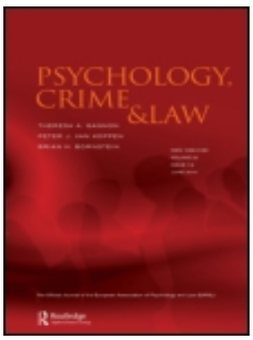

Psychology, Crime & Law

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

Prior knowledge influences interpretations of eyewitness confidence statements: 'The witness picked the suspect, they must be 100% sure%

Jesse H. Grabman & Chad S. Dodson

To cite this article: Jesse H. Grabman & Chad S. Dodson (2018): Prior knowledge influences interpretations of eyewitness confidence statements: 'The witness picked the suspect, they must be 100% sure%, Psychology, Crime & Law, DOI: 10.1080/1068316X.2018.1497167

To link to this article: https://doi.org/10.1080/1068316X.2018.1497167

Accepted author version posted online: 03 Jul 2018.
Published online: 12 Jul 2018.

PSYCHOLOGY, CRIME & LAW
https://doi.org/10.1080/1068316X.2018.1497167

Routledge
Taylor & Francis Group

Prior knowledge influences interpretations of eyewitness confidence statements: 'The witness picked the suspect, they must be 100% sure%

Jesse H. Grabman and Chad S. Dodson

Department of Psychology, University of Virginia, Charlottesville, VA, USA

ABSTRACT

When an eyewitness identifies a suspect from a lineup, it is important to know how certain they are about the decision. Even though eyewitnesses are likely to express certainty with words, past research shows that verbal confidence statements (e.g. 'I'm pretty sure') are prone to systematic misinterpretation. Until now, no one has examined how an evaluator's prior knowledge, such as which lineup member is the police suspect, influences their interpretation of eyewitness confidence about a lineup identification. Experiments 1 and 3 show that participants perceived the identical statement of confidence as meaning a higher and lower level of certainty, respectively, when the eyewitness's selection either matched or mismatched the police's suspect. Experiment 2 shows that these effects generally persist when the bias manipulation is manipulated between-subjects. Finally, Experiment 3 finds that clarifying the witness's statement with numeric information (e.g. I'm 80% sure) does not eliminate the influence of biasing information.

ARTICLE HISTORY
Received 23 January 2018
Accepted 26 June 2018

KEYWORDS
Context effects; confidence; photo lineups; eyewitness identification

Introduction

Imagine you are a police officer. After a few months investigating a string of recent robberies, you finally have a suspect. Following the recommended 'double-blind' procedure, another officer unaware of the identity of the suspect interviews an eyewitness to one of the crimes. The colleague administers a lineup consisting of the suspect and five additional individuals known to have not committed the crime. The eyewitness points to the suspect, and says, 'I'm very sure it's him'. How certain is the eyewitness? Would your answer change if you were the officer who did not know which lineup member was the suspect?

These are increasingly important questions in the criminal justice system. Unlike most laboratory studies, where confidence is collected on a numeric scale (e.g. 0–100% certain), police are advised by the United States Department of Justice to 'ask the witness to state, in his or her own words, how confident he or she is in the identification' (Yates, 2017). Evidence suggests that people generally report confidence verbally, rather than numerically (Brun & Teigen, 1988; Budescu & Karelitz, 2003; Erev & Cohen, 1990; Wallsten, Budescu,

CONTACT Jesse H. Grabman jhg5dq@virginia.edu Department of Psychology, University of Virginia, PO Box 400400, Charlottesville, VA 22904-4400, USA

Zwick, & Kemp, 1993). For example, 20 out of 23 mock witnesses in a lineup identification task provided verbal (rather than numeric) expressions of certainty when given instructions closely mirroring Department of Justice guidelines (i.e. 'in your own words, please explain how certain you are in your response') (Dodson & Dobolyi, 2015). Using identical prompts, the same authors recently found that 275 out of 342 participants expressed confidence verbally for every response to twelve mock lineups (Dobolyi & Dodson, in press).

The critical question is whether law enforcement professionals accurately interpret the intended meaning of verbal expressions of confidence. Unfortunately, much research outside the eyewitness domain indicates that the answer is no (e.g. Budescu & Wallsten, 1985; Budescu, Broomell, & Por, 2009; Budescu, Por, Broomell, & Smithson, 2014; Gurmankin, Baron, & Armstrong, 2004; Mullet & Rivet, 1991; Reagan, Mosteller, & Youtz, 1989). Further, expertise does not eliminate interpersonal variability in the interpretation of verbal expressions of certainty, even when verbal statements are drawn from a content area where raters have considerable familiarity (Beyth-Marom, 1982; Bryant & Norman, 1980; Nakao & Axelrod, 1983; Wallsten, Fillenbaum, & Cox, 1986). In one illustration of this difficulty, Nakao and Axelrod (1983) found that physicians showed just as much variability as non-physicians in their assessment of the intended probability value of 17 out of 22 verbal modifiers (e.g. 'Invariably') that are commonly used to express frequencies of medical events.

Few studies have examined whether eyewitness expressions of confidence are similarly likely to be misinterpreted (Cash & Lane, 2017; Dodson & Dobolyi, 2015, 2017). There is urgent need for examining this issue, given that an eyewitness's confidence guides many criminal justice decisions. For example, research indicates that the level of an eyewitness's confidence is likely the most important influence on jury decision-making (e.g. Bradfield & Wells, 2000). Moreover, in a synthesis of over 30 years of research, Wixted and Wells (2017) note a strong positive relationship between eyewitness accuracy and confidence at the time of an initial identification, assuming that investigators follow 'pristine eyewitness identification procedures' (e.g. one suspect per lineup; see Wixted & Wells, 2017 for a full review). However, one general assumption underlying this relationship is that the witness's confidence is properly understood by evaluators, which may not always be the case.

One major finding is that confidence interpretations become less consistent as contextual information is introduced (e.g. Brun & Teigen, 1988). This likely occurs because interpreters have varied conceptions of the utility and prior likelihood of the events the confidence phrases are modifying (Beyth-Marom, 1982). Recent work by Dodson and Dobolyi (2015) suggests that perceptions of certainty are influenced by how witnesses justify their selections from lineup. In this study, participants viewed twelve mock eyewitnesses' identifications of one of the six members of the lineup, and the eyewitness's statement of confidence about his/her identification (e.g. 'I am very certain. I remember his hair'). The critical manipulation was that some participants saw the confidence statement only ('I am very certain') whereas other participants saw the confidence statement with either a featural justification – one that referred to a visible feature about the suspect (e.g. 'I am very certain. I remember his hair.') or an unobservable justification – one that referred to a quality about the suspect that is unobservable to anyone but the eyewitness (e.g. 'I am very certain. He looks like a friend of mine.'). When asked to translate the verbal

confidence statements into a numeric value of certainty, participants rated confidence only statements similarly to those where the eyewitness provided an unobservable justification. However, featural statements were interpreted as meaning significantly lower values than both other conditions, especially when the eyewitness used language associated with high certainty (see Cash & Lane, 2017 for a replication and extension of this effect).

An open question is whether evaluators' judgements of confidence are also impacted by contextual factors unstated by (or even unknown to) the witness, such as whether the suspect confessed to the crime. These contextual influences are (to our knowledge) unexplored in evaluations of confidence statements, but do appear in other criminal justice research. For instance, Kassin, Dror, and Kukucka (2013) reviewed a variety of judgements in different forensic procedures that are affected by contextual knowledge. Polygraph examiners are more likely to interpret an interviewee's polygraph chart as deceptive when they are told that the interviewee later confessed to the crime than when told that someone else confessed (e.g. Elaad, Ginton, & Ben-Shakhar, 1994). Expertise likely does not mitigate the influence of external knowledge as multiple studies show the effects in both experienced polygraph examiners (Elaad et al., 1994), and fingerprint examiners (e.g. Dror & Charlton, 2006; see also Dror, Charlton, & Péron, 2006). One final key point that Kassin et al. (2013) emphasize is that contextual knowledge is more likely to influence judgments when the evidence is ambiguous rather than clear-cut.

