How Witnesses' Perceptions of Identification Decision Consequences Affect Their Choosing Patterns

Carmen A. Lucas

Bachelor of Psychology (Hons)

Thesis submitted in fulfilment of the requirements for the degree

Doctor of Philosophy

Faculty of Social and Behavioural Sciences

School of Psychology

Flinders University

Adelaide, Australia

January 2017

Table of Contents

Summary

This thesis describes five experiments that aimed to manipulate people's perceptions about the consequences associated with eyewitness identification decisions and test whether this affected their subsequent willingness to choose from a lineup. If consequence perceptions influence identification decisions, a serious applied problem arises. Due to the complexity of a criminal investigation people's consequence judgements are undoubtedly highly subjective and, therefore, potentially form a misguided basis for choosing from or rejecting the lineup. While there are both intuitive and theoretical grounds for thinking that identification decisions are influenced by perceptions regarding the associated consequences, only one experiment (to my knowledge) has been previously conducted to address this possibility, with inconclusive results. The lack of research in this area may be, at least in part, due to the fact that there are significant ethical and methodological obstacles associated with researching identification decisions in a context where people think that their decisions have meaningful repercussions. However, prompting people to consider the consequences of an identification decision in a laboratory context may provide an adequate paradigm for beginning to develop an understanding of this issue. Therefore, the experiments comprising this thesis implemented a hypothetical consequences paradigm to investigate the issue.

Information about the likelihood that mistakenly choosing from or rejecting a lineup could result in the worst case scenario (i.e., a wrongful conviction or the guilty party getting away) was used as the primary means to manipulate people's consequence perceptions—with the aim of influencing which mistake people thought would be worse to make. It was hypothesised that those who were led to believe that mistakenly rejecting a lineup is worse than mistakenly choosing from a lineup would be more likely to make positive identification decisions than those who were led to think the opposite, with the result being higher target-present and lower target-absent accuracy for the former. The results from the first four

experiments showed no difference in choosing (or accuracy) between the two consequence likelihood conditions. However, each provided a step forward in refining the manipulation of consequence perceptions and testing their influence in a hypothetical context. The fifth and final experiment showed a small significant difference in choosing, in the predicted direction. This difference translated into the hypothesised differences in accuracy for target-present and target-absent lineups. Moreover, descriptive evidence was obtained to suggest that the effect of perceived consequence likelihood on identification decisions was greater for people who were not very confident in which decision to make. These experiments form a basis for delineating the conditions under which hypothetical consequences might be reliably shown to bias identification decisions in future research. The theoretical and applied implications of the findings for the likely influence of consequences perceptions on actual identification decisions is discussed, along with potential future directions for research in this area.

Declaration

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

……..

Signature

……...

Date

Acknowledgements

Neil, without you I would have never considered doing a PhD—or managed to submit one. Your approach to conducting research is innovative, purposeful and inspiring. It has been an absolute honour to be your student and learn from you.

Mum and Dad, you have supported me through thick and thin and getting me through this PhD was no exception. I am so lucky to have you as my parents. Yarran, you were an absolute legend during write-up. I am so grateful for everything you did for me.

Em and Lizzie, you were always there to talk things through, listen to me vent, or distract me. Alex, you've been the best office mate I could have ever hoped to have. Bec, your commiserations, encouragement and wealth of knowledge about APA formatting and submission procedures were invaluable. Shout out to Nicole for running the lab and being so supportive over the years. And to Tomoko and Stacey for being my conference buddies.

Thank you to every person who volunteered their time to take part in the experiments I conducted. Without you this research would not have been possible. Also, thank you to all the people—too numerous to name—who assisted in some small way in the production of this thesis.

Finally (but most of all) thank you Amma, for being a constant inspiration and helping me find my way.

CHAPTER 1

Introduction

A large body of scientific literature highlights the fallibility of eyewitness identification decisions (Brewer & Wells, 2011). Further, research on this topic has led to recommendations for policy (e.g., Steblay, Dysart, & Wells, 2011; Wells et al., 2000; Wells & Quinlivan, 2009) and practice (e.g., Brewer & Palmer, 2010; Wells et al., 1998). The broad aim of these recommendations is, of course, to reduce the likelihood of errors that may contribute to miscarriages of justice. As evidenced by DNA exoneration cases where eyewitness evidence played a role in the conviction of innocent people (Innocence Project, 2016), this objective is important to pursue, particularly as there are still aspects of the identification decision process that are not well understood (Brewer & Wells, 2011; Charman & Wells, 2007).

One area that research has largely neglected relates to what witnesses might be thinking as they face making an identification decision and how this might influence their ultimate choice of whether to choose from or reject the lineup. Gaining an understanding of the interplay between people's cognitive (and metacognitive) processes, their memory quality and aspects of the decision making environment is arguably important from the perspectives of reducing the likelihood that a witness (a) makes a positive identification (sometimes of an innocent suspect) when the perpetrator is absent from the lineup, and (b) fails to identify the perpetrator when present in the lineup (Brewer, Weber, & Semmler, 2007). One example of cognitions that could accompany and influence an identification decision involves perceptions the witness may form about the possible consequences associated with that decision. In the experiments reported in this thesis I examine whether such perceptions—specifically, feeling that either mistakenly choosing from or mistakenly rejecting a lineup

would be a worse error to make—are associated with witnesses' tendency to choose from the lineup.

