



Detailed information mitigates confidence inflation

Ryosuke Iida^a, Yukio Itsukushima^a and Eric Y. Mah^b

Q1 ^aNihon University, Japan; ^bUniversity of Victoria, Canada

ABSTRACT

Jurors distrust eyewitness testimony when eyewitness confidence is inflated between the incident and identification in court. Nevertheless, jurors may view inflated-confidence testimony as reliable if the eyewitness gives a justification for the inflation. Researchers have not examined how this 'recovery of eyewitness credibility' is affected by specific features of the justification (i.e. degree of detail). In Experiment 1, we manipulated the degree of detail in post-confidence-inflation eyewitness justifications containing information *related* to a witnessed criminal. We examined the effects of such justifications on participants' ratings of the eyewitness testimony. Although highly detailed but inconsistent eyewitnesses who gave a related justification were not able to fully recover their credibility, we found that they showed reduced credibility loss relative to eyewitnesses who gave a less detailed justification or no justification. In a second experiment, we investigated the possibility that an eyewitness with inflated confidence could recover their credibility with a justification containing information *unrelated* to the criminal. Interestingly, we found that even when the justification was unrelated to the criminal, highly detailed but inconsistent eyewitnesses could mitigate some of their credibility loss. Implications for the mechanisms underlying eyewitness credibility recovery, and their ramifications for real-world cases are discussed.

ARTICLE HISTORY

Received 4 May 2020
Accepted 1 October 2020

KEYWORDS

Eyewitness confidence;
inconsistent eyewitness;
confidence inflation;
eyewitness testimony;
eyewitness credibility

Eyewitness confidence is one of the most important factors in judging eyewitness credibility. The US Supreme Court recommends that judges consider eyewitness confidence as an important factor when assessing an eyewitness' accuracy (*Neil v. Biggers, 1972*). Accordingly, research shows that many legal personnel and potential jurors believe that eyewitness confidence is a good predictor of eyewitness testimonial accuracy (e.g. Magnussen et al., 2010; Wise et al., 2009; Wise & Safer, 2010).

Due to the emphasis placed on eyewitness confidence, if an eyewitness' confidence has been inflated (i.e. increased from the time of identification to the time of testimony), their testimony may not be viewed as credible. Bradfield and McQuiston (2004) conducted two experiments where participant-jurors judged the credibility of an eyewitness in a fictitious trial transcript. In the transcript, the eyewitness stated their current confidence in their identification. After that, the defense attorney requested the record of the eyewitness' first testimony at the police office. There were three conditions: a

CONTACT Ryosuke Iida ✉ chry18001@g.nihon-u.ac.jp

Supplemental data for this article can be accessed <https://doi.org/10.1080/1068316X.2020.1849696>

© 2020 Informa UK Limited, trading as Taylor & Francis Group

control condition where the eyewitness' initial level of confidence was the same as their current one; an *inflation* condition in which the eyewitness had reported being 'not sure' at initial identification but 'positive' at the trial; and an *inflation + challenge* condition identical to the *inflation* condition except that defense attorney aggressively challenged the eyewitness' confidence inflation. In their first experiment (with White participants), the results indicated that participants in both *inflation* conditions rated the eyewitness as less credible than in control condition. However, unless inflated-confidence eyewitnesses were challenged, perceptions of their accuracy and the defendant's guilt were not substantially reduced. Though these results are interesting and suggest that confidence inflation may not be as damaging to eyewitness credibility as one might expect, the findings did not replicate in a 2nd experiment (with Hispanic participants), indicating that race may moderate confidence inflation effects (Bradfield & McQuiston, 2004).

Despite this, the previous findings suggest that testimony may not necessarily be disregarded as a result of confidence inflation. Consider the case of Bruce Godschalk, an innocent man who served 15 years for a rape he didn't commit. At trial, the accusing witness testified that she was very certain that he was the rapist, despite the fact that she wasn't certain at the time of identification (Garrett, 2011). Unfortunately, similar cases may be common (at least, 34 out of 161 trials with eyewitness testimony where the records of testimony were available, and potentially more that go unreported; Garrett, 2011). Indeed, researchers have found evidence that inconsistent eyewitnesses can recover their perceived credibility after confidence inflation (e.g. Jones et al., 2008). Jones et al. (2008) manipulated eyewitness confidence consistency and tested three confidence inflation conditions where the eyewitness gave different justifications for their confidence inflation. In the *strategic inflation* condition, the eyewitness explained, 'I want people to believe me. I want someone to be held accountable for what happened to me'. The eyewitness in the *memory contamination* condition said, 'I have been rehearsing my testimony a lot with my lawyers and the police to prepare for court. I have become more confident each time I have rehearsed my testimony'. In the *confidence epiphany* condition, the eyewitness stated, 'I was nervous at the time of the identification, but now I am confident. Since the identification, I have recalled other details of the event that have made me confident that I am correct'. Participants read one version of the explanation and rated the eyewitness on various dimensions (e.g. credibility, accuracy, etc.). The main finding of interest was that participant ratings of credibility and accuracy were similar in the confidence epiphany and control (i.e. no confidence inflation) conditions. In other words, the eyewitness in the confidence epiphany condition was able to recover credibility simply by stating that she remembered further information about the event. As Jones et al. (2008) suggested, 'if the juror believes that a witness has become more confident due to an insight about the situation, the juror may continue to attribute the witness' courtroom behavior to a desire to tell the truth'. Crucial here is the idea that the interpretation and attribution of the justification for the eyewitness' confidence inflation could influence judgements of eyewitness credibility.

Naturally, one might ask what justifications are accepted as reasonable when an eyewitness' confidence has been inflated. Other studies have investigated the recovery of eyewitness credibility. For example, Douglass and Jones (2013) examined the effects of justification format. Specifically, they compared the effects of written vs. videotaped post-inflation justifications. They found that participants were more cautious of

inflated-confidence testimony with videotaped evidence. However, no studies have examined what kinds of justifications result in credibility recovery. It is plausible that other aspects of the reasons given for confidence inflation, such as the degree of detail of that information, could influence the perceived eyewitness' credibility.

95 According to Bell and Loftus (1988, 1989), the degree of detail given in eyewitness testimony influences perceived eyewitness credibility and perceived probability of suspect guilt. To test this, they employed a 2 (Testimony type; *related vs unrelated*) \times 2 (Prosecution detail; *high vs low*) \times 2 (Defense detail; *high vs low*) between-participant design. Specifically, the eyewitness in the *related/high detail* condition testified, 'The man (the
100 suspect) went and got a box of Milk Duds and a can of Diet Pepsi'. The eyewitness testimony in the *unrelated/high detail* condition was that she 'saw the boy (non-suspect bystander) purchase a box of Milk Duds and a can of Diet Pepsi'. In the *low detail* condition, 'a box of Milk Duds and a can of Diet Pepsi' was replaced with 'a few store items'. Note that the operational definition of relatedness was whether the testimony was about the suspect (related) or the boy (unrelated). Bell and Loftus (1989) found that both the degree of detail and testimonial relatedness significantly affected the perceived probability of suspect guilt, with more detail and relatedness increasing the perceived probability. Degree of detail, but not testimonial relatedness, affected perceptions of eyewitness credibility (more detail = higher credibility). The authors (Bell & Loftus, 1989) also found that the influence of relatedness may depend on the degree of detail. Specifically, related prosecution details influenced judgements of guilt when defense detail was high, whereas unrelated prosecution details influenced judgements of guilt when defense detail was low. Finally, they found that the effects of detail on eyewitness credibility may be attenuated when details were unrelated (though the effect was marginal). These results suggest although the degree of testimonial detail may affect jurors' evaluations more than the degree of testimonial relatedness, both factors can play a role in perceptions of eyewitness credibility.

120 The current study

Extending these findings to the recovery of eyewitness credibility after confidence inflation, we conducted two experiments exploring the influence of the type of justification (*related vs unrelated*) and the degree of detail of that information (*high vs low*) on the recovery of eyewitness credibility. Our first experiment used only justifications with related information – our aim here was to first see if the stronger effects of detail with related information observed in Bell and Loftus (1989) replicate in the confidence inflation paradigm, before extending the scope of the research to the likely weaker (i.e. harder to observe) effects of unrelated information. In order to achieve this goal, we adapted the manipulation used in Bell and Loftus (1989) to the paradigm of confidence inflation studies. We used one consistent condition and three confidence inflation conditions. In the *consistent* condition, the eyewitness' confidence was the same at the time of identification and trial. In the three confidence inflation conditions, the eyewitnesses' confidence at trial was inflated from the time of the identification. The three confidence inflation conditions were divided into *no justification*, *low detail justification*, and
135 *high detail justification* conditions. The eyewitness in the *no justification* condition didn't provide any reasonable explanation for his confidence inflation. In the *low detail*

condition, the eyewitness said, 'After the identification, I remembered additional details about the incident. The suspect stole some items'. In the *high detail* condition 'some items' was replaced with 'two Clorets and one Mintia'. In terms of the relatedness of the justification, we chose to use Bell and Loftus (1989) *related* testimony for the reasons mentioned previously.