The 'post-identification feedback' literature suggests that investigators' contextual knowledge can directly alter the witness's actual level of confidence (see Steblay, Wells, & Douglass, 2014 for a recent meta-analysis). By their nature, lineups contain an individual that the police have identified as their suspect. Wells and Bradfield (1998) demonstrated that investigators can inflate eyewitnesses' confidence reports by validating their selection from a lineup (e.g. 'Good, you identified the suspect'), or (to a lesser extent) deflate certainty by invalidating the selection (e.g. 'Actually the suspect was [somebody else]'). Of particular note, after receiving confirmatory feedback, roughly half of the witnesses who chose an incorrect face from the lineup reported high levels of confidence (6 or 7 on a 7-point scale) (Wells & Bradfield, 1998). This could lead observers to erroneously place high credibility on these witnesses' testimony (Steblay et al., 2014; Wells & Bradfield, 1998). Concerns about investigator influence on witness's lineup decisions and confidence reports lead these authors (and others) to emphasize double-blind lineup administration procedures (e.g. Dror, Kukucka, Kassin, & Zapf, 2018; Kovera & Evelo, 2017; Steblay et al., 2014).

In addition to influencing eyewitness reports, prior knowledge of the police suspect could also affect evaluators' perceptions of eyewitness certainty. In fact, in explaining the post-identification feedback effect, Wells and Bradfield (1999) speculate that witnesses adjust confidence reports based on how a hypothetical outside observer would perceive the statement. A confidence expression, such as 'I'm pretty sure it's him,' may be interpreted as conveying a higher level of confidence when the statement refers to an individual who corresponds to the police's suspect than when it does not. This is important to document, because ultimately evaluators of these statements must make decisions about how to proceed in legal settings (e.g. police officers, judges, jurors).

The purpose of this study is to investigate whether prior knowledge of the police suspect impacts perceptions of eyewitness confidence. The results bear on current recommendations of best-practices for eliciting eyewitness confidence statements. If contextual information leads evaluators to misinterpret witnesses' intended level of certainty, current standards for 'pristine lineup conditions' (including double-blind administration) (Wixted & Wells, 2017) do little to mitigate misinterpretations, because most legal decisions (e.g. arrests, evidence review, indictment) are made with knowledge of the police suspect. These mistakes could lead evaluators to put undue emphasis on the testimony of witnesses who identify the police suspect, potentially increasing false convictions. Further, Wells, Yang, and Smalarz (2015) note that picking someone other than the suspect (i.e. 'fillers') with high confidence should be considered strong exculpatory evidence, meaning that downplaying the importance of disconfirmations could decrease exoneration rates.

Finally, we investigate whether knowledge of the police suspect moderates other contextual effects, such as the way eyewitnesses justify their selections from a lineup. Dobolyi and Dodson (in press) recently found that participants who picked a face from a lineup provided a featural justification for nearly half of their verbal confidence statements. Given the ubiquity of these expressions, manipulating an eyewitness's justification for a level of confidence may provide a more complete picture of the factors that influence interpretations of eyewitness confidence statements.

Experiment 1

Experiment 1 manipulates prior knowledge of the suspect within-subjects, and justification condition between-subjects. In some conditions, participants will know which lineup member is the police's suspect whereas in other conditions they will not have such knowledge. Given previous research, we expect that identical eyewitness statements of confidence will be evaluated as meaning higher or lower levels of confidence when the eyewitness has chosen someone from a lineup that either matches or mismatches, respectively, the police's suspect. Further, we hypothesize that this knowledge will moderate effects of eyewitness justification. Specifically, we predict that (a) *identifying the suspect will diminish the featural justification effect*, as this evidence supports that the feature the witness chose is diagnostic, and (b) *filler identifications will enhance (or have minimal influence on)* the featural justification effect, because this bolsters perceptions of low featural discriminability.

Method

Participants

We analyzed the data of 181 participants, located in the United States, who completed the experiment over the internet using Amazon's Mechanical Turk (mean age = 37.07, SD = 10.92, range = 20–74, 58.01% female, 84.53% White/Caucasian), and were randomly assigned to one of two justification conditions: (a) Confidence only ($n = 93$) and (b) Featural Justification ($n = 88$). The sample size was sufficient to detect moderate-sized effects with greater than 95% power at an alpha level of .05, according to G*POWER (Faul, Erdfelder, Lang, & Buchner, 2007).

Materials

All participants interpreted the verbal confidence of mock eyewitnesses responding to six 'fair' lineups (see Dobolyi & Dodson, 2013 for methodology of lineup creation). Fair lineups meet the conditions that no face is comparatively noticeable, and that individuals naïve to the target stimulus are equally likely to select each face (e.g. Gronlund, Carlson, Dailey, & Goodsell, 2009; Malpass, Tredoux, & McQuiston-Surrett, 2007). As shown in Figure 1, lineups consisted of the head and shoulders of six white males arranged in a 2×3 grid, exhibiting a neutral expression, and wearing a solid maroon colored t-shirt. In each case, one face was randomly selected to serve as the eyewitness's identified suspect and outlined with a red border. In the within-subjects prior knowledge manipulation, the word 'suspect' in white text represented the police suspect and was superimposed over either the eyewitness's selection (suspect-ID), another randomly selected face (filler-ID), or omitted (i.e. no police suspect, no-context).

Accompanying each lineup was the eyewitness's written verbal expression of confidence in their selection. Briefly, expressions of confidence were transcribed verbatim from participant responses to a previous eyewitness study, and reflected either 'moderate' (e.g. 'I am fairly confident') or 'high' (e.g. 'I am positive') levels of certainty (see Dodson & Dobolyi, 2015 for a full summary on expression generation). Additionally, statements in the 'featural' condition included justifications reflecting visible features of the suspect (e.g. 'I remember his nose') (see Table 1 for confidence statements and justification conditions).

Procedure

Figure 1 (panels a – c) shows examples of the task. At study onset, participants were randomly assigned to a single justification condition and viewed the same sequence of six lineups. We counterbalanced the order of all possible pairings of the prior knowledge (suspect-ID, filler-ID, no-context) and statement confidence (moderate, high) conditions across participants.

Given the goal of approximating how police officers interpret eyewitness statements, we started by broadly informing participants about the process police use to generate

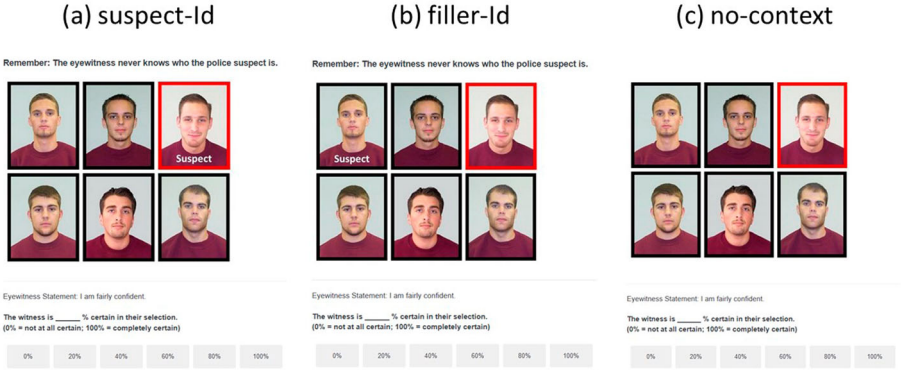

Figure 1. In Experiment 1, participants evaluated the eyewitness's intended numeric confidence expression under 2 justification conditions and when: (a) the eyewitness chose the same person as the police (suspect-Id), (b) the eyewitness chose someone other than the police suspect (filler-Id), and (c) when given no information about the suspect (no-context).