It seems intuitively likely that perceptions about which mistake would be worse to make might be considered by a witness and, in turn, affect their decision. More importantly, perceptual decision making theory (Green & Swets, 1966) stipulates that uneven rewards and (or) penalties associated with reports of whether a stimulus was previously encountered or not will shift willingness to make a positive response, thereby providing a theoretical basis for expecting that identification decisions might be shaped by consequence perceptions. One very obvious explanation for the almost complete neglect of this research question by eyewitness identification researchers is that practical and ethical considerations have meant that eyewitness memory research has largely been conducted in the laboratory context where it is difficult to manipulate identification decision consequences meaningfully. In the experiments reported in subsequent chapters, a "hypothetical consequences" paradigm is used to test the possible impact of perceived identification decision consequences on decision making. Specifically, I examined whether explicitly asking people in a laboratory study to think about the consequences associated with an actual identification decision might be able to offer some insight into how identification decisions are affected by the perceived consequences of the decision.

The Rationale for Thinking that Consequences Might Influence Identification Decisions

Beginning with the work of B. F. Skinner (e.g.,1953) it has been clearly established in the scientific literature that consequences are a key factor in shaping human behaviour—with people ultimately aiming to achieve positive outcomes and avoid negative ones. In the context of someone witnessing a crime, apprehending the perpetrator clearly represents the positive outcome to be achieved and the perpetrator getting away with the offense is the negative outcome to be avoided. Depending on the nature of the crime that is committed, the

witness might be more or less motivated to pursue the goal of apprehending the perpetrator. For example, if they notice a teenager shoplifting they might think that teaching them a lesson is in order (but might not even bother to report the crime). However, if they see someone commit an armed robbery, a sexual assault or murder, then they would presumably be much more likely to think that the person should be punished for their crime and be motivated do what they can to ensure that outcome. Under the latter circumstances, if a witness is asked to try to identify the culprit from a lineup they would be expected to agree to view the lineup and either (a) choose from the lineup if they recognised the perpetrator or (b) reject the lineup if they did not. The assumption that witnesses will behave in this way aligns with the purpose of an identification test, namely, to establish whether the police's suspect is the perpetrator or not (Charman & Wells, 2007). However, it is important to note that the witness needs to be clear on the fact that the culprit may or may not actually be in the lineup. It has been argued that the mere act of asking a witness to view a lineup may suggest that the culprit is likely to be in the lineup and lead to an increased propensity to choose (Brewer & Palmer, 2010; Wells, 1993). This underscores the importance of making it explicitly clear to witnesses that the perpetrator may not be present in a lineup and ensuring that there is a tangible way to reject the lineup (Brewer & Palmer, 2010; Steblay, 1997).

Supposing that the witness is clear on the fact that both choosing from and rejecting the lineup are informative responses, the choice might nevertheless end up being difficult to make—as suggested by the frequency with which people have been found to make identification errors. There might be several lineup members who look a lot like the perpetrator, so the witness may think it is probably one of them—but which one? Alternatively, maybe no one provides a very good match to the witness's memory, but the witness cannot be sure that none of the options is the perpetrator. Or it could be that someone stands out from the rest of the lineup members, but the witness is not sure whether s(he) is the

culprit. If any of these situations is experienced, the witness is likely to appreciate, if they had not previously, that there is a chance that they could make a mistake and that this could have serious repercussions. They might recognise that if they reject the lineup and the culprit was present, then the guilty party might get away and commit further crimes. Or, if they make a positive identification they might choose someone who is not actually the perpetrator. In the event that the perpetrator is not present they could inadvertently pick an innocent suspect and the outcome could be that the suspect ends up in jail. As both mistakenly choosing from and rejecting a lineup are associated with serious potential consequences, witnesses who are not certain in their decision should probably not choose from or reject the lineup. This conclusion is supported by the finding that allowing people to make use of a "don't know" option improves the accuracy of identification responses that are volunteered (Kekessie, 2013; Perfect & Weber, 2012; Warnick & Sanders, 1980; Weber & Perfect, 2012). However, it does not appear to be a current practice to provide people with the explicit response option to indicate that they do not know whether the target is present in the lineup or not (e.g., National Research Council, 2014) and it is unclear to what extent people feel free to respond in this way in the various identification procedures that are used. There are some tentative grounds for thinking that there may be a fairly low prevalence of spontaneous "don't know" responses (Köhnken & Maass, 1988; Sanders & Warnick, 1981; Warnick & Sanders, 1980; Weber & Perfect, 2012). A possible explanation for why people might have a tendency to choose from or reject a lineup, even when they are not sufficiently confident enough to do so, is that opting out of the decision is in direct conflict with the witness's likely motivation to provide evidence to help apprehend the perpetrator. It seems probable that by having agreed to view the lineup the witness feels more pressured to provide instructive information than to avoid making a mistake. This interpretation is consistent with the view that information is provided

in a social context and people are likely to choose to respond in what they perceive to be a socially acceptable way (Ackerman & Goldsmith, 2008).

