

Investigator Sensitivity to Alibi Witness Inconsistency after a Long Delay

Heather L. Price^{*†} and Leora C. Dahl[†]

In two studies, mock investigators conducted a computer-based investigation of a crime involving an alibi witness who varied in the consistency of his statements taken 5 years apart. Investigators showed evidence of skepticism of alibi witness statements in which major contradictions (activity, location) were present, and some skepticism of statements in which minor (activity) details were contradictory. Entirely consistent statements were judged favorably, and reduced perceptions of suspect guilt (Study 2). The age of the alibi witness did not impact judgments of suspect guilt when children (6 years) and adults (25 years) were compared (Study 1, $N = 254$), or when children of different ages were compared (6, 8, 11 years; Study 2, $N = 234$). The present data suggest that investigators were relatively more sensitive to considerations of accuracy than honesty. Copyright © 2017 John Wiley & Sons, Ltd.

INTRODUCTION

Criminal investigators must evaluate the credibility of each piece of evidence encountered during an investigation. How such evaluations are made has increasingly been of interest for researchers studying decision-making bias in forensic settings (e.g., forensic confirmation bias; see Kassin, Dror, & Kukucka, 2013). Evaluations of person evidence (i.e., evidence based on reports of observations and direct experiences) may be particularly elastic and subject to substantial investigator discretion (Ask, Rebelius, & Granhag, 2008) and thus provide a fruitful avenue for understanding investigator decision-making. There are several interrelated factors that may influence how bias can impact the evaluation of such elastic evidence, including investigator and witness motivation and the credibility or reliability of the evidence itself. Yet we do not know how such variables, in combination, will influence investigator judgments. In the present experiments, we explored mock investigator decisions related to a particular type of person evidence, alibi witness statements that varied in consistency/contradictions over a long period of time. Though inconsistencies in other types of evidence (e.g., eyewitness evidence) have been explored in the literature, the evaluation of inconsistent and/or contradictory alibi witness statements has been neglected.

Alibi witness statements can form a critical basis for a suspect's defense. If a suspect can prove that he or she was elsewhere at the time of the crime, logic should prevail and the suspect should be ruled out as the offender. However, the notion of proving one's

* Correspondence to: Heather L. Price, Canada Research Chair, Department of Psychology, Thompson Rivers University, 900 McGill Road, Kamloops, British Columbia, Canada, V2C 0C8. E-mail: hprice@tru.ca
†Psychology, Okanagan College, 1000 KLO Rd, Kelowna, British Columbia V1Y 4X8, Canada

innocence is not as straightforward as it initially appears. When the corroborating alibi information is of a physical nature (e.g., surveillance video), it is subject to less investigator discretion; however, when the corroborating evidence that supports a suspect's alibi is the statement of an alibi witness (a person who confirms that the suspect was elsewhere), that statement is subjected to the same credibility analysis as any other witness testimony. Yet the study of how evaluations of alibi witness statements are made is only a recent focus of researchers' attention.

Evaluating Person Evidence

The study of the influence of another type of person evidence, eyewitness evidence, on legal decision-making has been ongoing for decades and has clearly established that mock jurors find such evidence to be highly persuasive (e.g., Semmler, Brewer, & Douglass, 2012). However, there is at least one condition under which eyewitness evidence is viewed with skepticism by both mock jurors and police officers: when an eyewitness provides inconsistent evidence (see, e.g., Berman & Cutler, 1996; Berman, Narby, & Cutler, 1995; Brewer *et al.*, 1999; Krix, Sauerland, Lorei, & Rispen, 2015; Oeberst, 2012; Yuille & Cutshall, 1986). It is reasonable to assume that the negative perception of an inconsistent witness is not limited to people recalling witnessing a crime. People who report the same event in different ways must be wrong on at least one of the occasions, and thus we may presume that they are inaccurate and unreliable. However, the literature has firmly concluded that the relationship between accuracy and consistency is weak (e.g., Brewer *et al.*, 1999; Fisher & Cutler, 1995). Particularly when time has passed between the event in question and recall, some level of inaccuracy or inconsistency is to be expected in autobiographical memory reports, and specifically in alibi witness statements (Olson & Wells, 2004).

Evaluations of witness statement inconsistencies involve more than just an understanding of memory processes. A critical consideration in the interpretation of inconsistent eyewitness testimony is that eyewitnesses most often provide incriminating evidence, and thus it is likely that investigators will be more forgiving of minor inconsistencies if the witness provides evidence that confirms their suspicions (i.e., investigator bias; see Kassin *et al.*, 2013). Other forms of person evidence, however, may not receive an investigator's benefit of the doubt, particularly if that evidence is exculpatory.

Given that alibis are a form of exculpatory evidence, investigators may be particularly likely to critically evaluate such evidence if it is incompatible with their perception of the suspect. Indeed, some researchers have suggested that police investigators may approach alibi evidence with more skepticism than they would other evidence and may be more prone to look for inconsistencies and signs of deception in those providing such evidence (e.g., Burke & Turtle, 2003). Thus, inconsistencies in alibi witness statements may be weighed even more heavily, at least by investigators, than inconsistencies in eyewitness evidence.