Table 1. Confidence expressions used in Experiments 1–3. Depending on condition, participants in Experiment 3 saw only Verbal expressions, or Numeric expressions. In all experiments, half of participants saw confidence statements with the addition of a featural justification.

Confidence Level	ContextCondition	Expression – Verbal	Expression – Numeric (Experiment 3 Only)	Featural Justification
Moderate	suspect-ID	I am pretty certain.	I am 60% certain.	I remember his chin.
	no-context	I am fairly confident.	I am 60% confident.	I recall his hair.
	filler-ID	I am pretty sure.	I am 60% sure.	I remember his ears.
High	suspect-ID	I am very certain.	I am 100% certain.	His mouth is memorable.
	no-context	I am very confident.	I am 100% confident.	I remember the shape of his head.
	filler-ID	I am positive.	I am 100% positive.	I recall his eyes.

criminal lineups. We told them that sometimes, when police create a lineup of faces, they include one person that they think could be the criminal. The remaining faces in the lineup are innocent individuals that the police know did not commit the crime. We emphasized that the eyewitness is never told whether his/her choice matches the police's suspect.

For suspect-known lineups, participants pretended to be the police officer who created the lineup. For no-context lineups, participants pretended to be a police officer who did not have a particular suspect in mind. In all cases, the participant's task was to translate the written statements into a numerical value of certainty, by filling in the sentence, 'the witness is ___% certain in their selection,' using a scale ranging from 0% (not at all certain) – 100% (completely certain) in 20% increments.

To ensure participants understood all the instructions, and to preview the task, we first showed examples of suspect-ID and no-context lineups composed of six colorful smiley faces. In each case the eyewitness statement read, 'I know it's him'. Participants were instructed to pretend that the example eyewitness was completely certain about their decision. Finally, upon completing the task, we checked for comprehension by asking participants, 'when the eyewitness looked at the lineup, did he/she know who the police suspect was?' and 'other than the police suspect, is it possible that anyone else in the lineup committed the crime?' We did not analyze the data of 175 individuals who responded <80% on either one of the smiley lineups ($N = 127$), and/or answered 'yes' to either of the final manipulation checks ($N = 93$).

Results and discussion

We analyzed participants' estimates of the intended meaning of an eyewitness's verbal expression of confidence with a 2 (Justification: confidence only, featural; between) × 3 (Prior Knowledge: suspect-ID, filler-ID, no-context; within) × 2 (Confidence Level: high vs. moderate; within) mixed ANOVA. Replicating past results, a main effect of Justification condition, F (1, 179) = 9.76, $MS_e = 985.18$, $p = .002$, $\eta_p^2 = .052$, reflects participants perceiving eyewitnesses as less confident when the eyewitness's confidence-statement was accompanied by a featural justification ($M = 64.58$, $SD = 14.62$) than when the eyewitness provided a confidence statement only ($M = 70.53$, $SD = 10.84$). As hypothesized, a main effect of Prior Knowledge, F (2, 358) = 45.24, $MS_e = 364.43$, $p < .001$, $\eta_p^2 = .202$, is fueled by higher perceived confidence for statements about an identification that matched the

police's suspect ($M = 75.14$, SD = 14.93) than in the no-context condition (M = 65.75, SD = 15.53; t (180) = 7.14, $p < .001$, Cohen's $d = .53$, 95% CI [.37, .68]), which in turn showed higher perceived confidence than the filler-ID condition ($M = 62.04$, $SD = 20.38$; $t(180) = 2.85$, $p = .005$, Cohen's $d = .21$, 95% CI [.06, .36]). And, as expected, there was a main effect of Confidence, F (1, 179) = 489.59, $MS_e = 254.61$, $p < .001$, $\eta_p^2 = .732$, with increased perceived confidence for statements expressing high confidence ($M = 78.38$, $SD = 15.24$) than moderate confidence ($M = 56.91$, $SD = 14.04$).

These main effects should be interpreted in the context of a significant two-way interaction between Prior Knowledge and Confidence Level, F (2, 358) = 5.29, $MS_e = 203.38$, $p = .005$, $\eta_p^2 = .029$. As Figure 2 shows, the same moderate confidence statement is interpreted as meaning a higher numeric value when the statement refers to a suspect-ID lineup ($M = 65.52$, $SD = 17.78$) than a no-context lineup ($M = 55.91$, $SD = 18.46$), which in turn is rated higher than the filler-ID lineup ($M = 49.28$, $SD = 20.71$), $t(180) = 5.61$, $p < .001$, Cohen's $d = .42$, 95% CI [.26, .57] and $t(180) = 4.17$, $p < .001$, Cohen's $d = .31$, 95% CI [.16, .46]. For high confidence statements, we also found that participants perceived greater confidence for suspect-ID lineups ($M = 84.75$, $SD = 17.18$) than no-context lineups ($M = 75.58$, $SD = 17.18$), $t(180) = 6.33$, $p < .001$, Cohen's $d = .47$, 95% CI [.31, .62]. However, in contrast to the results for moderate statements, ratings for filler-ID lineups ($M = 74.81$, $SD = 27.28$) did not significantly differ from no-context lineups, $t(180) = .41$, $p = .684$, Cohen's $d = .03$, 95% CI [-.12, .18].

Unexpectedly, we did not observe any further interactions (all $ps \geq .070$, $\eta_p^2 \leq .015$). Contrary to our hypotheses, the Prior Knowledge manipulation did not significantly impact the effects of featural justification on perceived confidence.

Taken together, the results of Experiment 1 support the conclusion that knowledge of the police suspect influences perceptions of eyewitness confidence. Despite acknowledging that the witness did not know who the police suspect was, meaning that this information should not influence certainty expressions, participants perceived identical

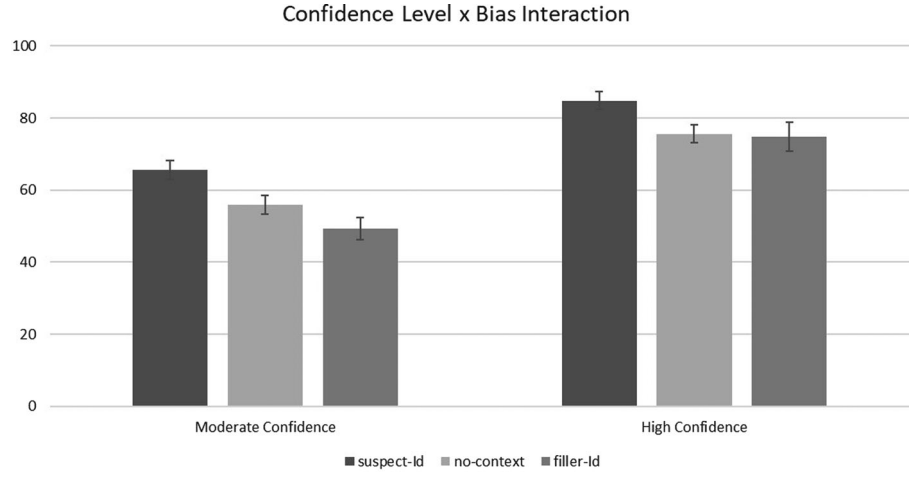

Figure 2. In Experiment 1, we found a significant two-way interaction between the Confidence and Context conditions. As seen on the graph, the effect of the Context manipulation is stronger for moderate confidence than high confidence statements. Error bars indicate 95% confidence intervals of the mean.

confidence statements as expressing lower (higher) values when the statement was about a member of the lineup that (mis-)matched the police suspect, than when the suspect was unknown. Logically this is an error.