So, when a witness's memory is unable to unequivocally confirm or guide their decision, but they feel compelled to either choose from or reject the lineup, it seems plausible that people might weigh up the potential seriousness of the errors they could make and allow their perception about which mistake would be worse to make to have a bearing on the final decision that is made. Thus, two witnesses experiencing a similar level of difficulty in making a decision might come to different conclusions about which course of action to pursue based on divergent consequence perceptions. One witness might decide that the threat of letting a guilty person get away warrants narrowing down several plausible options in order to make a pick, or identifying someone who is not a very good match to memory but clearly stands out from the other lineup members. In contrast, the other witness might decide that the possibility of identifying an innocent suspect does not justify picking under these circumstances. What this suggests for the accuracy of the identification decision is that, if the target is in the lineup, the first witness will be more likely to make the correct choice. While they might inadvertently pick a filler, the target should have a reasonable chance of being identified. If the target is not in the lineup, however, the second witness will have made the correct decision.

Despite the intuitive appeal of the view that consequences might influence the identification decision in the ways just indicated, formal identification decision theory (e.g., Charman & Wells, 2007; Clark 2003; Lindsay & Wells, 1985; Wixted & Mickes, 2014) is in its early stages of development (Wells & Olson, 2003) and has not accounted for this possibility, beyond acknowledging that non-memorial factors are likely to play a role in the decision process (Charman & Wells, 2007). However, it has been previously pointed out that research conducted within the framework of signal detection theory (SDT; Green & Swets,

1966)—a framework that has guided a significant body of perceptual decision making research—provides general support for the argument that consequences might affect identification decisions (Malpass & Devine, 1980). While there are a large number of theoretical accounts of judgement and decision making stipulating that consequences are integral to choice, SDT is perhaps the most relevant to the present context as it shows how consequences influence perceptual judgements, with a clear application to recognition memory. In the subsequent section, the key elements of SDT are outlined and the basis it provides for thinking that people's consequence perceptions could affect their willingness to make positive identification decisions is described.

Signal detection theory and consequence bias. At its core SDT describes the process of making decisions when "certain waveforms called *signals* may or may not be added to a random background disturbance call *noise*" (Green & Swets, 1966, p.7). In applying this to the recognition memory context, the paradigm for discerning signal from noise generally refers to people studying a list of stimuli (e.g., words, faces, pictures) and then being shown some or all of those stimuli along with other previously un-encountered stimuli in a test phase. SDT specifies that the test stimuli form two distributions along a continuum of signal strength. Previously seen, or "old," items will fall in a distribution that is higher up on the continuum than the distribution of the "new" items that were not previously encountered. This means that the old items are on average associated with a stronger match to memory than the new items. An identification decision is somewhat distinct from this as often only one stimulus is encoded and the test set involves either all previously unseen stimuli (in target-absent cases) or one previously seen face amongst several lures (in target-present cases).[1] However, it has been argued that the same general stipulations can be made

[1]The lineup members might be shown to the witness simultaneously, or one after another. Traditionally lineup presentation was simultaneous, but more recently a procedure was developed for presenting the lineup sequentially, which is associated with producing a shift in willingness to make a positive identification (Lindsay

about where the lineup members fall on a continuum of signal strength (Ebbesen & Flowe, 2002). That is, if the target face appears in a lineup it should on average (across samples of participants and stimuli) fall higher up on the continuum of signal strength than the fillers. In a target-absent lineup, as long as it has been fairly constructed (see Malpass, Tredoux, & McQuiston-Surrett, 2007), the innocent suspect should fall amongst the fillers.

In both a traditional recognition memory and an eyewitness identification context, the task of the decision maker is to decide whether a signal is or is not present. In a recognition paradigm this refers to repeated judgments of whether an item is old or new. In the identification decision context only one of the stimuli can be the target—at least in a single-suspect lineup, which has been a long-standing recommended practice (Wells & Turtle, 1986)—so the witness has to judge whether the best option in the lineup constitutes an old or a new response. SDT conceptualises this judgement as the setting of a response criterion along the continuum of signal strength. In the recognition paradigm, all of the items that fall above this criterion level will be judged as "old", indicating that they were previously encountered and all of the items below the cut-off will be judged as "new", indicating that they have never been seen before. Generally, the same principle applies in the eyewitness context. If all of the lineup members fall below the threshold where a person is willing to make a positive recognition response, the lineup will be rejected. If only one lineup member falls above their threshold, then they will be identified. However, technically it is also seems possible that multiple lineup members could fall above the person's threshold for a positive response. Because only one lineup member can be the target it becomes a bit unclear what a witness would do under these circumstances. It might be the case that this presents a situation that has been described by some witnesses, where they find themselves narrowing down the

& Wells, 1985; Palmer & Brewer, 2012; Steblay, Dysart, Fulero, & Lindsay, 2001; Steblay, Dysart, & Wells, 2011).

options to make a positive identification (Charman & Wells, 2007; Dunning & Stern, 1994). But, depending on how this information is interpreted by the person making the identification decision, it is not out of the question that this could also form a basis for rejecting the lineup. Despite not having a clear indication of what a witness would do under these circumstances, the concept of people having a threshold for choosing from or rejecting a lineup is not foreign in the eyewitness context (e.g., Clark, 2005; Palmer & Brewer, 2012).

