity of an individual to one's memory of a criminal is inflated by the presence of highly dissimilar fillers, but it also seems to be true that perceived similarity is deflated by the presence of highly similar fillers. The current study showed that when witnesses viewed a lineup composed of low dissimilarity (i.e., highly similar) fillers, their similarity ratings for the false targets were lower than control. This unanticipated, yet perfectly consistent finding, suggests that the dud effect is simply a specific manifestation of a more general principle: Similarity judgments are

determined, in part, by context. Given the importance that perceived similarity plays in eyewitness identification, it is surprising that this realization has been so slow in coming. In fact, only two other studies that we are aware of (Charman, Hyman Gregory, & Carlucci, 2009; McQuiston-Surrett, Douglass, & Burkhardt, 2008) have demonstrated the malleability of similarity judgments. Models of eyewitness behavior have tended to neglect the role of context on perceived similarity, often assuming that perceived similarity is purely a function of shared features between two faces plus random variation (e.g., the WITNESS model; Clark, 2003). These results suggest that these theories are missing an important piece of the picture: Perceived similarity is largely subjective, varying from context to context.

And if perceived similarity is dependent on context, why not other variables as well? As we have shown, confidence that one's identification was correct is similarly dependent on context. Although this point has been partly recognized in the literature—for example, it has been recognized that confidence varies depending on whether the witness received post-identification feedback (Wells & Bradfield, 1998), the public vs private nature of the feedback (Shaw, Appio, Zerr, & Pontoski, 2007), etc., the current studies demonstrate that this context is embedded within the structure of the lineup itself. When a witness claims to be 80% confident that the identified person was the criminal, it is important to realize that this claim is not purely a function of the witness's meta-cognitive abilities, but is tied to the specific lineup that was viewed. The confidence a witness expresses concerning an identification, in other words, is expected to vary, depending on the fillers present within that lineup. In the best-case scenario, this influence of lineup context on witness confidence is limited to the presence of highly dissimilar fillers, as demonstrated in the current study; on the other hand, research into the influences of lineup construction on confidence has likely just scratched the surface, and there may be many other contextual lineup factors that also influence confidence. For example, if the dud effect is a perceptual phenomenon, as the data tend to indicate, then the magnitude of the effect may be dependent not just on the contents of the lineup (i.e., the fillers), but it may also be dependent on the physical structure of the lineup itself—lineups composed of faces that are tightly bunched may create a stronger perceptual contrast, and thus a stronger dud effect, than lineups composed of faces that are not tightly bunched. Researchers, as well as those who provide expert testimony in court, should be aware as much as possible of how lineup context may affect witness confidence.

General Discussion

Results from four studies provide evidence that the addition of highly dissimilar lineup fillers—duds—to a lineup can increase the confidence with which a false identification is made, and that it can do so independently of witnesses' identification decision. Results from multiple studies suggest that this phenomenon is mediated by an increase in perceived similarity of non-dud lineup members to witnesses' memories of the criminal. In other words, the presence of highly dissimilar lineup members makes the similar-looking lineup members appear even *more* similar to the criminal. This, in turn, increases the confidence with which witnesses identify one of these similar-looking, but innocent, lineup members.

This finding that, under specific circumstances, the addition of alternatives can increase the confidence with which one chooses a response option, provides evidence that support theory, which claims that the addition of response options can never increase the perceived support for a response, is, at best, an incomplete account of how people assess their confidence in decisions. This effect was first noted by Windschitl and Chambers (2004), but was limited to non-perceptual and non-memorial decision-making tasks. The current manuscript extends the generality of this effect to cover a perceptual, memorial decision-making task. But the mechanism by which the presence of duds inflated witnesses' confidence in their lineup identifications (i.e., via an increase in perceived similarity) raises interesting questions about the nature of the dud-alternative effect more broadly. After all, the original Windschitl and Chambers (2004) studies did not involve perceptual tasks, and yet they nonetheless found the dud-alternative effect. There are at least two ways to account for this.

First, it is possible that the dud-alternative effect operates in multiple ways. The presence of duds might have multiple effects; the specific route via which the duds inflate confidence may be a consequence of the specific stimuli and methodological procedures used in that study. Duds may increase the number of favorable pair-wise comparisons one makes from general probability-based reasoning, but may also invoke a perceptual contrast when the task is more perceptual in nature.