However, a potential limitation of Experiment 1 is that Prior Knowledge is manipulated within-subjects. Participants could have altered their responses to no-context, suspect-ID, and filler-ID lineups, based on their expectation that there should be a difference between the three conditions. In Experiment 2, we address this concern by manipulating knowledge of the police suspect in a between-subjects manner.

Experiment 2

Experiment 2 directly replicates the method of Experiment 1, though this time prior knowledge is manipulated between-subjects. If task demand is the sole factor accounting for discrepancies between the suspect-known and no-context lineups, then in this experiment we would expect minimal perceived differences in confidence statement ratings between each of the conditions.

Method

Participants

All 476 participants who passed study manipulation checks were located in the United States and completed the experiment over the internet using Amazon mechanical Turk (mean age = 36.36 years, SD = 11.59, range = 18–75, 58.61% female, 81.93% White/Caucasian). Participants were randomly assigned to one of six conditions, each comprised of 78–81 participants, representing the 3×2 intersection between the Prior Knowledge (suspect-ID, filler-ID, no-context) and Justification (Statement Only, Featural) manipulations.

Procedure

All participants viewed the same order of two lineups, randomly chosen from the pool of six lineups in Experiment 1. We arbitrarily selected a high confidence and moderate confidence eyewitness statement to accompany each lineup, counterbalancing the order of presentation across participants. One face in each lineup was bordered in red, representing the eyewitness's selection, with both lineups corresponding to the same Prior Knowledge condition (either suspect-ID, filler-ID, or no-context). Participants in the suspect-ID and filler-ID conditions read the same instructions as in Experiment 1, without the paragraphs describing that sometimes police officers do not know the police suspect, and pretended to be the officer who created the lineup. Those in the no-context condition received no information about the lineup generation process, and simply imagined themselves to be police officers. All other aspects of the design and procedure are identical to Experiment 1. We excluded 259 individuals who either assigned values <80% for the smiley lineup ($N = 208$), and/or failed one or more of the manipulation checks ($N = 87$).

Results and discussion

We examined perceived confidence using a 2×3×2 mixed ANOVA with between-subjects factors of Justification (statement only, featural) and Prior Knowledge (suspect-ID, filler-ID,

no-context), and a within-subjects factor of Confidence Level (moderate, high). In line with Experiment 1, we observed a main effect of Justification, F (1,470) = 14.66, MS_e = 694.04, p < .001, η_p^2 = .030, with lower levels of perceived confidence when participants encountered confidence statements with an accompanying featural justification (M = 61.46, SD = 20.37) than a confidence statement only (M = 68.06, SD = 19.49). Unsurprisingly, we found an effect of Confidence Level, F (1, 470) = 485.11, MS_e = 206.01, p < .001, η_p^2 = .508, such that participants interpreted high confidence statements (M = 75.00, SD = 22.65) as meaning a higher value than moderate confidence statements (M = 54.50, SD = 21.70).

Interestingly, we observed a main effect of Prior Knowledge, F (2, 470) = 23.34, MS_e = 694.04, p < .001, η_p^2 = .090. As in Experiment 1, participants perceived the identical confidence statements as meaning a lower numeric value in the filler-ID condition (M = 56.52, SD = 21.20) than in either the no-context or suspect-ID conditions, t (294.82) = 5.41, p < .001, Cohen's d = .61, 95% CI [.38, .83] and t(309.56) = 5.84, p < .001, Cohen's d = .66, 95% CI [.43, .88], respectively. However, perceived confidence was comparable in the suspect-ID (M = 69.68, SD = 18.80) and no-context conditions (M = 68.00, SD 16.32), t (316) = .85, p = .394, Cohen's d = .10, 95% CI [-.12, .32].

We did not observe a three-way interaction, nor any significant two-way interactions with the Justification condition (all F's < 1.53, all $ps \geq .218$, all $\eta_p^2 \leq .006$). Additionally, in contrast to Experiment 1, we did not find a significant interaction between Confidence Level and Prior Knowledge, F (2, 470) = 1.09, MS_e = 206.01, p = .338, η_p^2 = .005.

Overall, we generally replicated Experiment 1, even when using a more conservative Between-subjects design. Participants in the filler-ID condition, based on their own knowledge of the police suspect, perceived the witness as having a lower level of confidence, relative to the no-context condition, even though the same confidence statement was used in both conditions. While we did not replicate the finding of amplification of perceived confidence in the suspect-ID condition, the effects are in the same direction. These results rule out the critique that the effects of context in Experiment 1 were solely caused by task demands induced by its within-subjects design. Additionally, arguably, the within-subjects design is a more ecologically valid test as police officers often administer and interpret responses to multiple lineups over the course of their careers.

Given that participants in both Experiments 1 and 2 showed susceptibility to contextual effects when they interpreted eyewitness statements, in Experiment 3 we test whether clarifying the witness's confidence level using numerical information can protect against these errors.

Experiment 3

One outstanding question in the judgement literature is the degree to which contextual effects are influenced by expressing confidence numerically instead of verbally. There are reasons to think that these effects will be weaker when witnesses clarify their level of confidence using numeric information. Research shows that context effects are most pernicious when evidence is ambiguous (Kassin et al., 2013). Multiple studies recommend that numerical probabilities are preferable to verbal probabilities because they are less vulnerable to flexible interpretations (e.g. Budescu et al., 2009, 2014; Nakao & Axelrod, 1983).

When interpreting eyewitness confidence, the ambiguity of verbal statements may lead investigators to place more weight on other accessible information, such as whether the witness selected the police suspect. In contrast, expressing eyewitness confidence numerically may serve to protect investigators from unintentionally misinterpreting the eyewitness's intended level of confidence, because there is less need to search for additional context behind the statement. For example, there should be little chance of misunderstanding the eyewitness's level of confidence when s/he expresses 100% confidence in the identification.

However, there is an alternative explanation. Eyewitnesses are viewed as inaccurate when they select someone from the lineup that is not the police's suspect, which in turn causes participants to discount and lower their perception of the eyewitness's level of certainty. This account is supported by recent evidence in studies of the justification bias. Cash and Lane (2017) found that participants perceived confidence statements with featural justifications as denoting both lower eyewitness certainty and *accuracy* than when confidence statements were presented alone. Moreover, Dodson and Dobolyi (2017) observed that even when participants were shown numeric expressions of confidence they still rated eyewitnesses as less accurate when the confidence statement was accompanied with a featural justification than when it appeared by itself.

Experiment 3 is the first test of whether expressing confidence numerically (rather than verbally) reduces the effects of prior knowledge on interpretations of eyewitness confidence statements. Moreover, as a conceptual replication of Dodson and Dobolyi (2017), we test whether these benefits extend to statements with featural justifications. If discordance is reduced in either case, then there may be enormous practical benefits to asking eyewitnesses for numeric expressions of confidence.

Method

Participants

In this experiment, 403 participants were randomly assigned to view numeric or verbal statements of certainty, with one of the two Justification conditions from Experiment 1 ($n = 101$ for all between-group combinations, excepting numeric x featural where $n = 100$). These participants, located in the United States, completed the experiment over the internet using Amazon's Mechanical Turk (mean age = 37.04, SD = 11.34, range = 18–72, 60.79% female, 84.12% Caucasian), and passed all manipulation checks. We calculated that there was > 95% power to detect moderate effects at an alpha level of .05.