SDT further stipulates that several factors influence where people's response criterion is set. These include the a priori probability of a signal being present, factors that affect discriminability and, importantly for the present focus, the overall expected value associated with the response options. Green and Swets (1966) stated that often it is the case that the costs associated with both mistakes that can be made (i.e., a miss or a false alarm) are equally bad and the benefits associated with both correct decisions (i.e., a hit or a correct rejection) are equally good, in which case the decision maker should be expected to just do their best to be accurate. But where the prospective rewards and penalties associated with each of the decision outcomes are uneven, this requires a rational decision maker to shift their decision criterion to maximise the positive outcomes and minimise the negative outcomes. If, for example, someone expects to be more heavily penalised for a miss (-10) than a false alarm (-5) and both correct rejections and hits are rewarded equally (+5), then it is in the interest of the decision maker to say an item is "old" when they are not adequately sure that they will not make the crucial mistake of missing an old item. Patterns of responding that are consistent with this prediction have been found in a large number of basic perception and recognition memory experiments (e.g., Curran, DeBuse, & Leynes, 2007; Green & Swets, 1966; Healy & Kubovy, 1978, 1981; Maddox & Bohil, 1998; Ulehla, 1966). In one example, Healy and Kubovy (1978) manipulated payoff schedules between blocks of (word) recognition decisions. In some blocks participants were given a balanced payoff schedule

(i.e., they were told that correct decisions would earn them 5 points and incorrect decisions would lose them 5 points) and in other blocks they were given an unbalanced payoff schedule favouring negative responses (i.e., they were told that a hit earned 1, whereas a correct rejection earned 3; and a miss was penalised 1 whereas a false positive was penalised 3). All participants were also told that the person with the highest score at the end would receive a cash prize. Results showed that in the blocks with the uneven payoff schedule (where "new" responses were both less heavily penalised when they were wrong and more highly rewarded when they were right), participants recorded a lower mean hit rate, as well as a lower false alarm rate, than the blocks where a balanced payoff schedule was implemented. However, people's performance was not optimal. That is, participants demonstrated a tendency to be conservative in that they did not shift their response criterion far enough to fully maximise their payoffs given their discrimination ability—a consistent finding in the relevant literature (VonWinterfeldt & Edwards, 1982). In spite of this, payoff manipulations have come to become regarded as failsafe way to shift people's response criterion level in recognition memory tasks (e.g., Benjamin, Diaz, & Wee, 2009; Kantner, Vettel, & Miller, 2015).

While this theoretical framework is a simplistic representation of the eyewitness identification context, it generally supports the expectation that when the consequences associated with an identification decision are such that picking an innocent suspect or failing to identify a guilty party is perceived to be worse, willingness to choose from a lineup will be affected. Specifically, witnesses who think mistakenly choosing from a lineup is worse would be less likely to positively identify someone from a lineup than witnesses who think mistakenly rejecting a lineup would be worse. This difference in choosing would in turn be expected to have a differential effect on accuracy, depending on whether or not the target is present in the lineup. For target-present cases accuracy would be expected to be higher in the

condition where people are more likely to make positive identifications and for target-absent cases accuracy would be higher when people are less likely to make positive identifications.

Consequences Bias and the Identification Decision Literature

Despite there being both intuitive and theoretical grounds for thinking that consequence perceptions might affect people's willingness to make a positive identification, there has been little research addressing this issue. Further, the one study (to my knowledge) considering the possibility that people might favour choosing from or rejecting a lineup based on consequence perceptions yielded inconclusive results. Malpass and Devine (1980) conducted an experiment where they staged a disruption and act of vandalism in a lecture theatre full of students before leading people who volunteered to make an identification decision (N = 65) to believe that the punishment for the offender would be fairly trivial (i.e., if caught the perpetrator would only receive on-campus discipline) or more serious (i.e., if caught the college would press criminal charges). They postulated that people would be more worried about implicating an innocent suspect in the more serious punishment condition than the less serious punishment condition, but that the concern about letting the guilty party walk would be the same in both conditions because in both cases the person would be getting away with a relatively trivial crime. Based on the same theoretical perspective outlined earlier (i.e., SDT, Green & Swets, 1966) they hypothesised that choosing would be lower in the serious punishment condition. However, the opposite pattern emerged. Across target presence conditions,[2] those in the serious penalty condition were significantly more likely to make a positive identification (83%) than those in the less serious condition (26%). Retrospectively, the authors speculated that perhaps some aspect of the situation (e.g., the vandal destroyed equipment that belonged to a well-liked professor) had led participants to adopt the mindset that they wanted to help "get the guy" responsible and this fuelled an increased propensity to

[2] Half of the participants were shown a target-present lineup and the other half were shown a target-absent lineup.

choose when the consequences were more serious— a possibility that does not bode well for an innocent suspect in a murder investigation where the death penalty could be enforced. However, people's consequence perceptions were not actually measured, so the veracity of this possibility was unable to be verified. Further, while the difference in choosing between the two punishment conditions was large, it was based on a relatively small sample size and, of course, it represents a one-off demonstration. Therefore, further research in this area is clearly necessary. However none has been forthcoming, leaving a substantial gap in our understanding of how eyewitness identification decisions are shaped by consequence perceptions. In the next section I argue that the lack of research on whether and to what extent people's consequence perceptions bias their identification decisions probably largely reflects the absence of a viable paradigm with which to investigate this issue systematically.