There are many aspects of eyewitness credibility that one can examine. For instance, the *Biggers* criteria suggest that triers of fact should evaluate eyewitnesses on the following aspects: certainty (i.e. confidence), accuracy of description, eyewitness attentiveness at time of witnessing, quality of view at time of witnessing, and time between the witnessed incident and initial identification (Neil v. Biggers, 1972). According to the summative hypothesis (Bradfield & Wells, 2000), each of the five criteria contributes some amount to an integrated overall impression of an eyewitness' credibility (e.g. consistency, reliability, and probability of suspect guilt). Thus, if the inconsistent eyewitness' justification influences the integrated (overall) impression of them that jurors hold, it should also influence individual judgements of the five criteria. Previous studies (Douglass & Jones, 2013; Jones et al., 2008) mainly examined integrated impressions (perceived eyewitness consistency, reliability, and probability of suspect guilt) although some individual judgement criteria (eyewitness accuracy) were also used. In both our experiments, we measured participant perceptions of eyewitness consistency, reliability, probability of guilt, accuracy, confidence, attentiveness, quality of view, and time between the witnessed incident and initial identification. However, here we will focus specifically on a subset of these characteristics that we deemed most important: eyewitness accuracy, consistency, reliability, and probability of suspect guilt.¹

There are at least two practical motivations for our research. First, understanding the recovery of eyewitness credibility would allow us to understand when, how, and why jurors come to trust an eyewitness whose confidence has been inflated. Understanding the circumstances under which jurors come to trust confidence-inflated eyewitnesses could allow us to prevent misguided trust. Second, if we consistently observe robust credibility recovery, we could provide further support for the adoption of procedures that maximize the probative value of eyewitness testimony. Such procedures put more emphasis on initial, uncontaminated eyewitness testimonies, where confidence is less likely to be inflated (Wixted & Wells, 2017). The results of previous studies on credibility recovery (Douglass & Jones, 2013; Jones et al., 2008) showed that participant-jurors viewed inconsistent eyewitnesses as reliable simply because eyewitnesses remembered novel information. However, in real cases, this 'novel information' could come not only from true memories, but also from contaminating sources such as investigators, co-witnesses, news media, and so on. Considering the influence of post-event information and the ease with which source monitoring errors can occur (Lindsay, 1994), it is extremely difficult to distinguish eyewitnesses' true memories and the products of memory contaminations. Understanding whether novel information is seen as probative could inform procedures aimed at preventing tragic errors caused by jurors falsely accepting contaminated testimony.

Experiment 1

In Experiment 1, we tested whether the degree of detail given with justifications that contained information directly related to the crime affected post-confidence-inflation

eyewitness credibility recovery. Specifically, we examined judgements of eyewitness credibility using four versions of a fictitious trial transcript: two experimental conditions (*low* and *high detail* justification conditions) and two control conditions (*consistent* and *no justification* conditions). Based on the results of Bell and Loftus (1989), we predicted that judgements of credibility would be highest in the (1) consistent and high detail conditions followed by (2) low detail and lastly (3) no justification conditions.

Methods

Participants and design

Participants in Experiment 1 included 136 Japanese students from the University of Tokyo (56 males, 80 females; Age $M = 19.46$, $SD = 1.14$). Thirty-four participants were randomly assigned to each of the four conditions (*consistent*, *no justification*, *low detail*, *high detail*). Post-hoc power analyses revealed that our power to detect 'small', 'medium', and 'large' differences between any two conditions (i.e. $d = .2$, $d = .5$, $d = .8$, respectively) was $\sim .12$, $\sim .53$, and $\sim .90$.² Thus, our sample was most equipped to detect medium-to-large differences. The *low* and *high detail* conditions were treated as experimental conditions, with the *consistent* and *no justification* conditions serving as controls. All participants received compensation worth 500 yen.

Materials and procedure

Participants (in groups of 2–6) were met by an experimenter, who explained that the study's purpose was to explore the influence of eyewitness testimony on judicial judgements. After that, the experimenter randomly distributed one of the four transcripts to participants. The experimenter was blind to experimental condition. Participants were given five minutes to read the transcript and then were given the judicial judgement questionnaire assessing various aspects of the eyewitness' credibility (and unlimited time to complete it).

Transcript

All four transcripts were two pages long and consisted of a brief summary, indictment, and court script. The transcripts used in this experiment were adapted from mock court scripts made by Japanese Ministry of Justice for judicial education purposes (Japanese Ministry of Justice, 2013). These transcripts were modified to manipulate eyewitness testimony and confidence as per our experimental design. Table 1 below shows an example of a transcript, with the critical manipulations bolded. These four transcripts were identical except for the eyewitness testimony in the court script portion. The upper half of first page included a brief summary of a convenience store robbery and the indictment against the suspect. The summary and indictment alleged that the suspect (Sato) committed the robbery with a kitchen knife, stealing 107,000 yen and some store items. The clerk (Suzuki), who was the only eyewitness, was hit in his face by the robber, who then fled the scene. After the robbery, police found the suspect with 105,000 yen 5 kilometers away from the convenience store. Although the suspect denied his involvement, he was accused of the theft and assault on the basis of the eyewitness identifying him in a subsequent lineup, and the fact that he had almost the exact amount of stolen money. The remaining part of the transcript was the court script. In the

Table 1. Eyewitness testimony: manipulations of confidence inflation and reasons.

Prosecutor	Who is the man that you identified?
Mr.Suzuki	The person is the robber.
Prosecutor	Are you sure?
Mr.Suzuki	I am 100% confident that he is the robber.
Prosecutor	That's all.
Judge	Lawyer, it's your turn.
Lawyer	How confident were you when you identified the suspect?
Mr.Suzuki	I testified that I was 100% confident that he was the robber.
Lawyer	Judge, please let me disclose the police record of his identification in order to confirm the accuracy of his testimony. Does the prosecution have any objections?
Prosecutor	No.
Judge	Go ahead.
Lawyer	Then I'll play the tape. (Tape starts)
Investigator	Please tell me whether or not the robber is one of the men in this lineup?
Mr.Suzuki	Well number three (the number of the suspect) is the robber.
[Mr.Suzuki]	[Well number five, no, maybe number three (the number of the suspect). Uh number three may be the robber.]
Investigator	How confident are you?
Mr.Suzuki	I am 100% confident that he is the robber.
[Mr.Suzuki]	[I am not sure that I'm 50% confident that he is the robber.] (Tape stops)
Lawyer	This is your testimony, isn't it?
Mr.Suzuki	Yes.
[Lawyer]	[According to the record, your confidence at the time of identification was 50%. Nevertheless, you stated that your confidence was 100% a few minutes ago. Why are you so confident now?]
[Mr.Suzuki]	[I'm not sure what you mean.]
[Lawyer]	[At the time of identification, you stated that you weren't sure the suspect was the robber. What makes you so confident now?]
[Mr.Suzuki]	[uh I don't know why, but I am confident that the suspect is the robber.]
[Mr.Suzuki]	<i>[After the identification, I remembered additional details about the incident. The suspect stole some items (low detail) The suspect stole two Clorets and one Mintia (high detail).]</i>
[Lawyer]	<i>[I understand. The additional testimony is consistent with the indictment.]</i>
Lawyer	That's all.

Note: Script included only in the consistent condition is denoted by bold text. Script included only in the three inflation conditions is denoted by square brackets. Within these, reasons for confidence inflation in the *low* and *high detail* conditions are denoted by italic text.

court script, the eyewitness first gave a full account of the incident and testified that the suspect was the robber, saying, 'I am 100% confident that he is the robber'. After the testimony, the defense lawyer asked the eyewitness what his confidence was during the initial lineup identification. The eyewitness responded: 'I testified that I was 100% confident that he was the robber'. Nevertheless, the defense lawyer challenged the eyewitness' claim by requesting police records. In the *consistent* condition, the records revealed that eyewitness chose the photo of the suspect and said: 'I am 100% confident that he is the robber'. However, the eyewitness in the remaining three confidence inflation conditions chose the suspect but testified: 'I am not sure that I am 50% confident that he is the robber'. In these three inflation conditions, the defense lawyer followed up by asking the eyewitness the reason why his confidence was inflated. The eyewitness in the *no justification* condition stated: 'uh I don't know why, but now I am confident that the suspect is the robber'. Meanwhile, the eyewitness in the *low* and *high detail* conditions gave additional justification, saying: 'After the identification, I remembered more details about the incident. The suspect stole

“some items” (*low detail*) or “two Clorets and one Mintia” (*high detail*) before he escaped’. Finally, the defense lawyer accepted the additional testimony as consistent with the indictment.

275 *Dependent variables*

After reading the transcript, participants completed the judicial judgement questionnaire, which consisted of either 10 or 11 items. As with previous studies, we asked the participants to rate the eyewitness’ confidence, consistency, accuracy, reliability, and the probability of the suspect’s guilt. In order to further explore the potential effects of justifications given for confidence inflation, we added four items based on the *Biggers* criteria for aspects that should be taken into account in judging eyewitness reliability (Neil v. Biggers, 1972); confidence, attentiveness, view quality, and time between the incident and identification. However, the *Biggers* criteria for ‘accuracy of suspect description’ was removed since it was largely redundant with our already-included accuracy criterion. The remaining three items were manipulation checks, where participants reported the eyewitness’ initial and final confidence in addition to the justification given for the confidence inflation. The justification manipulation check item was not included in the questionnaire for the *consistent* condition. Table 2 below lists our dependent variables and their scoring.