Materials and procedure

Materials and procedures are similar to Experiment 1, except for two changes. First, half of the participants were presented with numeric expressions of eyewitness confidence, rather than verbal as before (see Table 1 for numeric confidence statements and justifications). The statements remained nearly identical; however instead of using an adverb of degree (e.g. 'very') or an adjective (e.g. 'positive') to indicate certainty, mock eyewitnesses reported 60% confidence in cases of moderate certainty, and 100% when highly certain. The other half of participants completed the experiment using the same confidence statements from Experiment 1. To avoid suspicion, statements in the example lineups corresponded to the participant's numeric/verbal random group assignment.

Second, we altered the response scale to avoid mirroring the numeric statements of certainty. Participants indicated perceived confidence by using a slider in a visual analog scale, ranging from a low value of 'Not at all certain' to 'Completely Certain'. The scale was subdivided into 6 equal sections to analogize the results to Experiments 1 and 2. We did not include any visual demarcation points (e.g. tick marks) to avoid the potential numeric information of counting up the elements. All lineups, Prior Knowledge conditions, and counterbalancing remain identical to Experiment 1. A total of 297 individuals did not meet criteria for analysis – either assigning values <80% for at least one smiley lineup ($N = 149$) and/or responding with 'yes' to at least one manipulation check ($N = 199$).

Results and discussion

We used a mixed ANOVA to analyze the 2×2×3×2 factors of Format (verbal vs. numeric; between), Justification (statement only, featural; between), Prior Knowledge (suspect-ID, filler-ID, no-context; within), and Confidence Level (moderate, high; within). Although we did not find a four-way interaction, $F(2,798) = .43$, $MS_e = .46$, $p = .651$, $\eta_P^2 = .001$, there are several significant interactions and main effects.

Beginning with the main effects, we replicated previous findings for the Justification and Confidence Level manipulations, $F(1,399) = 15.67$, $MS_e = 2.73$, $p < .001$, $\eta_P^2 = .038$ and $F(1, 399) = 1171.09$, $MS_e = 1.10$, $p < .001$, $\eta_P^2 = .746$, respectively. Participants perceived more certainty for confidence only expressions ($M = 4.54$, $SD = .60$) than those with an accompanying featural justification ($M = 4.27$, $SD = .76$), and assigned increased values to high confidence statements ($M = 5.14$, SD .88) compared to moderate confidence statements ($M = 3.68$, $SD = .75$). A main effect of Prior Knowledge replicated Experiment 1, $F(1.58, 629.21) = 110.96$, $MS_e = 1.25$, $p < .001$, $\eta_P^2 = .218$), with participants perceiving greater confidence for suspect-ID lineups ($M = 4.77$, $SD = .77$) than no-context lineups ($M = 4.42$, $SD = .77$), $t(402) = 8.82$, $p < .001$, Cohen's $d = .44$, 95% CI [.34, .54], which in turn showed increased perceived confidence compared to filler-ID lineups ($M = 4.03$, $SD = 1.12$), $t(402) = 8.40$, $p < .001$, Cohen's $d = .42$, 95% CI [.32, .52]. Finally, we found a main effect for Format, with numeric expressions ($M = 4.51$, $SD = .66$) showing higher perceived confidence than verbal expressions ($M = 4.30$, $SD = .71$), $F(1,399) = 9.37$, $MS_e = 2.73$, $p = .002$, $\eta_P^2 = .023$.

The primary aim of this experiment was to investigate whether numeric, as compared to verbal, expressions of confidence reduce the effects of prior knowledge on judging an eyewitness's intended confidence. We found a significant three-way interaction between Confidence Level × Statement Format × Prior Knowledge, $F(2, 798) = 13.46$, $MS_e = .46$, $p < .001$, $\eta_P^2 = .033$), which we discuss below. Replicating Dodson and Dobolyi (2017), we did not observe significant three- or two-way interactions between the Format and Justification manipulations, suggesting that participants were just as susceptible to the featural justification effect in the numeric condition as in the verbal condition.

Confidence level × statement format × prior knowledge

Figure 3 displays the 3-way interaction between Confidence Level, Statement Format, and Prior Knowledge. The 3-way interaction is due to a stronger interaction between Format

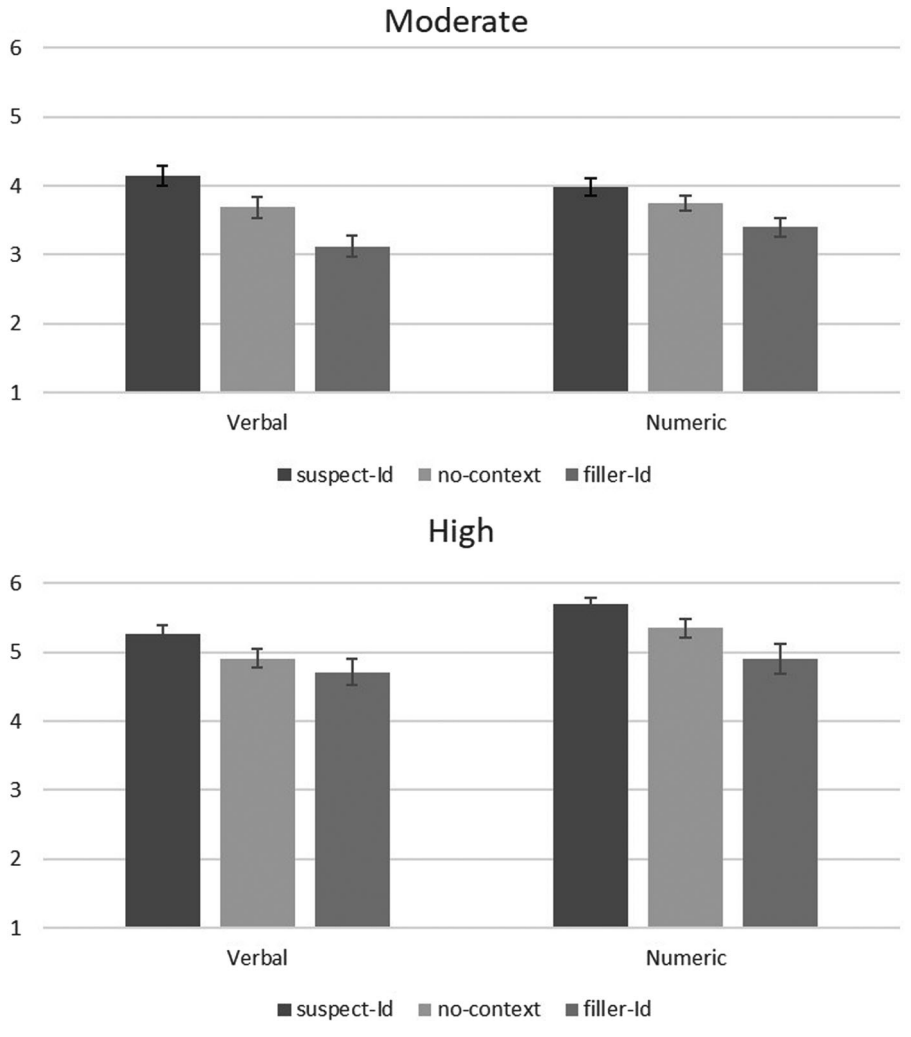

Figure 3. In Experiment 3, participants evaluated the eyewitness's intended confidence when expressed either verbally or numerically. Higher scores indicate greater perceived confidence. The figure displays the interaction between intended Confidence of the expression, Format of the expression, and the Context condition. Error bars indicate 95% confidence interval of the mean.

and Prior Knowledge for moderate confidence level statements, $F(2,798) = 10.88$, $MS_e = .46$, $p < .001$ $\eta_p^2 = .026$, compared to high confidence level statements, $F(2,798) = 4.37$, $MS_e = .46$, $p = .013$, $\eta_p^2 = .011$).