Second, there may exist a more general principle at play, a single superordinate principle that can explain both the increase in confidence seen in the Windschitl and Chambers (2004) studies as well as the increase in confidence seen in the current studies. The two different routes through which duds inflate confidence may simply be two different manifestations of this more general, underlying principle. This overarching principle would have to be broad enough to transcend the methodological differences between the

various studies demonstrating a dud-alternative effect. What could this principle be?

A Scaling Effect Explanation

The inflation of confidence seen in the Windschitl and Chambers (2004) studies as well as in the current manuscript's studies may, in fact, both be due to an artifactual scale effect. When someone must generate a similarity score between two faces on a subjective scale (such as a 1-to-7 scale), she must first, in effect, generate anchor points. What exactly does a score of '1' mean? If the criminal was a 21-year-old white male, surely a 80-year-old black male may be considered dissimilar enough to warrant a '1' on the similarity scale. But someone else may reason that a 80-year-old black female would be even more dissimilar, and if so, then the 80-year-old black male must warrant a higher similarity score than '1.' (And someone else may reason that a refrigerator is even more dissimilar than that….) The point is that each individual person must subjectively (and probably implicitly) decide for themselves what these anchors are. Because there is no objective way to determine these anchor points, they may be affected by external influences. One of these influences may be the presence of duds.

In the absence of duds, people may implicitly assign a moderately dissimilar person as being the lowest anchor point on the similarity scale. But when those duds are present, people may reason that the lowest anchor point must account for those highly dissimilar individuals, and thus they implicitly assign a highly dissimilar person as being the lower anchor point. The presence of duds thus creates a different subjective similarity scale. If we assume then that a given non-dud is 4 'units' of similarity away from a moderately dissimilar person, and 5 'units' of similarity away from a highly dissimilar person, then that non-dud will receive a similarity score of '5' from the non-dud lineup (because the lowest anchor point is a moderately dissimilar person), whereas the same non-dud will receive a similarity score of '6' when the duds are present (because the lowest anchor point is a highly dissimilar person). To the extent that confidence judgments are based, at least in part, on similarity judgments, duds will affect confidence similarly.

Both the perceptual contrast explanation and the scaling account make similar predictions regarding the effect of the addition of duds to a lineup on perceived similarity and confidence. Other data, however, might be able to distinguish between these two accounts. Take the effect of duds on choosing. One the one hand, a scaling effect, because it is purely an artifact of the subjectivity of similarity scales and not a psychological or perceptual effect, has difficulty accounting for the increase in choosing as seen in Study 2,

whereas the perceptual contrast hypothesis can easily explain it (i.e., inflated perceived similarity increases the likelihood that the lineup member will surpass a decision criterion). On the other hand, this effect failed to replicate in Studies 3 and 4. If duds are eventually shown to reliably increase choosing, that would seemingly provide strong evidence against the scaling explanation as the driving force behind the dud effect. Until that time, the scaling explanation seems viable.

There could be another way to distinguish between these two explanations. Note that the scaling account would produce the dud effect without changing witnesses' phenomenological experience of the perceptual similarity of the non-duds to the criminal. The difference between the scaling effect explanation and the perceptual contrast explanation then is whether the duds produce an actual change in perceptions of similarity (as predicted by the perceptual contrast account), or whether they leave the perceptions intact, but simply change the underlying scale on which similarity is measured. One way to differentiate between these two hypotheses is thus to use scales with more objective anchor points. Imagine that instead of allowing the witness to self-generate the anchor points, they are instead provided to the witness. The perceptual contrast explanation predicts that the presence of duds will still produce a dud effect, whereas the scaling effect predicts that forcing these anchors will eliminate the dud effect (since the scale is, in effect, locked in place).

Anomalous Data

In addition to the predicted dud effects, there were also some interesting secondary effects observed in the data. First, as mentioned in the discussion for Study 3, the effect of duds on confidence was dissociated from the effects of duds on other testimony-relevant measures. This is somewhat unusual in the eyewitness field; other manipulations (such as post-identification feedback; Wells & Bradfield, 1998) tend to affect all of these measures (although see Wells & Bradfield, 1999, for an exception of a prior thought manipulation that affected confidence but none of the other measures). This dissociation has implications for our understanding of how witnesses form their post-identification judgments (as discussed in the discussion section for Study 3), and thus warrants further research.