There are likely other important differences between investigator evaluations of eyewitness and alibi witness statements. When an eyewitness witnesses a crime, the witness may be more likely to be aware that a crime is being committed and can thus encode the event as important. In contrast, when an alibi witness experiences an event about which they will be later asked to provide a statement, he or she may be less likely

to encode the event as noteworthy. These differences lead to conflicting hypotheses about how bias may influence investigator evaluations of alibi witness testimony. On the one hand, alibis are viewed with skepticism and may be subject to heightened scrutiny by investigators. Conversely, given the encoding context, alibi witnesses may not be expected to provide the same level of detail as a person present at a crime, and if this is understood investigators may be more forgiving of a lack of detail or the presence of inconsistencies or contradictions.

Who Provides the Alibi Evidence

The body of work examining suspects who provide their own alibi information has clearly found that, even though evaluators tend to believe that a suspect's true alibi should be consistent over time (see Burke, Turtle, & Olson, 2007; Culhane & Hosch, 2012; Dysart & Strange, 2012) and that a changed alibi may be an indication of deception (Dysart & Strange, 2012), a consistent alibi is very difficult to provide (Leins & Charman, 2016; Olson & Charman, 2012; Strange, Dysart, & Loftus, 2014). Strange *et al.* (2014) recently characterized the provision of an alibi as a test of our autobiographical memory. Given the now very clear evidence about the malleability of autobiographical memory and reporting errors as an outcome of natural memorial process, memory researchers well understand that fallibility and inconsistency is to be expected and that such errors are often independent of intent to deceive; however, the general public appears much less aware of these memory "facts" (Benton *et al.*, 2006). The laboratory work specific to "suspects" providing their own alibis—and the inconsistencies observed in this work—led Strange *et al.* (2014) to conclude that alibi consistency should not be used as a proxy for accuracy.

The extant research into the credibility of inconsistent alibis focuses on alibis provided by suspects. However, the motivations behind (i) a suspect providing an alibi and (ii) alibi witnesses corroborating an alibi are likely to be different (e.g., they will likely vary in levels of self-interest). Yet they are still both forms of exonerating evidence and likely subject to similar levels of investigator scrutiny and cynicism. We do not know, however, if alibi witness statements will be evaluated similarly to a suspect's alibi statements.

A further consideration in the "who" of alibi evidence is the relationship between an alibi witness and the suspect. The most replicated finding in the alibi literature is that who provides the alibi evidence can be critical: those with familial relations to the defendant are perceived as less credible alibi witnesses than those unrelated to the defendant (e.g., Culhane & Hosch, 2004; Dahl & Price, 2012; Olson & Wells, 2004). This pattern is hypothesized to be the result of a perception of incentive to lie: someone close to the suspect may have motivation to lie to protect the suspect, whereas someone with little or no relationship to the suspect should be less motivated.

The alibi literature is not the only area in which witness honesty has been a focus. In the broader credibility assessment literature, witness credibility is often characterized as lying along two dimensions: honesty and accuracy (e.g., Ross, Jurden, Lindsay, & Keeney, 2003). Contrasting considerations of accuracy (e.g., cognitive ability, quality of viewing conditions) with judgments of honesty (e.g., motivations) are central to credibility judgments. For instance, when comparing witnesses of different ages, children are generally perceived as more honest witnesses (i.e., have less reason or ability to lie), whereas adults are perceived as more cognitively competent or accurate witnesses (i.e., have greater ability to perceive and report event details accurately)

(Bottoms, 1993; Connolly, Price, & Gordon, 2010; Ross *et al.*, 2003). The honesty dimension has recently been highlighted when comparing alibi witnesses of different ages. Under conditions that prime considerations of honesty, children appear to be particularly believable as alibi witnesses when compared with adults: Dahl and Price (2012) found clear evidence that a child is viewed as a more credible alibi witness than an adult when honesty is a salient issue in the alibi statement. This perception appears to be driven by the inherent belief that children have more difficulty telling convincing lies and may also be less motivated to lie for a suspect than adults (Ross *et al.*, 2003). However, in the work of Dahl and Price (2012), honesty was highlighted as a salient issue, the alibi witnesses only stated the alibi once, and accuracy was not emphasized. Thus, depending on which dimension is more salient to a particular witness's testimony, a child or an adult may be perceived as more credible. Of primary interest in the present work is conditions under which perceptions of honesty are contrasted with inconsistencies/accuracy; it is unclear which of these important factors will carry more weight with investigators.

THE PRESENT RESEARCH

In the present studies, we explored the influence of the consistency of child and adult alibi witness statements taken 5 years apart. In two studies, undergraduate participants played the role of an investigator in a criminal investigation, searched through a computer database of suspects, and reviewed a case summary and videotaped alibi witness statement(s). The alibi witness statements varied in the level of consistency between repeated statements as well as in the age of the alibi witness.

STUDY 1

Participants

Two hundred and fifty-four undergraduate participants between the ages of 17 and 62 (M age = 20.78 years, SD = 4.64; n = 53 males) participated individually for course credit. Participants were randomly assigned to experimental condition in a 2 (alibi witness age: 6 then 11 years, 25 then 30 years) \times 5 (alibi consistency: consistent, minor contradiction, major contradiction, early only, delayed only) between-subjects design. Cell sizes were approximately equal, both early only conditions had 24 participants and the adult minor and major contradiction conditions had 25 participants, while the remaining cells had 26 participants each).