290 *Results³*

Manipulation check

Three research assistants examined the answers for the questions that asked about initial and final eyewitness confidence (questions 9 & 10), and the justification given by the eyewitness (question 11). For this question, the participants in the *no justification* condition had to mention the absence of any justification for the eyewitness’ confidence inflation. For example, if a participant was in the *no justification* condition, the following answers would be considered correct: ‘The eyewitness didn’t give a justification for his confidence inflation’. or ‘The eyewitness said, “I don’t know”’. On the other hand, the participants in the two justification conditions had to mention that the eyewitness remembered additional details about the incident or that the eyewitness remembered the

305 **Table 2.** Eyewitness assessment questionnaire.

Order	Question	Answer type
1	How confident do you think the eyewitness was?	0 (Not at all confident) to 6 (Very confident)
2	How accurate do you think the eyewitness’ account was?	0 (Inaccurate) to 6 (Accurate)
3	How much attention do you think the eyewitness paid during the incident?	0 (No attention at all) to 6 (Complete attention)
4	How good do you think the eyewitness’ view was?	0 (Bad) to 6 (Good)
5	How many days do you think passed from the day of the incident to the day of the identification?	Open-ended
6	How consistent do you think the eyewitness’ testimony was?	0 (Inconsistent) to 6 (Consistent)
7	How reliable do you think the eyewitness was?	0 (Unreliable) to 6 (Reliable)
8	Considering the eyewitness testimony alone, to what extent do you think the suspect is guilty?	0-100%
9	How confident was the eyewitness at the police office?	0-100%
10	How confident was the eyewitness in court?	0-100%
11	How did the eyewitness explain his inconsistent confidence?	Open-ended

Note: The last question was not included in the questionnaire in the consistent condition.

items the suspect stole. Prior to analysis, we excluded 17 participants with at least one incorrect answer to any of these questions (See Supplementary Material A for details).⁴ ▲

Eyewitness assessments and judgement of probability of guilt

For each of our four main eyewitness assessment criteria (eyewitness accuracy, consistency, reliability, and the probability of suspect guilt)⁵, we conducted a one-way ANOVA comparing ratings of across our four reason conditions (per-ANOVA alpha level adjusted to .013 via Bonferroni correction), with follow-up Tukey's HSD/Games-Howell tests.

Recall that one of our hypotheses was that there would be *no* difference in ratings between consistent eyewitnesses and those who gave a highly detailed reason. Standard null hypothesis significance testing cannot provide evidence for a null effect ▲ only against it (Wagenmakers et al., 2018). Bayesian hypothesis testing is an alternative approach that can provide evidence for both the alternative hypothesis (H_1) and the null hypothesis (H_0) in the form of Bayes factors (BFs) that quantify the evidence for H_1 relative to H_0 (see Wagenmakers et al., 2018 for a comprehensive review). For example, BF_{10} of 3 indicates that given data are 3 times more likely under H_1 than H_0 , whereas a BF_{10} of 1/3 indicates that the given data are 3 times more likely under H_0 than H_1 . BF evidence ranges from *anecdotal* if the ratio in favor of one hypothesis is less than 3:1 to *extreme* if the ratio exceeds 100:1 (Wagenmakers et al., 2011).

In order to investigate any null differences as well as any differences, we supplemented our NHST analysis with Bayesian one-way ANOVAs (Rouder et al., 2012) and follow-up Bayesian independent-samples t -tests (Rouder et al., 2009) testing the effect of reason on each of our dependent variables. All analyses used 'default' prior specifications (see the respective papers for more details). As the results of those analyses, we focused on eyewitness accuracy, consistency, reliability, and probability of guilt where the effects of degree of detail appeared so that our results can be easy to interpret (see Supplementary Material Table S2 for all detailed pairwise comparison results on the rest four variables).

Moving now to the results, Figure 1 below depicts standardized condition means and 95% CIs for ratings of eyewitness accuracy, consistency, reliability, and probability of guilt.

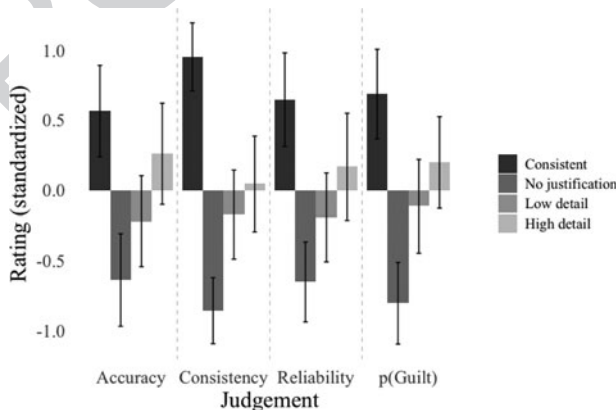


Figure 1. Experiment 1: Judgements of eyewitness accuracy, consistency, reliability, and probability of guilt, by justification condition. Error bars = 95% CIs (non-overlapping CIs indicate a significant pairwise difference).

All ANOVAs were significant, and Bayesian evidence for overall effects was extreme. To account for some instances of violated assumptions we conducted Welch's one-way tests and non-parametric Kruskal-Wallis tests alongside each parametric ANOVA. In all cases, the results agreed. We followed each significant ANOVA with pairwise comparisons via Tukey's HSD/Games-Howell tests. We summarize the results of these pairwise comparisons in Table 3 (including effect size estimates and Bayes Factors for pairwise comparisons and omnibus ANOVAs).⁶

For judgements of eyewitness accuracy, we see differences between the *consistent* condition and the *no justification* and *low detail* conditions, and a difference between the *no justification* and *high detail* condition. In the case of our null results, our Bayesian tests suggest that evidence is ambiguous at best in favor of either hypothesis. Importantly, we did not find evidence for a substantial difference between the *consistent* and *high detail* conditions, suggesting that recovery of credibility in terms of perceived accuracy may be possible. The fact that we did not observe substantial differences between the *high detail* and *low detail* conditions additionally suggests that even minimally detailed justifications may mitigate some loss of perceived accuracy.

However, for judgements of consistency, ratings in the *consistent* condition were higher than in all other conditions (Extreme Bayesian evidence). Additionally, ratings in the *no justification* condition were lower than in either detail condition (Very strong to extreme Bayesian evidence). The *low* and *high detail* conditions did not differ, but there was only anecdotal evidence against an effect. Thus, even if inconsistent eyewitnesses are able to recover some of their perceived accuracy, participants were still acutely aware of the inconsistency.

For judgements of reliability, ratings in the *consistent* condition were higher than in the *no justification* and *low detail* conditions (Very strong to extreme Bayesian evidence), and

Table 3. Experiment 1 pairwise comparisons: Bayes factors and effect sizes.

Comparison		Question			
		Accuracy	Consistency	Reliability	<i>p</i> (Guilt)
<i>Consistent vs. No justification</i>	<i>d</i>	1.38	2.84	1.57	1.83
	[95% CI]	[.75, 1.99]	[1.95, 3.73]	[.91, 2.21]	[1.13, 2.52]
	BF ₁₀	7347.85	3.08 * 10¹²	84,441.24	2.91 * 10⁶
<i>Consistent vs. Low detail</i>	<i>d</i>	.91	1.51	.97	.92
	[95% CI]	[.34, 1.46]	[.86, 2.14]	[.40, 1.53]	[.35, 1.48]
	BF ₁₀	32.62	40,036.6	61.7	37.5
<i>Consistent vs. High detail</i>	<i>d</i>	.33	1.13	.49	.56
	[95% CI]	[-.18, .83]	[.55, 1.71]	[-.03, 1.01]	[.03, 1.08]
	BF ₁₀	.52	445.81	1.21	1.85
<i>No justification vs. Low detail</i>	<i>d</i>	.49	.93	.57	.84
	[95% CI]	[-.05, 1.01]	[.36, 1.50]	[.04, 1.11]	[.27, 1.39]
	BF ₁₀	1.11	38.23	1.96	15.53
<i>No justification vs. High detail</i>	<i>d</i>	.97	1.15	.91	1.21
	[95% CI]	[.40, 1.53]	[.56, 1.73]	[.34, 1.46]	[.60, 1.80]
	BF ₁₀	63.33	429.24	32.16	935.35
<i>Low detail vs. High detail</i>	<i>d</i>	.52	.24	.39	.35
	[95% CI]	[-.01, 1.04]	[-.26, .76]	[-.13, .90]	[-.17, .86]
	BF ₁₀	1.39	.39	.65	.57
Overall ANOVA	BF ₁₀	3991.19	5.21 * 10¹⁰	11,757.67	1.18 * 10⁶

Note: 95% CIs for Cohen's *d*s based on uncorrected $\alpha = .05$, so they may not map onto pairwise comparison results. BF₁₀ values represent Bayes factors in favor of an effect/difference relative to no effect/difference. Values in bold represent greater than anecdotal evidence for an effect/difference (BF₁₀ > 3), and non-formatted values represent ambiguous evidence (.33 < BF₁₀ < 3). All Bayes factors were computed with default, uninformative priors.

ratings in the *no justification* condition were lower than those in the *high detail* condition (Very strong Bayesian evidence). Reliability ratings did not differ between the *consistent* and *high detail* conditions, the *low detail* and *high detail* conditions, or the *no justification* and *low detail* conditions, though Bayesian evidence in these cases was inconclusive. Surprisingly, despite the fact that participants were aware of eyewitness inconsistency, these results suggest that highly-detailed eyewitnesses were still able to recover some of their perceived reliability.