Second, one unanticipated but highly interesting finding was that biased instructions on their own increased witness confidence in a mistaken identification (Study 3). Despite decades of research on the effects of biased instructions on witnesses' tendency to make an identification, few other studies have examined the effect of biased instructions on confidence in a way that avoids confounding differences in instructions with differences in choosing rates. If a reliable effect, this finding suggests that biased instructions may have a doubly biasing effect – not only do they increase the rate of mistaken identifications (Malpass & Devine, 1981; Steblay, 1997), but they may also increase the confidence with which those mistaken identifications are made. Clearly, this effect warrants further research.

Third, Study 2 (but not Studies 3 or 4) showed an increase in choosing rates among witnesses shown a dud lineup. At this time, the reliability of these effects is unclear. Certainly it is normal to observe some jumpiness in the data when examining a new effect, as we do not know the exact conditions that promote or inhibit the effects. Future research should help to determine the reliability of these effects, and the conditions under which they will be observed. For example, an inflation in perceived similarity may result in increased choosing when witnesses make absolute judgments (because a given lineup member is more likely to surpass a witness's decision criterion), but not when witnesses make relative judgments (because the difference between the two non-duds increases equally, thus maintaining the magnitude of the difference between the best and next-best lineup choices; some conceptualizations of relative judgments postulate that the magnitude of this difference may have to surpass a criterion before a lineup member is identified; Clark, 2003). If, as suggested by some researchers, the tendency to use an absolute or a relative judgment strategy in turn varies as a function of memory quality (e.g., Brewer, Gordon, & Bond, 2000; Charman & Wells, 2007), then the effects of duds on choosing rates may likewise depend on such factors.

Fourth, witnesses who decided not to make an identification in Study 2 were more confident in their lineup rejection if they had viewed a dud lineup than if they had viewed a non-dud lineup. This anomalous datum is difficult to reconcile with our explanation for the dud effect: If the presence of duds increases the perceived similarity of the non-duds to the criminal, then presumably non-choosers should be *less* confident in their lineup rejections. Although there are numerous post-hoc explanations for this anomaly (e.g., perhaps witnesses who decide not to make an identification differ in certain relevant respects compared to witnesses who decide to make an identification), because no other studies collected non-choosers' confidence in their rejections, the reliability of this finding is unknown, and we prefer to approach this finding cautiously.

Practical Implications

The observed findings have important implications with respect to the creation of lineups as well as the post-hoc interpretation of witnesses' lineup decisions. For instance, one of the most common and obvious recommendations to decrease the likelihood of a false identification is to include a minimum number of fillers in the lineup (often five: e.g., Technical Working Group, 1999). Although this is undoubtedly a good recommendation in the majority of cases, its mandate could, under specific circumstances, backfire. A difficulty in meeting any arbitrary minimum number of similar-looking lineup members may lead police to insert dissimilar-looking people into a lineup, potentially putting the innocent suspect at *greater* risk than if the recommendation had not been made in the first place! And although Study 4 indicated that this backfire effect may tend to occur primarily when the additional lineup members are extremely dissimilar to the criminal, some eyewitness researchers (including the second author of this paper) have seen multiple examples of egregious lineups. To be clear, we are certainly not recommending that police officers NOT attempt to meet a standard for the minimum number of lineup fillers; rather, this potential backfire effect simply underscores the importance of choosing those fillers wisely.

The current results also have important implications for how eyewitness experts evaluate the quality of a lineup post-hoc. One method of evaluating the fairness of a lineup is to administer the lineup to a series of mock witnesses—people who have never seen the perpetrator—along with a description of the perpetrator, and to force these mock witnesses to attempt an identification of the suspect (Wells, Leippe, et al., 1979). A fair lineup is considered one in which the suspect is not chosen any more frequently than chance (e.g., roughly 16.7% of the time from a six-person lineup). If a suspect is chosen more often than by chance, the logic goes, then there must be something *other than one's memory for the criminal* that led mock witnesses to identify that person (which would usually be some sort of bias within the construction of the lineup itself). For example, if a suspect is identified from a six-person lineup by 50% of mock witnesses, it is considered to be biased against the suspect.

One can calculate the functional size of the lineup from this measure as being the reciprocal of the proportion of mock witnesses who identify the suspect. In our hypothetical case, the functional size of the lineup is thus $2/1 = 2$, which means that that particular six-person lineup is functionally equivalent to a fair two-person lineup. This means that a two-person lineup composed to two equally similar lineup members is functionally equivalent to a six-person lineup composed of those same two lineup members alongside four duds who are so dissimilar to the perpetrator that they do not even get identified by mock witnesses.