Materials and Procedure

Participants were informed that they would assume the role of a police officer and would conduct a mock investigation of a crime. Participants first read a mock police file that described a robbery and an eyewitness description of a male culprit described as approximately 50 years old. The description of the crime and culprit were based on descriptions obtained in a prior study (see Dahl, Lindsay, & Brimacombe, 2006). Participants then received instructions for using a computer database to search through

several potential suspects who had prior arrests on file. They were informed that the culprit might not be in the database, but that they must explore all suspects in the database before they made their decision to either reject all suspects or further investigate their chosen suspect. The database provided information about each suspect's physical description, prior criminal record, current employment, and registered vehicles. The database was rigged to make one of the suspects the best fit based on physical description (i.e., matched a general physical description) and prior criminal record (previously convicted of committing a burglary). Sixteen participants were excluded from the study for not selecting the appropriate suspect (from the complete N of 270).

Next, investigators viewed either one or two alibi videos. When investigators viewed two videos, the videos depicted the same actor, described as the son of the suspect, providing alibi statements that were filmed 5 years apart (aged 6 then 11 years or 25 then 30 years). Participants were informed via the video that the first statement was filmed less than a week after the day in question and the second was filmed 5 years after the day in question. The videos varied in the consistency between the two alibi statements. The statements were characterized as follows. (1) Consistent: general activities covering entire day. (2) Minor contradiction: differs from original alibi on activity, but location remains the same (i.e., at the suspect's house watching movies or watching hockey playoffs). (3) Major contradictions: differs from original alibi on activity *and* location (i.e., first reported at suspect's house watching movies, later reported camping in the woods with the suspect). Participants who viewed two videos completed a distractor task between viewing the videos (a five-minute logic puzzle) to prevent direct comparison of alibi scripts. To provide control conditions, some investigators only viewed one video: (4) only the early alibi (taken less than a week after the day in question) or (5) only the delayed alibi (taken 5 years after the first statement). Thus, this study was a 2 (alibi witness age: 6 then 11 years, 25 then 30 years) \times 5 (consistency: consistent, minor contradiction, major contradiction, early only, delayed only) between-subjects design. Participants then rated the probability that their suspect committed the crime (from 1 to 100%, with higher numbers indicating a higher degree of certainty) and indicated whether or not they would be willing to arrest the suspect with the current information, and if not what evidence they would require to make an arrest. Finally, participants were asked to rate the credibility, honesty, and accuracy of the alibi witness on a scale from 1 (not at all) to 6 (very).

Results

Probability Suspect Committed the Crime

Participant ratings of the probability that the suspect committed the crime were entered into a 2 (alibi witness age) \times 5 (consistency) analysis of variance (ANOVA). There was no main effect of alibi witness age, $F(1, 253) = 1.21$, $p = .27$, $\eta_p^2 = .01$, nor an interaction between alibi witness age and consistency, $F(4, 244) = 0.81$, $p = .52$, $\eta_p^2 = .01$, but there was a main effect of alibi statement consistency, $F(4, 253) = 8.53$, $p < .001$, $\eta_p^2 = .12$. Post hoc LSD tests indicated that viewing the major contradiction resulted in higher guilt ratings than all other conditions (p values $< .01$). No other condition comparisons were statistically significant, p values $> .11$. Refer to Table 1 for descriptive data across conditions.

Table 1. Probability that the suspect committed the crime in Study 1

Statement consistency	Alibi witness age						Total
	6 then 11 years			25 then 30 years			
	Mean (SD)	95% CI		Mean (SD)	95% CI		
		LL	UL		LL	UL	
Early only	64.58 (22.30)	56.29	72.88	72.71 (17.86)	64.41	81.00	68.65 (20.41)
Delayed only	63.73 (22.67)	55.76	71.70	71.27 (21.48)	63.30	79.24	67.50 (22.20)
Consistent	63.23 (22.28)	55.26	71.20	61.58 (19.39)	53.61	69.55	62.40 (21.26)
Minor contradiction	70.50 (24.82)	62.53	78.47	67.16 (20.72)	59.03	75.29	68.86 (22.74)
Major contradictions	82.96 (12.15)	74.99	90.93	86.52 (18.64)	78.39	94.65	84.71 (15.61)
Total	67.50 (22.20)	65.41	72.56	70.40 (21.80)	68.23	75.47	

Note: Probability ratings ranged from 1 to 100%.

Final Arrest Decision

Age, consistency, and the age \times consistency interaction were entered into a logistic regression model to predict arrest decisions. The model was significant, $\chi^2(9) = 21.91$, $p = .01$, Nagelkerke $R^2 = .11$, and consistency was a significant predictor of arrest decision, $p < .05$. Final arrest decisions differed across alibi witness statement conditions, with statements that had major contradictions resulting in significantly higher arrest rates (67%) than all other conditions (all other rates $< 40\%$; follow-up z tests, z values > 2.77 , p values $< .001$). There was no effect of age and no interaction between age and consistency, p values $> .33$.

Credibility Evaluations of Alibi Witnesses and Eyewitnesses

Supporting the pattern of findings observed in response to the question of the probability that the suspect committed the crime, when asked to rate the overall credibility of the alibi witness, there was no effect of alibi witness age, $F(1, 253) = 1.71$, $p = .19$, $\eta_p^2 < .01$, and no interaction between witness age and statement consistency, $F(4, 244) = 0.16$, $p = .96$, $\eta_p^2 < .01$, but there was a strong main effect of consistency on ratings of alibi witness credibility, $F(4, 244) = 5.59$, $p < .001$, $\eta_p^2 = .09$. Post hoc LSD tests indicated that the major contradiction statements were rated as significantly less credible than all other conditions (p values $< .02$) except for the delay only alibi ($p = .12$), and that the consistent alibi statements were rated as significantly more credible than the major contradiction and delay conditions (p values $< .02$), but not the minor contradiction and early only conditions (p values $> .31$). Refer to Table 2 for descriptive data.