A Brief History of Identification Decision Research and the Issue of Ecological Validity

In the early years of research on identification decision making, researchers attempted to simulate the circumstances under which actual identification evidence is obtained as closely as possible. They did so by staging crimes in front of unsuspecting "witnesses" and, in a few cases, maintained the pretence that a real crime was being investigated until identification decisions were made (Wells, 1993). It was at this time that the potentially biasing influence of consequence perceptions in identification decisions was first considered (Malpass & Devine, 1980). However, later years saw a marked shift to laboratory research, a shift that can be attributed to several factors. Firstly, researchers only staged trivial crimes (e.g., petty theft, vandalism, test cheating), as exposing people to anything more serious was clearly considered to be unethical. This led to the criticism that, despite embodying some degree of realism, the "uninformed paradigm" (Murray & Wells, 1982) may have fallen short of creating circumstances that would be seen by the witnesses as consequential enough to effectively simulate real identification decision making (Foster, Libkuman, Schooler, &

Loftus, 1994). Secondly, staging criminal investigations was found to be associated with substantial practical problems. Managing to deceive participants from the time of staging the crime until they made an identification decision proved to be quite demanding for researchers, and not something that could be done on a regular basis (Murray & Wells, 1982). Even in cases where participants were debriefed immediately following the crime there were potential pitfalls. For example, Leippe, Wells, and Ostrom (1978) staged a theft while participants were waiting to take part in a different study and reported that several participants had to be debriefed during the event because they tried to intervene. Thirdly, the few experiments that explicitly compared identification decision making for people who thought they were making a real decision and people who knew they were not (Köhnken & Maass, 1988; Murray & Wells, 1982; Sanders & Warnick, 1981) provided no clear indication that people made their decisions any differently under the two conditions. That is, the results yielded no evidence to suggest that thinking the decision has actual consequences leads to an overall increase (or decrease) in accuracy and no convincing indications that this might moderate the effect of other variables.[3] While these experiments were far from conclusive, they did not appear to provide any compelling reason why researchers should not conduct their research in the more practical and convenient setting of the laboratory.

More recently, the issue of whether laboratory findings generalise to the applied context was revisited in several attempts to corroborate established effects with archival data from real identification tests. However, this approach has also turned out to be problematic in terms of the insight into real identification decisions it is able to provide. A recent review (Horry, Halford, Brewer, Milne, & Bull, 2014) showed that these studies have yielded mostly

[3] Note, however, that one study (Köhnken & Maass, 1988) did show tentative evidence that the effect of lineup instructions which pressured the witness to make a positive identification did not have as much influence on people's choices when they thought they were making real decisions. In support of this, a later meta-analysis of studies that investigated the influence of lineup instructions on choosing behaviour (Steblay, 1997) found that the effect sizes for studies conducted under realistic conditions were smaller than studies conducted in a laboratory environment.

inconsistent findings and often suffered from methodological drawbacks such has having inappropriately small sample sizes (Behrman & Davey, 2001; Tollestrup, Turtle, & Yuille, 1994), making use of unsuitable analysis techniques (Behrman & Davey, 2001; Memon, Havard, Clifford, Gabbert, & Watt, 2011; Tollestrup et al.,1994) and including cases where the witness was familiar with the suspect (Behrman & Davey, 2001; Pike, Brace, & Kynan, 2002; Tollestrup et al.,1994; Wright & McDaid, 1996). Further, it was pointed out that, even if researchers do everything right, these data are of limited value due to a number of unavoidable methodological constraints. For example, often no information regarding the variables of interest is stored, it is almost impossible to establish for sure whether the suspect in the case was actually guilty and predictors are often confounded with one another (Horry et al., 2014).

Despite the limitations of data obtained from real-world settings and the continued awareness that the generalisability of findings from laboratory studies may be queried, research on various aspects of eyewitness identification decision making has flourished, furthering an understanding of, and the steps required to combat, the issues involved with this form of evidence. But the fact that investigating identification decisions under naturalistic conditions has proven to be problematic presents a plausible explanation for why the potential effect of consequence perceptions on willingness to choose from a lineup has not been addressed any further than the inaugural experiment (at least in the published literature to date). This issue is one of the few confronting the field that, at face value, appears to rely heavily on the investigation being embedded in a realistic context, a consideration which has probably deterred researchers from pursuing this line of inquiry. However, this does not necessarily mean that there is in fact no way of addressing this question. Another area of research that ostensibly comes up against a similar obstacle involves establishing how the physiological and emotional stress associated with witnessing, or being the victim of, a crime