But herein lies the problem. Because although a functional size analysis indicates that these two hypothetical lineups are functionally equivalent, the current results empirically demonstrate that this is in fact not the case. The dud-present, six-person lineup is actually more biased than the dud-absent two-person lineup, because the presence of duds inflates the confidence with which witnesses identify an innocent person. Other measures of lineup size—such as effective size (Malpass, 1981)—also do not pick up this type of bias. In other words, the most commonly used methods of ascertaining lineup bias post-hoc are blind to certain types of bias. As a consequence, to the extent that the dud effect occurs in real-world lineups, post-hoc assessments of lineup bias should actually tend to *underestimate* the amount of bias actually present within a lineup. Eyewitness researchers who testify in court about the bias within a given lineup should be cognizant of this possibility.

Theoretical Implications

Presumably, a witness's decision to identify a specific lineup member and provide a corresponding confidence judgment are based on a number of factors, one of the most important of which is likely the perceived similarity between the lineup member and the witness's memory of the criminal. And yet despite the importance that this similarity likely plays in driving witnesses' decision processes, very little research has examined factors that may influence it (but see Charman et al., 2009, for an example of the malleability of similarity judgments among mock detectives and mock jurors). The current results suggest that perceived similarity may be an important mediator of the witnesses' decisions, and that consequently, variables that alter similarity in turn alter these decisions. In fact, some theoretical models of eyewitness behavior already incorporate a role for this similarity (e.g., the WITNESS model; Clark, 2003). It is unfortunate then that there has been little research examining factors that influence perceived similarity. Certainly, it would benefit the field to discover such factors.

Predictions about the existence of the dud effect originated by extending findings from the basic decision-making literature to a lineup context. This general strategy of incorporating existing eyewitness phenomena into pre-existing theoretical frameworks, often underutilized, should benefit the field by leading to deeper insights about the cognitive processes of witnesses as well as leading to the discovery of other variables that may influence their

decisions. After all, it is not as if witnesses rely on a different subset of cognitive processes than non-witnesses; humans certainly did not evolve modular brain areas to perform lineup identifications! And yet our theories of eyewitness behavior—when they even exist—are often divorced from the general decision-making literature. Instead, our theories of eyewitness behavior are often created de novo to explain findings specific to the eyewitness area. Although creating a theory specifically to address a particular phenomenon has its place and is not necessarily problematic, its overuse does have potential drawbacks, including (1) potentially insulating the field, (2) mistakenly implying that witness behavior is fundamentally different from other general behaviors, and (3) rendering us blind to procedural improvements and problems that are hinted at by general psychological principles. One way to avoid these potential problems is to develop theories for our specific sub-area by co-opting basic cognitive and social theories and then adapting them to the specific problem at hand. Although a few researchers have attempted to adapt general psychological theories to explain more specific eyewitness phenomena (e.g., instance theory: Charman & Wells, 2007; signal detection theory: Meissner, Tredoux, Parker, & MacLin, 2005), more bridges between eyewitness phenomena and basic decision-making processes would undoubtedly be beneficial.

Recommendations

The most obvious recommendation that falls out of this research is that lineup-constructors should not add highly dissimilar people to their lineup. Of course, given that the lineup-constructor does not know who the criminal is (indeed, that is the very purpose of the lineup), this may be easier said than done. Fortunately, our research suggests that the dud effect may be limited to lineups that contain highly dissimilar duds (indeed, in Study 4 we only found it when the criminal was a white male and the duds were composed of two Black men and an Asian man!). Presumably then, matching the fillers to a description of the criminal given by the witness should be sufficient to avoid the effect. Then again, Studies 1 and 2 found the effect with less egregious (although still heavily biased!) lineups. Obviously more research is needed to determine the conditions that promote or inhibit the dud effect; suffice to say, the addition of dissimilar fillers should be avoided.