Participants were also asked to rate the accuracy and honesty of the alibi witnesses. For ratings of alibi witness accuracy, there was no main effect of age nor an interaction between age and consistency, F values < 2.94 , p values $> .09$, but there was a strong main effect of consistency, $F(4, 244) = 11.78$, $p < .001$, $\eta_p^2 < .16$, such that consistent alibi witnesses were rated as more accurate than all other alibi witnesses, p values $< .01$, and major contradiction alibi witnesses were rated as less accurate than all other alibi witnesses, p values $< .01$. No other accuracy comparisons differed significantly.

Table 2. Alibi witness statement credibility ratings in Study 1

Statement consistency	Alibi witness age						Total
	6 then 11 years			25 then 30 years			
	Mean (SD)	95% CI		Mean (SD)	95% CI		
	LL	UL		LL	UL		
Early only	2.46 (1.22)	2.02	2.90	2.38 (1.28)	1.94	2.81	2.42 (1.24)
Delayed only	2.12 (0.65)	1.70	2.54	2.04 (0.82)	1.62	2.46	2.08 (0.74)
Consistent	2.69 (1.16)	2.72	3.11	2.58 (1.10)	2.16	3.00	2.63 (1.21)
Minor contradiction	2.65 (1.26)	2.23	3.07	2.32 (0.99)	1.89	2.75	2.49 (1.14)
Major contradictions	1.88 (1.14)	1.46	2.31	1.60 (1.12)	1.17	2.03	1.75 (1.13)
Total	2.27 (1.12)	2.17	2.55	2.18 (1.11)	1.99	2.37	

Note: Credibility ratings ranged from 1 to 6.

For ratings of honesty, there was again a main effect of consistency, $F(4, 244) = 14.09$, $p < .001$, $\eta_p^2 = .19$, and also a main effect of age, $F(1, 253) = 22.12$, $p < .001$, $\eta_p^2 = .08$, but no interaction between the two variables, $F(4, 244) = 0.92$, $p = .45$, $\eta_p^2 = .02$. Children were perceived as significantly more honest ($M = 2.91$, $SD = 1.22$) than adults ($M = 2.29$, $SD = 1.10$). The main effect of consistency was driven by the major contradiction alibi witnesses, which were rated as significantly less honest than all other consistency conditions, p values $< .01$. The only other observed difference was that participants in the delayed only condition rated the witness as significantly less honest than participants in the minor contradiction condition ($p = .04$).

Consistency Check

Participants were asked to indicate whether they noticed any differences between alibi witness statements and, if so, how concerned they were about the differences (from 1 to 7, with 7 most concerned). All participants in the major contradictions condition reported noticing differences, 92% in the minor contradiction condition did, and 92% of participants in the consistent alibi condition reported noticing differences. Although scripts in the consistent condition did not differ substantively between the two videos, participants may have responded to this question considering the appearance of the actor or other interviewer characteristics. In response to the question about level of concern about the inconsistencies, there was a significant difference across consistency conditions, $F(2, 145) = 34.17$, $p < .001$, $\eta_p^2 = .32$, with post hoc LSD tests indicating that all comparisons differed significantly (p values $< .05$). Participants who witnessed an alibi witness with major contradictions were most concerned about the inconsistencies ($M = 6.25$, $SD = 1.25$), followed by the minor contradiction condition ($M = 4.62$, $SD = 1.60$), and then the consistent condition ($M = 4.04$, $SD = 1.30$).

Study 1 Discussion

Mock investigators were sensitive to inconsistencies in alibi witness statements, as they have been with inconsistencies in eyewitness testimony (see, e.g., Brewer *et al.*, 1999). It was particularly interesting to note that participants clearly differentiated degrees of inconsistency. Contradictions regarding activity and location of an alibi witness

statement were highly problematic for investigators, and the suspect in this major contradiction condition was rated as more likely to be guilty and more likely to be arrested, and the alibi witness in this condition was rated as less credible, accurate, and honest than in all other conditions. In line with what one would hope based on our understanding of memory processes, participants were more forgiving of a minor contradiction in alibi witness statements provided across a five-year gap and rated witnesses with such minor contradictions similarly to witnesses who were entirely consistent.

The lack of difference, on any measure, between the early alibi statement and the delayed alibi statement is also worthy of note. Statements taken either less than a week after the day in question or 5 years after the first statement would certainly come from memories of different strength. However, it is important to note that mock investigators were not explicitly informed that this was the only statement taken from the witness. It may have been unclear to the mock investigators if the delayed alibi was simply a re-statement of a previously taken statement. At face value, the lack of difference between these two conditions indicates that mock investigators did not consider the impact of the passage of time on memory. In the judgments of credibility and honesty, there were hints that most investigators were skeptical of the motivation behind the delayed alibi witness, which may be worthy of further investigation.

In a further interesting comparison, consistent statements were rated similarly to both single statement conditions (early only, delayed only). This finding may indicate that mock investigators expected that two statements by the same witness would be consistent as a baseline, and only demoted a witness's statement pair when there were signs of weakness or unreliability. Put another way, mock investigators may believe that a second statement provided by the same witness should be a re-statement of evidence already received, and thus they do not credit the witness for consistency. Investigators may attribute consistency to non-memorial processes such as reviewing the previously provided statement or remembering what had been previously said to investigators (rather than recalling the events themselves).