For judgements of $p(\text{Guilt})$, participant judgements were higher in the *consistent* condition than in the *no justification* and *low detail* conditions (Very strong to extreme Bayesian evidence). Additionally, estimates in the *no justification* condition were lower than those in the *low* and *high detail* conditions (Strong to extreme Bayesian evidence). Guilt probability estimates did not differ between the *consistent* and *high detail* conditions, or between the *low* and *high detail* conditions, though again Bayesian evidence was inconclusive. As with reliability, we see some evidence for credibility recovery with a highly detailed justification.

Discussion

In general, our results were consistent with our predictions—judgements of credibility were highest in the (1) consistent and high detail conditions followed by (2) low detail and lastly (3) no justification conditions. Specifically, our results suggest that if an eyewitness gives a justification for confidence inflation that includes *related* information directly pertaining to the crime (high or low in detail), they are perceived more positively than an eyewitness who gives no reason. The degree of detail also seems to matter—although consistent eyewitnesses were viewed as more accurate, consistent, reliable, and suggestive of suspect guilt than eyewitnesses who gave no justification or provided a *low detail* justification, the same was not true of eyewitnesses who provided a *high detail* justification. Contrary to our general predictions about credibility, but perhaps unsurprisingly, consistent eyewitnesses were viewed as more consistent than eyewitnesses who gave a highly detailed justification for their confidence inflation. However, as we predicted, we found evidence (albeit weak) against differences in perceptions of accuracy between *consistent* testimonies and *highly detailed but inconsistent* testimonies, and only weak evidence for differences in perceptions of reliability and suspect guilt. Together with our effect size estimates, this weak Bayesian evidence suggests that differences in ratings of consistent witnesses and those that give a highly detailed justification are either null or small. This finding of ‘recovered eyewitness credibility’ (or at least ‘credibility loss mitigation’) with highly detailed justifications is striking, given the importance of assessments of reliability and probability of guilt in real-world cases. Of course, we note that we did not observe compelling evidence for *complete* credibility recovery, even with highly-detailed justifications.

We did not observe evidence for substantial differences between the *high* and *low detail* conditions. However, we note that differences between these conditions may have been too small to detect with our sample (in particular, detecting substantial Bayesian evidence for small effects can require massive sample sizes; Rouder et al., 2009). Descriptively, our results suggest that there may be small but consistent differences in ratings between an eyewitness who gives a minimally detailed related justification and

one who gives a highly detailed related justification. Additionally, Bayesian evidence for differences between the *consistent* and *low detail* conditions was generally more compelling than evidence for differences between the *consistent* and *high detail* conditions, providing further suggestive evidence for the potential importance of level of detail.

455 Why might eyewitnesses who provide a highly detailed justification for confidence inflation 'recover' their credibility? This recovery might be explained in terms of ratings of accuracy, where evidence suggested either null or small differences between consistent and highly detailed eyewitnesses. That is, participants believed that the eyewitness was more accurate when the eyewitness provided a highly detailed justification, which, 460 in turn led to some recovered credibility. Similar relationships between perceived accuracy and reliability have been observed in prior research (Jones et al., 2008; Experiments 2 & 3). For example, Jones et al. (2008) observed similar accuracy in consistent eyewitnesses and inflated confidence eyewitnesses who recovered their credibility, but lower accuracy relative to consistent eyewitnesses when credibility was not recovered. This is 465 consistent with Bradfield and Wells summative hypothesis (2000). According to the summative hypothesis, each variable (i.e. *Biggers* criteria) contributes some amount to the overall impression of the credibility of the eyewitness and the amount contributed by one variable does not depend on the levels of the other variables.

470 Experiment 2

We found observable effects of Bell and Loftus (1989) detail manipulation on perceptions of confidence-inflated eyewitnesses, suggesting that the detail/relatedness paradigm can be meaningfully adapted to the confidence inflation paradigm. As a result, we turned our 475 investigation next to the influence of detail on justification containing information *not* directly related to the witnessed criminal. The only-partial recovery of eyewitness credibility that we observed in Experiment 1 led us to predict that a confidence-inflated eyewitness would not be able to attenuate their credibility loss when they justified their confidence inflation with information *unrelated* to the witnessed criminal. According to Bell and 480 Loftus (1989), the effect of testimony detail may be limited when the testimony involves unrelated information (e.g. in their study, the store items are dropped by a bystander prior to the crime). Their results suggest that unrelated testimony may not boost eyewitness credibility even if given with a highly detailed testimony. However, this prediction has not been directly tested in the context of confidence inflation and *recovery* of credibility.

485 In all prior confidence inflation eyewitness credibility experiments, the justifications for confidence inflation involved information directly related to the witnessed criminal (Bradfield & Jones, 2013; Jones et al., 2008). Although these studies found evidence for 490 credibility recovery, it is not clear whether credibility recovery is possible when justifications are based on information unrelated to the criminal. Based on the non-confidence-inflation work from which we drew our manipulation (Bell & Loftus, 1989), we hypothesized that regardless of the level of detail, eyewitnesses who gave justifications with unrelated information would not recover their credibility. In order to test this hypothesis, we conducted a second experiment examining the effects of unrelated justifications on eyewitness credibility after confidence inflation.

495 To do this, we replaced 'the suspect stole' in the eyewitness justifications in Experiment 1 with 'the woman who came into the store before the robbery bought'. For example, the

justification in the *low detail* condition was 'The woman who came into the store before the robbery bought some items'. As a result, the eyewitness' justification for his confidence inflation wasn't directly related to the suspect or the crime. According to Bell and Loftus (1989), the effect of *unrelated* justifications would be weaker than that of *related* justifications. Specifically, we predicted that even the eyewitness in the *high detail* condition should be rated as less credible than the eyewitness in the *consistent* condition (but acknowledged that there might also be no difference). Similarly, we predicted that the eyewitness in the *no justification* and *low detail* conditions would be rated as less credible (on all dependent variables) than the eyewitness in the *consistent* condition. We had no specific predictions regarding potential differences between the three justification conditions.

Because we tested the effects of related and unrelated justifications across experiments, firm conclusions cannot be drawn about the effects of the testimonial relatedness. However, our primary objective with the second experiment was not to compare the relative effects of related/unrelated justifications, but rather to establish whether or not an unrelated justification could mitigate credibility loss caused by eyewitness confidence inflation.

Methods

Participants and design

Participants in Experiment 2 included 132 Japanese students from the University of Tokyo (39 males, 93 females; Age $M = 19.41$, $SD = 1.38$). Thirty-three participants were randomly assigned to each of the four conditions (*consistent*, *no justification*, *low detail*, *high detail*). Similar to Experiment 1, our sample was most equipped to detect medium-to-large condition differences. The *low* and *high detail* conditions were treated as experimental conditions, with the *consistent* and *no justification* conditions serving as controls. All participants received compensation worth 500 yen.

Materials and procedure

Experiment 2's procedure was identical to Experiment 1, with the only difference being the relatedness of the justification.

Results

Manipulation check

Like Experiment 1, three research assistants examined the answers for the questions that asked about initial and final eyewitness confidence (question 9 & 10), and the justification given by the eyewitness (question 11). The same criteria from Experiment 1 were used to exclude 16 participants prior to analysis (See Supplementary Material B for details).⁷

Eyewitness assessments and judgement of probability of guilt

As with Experiment 1, we examined the effects of (unrelated) justifications of varying detail on eyewitness credibility recovery after confidence inflation. We adopted the same analytic strategy (one-way ANOVAs, pairwise comparisons via Tukey's HSD/Games-Howell, and Bayesian one-way ANOVAs and follow-up Bayesian independent-

samples *t*-tests). Bayesian one-way ANOVAs and follow-up Bayesian independent-samples *t*-tests (see Supplementary Material Table S3 for the results of NHST analyses). Analyses were conducted separately for each of our four main eyewitness credibility variables: eyewitness accuracy, consistency, reliability, and the probability of guilt. Figure 2 below depicts condition means and 95% CIs for these variables.

For these four variables, all overall ANOVAs were significant, and we found extreme Bayesian evidence for overall effects. We followed these ANOVAs with follow-up pairwise comparisons among our justification detail conditions (see Table 4 for detailed pairwise comparison results).⁸ For accuracy, we observed significant differences between *consistent* eyewitnesses and all other conditions. We did observe some Bayesian evidence for higher ratings for *high detail* eyewitnesses relative to *no justification* eyewitnesses. We also found substantial Bayesian evidence against a difference between the *low* and *high detail* conditions, suggesting that the degree of detail given with an unrelated justification does not affect judgement of eyewitness accuracy (see Supplementary Material Tables S3 and S4 for detailed ANOVA and pairwise comparison results, including effect size estimates and Bayes Factors for all pairwise comparisons).