More generally, this research undermines the implicit assumption that the addition of dissimilar lineup members cannot harm the lineup. It is likely this underlying assumption that has led police to sometimes throw highly dissimilar fillers into a lineup, under the belief that even if they were not beneficial to the lineup, at least they could not harm it. As the present research demonstrates, highly dissimilar fillers can harm the lineup. Not only does this have implications for how lineups should be constructed, but it also means that researchers providing expert testimony should be aware of this bias when assessing the quality of the lineup post-hoc. The presence of duds should lead an expert to be suspicious not only of any identification from that lineup, but also of the witness's confidence statement following that identification.
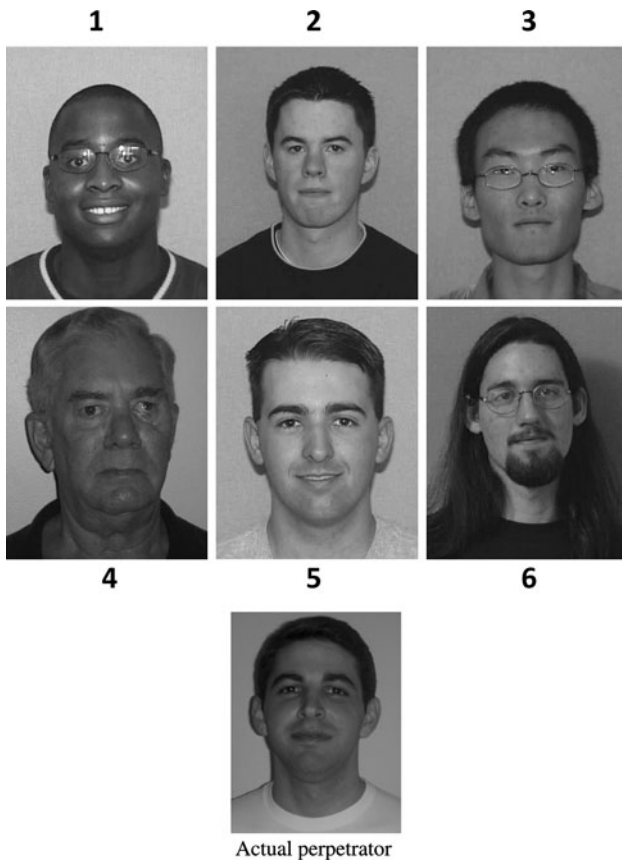
Appendix A

Dud lineup and actual perpetrator from Experiments 1 and 2. Note that the non-dud version of this lineup consisted of only lineup members 2 and 5.
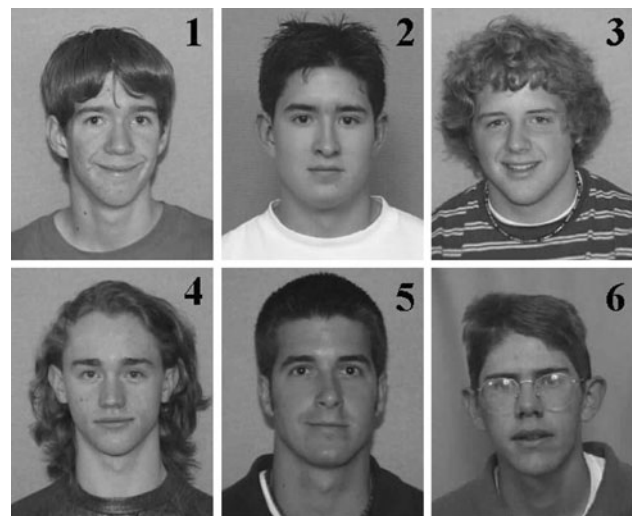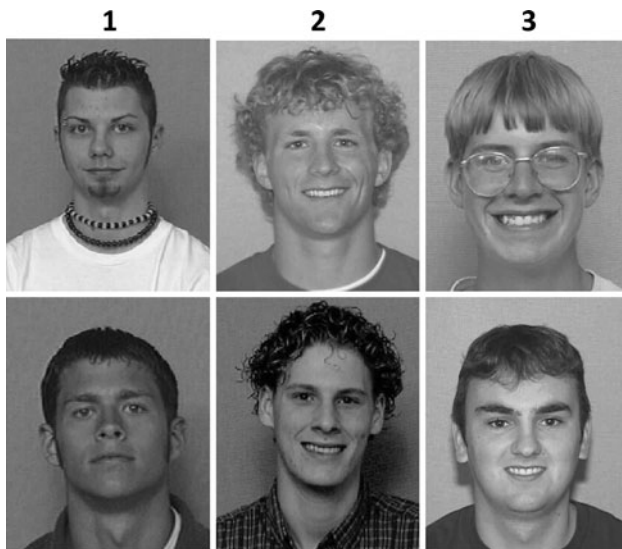