Mock investigator ratings of varying degrees of consistency in alibi witness statements were similar to patterns observed in the eyewitness literature and indicate that parallels may exist between evaluations of the two types of person evidence. However, though we established that consistency mattered, it is still unclear *why* consistency mattered. One way in which we can explore the mechanisms of how consistencies impact perceptions of guilt is through manipulating conditions that emphasize relative accuracy and honesty. That is, if conditions that prime considerations of accuracy are presented, perhaps consistency, or a lack thereof, will matter more. Conversely, if conditions that prime consideration of honesty are included, perhaps consistency will matter relatively less. We included both adults and children as alibi witnesses in Study 1, but found no differences between these two age groups. It may be the case that inconsistencies were so salient to mock investigators that age of the alibi witness was essentially washed out. Supporting this possibility, though children were rated as more honest than adults, overall suspect guilt ratings did not differ.

Although no age differences were observed in Study 1, in which a very young child was compared with an adult, thus maximizing potential differences in honesty perception, in Study 2 we sought to explore a wider range of child ages. Prior research has demonstrated that the most credible child witness is approximately 8 years old (Wright, Hanoteau, Parkinson, & Tatham, 2010), likely because this age balances the need for both honesty (too young to lie convincingly) and accuracy (old enough to be capable

of reporting reliable information). Thus, we selected an 8-year-old child as our mid-point. We then presented alibi witnesses that were a couple of years on either side of this age (6 and 11 years) to explore how a relative perception of honesty (6 years) and accuracy (11 years) might impact evaluations of alibi statements that varied in consistency. In Study 2, we opted to include only the conditions that involved two interviews from Study 1 (excluded the single-statement conditions). We did not develop a specific hypothesis about the effect of contradictions on judgments of child alibi witnesses of varying ages. On the one hand, we could predict an additive effect of contradictions and concerns about accuracy based on age. Conversely, we could also predict that any additive effect could be neutralized when a witness was also considered to be high in honesty. In the second study, we were interested in which of these factors, contradictions or the relative perception of children's abilities, would be weighed more heavily by mock investigators.

STUDY 2

Participants and Design

Two hundred and thirty-four undergraduate participants between the ages of 17 and 62 years (M age = 20.78 years, SD = 4.64; n = 53 males) participated individually for course credit. The experiment was the same as in Study 1, with three exceptions. (i) Three consistency conditions were included: major contradictions, minor contradiction, and consistent. (ii) Adult alibi witnesses were excluded, and three different ages of child alibi witnesses were examined (6 then 11 years; 8 then 13 years; 11 then 16 years). (iii) Participants were asked to rate the probability that the suspect committed the crime at two time points, once before and once after viewing the alibi witness videos (rather than one time point as in Study 1). Thus, this was a 3 (consistency) \times 3 (alibi witness age) \times 2 (pre-, post-alibi videos) design with consistency and age as between-subjects variables and 26 participants per cell.

Results

Probability Suspect Committed the Crime

Participant ratings of the likelihood that the suspect committed the crime were entered into a 2 (pre-, post-viewing) \times 3 (consistency) \times 3 (age) mixed model ANOVA. Descriptive data are presented in Table 3. As with Study 1, there was no effect of age, $F(2, 224) = 1.07$, $p = .34$, $\eta_p^2 = .01$, nor an interaction between alibi witness age and consistency, $F(4, 224) = 0.58$, $p = .68$, $\eta_p^2 = .01$. There was also a marginal main effect of pre-, post- ratings, $F(1, 224) = 3.85$, $p = .05$, $\eta_p^2 = .02$. However, there was again a strong effect of consistency, $F(2, 224) = 15.30$, $p < .001$, $\eta_p^2 = .12$, on guilt ratings. This main effect was qualified by an interaction between pre-, post- ratings and consistency, $F(2, 224) = 44.64$, $p < .001$, $\eta_p^2 = .29$.

To ensure that pre-evidence ratings did not differ across consistency conditions, despite random assignment, we first performed a one-way ANOVA to explore pre-viewing ratings across consistency conditions. There was no difference across consistency conditions, $F(2, 233) = 0.98$, $p = .38$, $\eta_p^2 = .01$. We then explored the change in

Table 3. Probability that the suspect committed the crime in Study 2

Statement consistency	Alibi witness age											
	6 then 11 years				8 then 13 years				11 then 16 years			
	Mean (SD)	LL	UL	95% CI	Mean (SD)	LL	UL	95% CI	Mean (SD)	LL	UL	95% CI
Consistent	Pre	61.85 (16.69)	53.76	69.94	57.50 (24.01)	49.41	65.59	65.38 (18.16)	57.29	73.48		
	Post	49.60 (20.81)	40.60	58.59	49.92 (27.97)	40.83	58.92	51.54 (16.66)	42.54	60.53		
	Change	-12.25 (19.78)	-18.98	-5.52	-7.58 (16.93)	-14.31	-0.84	-13.85 (17.22)	-20.58	-7.11		
Minor contradiction	Pre	68.46 (20.73)	60.37	76.55	57.92 (18.97)	49.83	66.01	64.73 (24.29)	56.64	72.82		
	Post	67.38 (30.14)	58.39	76.38	60.00 (23.21)	51.01	69.00	71.92 (25.02)	62.93	80.92		
	Change	-1.08 (18.68)	-7.81	5.66	2.08 (16.92)	-4.66	8.81	7.19 (16.19)	0.46	13.93		
Major contradictions	Pre	67.31 (19.43)	59.22	75.40	64.85 (24.00)	57.19	73.69	66.58 (20.45)	58.49	74.67		
	Post	83.58 (14.02)	74.58	92.57	82.76 (18.26)	73.59	91.93	78.62 (27.77)	69.62	87.61		
	Change	16.27 (18.06)	9.54	23.00	17.32 (17.30)	10.45	24.19	12.04 (15.33)	5.31	18.77		
Total												