(or being a law enforcement official engaged in a hostile situation) affects people's ability to recall information or make an identification decision. On the face of it, investigating this issue appears to be dependent on creating conditions where a person thinks that a real crime is being committed (which encounters the aforementioned practical and ethical issues involved with staging a crime that would lead to requisite levels of stress). However, researchers have addressed this issue in a variety of different ways. For example, training programs where people took part in enactments and simulations of high-stress situations have been used to later test people's recall and recognition memory (Ihlebæk, Løve, Eilertsen, & Magnussen, 2003; Morgan et al., 2004; Stanny & Johnson, 1999). On a less realistic level, people's orientation towards emotionally evocative to-be-remembered stimuli has sometimes been prompted by videos or slides showing a horrific accident (e.g., Loftus & Burns, 1982) or an assailant holding a weapon (e.g., Pickel, 1998, 1999). Physiological arousal, which is often part of being involved in such situations, has been induced by having participants engage in strenuous exercise (e.g., Hope, Lewinski, Dixon, Blocksidge, & Gabbert, 2012) or receive electric shocks (e.g., Brigham, Maass, Martinez, & Wittenberger, 1983). In one experiment genuine anxiety was induced in people who self-identified as having a fear of needles by leading them to believe that they were about to receive an injection, before testing them on their recognition of people they encountered leading up to the anticipated injection (Maass & Köhnken, 1989). Meta-analyses of these studies tentatively suggest that to the degree that stimuli cause an orienting response they have the potential to increase the accuracy of people's responses (Christianson, 1992), although the presence of a weapon might distract attention from and impair memory for other details of an event (Fawcett, Russell, Peace, & Christie, 2013; Steblay, 1992). At high levels of anxiety and physiological arousal, however, it seems that people's memories become impaired (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). Similar to the way that a variety of different approaches have been used to

investigate the effect that witnessing a crime can have on people's memories, the influence of consequence perceptions on willingness to choose from a lineup might also be able to be (at least tentatively) answered by implementing a variety of inventive methods. In the experiments reported in subsequent chapters I considered what might occur if someone knows they are making an identification decision that is not real, but they are explicitly prompted to think about the consequences associated with actual identification decisions. Is it possible they would anticipate those consequences in their decision in a manner that is sufficiently similar to the way their perception of these consequences would influence a real decision?

The Rationale for a Hypothetical Consequences Identification Decision Paradigm

Hypothetical decision making paradigms have been implemented in various areas of inquiry with an applied focus, where observations under realistic conditions are not practical or feasible to obtain. These include areas such as economic decision making (e.g., Schwartz, Bruine de Bruin, Fischhoff, & Lave, 2015; Stevens et al., 2015), moral choices (Greene, Somerville, Nystrom, Darley, & Cohen, 2002), social interaction (e.g., Levine, Cassidy, Brazier, & Reicher, 2002) and medical decision making (e.g., Petrova, Garcia-Retamero, Catena, van der Pligt, 2016). In conjunction with using hypothetical decisions as a research tool, the question of whether this approach provides valid indications of actual behaviour invariably arises. A limited amount of research has actually compared the same decisions in a real and hypothetical context. While results of these studies have sometimes showed no difference between real and hypothetical choices (Kühberger, Schulte-Mecklenbeck, & Perner, 2002; Johnson & Bickel, 2002; Locey, Jones, & Rachlin, 2011), some have provided evidence of systematic differences in people's cognitive processes (e.g., Kang, Rangel, Camus, & Camerer, 2012; Taylor, 2013) and decision making behaviour (e.g., Vlaev, 2012) depending on the reality of the decision. However, despite widespread use of hypothetical

research paradigms, there exists a relatively feeble understanding of when and why (or why not) hypothetical decisions might actually do a good job of approximating real ones. This is understandable because, for the same reason that hypothetical paradigms are often used (i.e., people's actual decisions are too difficult to investigate), comparing real and imagined decisions and establishing what similarities and differences exist between them is challenging and often impossible. Nevertheless, some broad rules of thumb have emerged for judging whether such an approach might be useful in a particular context. For example, it has been argued that it is important to carefully consider whether a particular hypothetical decision making context is both likely to motivate and allow people to behave as they would in reality (Eastwick, Hunt, & Neff, 2013). Reasons for thinking that these requirements might be met by encouraging mock-witnesses to think about the consequences of identification decisions are outlined below.

While it is the hope of psychological researchers that the majority of people who volunteer to take part in their experiments are motivated to take seriously the tasks they are asked to engage in, most would also agree that it is also a reasonable assumption that at least some participants will not be particularly invested in the task beyond fulfilling the minimum requirements involved with following the instructions they are given and performing to a reasonable standard. It might be the case that some participants who are asked to make what is, practically speaking, a fake identification decision, will take the task seriously enough to try to image what it would be like to be a witness in an actual criminal investigation when making their choice. But others are likely to only respond in a way that they feel is justified given the artificial environment they find themselves in. Nevertheless, participants guided by either of these motivations may make a reasonable attempt to account for the consequences of real identification decisions in their choice if they are explicitly prompted to think about

them, as such prompting should indicate the relevance of the consequences to the task at hand.