However, results were somewhat different for ratings of eyewitness consistency, reliability and probability of suspect guilt. For ratings of consistency, we again found that ratings in the *consistent* condition were significantly higher than those in all other conditions (extreme Bayesian evidence). However, we also observed a significant difference between the *high detail* and *no justification* conditions (substantial Bayesian evidence) – some evidence that a highly detailed unrelated justification affects juror perceptions. For ratings of reliability, ratings in the *consistent* condition were again significantly higher than ratings in the *no justification* and *low detail* condition (extreme Bayesian evidence), but not significantly higher than ratings in the *high detail* condition (only anecdotal Bayesian evidence for a difference). Additionally, ratings of reliability were significantly higher in the *high detail* condition relative to the *no justification* condition (substantial Bayesian evidence). Finally, a similar pattern was observed for probability of suspect guilt – significantly higher estimates in the *consistent* condition

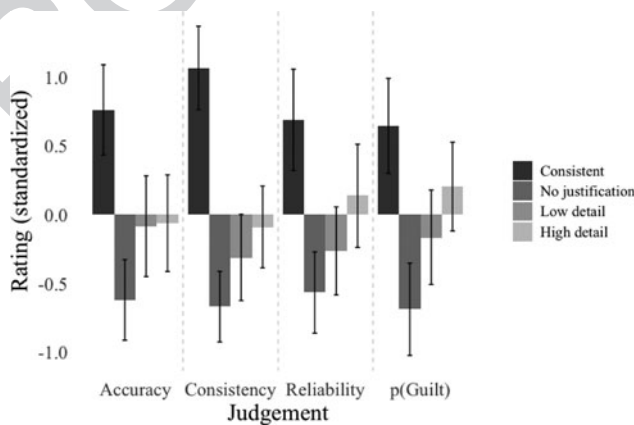


Figure 2. Experiment 2: Judgements of eyewitness accuracy, consistency, reliability, and probability of guilt, by justification condition. Error bars = 95% CIs (non-overlapping CIs indicate a significant pairwise difference).

Table 4. Experiment 2 pairwise comparisons: Bayes factors and effect sizes.

Comparison		Question			
		Accuracy	Consistency	Reliability	$p(\text{Guilt})$
590 <i>Consistent vs. No justification</i>	d	1.68	2.34	1.42	1.48
	[95% CI]	[1.00, 2.35]	[1.54, 3.13]	[.78, 2.05]	[.83, 2.12]
	BF_{10}	$3.06 * 10^5$	$2.32 * 10^9$	10,880.90	22,624.35
<i>Consistent vs. Low detail</i>	d	.93	1.70	1.05	.90
	[95% CI]	[.35, 1.49]	[1.01, 2.37]	[.46, 1.63]	[.32, 1.45]
	BF_{10}	35.44	$40.0 * 10^5$	131.19	26.40
<i>Consistent vs. High detail</i>	d	.92	1.47	.57	.50
	[95% CI]	[.35, 1.48]	[.82, 2.10]	[.03, 1.10]	[−.03, 1.03]
	BF_{10}	34.27	17,912.92	1.78	1.20
595 <i>No justification vs. Low detail</i>	d	.61	.47	.37	.59
	[95% CI]	[.07, 1.15]	[−.06, 1.00]	[−.15, .89]	[.04, 1.12]
	BF_{10}	2.46	1.02	.61	2.03
<i>No justification vs. High detail</i>	d	.65	.79	.79	1.03
	[95% CI]	[.11, 1.19]	[.23, 1.34]	[.23, 1.34]	[.44, 1.61]
	BF_{10}	3.33	10.34	10.29	105.03
600 <i>Low detail vs. High detail</i>	d	.02	.28	.44	.42
	[95% CI]	[−.49, .54]	[−.24, .80]	[−.09, .96]	[−.11, .95]
	BF_{10}	.27	.43	.85	.78
Overall ANOVA	BF_{10}	24,866.75	$3.54 * 10^{10}$	5,569.58	19,004.87

Note: 95% CIs for Cohen's d s based on uncorrected $\alpha = .05$, so they may not map directly onto pairwise comparison results. BF_{10} values represent Bayes factors in favor of an effect/difference relative to no effect/difference. Values in bold represent greater than anecdotal evidence for an effect/difference ($BF_{10} > 3$), underlined values represent greater than anecdotal evidence against an effect/difference ($BF_{10} < .33$), and non-formatted values represent ambiguous evidence ($.33 < BF_{10} < 3$). All Bayes factors were computed with default, uninformative priors.

relative to the *no justification* and *low detail* conditions (strong to extreme Bayesian evidence), no significant difference between the *consistent* and *high detail* conditions (only anecdotal Bayesian evidence for a difference), and a significant difference between the *no justification* and *high detail* conditions (extreme Bayesian evidence).

Exploratory analyses

Unlike Experiment 1, we observed a discrepancy between judgements of accuracy and those of reliability and probability of guilt – highly detailed eyewitnesses appeared to recover credibility in terms of reliability and probability of guilt, but not in terms of perceived accuracy. This raised the possibility that participants judgments of reliability and probability of guilt may have been made independently of judgments of accuracy. To investigate this possibility, we conducted exploratory correlation and regression analyses relating these judgments, for both experiments. We found that ratings of accuracy were correlated with those of reliability and probability of guilt (see Tables 5 and 6), suggesting that participant judgments were not wholly independent. In addition, judgments of perceived accuracy predicted those of reliability and probability of guilt (see Table 7 for a summary of the regression results and an overall summary of results for both experiments). However, we note that these analyses were exploratory and that our sample was not powered for these analyses. As a result, we interpret these results with caution – the main objective was to broadly investigate the surprising possibility that judgments we would expect to be related were in fact unrelated (e.g. rather than making detailed claims about the strength of the relationships).

Table 5. Experiment 1: correlations between select judicial judgments, by condition.

Judgment	Condition	Confidence	Accuracy	Attention	View	Time	Consistent	Reliable
Accuracy	<i>Consistent</i>	.41** (.02)		.43* (.02)	.56*** (.001)	-.18 (.35)	.39* (.03)	
	<i>No justification</i>	.05 (.81)		.43* (.02)	-.19 (.33)	-.13 (.51)	.58*** (<.001)	
	<i>Low detail</i>	.23 (.24)		.52** (.004)	.54** (.003)	-.04 (.85)	.73*** (<.001)	
	<i>High detail</i>	.64*** (<.001)		.28 (.12)	.36* (.04)	-.09 (.68)	.55*** (.001)	
Reliable	<i>Consistent</i>	.12 (.51)	.78*** (<.001)	.32*** (.08)	.68*** (<.001)	-.04 (.84)	.26 (.16)	
	<i>No justification</i>	-.07 (.72)	.60*** (<.001)	.18 (.36)	.01 (.97)	-.05 (.80)	.66*** (<.001)	
	<i>Low detail</i>	.09 (.64)	.78*** (<.001)	.57** (.001)	.61*** (<.001)	-.07 (.71)	.74*** (<.001)	
	<i>High detail</i>	.58*** (<.001)	.67*** (<.001)	.35 (.06)	.45* (.01)	-.40* (.03)	.88*** (<.001)	
<i>p</i> (Guilt)	<i>Consistent</i>	.10 (.58)	.59*** (<.001)	.41** (.02)	.57*** (<.001)	-.10 (.60)	.53** (.003)	.61*** (<.001)
	<i>No justification</i>	-.02 (.94)	.55** (.002)	.54** (.002)	.04 (.82)	.05 (.79)	.14 (.48)	.37 (.05)
	<i>Low detail</i>	.23 (.24)	.64*** (<.001)	.49** (.007)	.46* (.01)	-.02 (.93)	.57*** (.001)	.73*** (<.001)
	<i>High detail</i>	.49** (.005)	.48** (.007)	.53** (.002)	.60*** (<.001)	-.36* (.04)	.65*** (<.001)	.82*** (<.001)

Note: Highly exploratory, underpowered, and after Type I error control most of these are probably not significant.

Table 6. Experiment 2: correlations between select judicial judgments, by condition.

Judgment	Condition	Confidence	Accuracy	Attention	View	Time	Consistent	Reliable
Accuracy	<i>Consistent</i>	.30 (.11)		.76*** (<.001)	.72*** (<.001)	-.21 (.27)	.30 (.12)	
	<i>No justification</i>	.25 (.19)		.41* (.03)	.30 (.11)	-.08 (.68)	.47* (.01)	
	<i>Low detail</i>	.41* (.02)		.49** (.007)	.46* (.01)	.03 (.87)	.77*** (<.001)	
	<i>High detail</i>	.31 (.10)		.63*** (<.001)	.34 (.07)	-.22 (.25)	.48** (.008)	
Reliable	<i>Consistent</i>	.10 (.62)	.73*** (<.001)	.60*** (<.001)	.77*** (<.001)	-.24 (.20)	.30 (.12)	
	<i>No justification</i>	.11 (.58)	.73*** (<.001)	.56** (.002)	.49** (.007)	.08 (.69)	.54** (.002)	
	<i>Low detail</i>	.37 (.05)	.83*** (<.001)	.36 (.06)	.34 (.07)	.04 (.85)	.77*** (<.001)	
	<i>High detail</i>	.34 (.07)	.85*** (<.001)	.69*** (<.001)	.47** (.009)	-.47* (.01)	.61*** (<.001)	
<i>p</i> (Guilt)	<i>Consistent</i>	.33 (.08)	.58*** (<.001)	.54** (.003)	.69*** (<.001)	-.12 (.53)	.38* (.04)	.76*** (<.001)
	<i>No justification</i>	.16 (.41)	.48** (.009)	.50** (.005)	.50** (.006)	-.12 (.53)	.38* (.04)	.44* (.02)
	<i>Low detail</i>	.47** (.009)	.57** (.001)	.38* (.04)	.32 (.09)	-.45* (.01)	.64*** (<.001)	.67*** (<.001)
	<i>High detail</i>	.46* (.01)	.66*** (<.001)	.65*** (<.001)	.51** (.005)	-.57** (.001)	.53** (.003)	.80*** (<.001)

Note: Highly exploratory, underpowered, and after Type I error control most of these are probably not significant.

680

685

690

695

700

705

710

715

720

Table 7. Recovery of credibility by experiment and condition.