Beginning with moderate confidence statements, we found an effect of Prior Knowledge in both Format conditions. As seen in the top half of Figure 3, participants perceived both verbal and numeric expressions of confidence as meaning a higher value when the expression referred to a suspect-ID lineup (numeric $M = 3.98$, SD = .88; verbal $M = 4.14$, $SD = 1.08$) than a no-context lineup (numeric $M = 3.75$, $SD = .76$, $t(200) = 3.35$, $p = .001$, Cohen's $d = .24$, 95% CI [.10, .38]; verbal $M = 3.69$, $SD = 1.09$, $t(201) = 5.37$, $p < .001$, Cohen's $d = .38$, 95% CI [.23, .52]), which in turn were rated significantly higher than

expressions of confidence that referred to a filler-ID lineup (numeric $M = 3.40$, $SD = .98$, $t(200) = 5.18$, $p < .001$, Cohen's $d = .36$, 95% CI [.22, .51]; verbal $M = 3.12$, $SD = 1.14$, $t(201) = 6.89$, $p < .001$, Cohen's $d = .48$, 95% CI [.34, .63]). To determine if there is a meaningful attenuation of Prior Knowledge effects in the numeric compared to the verbal conditions, we computed difference scores between the no-context condition and the suspect-known conditions and compared these scores across statement Format. The comparison in perceived confidence between the suspect-ID and no-context conditions showed a smaller difference in the numeric format ($M = .23$, $SD = .97$) than in the verbal format conditions ($M = .46$, $SD = 1.21$), $t(383.92) = 2.08$, $p = .038$, Cohen's $d = .21$, 95% CI [.01, .40]. Similarly, the numeric condition ($M = .35$, $SD = .95$) produced a smaller difference in perceived confidence between the filler-ID and no-context conditions, relative to the verbal condition ($M = .57$, $SD = 1.18$), $t(385.34) = 2.08$, $p = .039$, Cohen's $d = .21$, 95% CI [.01, .40]. Overall, supporting our hypothesis, there are smaller effects of Prior Knowledge in the numeric condition than in the verbal condition when participants evaluate moderate levels of confidence.

The bottom half of Figure 3 shows the effects of Prior Knowledge for high confidence statements in both Format conditions. For both verbal and numeric statements of confidence, participants perceived the statements as meaning a higher value when they referred to suspect-ID lineups (numeric $M = 5.69$, $SD = .73$; verbal $M = 5.26$, $SD = .95$) than to no-context lineups (numeric $M = 5.35$, $SD = 1.00$, $t(200) = 4.99$, $p < .001$, Cohen's $d = .35$, 95% CI [.21, .49]; verbal $M = 4.91$, $SD = 1.02$, $t(201) = 5.39$, $p < .001$, Cohen's $d = .38$, 95% CI [.24, .52]), which exceeded ratings for filler-ID lineups (numeric $M = 4.90$, $SD = 1.53$, $t(200) = 5.04$, $p < .001$, Cohen's $d = .36$, 95% CI [.21, .50]; verbal $M = 4.71$, $SD = 1.37$, $t(201) = 2.22$, $p = .027$, Cohen's $d = .16$, 95% CI [.02, .29]). As we did for the moderate confidence statements, we computed difference scores between the no-context condition and the suspect-ID and filler-ID conditions in order to examine our hypothesis that the numeric format will lessen the effect of Prior Knowledge as compared to the verbal format. There was little effect of numeric vs. verbal format on differences in perceived confidence in the comparisons between (a) the no-context and suspect-ID conditions ($M = .34$, $SD = .96$ vs. $M = .35$, $SD = .91$, respectively), $t(401) = .088$, $p = .930$, Cohen's $d = .01$, 95% CI [-.19, .20), and (b) the no-context and filler-ID conditions ($M = .45$, $SD = 1.27$ vs. $M = .20$, $SD = 1.30$, respectively), $t(401) = -1.95$, $p = .052$, Cohen's $d = .19$, 95% CI [.00, .39]). In sum, when participants evaluate high confidence statements, there are similar effects of Prior Knowledge on perceived confidence in both the numeric and verbal conditions.

Prior knowledge × justification

Finally, there was a significant two-way interaction between the Context and Justification manipulations, $F(1.58, 629.21) = 5.00$, $MS_e = 1.25$, $p = .012$, $\eta_P^2 = .012$. As seen in Figure 4, we found near-equal featural justification effects for both suspect-ID (statement only $M = 4.95$, $SD = .59$ vs. featural $M = 4.58$, $SD = .88$) and no-context lineups (statement only $M = 4.60$, $SD = .66$ vs. featural $M = 4.25$, $SD = .84$), $t(379.66) = 4.65$, $p < .001$, Cohen's $d = .46$, 95% CI [.26, .66] and $t(349.92) = 4.87$, $p < .001$, Cohen's $d = .48$, 95% CI [.29, .68], respectively. However, deviating slightly from Experiments 1 and 2, we found no significant difference between the statement only and featural conditions for filler-ID lineups, $t(401) = .78$, $p = .436$, Cohen's $d = .08$, 95% CI [-.12, .27].

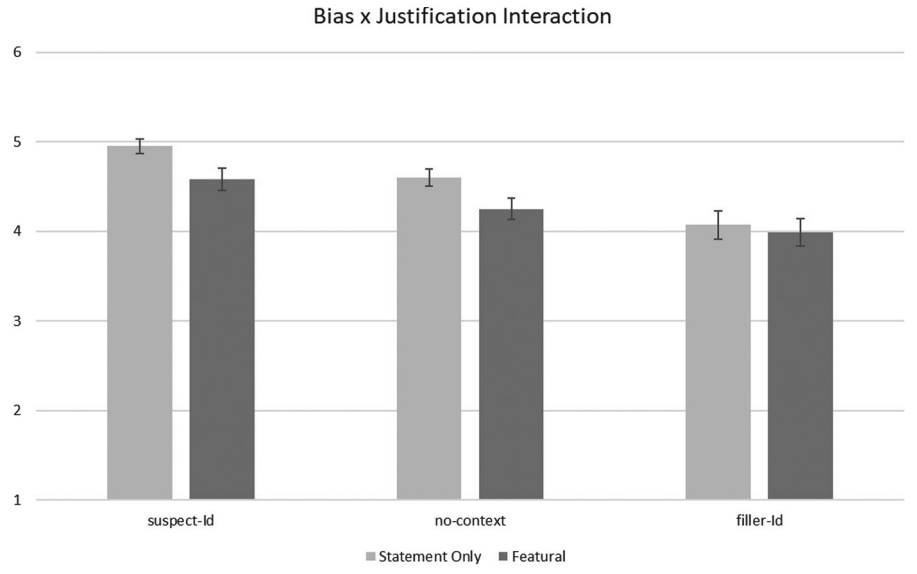

Figure 4. In Experiment 3, we found a significant interaction between the Context and Justification manipulations. For both suspect-Id and no-context lineups, participants judged the identical expression of confidence as meaning a lower value when it was accompanied by featural information than when the confidence statement was presented alone. By contrast, there are no differences between the statement only and featural conditions for filler-Id lineups. Error bars indicate 95% confidence intervals of the mean.

General discussion

For a moment, again imagine yourself as one of the investigators in the opening of this manuscript. After noting the eyewitness's confidence about who they selected from the lineup, you take the short walk to your superior's office. 'I see that the witness selected Face 1 from the lineup. How confident did he seem?' she asks. The results of three studies suggest that prior knowledge of the suspect will influence perceptions about the witness's certainty.