Note: Probability ratings ranged from 1 to 100%.

pre- and post-viewing ratings in each consistency condition with paired sample t tests. Consistent alibi statements significantly reduced perceptions of suspect guilt, $t(77) = 5.51, p < .001$, Cohen's $d = -0.54$, whereas minor contradictions did not change perceptions of suspect guilt, $t(77) = -1.39, p = .17$, Cohen's $d = 0.11$, and major contradictions significantly increased perceptions of suspect guilt, $t(76) = -7.90, p < .001$, Cohen's $d = 0.74$.

Final Arrest Decision

Age, consistency, and the age \times consistency interaction were entered into a logistic regression model to predict arrest decisions. The model was statistically significant, $\chi^2(8) = 43.37, p < .01$, Nagelkerke $R^2 = .23$, and consistency was a significant predictor of arrest decision, $p < .01$. Final arrest decisions again differed across alibi witness statement conditions, with participants most likely to decide to arrest the suspect when the alibi witness contradicted himself on major details (68%) compared with minor details (37%) and, finally, a consistent alibi (22%), z values >2.00 , p values $< .05$. There was no effect of age and no interaction between age and consistency, p values $> .13$.

Credibility Evaluations of Alibi Witnesses and Eyewitnesses

As with Study 1, when participants rated the overall credibility of the alibi witness, there was no main effect of age, $F(2, 232) = 0.12, p = .89, \eta_p^2 = .001$, and no interaction between age and consistency, $F(4, 224) = 0.04, p = .99, \eta_p^2 = .001$, but there was a strong effect of consistency on alibi witness credibility evaluations, $F(2, 232) = 15.22, p < .001, \eta_p^2 = .12$. Post hoc LSD tests indicated that consistent alibi witnesses were rated as significantly more credible ($M = 3.06, SD = 1.33$) than an alibi witness who contradicted himself on minor details ($M = 2.62, SD = 1.42$), who in turn were rated as significantly more credible than an alibi witness who contradicted himself on major details ($M = 1.88, SD = 1.20$; all p values $< .05$).

Participants were also asked to rate the accuracy and honesty of the alibi witnesses. For ratings of alibi witness accuracy, there was no main effect of age, $F(2, 232) = 0.51, p = .60, \eta_p^2 = .01$, or interaction between age and consistency, $F(4, 224) = 0.39, p = .81, \eta_p^2 = .01$. There was, however, a strong main effect of consistency, $F(2, 232) = 34.13, p < .01, \eta_p^2 = .23$. Not surprisingly, consistent witnesses were perceived as most accurate, followed by minor contradiction and then major contradiction witnesses (all p values $< .01$).

For ratings of alibi witness honesty, there was no effect of age nor an interaction between age and consistency (F values $<1.00, p$ values $> .5$). However, there was a main effect of consistency, $F(2, 232) = 21.26, p < .01, \eta_p^2 = .16$. Consistent witnesses were rated as most honest, followed by minor contradiction and major contradiction witnesses (all p values $< .01$).

Study 2 Discussion

Largely consistent with Study 1, there were no age differences across any dependent variable, thus supporting the notion that accuracy had a stronger influence on mock investigator decision-making than relative honesty. Study 2 was also consistent with Study 1 in the importance investigators placed on contradictions. Interestingly, participants in

Study 2 differentiated between major and minor contradictions on every dependent measure. Mock investigators were very suspicious of major contradictions, and to a lesser extent a minor contradiction. The pre-, post- assessment of guilt clearly demonstrates this pattern: viewing an alibi witness with major contradictions in his statement increased perception of suspect guilt, viewing an alibi witness with a minor statement contradiction did not change perceptions of suspect guilt, and viewing an alibi witness who provided two entirely consistent statements reduced perceptions of suspect guilt.

The gradient in probability of guilt ratings observed in Study 2 presents a clear picture to legal personnel who work with witnesses. Even though a memory expert would anticipate at least some variation in a truthful recounting of an event after 5 years, evaluators of these statements are likely to be less forgiving.

GENERAL DISCUSSION

The findings from the present studies indicate that investigators are skeptical of major contradictions between alibi witness statements taken 5 years apart. There was also evidence that a minor contradiction is concerning to mock investigators, relative to entirely consistent statements, but it was certainly less so than major contradictions. In addition, contradictions in alibi witness evidence were so powerful that they overshadowed an alibi witness characteristic that has previously had a strong impact on investigator decisions: the age of the alibi witness.

(a Lack of) Alibi Witness Age Differences

We had hypothesized that investigators may be particularly skeptical about alibi witness evidence because it is a form of exculpatory evidence (Burke & Turtle, 2003). However, though investigators may have approached the evidence with general skepticism, they did not differentiate among the value of statements provided by witnesses of different ages, which indicates that they did not seize all available justifications for casting doubt on the alibi witness.