Experiment	Confidence	Accuracy	Attention	View	Time	Consistent	Reliable	Guilt	Accuracy predicts reliability	Accuracy predicts $p(\text{Guilt})$
1: Consistent	–	–	–	–	–	–	–	–	***, $BF > 100$	***, $BF = 45$
1: No justification	Definitely No	Definitely No	Definitely No	Probably No	Maybe Yes	Definitely No	Definitely No	Definitely No	***, $BF = 49$	** , $BF = 18$
1: Low detail	Definitely No	Definitely No	Definitely No	Maybe No	Maybe Yes	Definitely No	Definitely No	Definitely No	***, $BF > 100$	***, $BF > 100$
1: High detail	Definitely No	Maybe Yes	Definitely No	Probably Yes	Probably Yes	Definitely No	Maybe No	Maybe No	***, $BF > 100$	** , $BF = 7$
2: Consistent	–	–	–	–	–	–	–	–	***, $BF > 100$	***, $BF = 34$
2: No justification	Definitely No	Definitely No	Definitely No	Definitely No	Probably Yes	Definitely No	Definitely No	Definitely No	***, $BF > 100$	** , $BF = 6$
2: Low detail	Definitely No	Definitely No	Definitely No	Definitely No	Probably Yes	Definitely No	Definitely No	Definitely No	***, $BF > 100$	** , $BF = 27$
2: High detail	Definitely No	Definitely No	Definitely No	Definitely No	Maybe Yes	Definitely No	Maybe No	Maybe No	***, $BF > 100$	***, $BF > 100$

Note: 'Definitely' = $BF > 10$, 'Probably' = $3 < BF < 10$, 'Maybe' = $.33 > BF > 3$. Confidence was not be recovered, regardless of relatedness. Accuracy may be recovered, if a detailed central justification was given. Attention was not recovered, regardless of relatedness. View quality was probably recovered, if a detailed related justification was given. Credibility losses somewhat were mitigated by related justification. Time estimates were probably recovered, if detailed related justification was given, or if an unrelated justification was given (???). Cautious of this result. Consistency was not recovered, regardless of relatedness. Reliability credibility losses were somewhat mitigated if related justification was given. Guilt credibility losses were somewhat mitigated if related justification was given. No matter the reason is given for confidence inflation, perceptions of accuracy likely inform perceptions of reliability. No matter the reason is given for confidence inflation, perceptions of accuracy likely inform perceptions of suspect guilt.

Discussion

We predicted, that confidence-inflated inconsistent eyewitnesses who gave a justification *unrelated* to the crime would be viewed as less credible than consistent eyewitnesses [▲] regardless of the level of detail. In support of this prediction, we observed strong evidence that eyewitnesses who gave unrelated justifications for confidence inflation were not able to recover their credibility, at least in terms of perceptions of accuracy and consistency. However, this was not the case for all aspects of credibility. Given the results of Experiment 1, we predicted that highly-detailed inconsistent eyewitnesses who gave an unrelated reason might be able to recover some credibility. Indeed, for the critical judgements of consistency, reliability, and probability of suspect guilt, the eyewitness who gave highly detailed unrelated information was viewed as more credible than the eyewitness who gave no information. Most surprisingly, for judgements of reliability and guilt, we did not find compelling evidence for a difference between perceptions of consistent and highly detailed eyewitnesses. Therefore, we would focus on this recovery of eyewitness credibility in high detail condition.

Our results for these judgements suggest that if there are any differences between consistent and highly detailed eyewitnesses, these differences are likely null or small. Thus, it appears that even when the information provided with the justification is unrelated and not directly relevant to critical crime details, eyewitnesses can recover at least some of their credibility in the eyes of the jury, if the information is highly detailed. As with Experiment 1, our descriptive results and weak Bayesian evidence suggest that this recovery is not complete (relative to consistent eyewitnesses). However, the implication that eyewitnesses can recover reliability relevant to related crime details (e.g. probability of suspect guilt) by giving information not related to the related crime details is troubling. Just because an eyewitness can accurately describe some details of the crime does not imply that they accurately remembered others (Wells & Leippe, 1981). Despite this, our results suggest that jurors may believe this to some degree.

Unlike in Experiment 1, we cannot explain this recovery of reliability in terms of judgements of eyewitness accuracy [▲] – the highly detailed but inconsistent eyewitness was viewed as less accurate than the consistent eyewitness. And yet, in both experiments, ratings of eyewitness reliability and the probability of the suspect's guilt were similar for consistent and highly detailed eyewitnesses. One sobering explanation for these results is that the summative hypothesis is false; judgements about eyewitness reliability and the suspect's guilt are made independently of judgements of eyewitness accuracy. However, exploratory correlation and regression analyses revealed that in all conditions in both experiments, ratings of accuracy consistently predicted ratings of reliability and probability of suspect guilt. Furthermore, ratings of accuracy, reliability, and suspect guilt tended to be interrelated, and related to judgements of consistency. Therefore, we think it unlikely that our participants' evaluations for the eyewitness reliability and the probability of guilt were not based on the ratings of accuracy. This interpretation is not based solely on the result of our exploratory analysis [▲] – other empirical research also supports the summative hypothesis (Bradfield & McQuiston, 2004; Bradfield & Wells, 2000; Jones et al., 2008; Lindsay, 1994) [▲]

Instead, we suggest that jurors likely do use the judgements of various aspects of credibility to inform one another (e.g. using perceptions of accuracy to inform perceptions of

reliability and suspect guilt), but when a highly detailed related or unrelated justification is given, the very fact that a detailed justification was given (regardless of relatedness) results in some discounting of confidence inflation. This would accord with social psychology research on the influence of 'placebic' information (Langer et al., 1978). Again, limited power restricts the conclusions we can draw here.

General discussion

Our experiments show that the degree of detail about a crime (as opposed to about a criminal) given with justifications for confidence inflation can influence the recovery of eyewitness credibility. In both experiments, the consistent eyewitness was viewed as more credible than eyewitnesses in the *no justification* and *low detail* conditions. However, we found some evidence for credibility recovery when eyewitnesses gave highly detailed justifications. In both experiments, when the inconsistent eyewitness gave a highly detailed justification, the eyewitness' perceived reliability, and the probability of suspect guilt were comparable to the consistent eyewitness. It is worth noting that although eyewitnesses who gave a highly detailed justification recovered some of their credibility, evidence for *full* recovery of credibility was weak. For most of the judgments we examined, our results favored either *no* recovery or *incomplete* recovery. The fact that condition differences were especially pronounced for ratings of consistency (i.e. *consistent* eyewitnesses > *low detail* eyewitnesses = *high detail* eyewitnesses > *no justification* eyewitnesses) suggests that participants were sensitive to the confidence inflation. As a result, even when a highly detailed justification is provided, it may be difficult to overcome the damage to credibility caused by a confidence inflation. At best, a highly detailed justification may mitigate some of this lost credibility.

Though this recovery was not complete, it appears that highly detailed justifications may at least mitigate some of the credibility loss that accompanies confidence inflation. Supporting this, our Bayesian analyses suggested that if there are differences in perceived reliability and suspect guilt between consistent and highly detailed inconsistent eyewitnesses, these differences are likely small or null.

We found that eyewitnesses who gave both related and unrelated information with their justification were able to recover some of their credibility. Highly detailed eyewitnesses who gave a related justification recovered credibility, showing accuracy, reliability, and probability of suspect guilt ratings similar to consistent eyewitnesses. Similarly, highly detailed eyewitnesses who gave an unrelated justification showed similar reliability and probability of suspect guilt ratings similar to consistent eyewitnesses. We qualify these results with two considerations. First, in both experiments, inconsistent eyewitnesses were not able to recover credibility in terms of perceptions of consistency, suggesting that participants did not completely discount the confidence inflation. Second, Bayesian evidence for the apparent credibility recovery was weak in both experiments.

One possible explanation for our results is the summative hypothesis (Bradfield & Wells, 2000). Specifically, the eyewitnesses' highly detailed justification would increase perceived credibility on several individual dimensions. These dimension-specific credibility increases would then 'sum together' and influence more general, integrated judgments such as reliability and probability of suspect guilt. In fact, Bradfield and Wells (2000) found that perceptions of eyewitness accuracy, attention, and confidence independently

influenced overall perceived credibility. The summative hypothesis is also supported by the results of our Experiment 1. In Experiment 1, the inconsistent but highly detailed eyewitness had comparable credibility ratings to the consistent eyewitness on both an individual dimension (accuracy) and the more integrated judgements (reliability and probability of suspect guilt). Based on the summative hypothesis, the eyewitness' highly detailed justification may have increased participant-juror perceptions of accuracy, and this individual increase was reflected in increased perceived credibility in terms of eyewitness reliability and probability of suspect guilt. The absence of differences in perceived attentiveness between the *high detail vs. no justification* conditions in the present study might be due to the fact that, unlike Bradfield and Wells (2000), we did not directly manipulate eyewitness attention. Nonetheless, the summative hypothesis explains why varying levels of detail may affect some, but not other aspects of eyewitness credibility.

However, what explanation could be plausible to interpret the results of Experiment 2 in light of the summative hypothesis? In this experiment, credibility in terms of accuracy was *not* recovered for highly detailed eyewitnesses. And yet, for these eyewitnesses, credibility in terms of eyewitness reliability and probability of suspect guilt was recovered to a similar degree as in Experiment 1. One might argue that the summative hypothesis is false, and that when an eyewitness is inconsistent, integrated judgements of eyewitness assessment (eyewitness reliability and probability of suspect guilt) might be made independently of the judgements of more specific dimensions (e.g. eyewitness accuracy). However, given the results of our first experiment and previous studies showing evidence for summative processing (e.g. Jones et al., 2008), we think it is likely that at least *some* summative processing is occurring.