Actual perpetrator

Appendix B

Dud lineup and actual perpetrator from Experiment 3. Note that the non-dud lineup consisted of only lineup members 2 and 5.
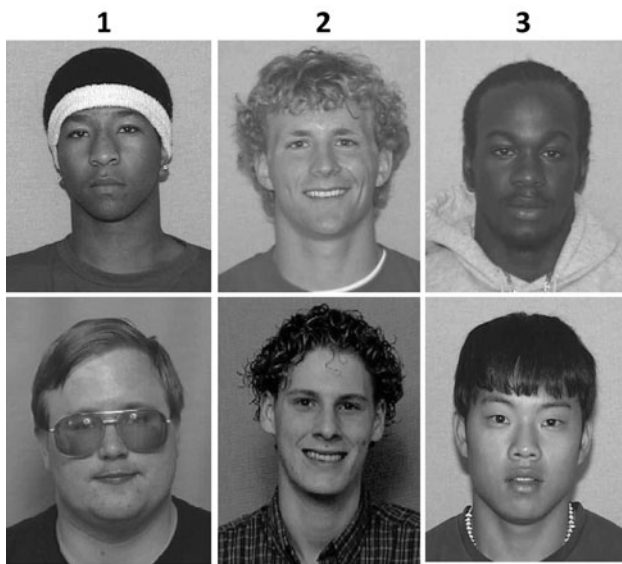

Actual perpetrator

Appendix C

Dud lineups and actual perpetrator from Experiment 4. Note that the non-dud version of the lineups consisted of lineup members 2 and 5.

Low dissimilarity lineup

Moderate dissimilarity lineup

Moderately high dissimilarity lineup

High dissimilarity lineup

Actual perpetrator

References

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182. doi:10.1037/0022-3514.51.6.1173.

Bem, D. J. (1967). Self-perception: An alternative interpretation of the cognitive dissonance phenomena. *Psychological Review, 74*, 183–200. doi:10.1037/h0024835.

Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. New York: Wiley. doi:10.1037/11622-000.

Brewer, N., Gordon, M., & Bond, N. (2000). Effect of photoarray exposure duration on eyewitness identification accuracy and processing strategy. *Psychology, Crime, & Law, 6*, 21–32. doi:10.1080/10683160008410829.

Charman, S. D., Carlucci, M., Vallano, J., & Hyman Gregory, A. (2010). The selective cue integration framework: A theory of postidentification witness confidence assessment. *Journal of Experimental Psychology: Applied, 16*, 204–218. doi:10.1037/a0019495.

Charman, S. D., Hyman Gregory, A., & Carlucci, M. (2009). Exploring the diagnostic utility of facial composites: Beliefs of guilt can bias perceived similarity between composite and suspect. *Journal of Experimental Psychology: Applied, 15*, 76–90. doi:10.1037/a0014682.

Charman, S. D., & Wells, G. L. (2007). Applied lineup theory. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.) *Handbook of eyewitness psychology: Memory for people* (pp. 219–254). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved from www.csa.com.

Clark, S. E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology, 17*, 629–654. doi:10.1002/acp.891.

Cutler, B., & Penrod, S. (1989). Forensically relevant moderators of the relation between eyewitness identification accuracy and confidence. *Journal of Applied Psychology, 74*, 650–652. doi:10.1037/0021-9010.74.4.650.

Cutler, B., Penrod, S., & Dexter, H. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior, 14*, 185–191. doi:10.1007/BF01062972.

Cutler, B., Penrod, S., & Stuve, T. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behavior, 12*, 41–55. doi:10.1007/BF01064273.

Douglass, A. B., & Steblay, N. (2006). Memory distortion in eyewitnesses: A meta-analysis of the post-identification feedback effect. *Applied Cognitive Psychology, 20*, 859–869. doi:10.1002/acp.1237.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58*, 203–210. doi:10.1037/h0041593.

Harmon-Jones, E., & Harmon-Jones, C. (2002). Testing the action-based model of cognitive dissonance: The effect of action orientation on postdecisional attitudes. *Personality and Social Psychology Bulletin, 28*, 711–723. doi:10.1177/0146167202289001.

Liberman, N., & Förster, J. (2006). Inferences from decision difficulty. *Journal of Experimental Social Psychology, 42*, 290–301. doi:10.1016/j.jesp.2005.04.007.