In Study 2, we anticipated that presenting evidence that primed relative differences in perceptions of honesty (younger children) and accuracy (older children) would influence mock investigators' evaluations of suspect guilt. However, we found no effects of age on any of the dependent variables. The lack of age effects may be a result of at least a couple of possible reasons. The first possibility is that our age range was not sufficiently wide to clearly differentiate the children of different ages and induce priming of either honesty or accuracy. The children in Study 2 spanned an age range of 5 years at the end points, but age was not explicitly provided to participants, and when not presented side by side (age was a between-subjects manipulation) their age differences may have simply been too narrow to elicit differences in the salience of honesty as an important witness trait. The inclusion of a second alibi statement taken at 11, 13, and 16 years in particular may have minimized any age effects. Indeed, in Study 2, when the comparison was between children of different ages, there were no differences in witness honesty (or accuracy) ratings. Prior research has found child alibi witnesses to be more credible than adult alibi witnesses when there is a close personal relationship between the suspect and the alibi witness (such as the suspect's son in the present research; Dahl

& Price, 2012). It appears that this implied honesty is distributed across children of all ages presented in the current work.

A second possibility, supported by the lack of differences in Study 1, is that contradictions overrode any potential differences between witnesses of different ages and the accompanying accuracy–honesty tradeoff. This possibility is supported by the ratings of alibi witness accuracy and honesty in Study 1, the study in which the age difference between alibi witnesses was maximized. When children were compared with adults, children were rated as significantly more honest than adults (there were no differences in accuracy), but this difference did not translate into lower ratings of suspect guilt.

Balancing Honesty and Accuracy

Of course, it is unlikely that perceptions of honesty will override drastic errors of accuracy. With the present manipulations, alibi witnesses with major statement contradictions (in location and activity) were devalued, regardless of the perception of relative honesty. This is an important finding because it suggests that, regardless of the motivation, investigators will balance honesty and accuracy to arrive at a global evaluation of the evidentiary value. Thus, even for a witness who is perceived as highly honest (e.g., a young child), a baseline level of accuracy still must be met in order for that honest evidence to be considered valuable.

In future research, it is well worth further highlighting honesty as a salient feature so that it may be more directly contrasted with different levels of accuracy to see when the balance tips in favor of either characteristic. In addition to manipulating salience of honesty through age of the alibi witness or other individual characteristics (such as the previously powerful manipulations of alibi witness–suspect relationship), honesty could be manipulated through a change to a witness’s statement that is perceived as more (or less) strongly supportive of innocence. If an alibi witness, who is already the subject of investigator skepticism, changes his statement in such a way that it would provide incontrovertible evidence (if believed) of innocence, this witness may be perceived as changing his story to suit the suspect’s needs. Culhane and Hosch (2012) recently demonstrated that such changes in a *suspect’s* alibi statement suffer from this fate: when a suspect simply changed an alibi—even to a stronger alibi of the same nature (e.g., by adding a corroborator)—changes led to higher ratings of suspect guilt. Thus, it is not the case that it is only inconsistent alibis that change in detail, but also alibis that change without affecting the alibi details themselves that are viewed cynically.

Finally, it is also worth considering manipulating the strength of all evidence presented in future studies. Overall alibi witness credibility ratings were low, and guilt ratings were high, suggesting that alibi witnesses were generally not believed and suspects were generally perceived as guilty. Alibi witness evidence is not a particularly persuasive form of evidence, but can likely be strengthened by, for example, the inclusion of additional detail in the statement, the lack of a prior relationship between suspect and alibi witness, or the presence of corroborating evidence.

CONCLUSION

The present studies confirm that accuracy is a powerful driver of judgments of alibi witness statements. Although honesty and accuracy have not been similarly contrasted in

evaluations of eyewitness evidence, the present work suggests that there may be parallels among evaluations of these two types of person evidence. It is clear that the context, manipulated by the motivations of the witness, is critical to the influence of alibi witness statements on perceptions of the suspect. Thus, future work exploring the persuasiveness of alibi witness statements must systematically and thoughtfully consider the context under which the evidence is presented.

ACKNOWLEDGMENTS

This research was supported by a Social Sciences and Humanities Research Council of Canada Insight Grant to the first author. The authors thank Brittany Whiting, Amanda Oliver, Jordan Nixdorf, and Caitlin Hunter for assistance with data collection.