We propose the following tentative explanation for our findings: Jurors have less favorable perceptions of confidence-inflated eyewitnesses, which affects their judgements of various dimensions of eyewitness credibility. These judgements in turn inform more integrated judgements of overall eyewitness credibility. However, independently of these judgements, the very fact that an eyewitness provided a highly detailed justification mitigates some of the overall credibility loss. In actual cases, eyewitnesses credibility is often dubious (Garrett, 2011). Nevertheless, testimonies have often been accepted despite eyewitness confidence inflation, inaccurate descriptions, lack of attentiveness, poor view quality, or long delays between the incident and identification. In such cases, justifications of varying levels of detail may have served to mitigate some of the lost credibility.

On the other hand, participants may have based their judgments on aspects of the testimony that we did not explicitly measure. For example, previous research has shown that the perceived *vividness* of eyewitness testimony influences mock jurors' judgments of defendant guilt (Reyes et al., 1980; Shedler & Manis, 1986). In our experiment, it is possible that our detail manipulation exerted its effects as a result of changes in perceived vividness. Furthermore, our participants might have inferred the eyewitness' level of arousal from the detail of the eyewitness justification (for example, the eyewitness could provide highly detail accounts because the eyewitness would be calm enough to). Bornstein et al. (2008) found that the participant-jurors' judgements corresponded to their beliefs in the effects of arousal on memory; mock jurors who thought emotional arousal would enhance memory were more likely to render a guilty verdict when the eyewitness was characterized as high in arousal, while the reverse was true for mock jurors

who believed that emotional arousal hinders memory. Therefore, it might be possible that we observed the difference of eyewitnesses' level of arousal along the manipulation of justification detail.

Alternatively, jurors may be somewhat tolerant of eyewitness' error or seek other aspects of eyewitness testimony (such as reasoning for dubious credibility) when an eyewitness is inconsistent. As suggested by Jones et al. (2008), a detailed justification may convince skeptical jurors that the eyewitness is still motivated to tell the truth (though we did not assess this dimension directly), which results in greater credibility recovery. Importantly, we do note that in our experiment, even highly detailed justifications did not result in a complete recovery of credibility. In other words, even when eyewitnesses give a justification for confidence inflation, jurors are likely still sensitive to potential deficits in credibility. In real cases of wrongful convictions, other external evidence (inadequate forensic evidence, jailhouse informants, and so on) can also corroborate or increase reliance on inconsistent testimonies (Garrett, 2011). Thus, although we did not observe full recovery of eyewitness credibility in this experiment, real-world cases of apparent credibility recovery might be explained by non-testimonial factors. Despite these possibilities, our objective was primarily to test the effects of detail (a potentially easier dimension to measure than vividness, arousal, etc.) independent of the specific mechanisms underlying such effects. As we have demonstrated, justification detail does seem to influence perceptions of jurors.

In terms of practical applications for trial procedures, we join others (e.g. Wixted et al., 2015) in recommending that greater emphasis should be placed on initial eyewitness testimonies. We now know that initial eyewitness confidence under pristine, uncontaminated testing conditions reliably predicts eyewitness accuracy (Wixted et al., 2015; Wixted & Wells, 2017). However, eyewitnesses in real trials have often been exposed to investigator suggestions, conversations with co-eyewitness, news about the related event, and other factors that could contaminate the testimony. The present study as well as the previous studies (Douglass & Jones, 2013; Jones et al., 2008) suggest that in situations like these, 'contaminated' eyewitnesses can recover their credibility by giving justifications for confidence inflation. In current procedures where special attention is often not given to initial testimonies (Wixted & Wells, 2017), there exists a risk that an inconsistent eyewitness who gives a highly detailed justification could have undue influence on a trial judgement— even after the testimony has been contaminated and the predictive power of the initial eyewitness confidence weakened.

Based on the results of our studies and others, we have several recommendations for procedural changes. One possibility would be to conduct direct examinations at the time of initial eyewitness testimony, submitting the examination as trace evidence, like DNA samples. Whether fully introducing such procedure is possible or not, weighting a record of initial eyewitness testimony like trace evidence could enhance the probative value of eyewitness testimony. This is not to say that delayed testimony is inherently inaccurate or non-probative— Gilbert and Fisher (2006) showed that reminiscent (no answer at Time 1, but witness provided an answer at Time 2) accuracy was comparable to initial + reminiscent (same answer at Time 1 and Time 2) accuracy. Nevertheless, it is extremely difficult to distinguish accurate and contaminated memories, given what we know about the malleability of memory (e.g. misinformation effect; Loftus, 1975). Therefore, a record of initial 'pristine' eyewitness testimony is critical for evaluating eyewitness

reliability. Until research reveals whether and how we can distinguish between accurate versus inaccurate delayed testimony (or justified versus unjustified confidence inflation), the best practice might be to warn jurors to focus on eyewitnesses' initial statements and to be skeptical of later testimony.

905 Our results also show that credibility can be partially recovered when highly detailed (and potentially even less-detailed) justifications for confidence inflation are given. Particularly troubling is the fact that some credibility can be recovered even when the justification is unrelated to related crime details. These results should prompt greater scrutiny of justifications for confidence inflation, and potentially the introduction of formal cautions instructing jurors how to treat justifications. For instance, jurors might be cautioned
910 that just because an eyewitness later remembers additional unrelated details, this does not mean that their memory for critical related details has also improved.

Finally, there exist several limitations in our study. First, our sample size was not sufficient to detect subtler effects of justification detail and justification relatedness
915 (e.g. whether highly detailed justifications resulted in full credibility recovery or whether there were small differences between highly detailed inconsistent eyewitnesses and consistent eyewitnesses). Second, any conclusions about the relative effects of related and unrelated justifications are limited because we manipulated testimonial relatedness across separate experiments. In order to draw any strong conclusions about relative effects, one would need to manipulate relatedness within a single experiment.
920 Despite this limitation, our research provides important insight into features that might affect credibility recovery, namely, that eyewitnesses can recover some credibility by giving a highly detailed *related* or *unrelated* justification for confidence inflation.

Third, our manipulation of inconsistency, though it was similar to manipulations used
925 in previous studies, includes a potential internal validity threat. Specifically, in the inconsistent confidence conditions, the witness testified that their confidence at the time of the identification was 100%, and it is then revealed by records that it was 50%. Thus, we and other researchers manipulated not only eyewitness consistency but may have also inadvertently manipulated perceptions of eyewitness honesty. We concede that this is a possibility, and one that may limit the conclusions we draw about theoretical mechanisms
930 underlying credibility recovery. However, even if participants thought the witness was dishonest (instead of or in addition to being overconfident), we still found that witnesses were able to recover some of their credibility. This is especially troubling because it implies that eyewitnesses viewed as either inconsistent or dishonest could recover their credibility.
935 However, without measuring ratings of eyewitness honesty, we can't know for sure. It is also possible that less pronounced confidence inflation (e.g. 100% at trial and 75% at testimony) produces different patterns of results. For example, it is plausible in situations like these that less detailed justifications might allow for credibility recovery. An experiment manipulating degree of confidence inflation (e.g. 0% vs. 100%, 50% vs. 100%, 75% vs.
940 100%) along with relatedness and detail of justification would provide further insight into this possibility. Such a future experiment would also be valuable for determining whether confidence inflation is viewed as more continuous (i.e. linear decreases in credibility recovery) or categorical (i.e. qualitative shifts in participant perceptions).

945 Fourth, it may be argued that our operationalization of relatedness does not truly capture the distinction between *related* and *unrelated* information. Both our relatedness manipulations were technically unrelated to memory for the culprit directly relevant to

the identification. For instance, remembering what items the culprit stole does not necessarily mean one has a good memory for the culprit themselves. Further complicating thing is the fact that the definition of 'relatedness' can also change from crime to crime (e.g. in some cases, stolen items may be directly relevant to testimony). We adopted the relatedness manipulations that we did for two reasons: (1) this manipulation was used in prior research (Bell & Loftus, 1989) that we hoped to extend to the confidence inflation paradigm, and (2) we think that our manipulation is especially interesting from a practical standpoint. That is, if justifications not technically related to identification-relevant memory can lead to credibility recovery, this suggests that triers of fact may be influenced by a broad scope of justifications, including those that are not diagnostic of eyewitness credibility. It may be possible witnesses with highly detailed justifications are viewed as having 'good' memory in general (due to the detail given), and that jurors might extrapolate from this that the witness is credible. Nevertheless, the fact that both our relatedness manipulations were similar in this non-diagnosticity might explain the similar patterns of results we observed across experiments. Our results are consistent with this possibility, but still provide evidence that the mere provision of a highly detailed justification can soften negative juror perceptions of inconsistent eyewitnesses. However, our operationalization of relatedness did not involve information that was truly directly related to the identification of the culprit. As such, we make no claims about the generalizability of these results to justifications with information more directly related to the culprit's identity. Such justifications might follow patterns similar to those observed in research on the Featural Justification Effect (Cash & Lane, 2017; Dodson & Dobolyi, 2015, 2017). According to research on this effect, eyewitness confidence may be interpreted differently by jurors depending on the justification given for the confidence. Interestingly, justifications that cannot be evaluated (e.g. 'I'm certain it's him') may result in higher perceived confidence than justification that can be evaluated in terms of diagnosticity (e.g. 'I remember his chin'; Cash & Lane, 2017; Dodson & Dobolyi, 2015). Our research was not specifically designed with the Featural Justification Effect in mind, and did not involve features of the criminal. However, this effect could plausibly explain credibility recovery, as the high-detail justifications given by inconsistent eyewitnesses may have been difficult to evaluate in terms of diagnosticity. Future research could look at the influence of related and unrelated information about the crime and the perpetrator in a single study. Such research could also establish whether the Featural Justification Effect can be extended to non-featural aspects and confidence inflation.