The main question is then whether people are in fact able to emulate their willingness to choose from a lineup with these consequences in mind in a similar way to how they would do this if they were actually confronted with a real lineup. Often when behaviour in hypothetical research paradigms does not appear to reflect people's actual behaviour, this is attributed to what has become known as a "hot-cold empathy gap". That is, when people are asked to report how they would behave in a particular scenario they are sometimes unable to appreciate what their actual priority between two separate motivations is likely to be when they are only imagining them (Loewenstein, 2005). For example, there is a consistent tendency for people to overestimate the value of a product when they are not actually trading off their desire to own the product with having to spend money to obtain it (Murphy, Allen, Steven, & Weatherhead, 2005). This situation is perhaps comparable to asking people to estimate prospectively how willing they would be to choose from a lineup given the outcomes that this decision could lead to. Under these circumstances it seems reasonable to expect that they may overestimate the degree to which the consequences would actually affect their decision because they are not aware of the influence that actually having the lineup in front of them would have. However, people making laboratory based identification decisions are still required to make the perceptual-memorial judgement that they would make when making a decision about a real lineup, so this aspect of the decision is as realistic as possible and distinguishes it from most situations where a hot cold empathy gap has been found. Only the consequences of the decision are hypothetical. It has been argued that people can account for hypothetical consequences in their decisions similarly to real consequences, because it involves the same cognitive process—of imagining potential future outcomes (Kühberger et al., 2002). Therefore, it is conceivable that, when prompted to consider the

consequences their identification decision could have if they were a witness in an actual criminal investigation, people may be able to adopt a choosing strategy that approximates what they would do if they were making a decision where those consequences could actually occur.

Having made the tentative argument for using a hypothetical paradigm to investigate the relationship between consequences and identification decisions, it is acknowledged (as alluded to earlier) that such an approach limits what can be gleaned from the results. In the worst case scenario, null effects are obtained using a hypothetical paradigm and it is unclear whether the methodology is unsuitable for showing the predicted relationship or whether the predicted relationship does not exist. But even if hypothesised effects are found, which would be more a promising index of the usefulness of the research paradigm, it would not be possible to know whether or to what degree the results represent what would occur for real identification decisions. However, given the difficulties that have been associated with researching lineup choosing behaviour in a context where people think they are (or actually are) making real identification decisions, determining whether basic predictions regarding consequence perceptions and these decisions could be demonstrated in a hypothetical setting seems to be a prudent first step. This approach provides a chance to think critically about the issue at hand and work on developing other aspects of the method—such as the manipulation of people's consequence perceptions—in a convenient setting. If support for basic predictions regarding the relationship between consequence perceptions and identification decisions could be shown in a hypothetical setting, this would indicate that questions of a similar nature could be tentatively researched in this way. And perhaps such findings would also justify expending the resources to attempt to validate the findings in a more ecologically valid context. Similar progressions have occurred in other areas (e.g., Cohen-Hatton & Honey, 2015; Davis et al., 1999; Sussman, Sharma, & Alter, 2015), where the researchers started out

using hypothetical scenarios to investigate a hypothesis. Once receiving encouraging findings, they sought to corroborate those findings using increasingly naturalistic paradigms.

In the following chapters, five experiments are reported in which variations on a hypothetical consequences paradigm were implemented to investigate basic predictions about whether consequence perceptions bias people's identification decisions. Experiments 1–4 addressed various methodological considerations involved with manipulating and measuring identification consequence perceptions and using a hypothetical paradigm to test the hypotheses that were made. Experiment 5 tested whether the influence of consequence perceptions on identification decisions depends on how confident someone is in making a choice from a lineup.

CHAPTER 2

In Chapter 1 the argument was advanced that the negative consequences associated with incorrect eyewitness identification decisions might form a basis for a witness being more or less willing to choose from a lineup, depending on which identification decision outcome was seen as worse than the other. Further, it was suggested that prompting participants in a laboratory study to imagine their identification decisions having real-life repercussions may be sufficient to lead people to consider perceptions of possible consequences when deciding whether or not to choose from the lineup. Experiments 1 and 2 tested whether choosing behaviour in a laboratory experiment was affected by the interaction between the overall severity of consequences associated with identification decisions and another variable, the perceived likelihood of implicating an innocent suspect.

Experiment 1

The decision regarding how to test the biasing effect of consequence perceptions on identification decisions in this initial experiment was based on the following rationale. First, it did not seem feasible to manipulate whether mistakenly providing evidence that might lead to a wrongful conviction, or letting a guilty party walk by failing to identify them, is perceived to be worse. Such reasoning is based on that fact that people's perceptions regarding justice have an ideological basis (Carroll, Perkowitz, & Lurigio, 1987; Miller, 1973), that by definition means such perceptions are likely to be relatively inflexible (Gerring, 1997; Miller, 1973). However, regardless of whether someone is more inclined to protect the innocent or seek retribution against the guilty, the gravity of both identification errors that can be made increases with the seriousness of the crime committed. A person getting away with a fairly trivial crime, such as petty theft, likely presents a less serious situation than a murderer remaining unpunished, because the threat to society is much less. Similarly, wrongfully convicting someone of a minor offence would be unjust, but still much