Across manipulations, Experiments 1 and 3 show that participants perceived the identical statement of confidence as meaning a higher level of certainty when the eyewitness's selection from the lineup matched the police suspect, relative to a no-context condition. Similarly, in all experiments (excepting 'High Confidence' statements in Experiment 1), participants perceived the same confidence statement as meaning a lower value when the eyewitness's selection differed from the police suspect, again relative to a no-context condition. These effects reflect judgment errors. Since participants acknowledged that the eyewitness is unaware of which lineup member is the police's suspect it is impossible for the eyewitness's level of certainty to be influenced by whether or not their choice happens to coincide with the police's suspect. Finally, results suggest that prior knowledge may add to pre-existing contextual effects, such as how the witness justifies their selection from a police lineup. Though, there is a notable exception in Experiment 3, wherein the featural justification effect was eliminated for filler-ID lineups.

Is it possible to attenuate the effects of prior knowledge? We investigated whether numeric expressions of confidence would be less susceptible than verbal expressions to contextual influences because of their greater interpretive clarity. Supporting this hypothesis, Experiment 3 showed that participants were less influenced by knowledge of the police suspect when the witness expressed moderate confidence statements numerically rather than verbally, demonstrating some protective effect of a numeric format. However, the differences observed between the format conditions is small (about .1 SD units), and contextual knowledge still impacted perceptions of certainty for moderate statements. Moreover, when eyewitnesses imparted high confidence, there was little benefit to expressing confidence numerically, as Experiment 3 showed comparable effects of confirmation and disconfirmation on statement evaluations for both confidence formats. This is discouraging, because high confidence statements are most likely to be used to guide legal decision making. As a whole, these results suggest that there may be some benefit to asking witnesses to clarify their confidence using a numerical indicator, though this alone is unlikely to fully cancel out contextual effects.

We view the findings of these experiments with a caveat. We did not observe a significant difference between suspect-ID and no-context lineups when the Context condition was manipulated between participants in Experiment 2, though participants continued to perceive less confidence for filler-ID lineups. We theorize that the consistent differences between the filler-Id and no-context conditions is due to a 'presumption of calibration' (Tenney, Spellman, & MacCoun, 2008), wherein participants view the eyewitness as well-calibrated (or justified) in their lineup decisions (especially if highly confident), unless given reason to think otherwise. Selection of a lineup filler is a strong signal that the eyewitness is poorly-calibrated, thus resulting in lower perceived confidence across all experiments, relative to the no-context control condition. However, task demand may explain why there are discordant findings for the suspect-Id and no-context lineup comparisons across experimental designs. In the between-subject condition, there is nothing to suggest that the witness's selection for no-context lineups differs from the police suspect, meaning that participants simply treated these lineups as if they are the same as the suspect-Id lineups. In contrast, when participants viewed all Context conditions in Experiments 1 and 3 (i.e. within-subjects design), they may have inferred that there should be a clear ordering of witness accuracy, resulting in bolstered confidence ratings for the suspect-ID condition as compared to the no-context condition.

We would like to emphasize that we believe the within-subjects design is the more valid test of real-world practice, given that many legal roles will consider evidence in the light of their experience with previous cases. However, one might argue that the effects of knowing the suspect are relatively modest, with Cohen's *d*s for statistically significant findings ranging from .16 to .61. On the other hand, there are reasons to think that contextual effects will have an even larger influence in the real-world as compared to what we observed in these experiments. Our participants viewed lineups with an arbitrarily assigned police suspect. This differs from real-world practice where police officers construct lineups with *their own* suspect in mind. Research on the confirmation bias suggests that the tendency to interpret evidence as consistent with our beliefs will be stronger in the latter case where there is (or should be) evidence supporting a particular person as the suspect as compared to the former case where the suspect was simply labeled as the police's suspect (Bastardi, Uhlmann, & Ross, 2011; Klayman & Ha, 1987; Kunda,

1990; Nickerson, 1998). A recent survey of forensic science examiners found that many professionals show a bias blind spot, acknowledging that others may be affected by confirmation bias, but seeing themselves as impervious to its effects, or able to combat them by setting aside prior beliefs and expectations (Kukucka, Kassin, Zapf, & Dror, 2017). Experience in the field did little to protect against this assumption, while training on cognitive biases helped minimally. We predict that contextual effects will be even more pronounced when participants hold strong prior beliefs about suspect guilt before evaluating eyewitness statements.

Additionally, this study placed participants in the role of a police officer, yet we believe that the observed effects could generalize to decisions in other legal roles. For example, in deciding whether to admit eyewitness testimony, judges often refer to criteria set forth by the United States Supreme Court in *Neil v. Biggers* (1972) and expanded in *Manson v. Brathwaite* (1977) (see Wells & Quinlivan, 2009 for a review). These include whether the witness (a) had ample opportunity to view the individual, (b) paid sufficient attention, (c) provided a detailed description, (d) made the identification promptly after the event, and – crucially – (e) professed certainty about the identification. In assessing a case, judges are generally aware of the suspect, so they may be more likely to ascribe reliability to a witness statement when the identification matches the suspect than when it does not. This would result in more evidence implicating the suspect making it to trial than is warranted by the witness's certainty. Future studies should address whether experienced legal decision makers are similarly prone to these contextual effects.

Finally, our results suggest that current recommendations for assessing eyewitness confidence are incomplete. Present procedures, such as double-blind lineup administration, prevent investigators from influencing eyewitness reports (and should continue to be used!). However, they do not mitigate the effects of prior knowledge on interpretations of eyewitness statements. To reiterate, it is probable that legal decision makers are aware of the suspect in advance of hearing the witness's confidence statement. Are there any ways to ameliorate contextual effects?

One potential method comes from the post-identification feedback literature, wherein witnesses are asked to think about their certainty level after making an identification, but *before* receiving dis-/confirmatory feedback (e.g. 'Good job, you identified the suspect') from an investigator (Neuschatz et al., 2007; Quinlivan et al., 2009; Wells & Bradfield, 1998, 1999). This 'confidence prophylactic' is shown to dampen the impact of feedback in the short-term (Neuschatz et al., 2007; Quinlivan et al., 2009; Wells & Bradfield, 1998, 1999), though is less effective in preventing long-term effects (Neuschatz et al., 2007; Quinlivan et al., 2009). Nevertheless, it may be beneficial to have non-blinded evaluators reflect on the confidence statement of the eyewitness, before showing them the witness's identification. Unlike witnesses, who may need to recall their certainty weeks or months after the fact for a trial, evaluators usually make decisions on a shorter time-scale, meaning that timing the intervention is less crucial. Undoubtedly, this will result in some difficulties related to other current recommended practices (e.g. video recording the lineup sessions), but making use of video editing and other technologies could mitigate these issues. Research should explore if this and other methods can limit the effects of prior knowledge on interpretations of eyewitness statements.

In summary, three experiments demonstrated that participants altered perceptions of verbal confidence ratings when they knew if a witness's identification matched the police

suspect. Clarifying the witness's statement by using a numeric judgement of certainty may provide some protection against the biasing influences of prior knowledge when evaluators are interpreting moderate levels of confidence, but not when they interpret high levels of confidence. Finally, future efforts should be undertaken to explore contextual effects, and examine potential methods to limit biasing interpretations based on prior knowledge of the police suspect. We are certain that researchers will rise to this challenge, but leave it up to the reader to interpret the exact level of 'certainty'.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Division of National Science Foundation: [Grant Number 1632174].

ORCID

Jesse H. Grabman http://orcid.org/0000-0002-2526-1085

References

Bastardi, A., Uhlmann, E. L., & Ross, L. (2011). Wishful thinking: Belief, desire, and the motivated evaluation of scientific evidence. *Psychological Science*, *22*(6), 731–732.