REFERENCES

- Ask, K., Rebelius, A., & Granhag, P. A. (2008). The “elasticity” of criminal evidence: A moderator of investigator bias. *Applied Cognitive Psychology, 22*, 1245–1259. doi:10.1002/acp.1432.
- Benton, T. R., Ross, D. F., Bradshaw, E., Thomas, W. N., & Bradshaw, G. S. (2006). Eyewitness memory is still not common sense: Comparing jurors, judges and law enforcement to eyewitness experts. *Applied Cognitive Psychology, 20*, 427–443. doi:10.1002/acp.1171.
- Berman, G. L., & Cutler, B. L. (1996). Effects of inconsistencies in eyewitness testimony and mock-juror judgements. *Journal of Applied Psychology, 81*, 170–177. doi:10.1037/0021-9010.81.2.170.
- Berman, G. L., Narby, D. J., & Cutler, B. L. (1995). Effects of inconsistent eyewitness statements on mock-jurors’ evaluations of the eyewitness, perceptions of defendant culpability and verdicts. *Law and Human Behavior, 19*, 79–88. doi:10.1007/BF01499074.
- Bottoms, B. L. (1993). Individual differences in perceptions of child sexual assault victims. In Goodman, G. S., & Bottoms, B. L. (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 229–261). New York: Guilford.
- Brewer, N., Potter, B., Fisher, R. P., Bond, N., & Luszcz, M. A. (1999). Beliefs and data on the relationships between consistency and accuracy on eyewitness testimony. *Applied Cognitive Psychology, 13*, 297–313. doi:10.1002/(SICI)1099-0720(199908)13:4<297::AID-ACP578>3.0.CO;2-S.
- Burke, T. M., & Turtle, J. W. (2003). Alibi evidence in criminal investigations and trials: Psychological and legal factors. *Canadian Journal of Police and Security Services, 1*, 193–201.
- Burke, T., Turtle, J. W., & Olson, E. (2007). Alibi evidence in criminal investigations and trials. In Toglia, M., Read, J. D., Ross, D. F., & Lindsay, R. C. L. (Eds.), *Adult eyewitness testimony: A handbook* (pp. 157–174). Hillsdale, NJ: Erlbaum.
- Connolly, D. A., Price, H. L., & Gordon, H. M. (2010). Judicial decision-making in timely and delayed prosecutions of child sexual abuse in Canada: A study of honesty and cognitive ability in assessments of credibility. *Psychology, Public Policy, and Law, 16*, 177–199. doi:10.1037/a0019050.
- Culhane, S. E., & Hosch, H. M. (2004). An alibi witness’s influence on jurors’ verdicts. *Journal of Applied Social Psychology, 34*, 1604–1616. doi:10.1111/j.1559-1816.2004.tb02789.x.
- Culhane, S. E., & Hosch, H. M. (2012). Changing your alibi: Current law enforcement, future law enforcement, and layperson reactions. *Criminal Justice and Behavior, 39*, 958–977. doi:10.1177/0093854812438185.
- Dahl, L. C., Lindsay, D. S., & Brimacombe, C. A. E. (2006). Investigating investigators: Examining witnesses’ influence on investigators. *Law and Human Behavior, 30*, 707–732. doi:10.1007/s10979-008-9151-y.
- Dahl, L. C., & Price, H. L. (2012). “He couldn’t have done it, he was with me!”: The impact of alibi witness age and relationship. *Applied Cognitive Psychology, 26*, 475–481. doi:10.1002/acp.2821.
- Dysart, J. E., & Strange, D. (2012). Beliefs about alibis and alibi investigations: A survey of law enforcement. *Psychology, Crime & Law, 18*, 11–25. doi:10.1080/1068316X.2011.562867.
- Fisher, R. P., & Cutler, B. L. (1995). The relation between consistency and accuracy of eyewitness testimony. In Davies, G., Lloyd-Bostock, S., McMurrin, M., & Wilson, C. (Eds.), *Law and criminal justice: International developments in research and practice* (pp. 121–136). Berlin: De Gruyter.
- Kassin, S. M., Dror, I., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition, 2*, 42–52. doi:10.1016/j.jarmac.2013.01.001.

- Krix, A. C., Sauerland, M., Lorei, C., & Rispens, I. (2015). Consistency across repeated eyewitness interviews: Contrasting police detectives' beliefs with actual eyewitness performance. *PloS One*, *10*(2), e0118641. doi:10.1371/journal.pone.0118641.
- Leins, D. A., & Charman, S. D. (2016). Schema reliance and innocent alibi generation. *Legal and Criminological Psychology*, *21*, 111–126. doi:10.1111/lcrp.12035.
- Oeberst, A. (2012). If anything else comes to mind... better keep it to yourself? Delayed recall is discrediting—unjustifiably. *Law and Human Behavior*, *36*, 266–274. doi:10.1037/h0093966.
- Olson, E. A., & Charman, S. D. (2012). “But can you prove it?” Examining the quality of innocent suspects' alibis. *Psychology Crime and Law*, *18*, 453–471. doi:10.1080/1068316X.2010.505567.
- Olson, E. A., & Wells, G. L. (2004). What makes a good alibi? A proposed taxonomy. *Law and Human Behavior*, *28*, 157–176. doi:10.1023/B:LAHU.0000022320.47112.d3.
- Ross, D. F., Jurden, F. H., Lindsay, R. C. L., & Keeney, J. M. (2003). Replication and limitations of a two-factor model of child witness credibility. *Journal of Applied Social Psychology*, *33*, 418–431. doi:10.1111/j.1559-1816.2003.tb01903.x.
- Semmler, C., Brewer, N., & Douglass, A. B. (2012). Jurors believe eyewitnesses. In Cutler, B. L. (Ed.), *Conviction of the innocent: Lessons from psychological research* (pp. 185–209). Washington, DC: American Psychological Association.
- Strange, D., Dysart, J., & Loftus, E. F. (2014). Why errors in alibis are not necessarily evidence of guilt. *Zeitschrift für Psychologie*, *222*, 82–89. doi:10.1027/2151-2604/a000169.
- Wright, D. B., Hanoteau, F., Parkinson, C., & Tatham, A. (2010). Perceptions about memory reliability and honesty for children of 3–18 years old. *Legal and Criminological Psychology*, *15*, 195–207. doi:10.1348/135532508X400347.
- Yuille, J. C., & Cutshall, J. L. (1986). A case study of an eyewitness memory of a crime. *Journal of Applied Psychology*, *71*, 291–301. doi:10.1037/0021-9010.71.2.291.