Fifth and similarly, it is possible that participants viewed the unrelated testimonial information (i.e. the information about the woman who came into the store before the robber) as related. Participants may have thought that the woman was an accomplice, or otherwise relevant to the crime. We think this possibility is unlikely (as the woman is not mentioned anywhere in the testimony or interview). Even if participants viewed the unrelated information as related, the alternate operationalization of relatedness discussed here defines both of our conditions as unrelated. Despite this, we found evidence of credibility recovery.

Sixth, it remains unclear whether eyewitnesses' detailed justifications for confidence inflation would influence real judgements of guilt. In the present study, participants assessed the probability of the suspect's guilt using only the testimony. In real cases, non-testimonial evidence also informs judgements of guilt, and real jurors have more

motivation to closely scrutinize testimony. As a result, in real cases, highly detailed justifications may not result in the recovery of credibility that we observed here. In addition, jurors' judgements of probability of guilt based on eyewitness testimony may not translate directly into corresponding final verdicts. For example, Loftus (1975) found that even after hearing the discredited witness, 68% of participant-jurors voted for guilty (compared to 72% who heard a credible witness). Furthermore, Clancy and Bull (2015) found the opposite; eyewitness credibility influenced participant judgements of probability of guilt, but not their final verdict. Thus, work extending our manipulation of credibility recovery beyond the somewhat abstract judgements of credibility to concrete final verdicts is a logical next step. Jurors in real trials are supposed to be made aware of factors that can taint eyewitness testimony (e.g. post-identification feedback), Thus, another valuable avenue of research could examine the effects of warnings on perceptions of confidence-inflated eyewitnesses juror final verdicts. Finally, our examination of the recovery of eyewitness credibility was restricted to a limited situation because we manipulated only eyewitness confidence and description, and employed only one court script.

Despite these limitations, our findings have important practical implications, they suggest that inconsistent eyewitnesses can regain credibility in the eyes of the jury, for largely unjustified reasons. The amount of credibility recovered may depend partly on the amount of detail given with the justification for confidence inflation. This credibility recovery may occur independently of judgements based on specific testimonial characteristics (e.g. the mere presence of a highly detailed justification may result in recovered credibility). Ultimately, we suggest that jurors should be made aware of (1) the importance and probative value of initial testimony and (2) the effects of post-inflation justifications and how best to evaluate them.

Notes

1. We report analyses for the remaining characteristics in the Supplementary Material.
2. Determined via power simulations (10,000 simulated *t*-tests per effect size, with *N* fixed at 68).
3. Our data is available on OSF (<https://osf.io/59w3d/>).
4. Inclusion the excluded data made little difference the result that we reported below.
5. As mentioned above, we also conducted exploratory analyses of four additional aspects of eyewitness credibility. We report results for these in the Supplementary Material Tables S1 and S2.
6. See also Supplementary Material Tables S1 and S2 for detailed NHST ANOVA and pairwise results for all eight eyewitness characteristics.
7. Inclusion of these participants in analyses did not change the overall pattern of results.
8. See also Supplementary Material Tables S3 and S4 for detailed NHST ANOVA and pairwise results for all eight eyewitness characteristics.

Disclosure statement

Q3 No potential conflict of interest was reported by the author(s).

Q4 References

- Bell, B. E., & Loftus, E. F. (1988). Degree of detail of eyewitness testimony and mock juror judgements. *Journal of Applied Social Psychology*, 18(14), 1171–1192. <https://doi.org/10.1111/j.1559-1816.1988.tb01200.x>

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology*, 56(5), 669–679. <https://doi.org/10.1037/0022-3514.56.5.669>

Bornstein, B. H., O'Bryant, S., & Zickafoose, D. J. (2008). Intuitions about arousal and eyewitness memory: Effects on mock juror's judgments. *Law & Psychology Review*, 32, 109–133.

Q5

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. <https://doi.org/10.1023/A:1005523129437>

Bradfield, A., & McQuiston, D. E. (2004). When does evidence of eyewitness confidence inflation affect judgements in a criminal trial? *Law and Human Behavior*, 28(4), 369–387. <https://doi.org/10.1023/B:LAHU.0000039331.54147.ff>

Cash, D. K., & Lane, S. M. (2017). Context influences interpretation of eyewitness confidence statements. *Law and Human Behavior*, 41(2), 180–190. <https://doi.org/10.1037/lhb0000216>

Clancy, D., & Bull, R. (2015). The effect on mock-juror Decision-making of Power-of-Speech within eyewitness testimony and Types of Scientific Evidence. *Psychiatry, Psychology and Law*, 22(3), 425–435. <https://doi.org/10.1080/13218719.2014.960029>

Dodson, C. S., & Dobolyi, D. G. (2015). Misinterpreting eyewitness expressions of confidence: The featural justification effect. *Law and Human Behavior*, 39(3), 266–280. <https://doi.org/10.1037/lhb0000120>

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. <https://doi.org/10.1080/1068316X.2017.1284220>

Douglass, A. B., & Jones, E. E. (2013). Confidence inflation in eyewitnesses: Seeing is not believing. *Legal and Criminological Psychology*, 18(1), 152–167. <https://doi.org/10.1111/j.2044-8333.2011.02031.x>

Garrett, B. L. (2011). *Convicting the innocent: Where criminal prosecutions go wrong*. Harvard University Press.

Gilbert, J. A. E., & Fisher, R. P. (2006). The effects of varied retrieval cues on reminiscence in eyewitness memory. *Applied Cognitive Psychology*, 20(6), 723–739. <https://doi.org/10.1002/acp.1232>

Japanese Ministry of Justice. (2013). In Japanese. <http://www.moj.go.jp/content/001180941.pdf>

Jones, E. E., Williams, K. D., & Brewer, N. (2008). "I had a confidence epiphany!": Obstacles to combating post-identification confidence inflation. *Law and Human Behavior*, 32(2), 164–176. <https://doi.org/10.1007/s10979-007-9101-0>

Langer, E., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology*, 36(6), 635–642. <https://doi.org/10.1037/0022-3514.36.6.635>

Lindsay, D. S. (1994). Memory source monitoring and eyewitness testimony. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments*. Cambridge University Press.

Q6

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7(4), 560–572. [https://doi.org/10.1016/0010-0285\(75\)90023-7](https://doi.org/10.1016/0010-0285(75)90023-7)

Magnussen, S., Melinder, A., Stridbeck, U., & Raja, A. Q. (2010). Beliefs about factors affecting the reliability of eyewitness testimony: A comparison of judges, jurors and the general public. *Applied Cognitive Psychology*, 24(1), 122–133. <https://doi.org/10.1002/acp.1550>

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

Reyes, R. M., Thompson, W. C., & Bower, G. H. (1980). Judgmental biases resulting from differing availabilities of arguments. *Journal of Personality and Social Psychology*, 39(1), 2–12. <https://doi.org/10.1037/0022-3514.39.1.2>

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56(5), 356–374. <https://doi.org/10.1016/j.jmp.2012.08.001>

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–237. <https://doi.org/10.3758/PBR.16.2.225>

Shedler, J., & Manis, M. (1986). Can the availability heuristic explain vividness effects? *Journal of Personality and Social Psychology*, 51(1), 26–36. <https://doi.org/10.1037/0022-3514.51.1.26>

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology: Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, 25(1), 35–57. <https://doi.org/10.3758/s13423-017-1343-3>

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & Van Der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: comment on Bem (2011). *Journal of Personality and Social Psychology*, 100(3), 426–432. <https://doi.org/10.1037/a0022790>

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.

1085
1090 Q7 <https://doi.org/10.1037/0021-9010.83.3.360>

Wells, G. L., & Leippe, M. R. (1981). How do triers of fact infer the accuracy of eyewitness identifications? Using memory for peripheral detail can be misleading. *Journal of Applied Psychology*, 66(6), 682–687. <https://doi.org/10.1037/0021-9010.66.6.682>

Wise, R. A., Pawlenko, N. B., Safer, M. A., & Meyer, D. (2009). What US prosecutors and defence attorneys know and believe about eyewitness testimony. *Applied Cognitive Psychology*, 23(9), 1266–1281. <https://doi.org/10.1002/acp.1530>

Wise, R. A., & Safer, M. A. (2010). A comparison of what U.S. judges and students know and believe about eyewitness testimony. *Journal of Applied Social Psychology*, 40(6), 1400–1422. <https://doi.org/10.1111/j.1559-1816.2010.00623.x>

Wixted, J. T., Mickes, L., Clark, S. E., Gronlund, S. D., & Roediger, H. L. III. (2015). Initial eyewitness confidence reliably predicts eyewitness identification accuracy. *American Psychologist*, 70(6), 515–526. <https://doi.org/10.1037/a0039510>

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. <https://doi.org/10.1177/1529100616686966>

1105

1110

1115

1120

1125