Beyth-Marom, R. (1982). How probable is probable? A numerical translation of verbal probability expressions. *Journal of Forecasting*, *1*(3), 257–269.

Bradfield, A. L., & Wells, G. L. (2000). The perceived validity of eyewitness identification testimony: A test of the five Biggers criteria. *Law and Human Behavior*, *24*(5), 581–594.

Brun, W., & Teigen, K. H. (1988). Verbal probabilities: Ambiguous, context-dependent, or both? *Organizational Behavior and Human Decision Processes*, *41*(3), 390–404.

Bryant, G. D., & Norman, G. R. (1980). Expressions of probability: Words and numbers. *The New England Journal of Medicine*, *302*(7), 411–411.

Budescu, D. V., Broomell, S., & Por, H. H. (2009). Improving communication of uncertainty in the reports of the intergovernmental panel on climate change. *Psychological Science*, *20*(3), 299–308.

Budescu, D. V., & Karelitz, T. M. (2003, July). Inter-personal communication of precise and imprecise subjective probabilities. In *ISIPTA* (pp. 91–105).

Budescu, D. V., Por, H. H., Broomell, S. B., & Smithson, M. (2014). The interpretation of IPCC probabilistic statements around the world. *Nature Climate Change*, *4*(6), 508–512.

Budescu, D. V., & Wallsten, T. S. (1985). Consistency in interpretation of probabilistic phrases. *Organizational Behavior and Human Decision Processes*, *36*(3), 391–405.

Cash, D. K., & Lane, S. M. (2017). Context influences interpretation of eyewitness confidence statements. *Law and Human Behavior*, *41*(2), 180–190.

Dobolyi, D. G., & Dodson, C. S. (2013). Eyewitness confidence in simultaneous and sequential lineups: A criterion shift account for sequential mistaken identification overconfidence. *Journal of Experimental Psychology: Applied*, *19*(4), 345–357.

Dobolyi, D. G., & Dodson, C. S. (in press). Actual vs. perceived eyewitness accuracy and confidence and the featural justification effect. *Journal of Experimental Psychology: Applied*.

Dodson, C. S., & Dobolyi, D. G. (2015). Misinterpreting eyewitness expressions of confidence: The featural justification effect. *Law and Human Behavior*, *39*(3), 266–280.

Dodson, C. S., & Dobolyi, D. G. (2017). Judging guilt and accuracy: Highly confident eyewitnesses are discounted when they provide featural justifications. *Psychology, Crime & Law*, *23*(5), 487–508.

Dror, I. E., & Charlton, D. (2006). Why experts make errors. *Journal of Forensic Identification*, *56*(4), 600–616.

Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International*, *156*(1), 74–78.

Dror, I. E., Kukucka, J., Kassin, S. M., & Zapf, P. A. (2018). No one is immune to contextual bias – not even forensic pathologists. *Journal of Applied Research in Memory and Cognition*. doi:10.1016/j.jarmac.2018.03.005

Elaad, E., Ginton, A., & Ben-Shakhar, G. (1994). The effects of prior expectations and outcome knowledge on polygraph examiners' decisions. *Journal of Behavioral Decision Making*, *7*(4), 279–292.

Erev, I., & Cohen, B. L. (1990). Verbal versus numerical probabilities: Efficiency, biases, and the preference paradox. *Organizational Behavior and Human Decision Processes*, *45*(1), 1–18.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Gronlund, S. D., Carlson, C. A., Dailey, S. B., & Goodsell, C. A. (2009). Robustness of the sequential lineup advantage. *Journal of Experimental Psychology: Applied*, *15*(2), 140–152.

Gurmankin, A. D., Baron, J., & Armstrong, K. (2004). Intended message versus message received in hypothetical physician risk communications: Exploring the gap. *Risk Analysis*, *24*(5), 1337–1347.

Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, *2*(1), 42–52.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*(2), 211–228.

Kovera, M. B., & Evelo, A. J. (2017). The case for double-blind lineup administration. *Psychology, Public Policy, and Law*, *23*(4), 421–437.

Kukucka, J. K., Kassin, S. M., Zapf, P. A., & Dror, I. E. (2017). Cognitive bias and blindness: A global survey of forensic science examiners. *Journal of Applied Research in Memory and Cognition*, *6*(4), 452–459.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498.

Malpass, R. S., Tredoux, C. G., & McQuiston-Surrett, D. (2007). Lineup construction and lineup fairness. In *The handbook of eyewitness psychology, Vol II: Memory for people* (pp. 155–178). Mahwah, NJ: Lawrence Erlbaum.

Manson v. Brathwaite. (1977). 432 U.S. 98.

Mullet, E., & Rivet, I. (1991). Comprehension of verbal probability expressions in children and adolescents. *Language & Communication*, *11*(3), 217–225.

Nakao, M. A., & Axelrod, S. (1983). Numbers are better than words: Verbal specifications of frequency have no place in medicine. *The American Journal of Medicine*, *74*(6), 1061–1065.

Neil v. Biggers. (1972). 409 U.S. 188.

Neuschatz, J. S., Lawson, D. S., Fairless, A. H., Powers, R. A., Neuschatz, J. S., Goodsell, C. A., & Toglia, M. P. (2007). The mitigating effects of suspicion on post-identification feedback and on retrospective eyewitness memory. *Law and Human Behavior*, *31*(3), 231–247.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.

Quinlivan, D. S., Neuschatz, J. S., Jimenez, A., Cling, A. D., Douglass, A. B., & Goodsell, C. A. (2009). Do prophylactics prevent inflation? Post-identification feedback and the effectiveness of procedures to protect against confidence-inflation in earwitnesses. *Law and Human Behavior*, *33*(2), 111–121.

Reagan, R. T., Mosteller, F., & Youtz, C. (1989). Quantitative meanings of verbal probability expressions. *Journal of Applied Psychology*, *74*(3), 433–442.

Steblay, N. K., Wells, G. L., & Douglass, A. B. (2014). The eyewitness post identification feedback effect 15 years later: Theoretical and policy implications. *Psychology, Public Policy, and Law*, *20*(1), 1–18.

Tenney, E. R., Spellman, B. A., & MacCoun, R. J. (2008). The benefits of knowing what you know (and what you don't): How calibration affects credibility. *Journal of Experimental Social Psychology*, *44*(5), 1368–1375.

Wallsten, T. S., Budescu, D. V., Zwick, R., & Kemp, S. M. (1993). Preferences and reasons for communicating probabilistic information in verbal or numerical terms. *Bulletin of the Psychonomic Society*, *31*(2), 135–138.

Wallsten, T. S., Fillenbaum, S., & Cox, J. A. (1986). Base rate effects on the interpretations of probability and frequency expressions. *Journal of Memory and Language*, *25*(5), 571–587.

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*(3), 360–376.

Wells, G. L., & Bradfield, A. L. (1999). Distortions in eyewitnesses' recollections: Can the postidentification-feedback effect be moderated? *Psychological Science*, *10*(2), 138–144.

Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior*, *33*(1), 1–24.

Wells, G. L., Yang, Y., & Smalarz, L. (2015). Eyewitness identification: Bayesian information gain, base-rate effect equivalency curves, and reasonable suspicion. *Law and Human Behavior*, *39*(2), 99–122.

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, *18*(1), 10–65.

Yates, S. Q. (2017, January 1). Memorandum for heads of department law enforcement components all department prosecutors. *Subject: Eyewitness identification: Procedures for conducting photo arrays*. Retrieved from https://www.justice.gov/archives/opa/press-release/file/923201/download