Lindsay, R. C. L., & Wells, G. L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior, 4*, 303–313. doi:10.1007/BF01040622.

Lindsay, R. C. L., Wells, G. L., & O'Connor, F. (1989). Mock juror belief of accurate and inaccurate eyewitnesses: A replication. *Law and Human Behavior, 13*, 333–340. doi:10.1007/BF01067033.

Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behavior, 5*, 299–309. doi:10.1007/BF01044945.

Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*, 482–489. doi:10.1037/0021-9010.66.4.482.

McQuiston-Surrett, D., Douglass, A., & Burkhardt, S. G. (2008). Evaluation of facial composite evidence depends on the presence of other case factors. *Legal and Criminological Psychology, 13*, 279–298. doi:10.1348/135532507X214192.

Meissner, C. A., Tredoux, C. J., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups. *Memory & Cognition, 33*, 783–792.

Neuschatz, J. S., Preston, E. L., Burkett, A. D., Toglia, M. R., Lampinen, J. M., Neuschatz, J. S., …Goodsell, C. A. (2005). The effects of post-identification feedback and age on retrospective eyewitness memory. *Applied Cognitive Psychology, 19*, 435–453. doi:10.1002/acp.1084.

Shaw, J. S., III (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied, 2*, 126–146. doi:10.1037/1076-898X.2.2.126.

Shaw, J. S., III, Appio, L. M., Zerr, T. K., & Pontoski, K. E. (2007). Public eyewitness confidence can be influenced by the presence of other witnesses. *Law and Human Behavior, 31*, 629–652. doi:10.1007/s10979-006-9080-6.

Shaw, J. S., III, & McClure, K. A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behavior, 20*, 629–654. doi:10.1007/BF01499235.

Skagerberg, E. M. (2007). Co-witness feedback in line-ups. *Applied Cognitive Psychology, 21*, 489–497. doi:10.1002/acp.1285.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodology 1982* (pp. 290–312). Washington, DC: American Sociological Association.

Sporer, S., Penrod, S., Read, D., & Cutler, B. L. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118*, 315–327. doi:10.1037/0033-2909.118.3.315.

Steblay, N. M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior, 21*, 283–298. doi:10.1023/A:1024890732059.

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement*. Washington, DC: United States Department of Justice, Office of Justice Programs.

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review, 101*, 547–567. doi:10.1037/0033-295X.101.4.547.

Wells, G. L., & Bradfield, A. L. (1998). "Good, you identified the suspect:" Feedback to eyewitnesses distorts their reports of the witnessing experience. *Journal of Applied Psychology, 83*, 360–376. doi:10.1037/0021-9010.83.3.360.

Wells, G. L., & Bradfield, A. L. (1999). Distortions in eyewitnesses' recollections: Can the postidentification feedback effect be moderated? *Psychological Science, 10*, 138–144. doi:10.1111/1467-9280.00121.

Wells, G. L., Ferguson, T. J., & Lindsay, R. C. L. (1981). The tractability of eyewitness confidence and its implication for triers of fact. *Journal of Applied Psychology, 66*, 688–696. doi:10.1037/0021-9010.66.6.688.

Wells, G. L., Leippe, M. R., & Ostrom, T. M. (1979). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior, 3*, 285–293. doi:10.1007/BF01039807.

Wells, G. L., Lindsay, R. C. L., & Ferguson, T. J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology, 64*, 440–448. doi:10.1037/0021-9010.64.4.440.

Wells, G. L., & Murray, D. (1984). Eyewitness confidence. In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives*. New York: Cambridge University Press.

Wells, G. L., Olson, E., & Charman, S. (2003). Distorted retrospective eyewitness reports as functions of feedback and delay. *Journal of Experimental Psychology: Applied, 9*, 42–52. doi:10.1037/1076-898X.9.1.42.

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). On the selection of distractors for eyewitness lineups. *Journal of Applied Psychology, 78*, 835–844. doi:10.1037/0021-9010.78.5.835.

Whitley, B. E., Jr., & Greenberg, M. S. (1986). The role of eyewitness confidence in juror perceptions of credibility. *Journal of Applied Social Psychology, 16*, 387–409. doi:10.1111/j.1559-1816.1986.tb01148.x.

Windschitl, P. D., & Chambers, J. R. (2004). The dud-alternative effect in likelihood judgment. *Journal of Experimental Psychology, 30*, 198–215. doi:10.1037/0278-7393.30.1.198.