less grim than condemning someone for a particularly heinous crime, as the impact on the individual's life would be far more severe in the latter situation. Therefore, the seriousness of the consequences associated with identification errors could perhaps be manipulated by varying the nature of the crime being investigated. But on its own this manipulation would not be expected to affect people's choosing behaviour, because it creates a difference in the importance of avoiding any error rather than creating a difference in whether mistakenly picking from or rejecting the lineup would be worse. However, what an increase in perceived consequence seriousness might be expected to influence, is the degree to which other variables affect willingness to choose from a lineup. One line of research in the identification literature tentatively supports this contention. A large number of experiments have shown that when lineup instructions fail to warn that the target may or may not be in the lineup (typically referred to as biased lineup instructions), or explicitly suggest that the target is in the lineup, people are more likely to make positive identification decisions (e.g., Greathouse & Kovera, 2009; Hall & Ostrom, 1976; Malpass & Devine, 1981; Pozzulo & Dempsey, 2006; Thompson & Johnson, 2008)—an effect that emerges both when the target is in the lineup and when they are not (Clark, 2005). However, a meta-analysis of earlier studies testing the effect of biased lineup instructions on choosing behaviour found that, compared with laboratory research, their effect was not as pronounced in cases where people thought they were making real identification decisions (Steblay, 1997). That is, when people thought their decision had actual consequences the effect of biased instructions was not as large as when people knew their decision had no consequences. One suggested explanation for this pattern of results is that biased lineup instructions produce pressure to make a positive identification decision—pressure that people resist when the importance of making a correct decision (and avoiding a critical error) is higher (Köhnken & Maass, 1988). This interpretation suggests that when it is intimated to people that the target is in the lineup they do not necessarily

believe this to be true, but conforming to social pressure is preferable to being right when being wrong has no further repercussions. If people actually believe that the target is present in a lineup, this belief should act as a strong indicator of whether or not making a positive identification will lead to a critical error. Therefore, in cases where the importance of avoiding a critical error is high this belief would be expected to be more influential in the decision making process than when the importance of avoiding a critical error is low. That is, the effect that anticipated target presence has on choosing should be greater when a more serious crime has been committed. Based on this premise, I sought to test whether increasing the belief that an innocent suspect might be identified when choosing from a lineup would decrease willingness to make a positive identification and, further, whether this effect would be greater for decisions associated with more serious consequences.

In order to manipulate people's impression of how likely an innocent suspect identification would be, participants were primed (prior to making their own identification decisions) with a scenario where either a rightful or a wrongful identification—and subsequent conviction—of a suspect occurred. In estimating the probability with which an event occurs, people have been found to rely on how mentally accessible an example of the event in question is (Tversky & Kahneman, 1973). Moreover, it has been found that merely imagining an event occurring, under uncertain circumstances, can lead it to seem more likely to eventuate. In one experiment, conducted immediately prior to a US election participants were randomly allocated to imagine either one candidate or the other winning the presidency and then asked to estimate who would be more likely to win the following day. Accounting for people's voting preferences, results showed that people rated the candidate they had imagined becoming president as more likely to win (Carroll, 1978). In a second experiment, similar findings were observed when participants were asked to imagine their university's football team either having a poor or good season. Those who imagined a good season were

more likely to predict a major bowl bid than those who imagined a poor season (62.9% vs. 39.5%, Carroll, 1978). These findings suggest that people who are led to consider the prospect of mistakenly implicating an innocent suspect will think that outcome is more likely to eventuate than otherwise. As a result, those in the "wrongful conviction" condition were expected to be more worried (than those in the "rightful conviction" condition) about implicating an innocent suspect and less likely to make positive identification decisions from both target-present and target-absent lineups. To examine the role that consequence perceptions might play in the identification process, each identification decision was associated with a crime that was either low in seriousness or high in seriousness. It was expected that the shift in choosing attributable to innocent suspect identification likelihood would be greater for decisions associated with more serious crimes. Having no reason to expect otherwise, the decrease in choosing in the wrongful conviction condition was anticipated to occur regardless of whether the target appeared in the lineup or not, resulting in lower target-present accuracy and higher target-absent accuracy relative to the rightful conviction condition.

Method

Design and Participants

A 2 (case outcome: wrongful conviction, rightful conviction) × 2 (crime seriousness: serious, not serious) × 2 (target presence: present, absent) mixed design was used. Participants were randomly assigned to the wrongful or rightful conviction condition and made identification decisions from several lineups across which target presence and crime seriousness varied. As there was no basis for estimating how large the hypothesised effects might be, a sample size of 30 subjects for each level of the between-subjects manipulation was chosen to at least provide an appropriately powered test of the manipulation itself (Wilson VanVoorhis & Morgan, 2007). Sixty volunteers (40 female and 20 male) took part in

the experiment, receiving either an honorarium of $15 or first year course credit for their participation. Their ages ranged from 17 to 61 years (M = 22.18 SD = 6.70).

Materials

Mini-lineup procedure and stimuli. The hypothesised effects were tested using a mini-lineup procedure (Sauer, Brewer, & Weber, 2008; Weber & Brewer, 2006; Weber & Varga, 2012), a modified recognition memory task in which participants view a serious of faces each presented with a descriptive cue (e.g., a person's name), before again being shown each cue and judging whether the face previously paired with that cue is present in a lineup of faces. While somewhat removed from the traditional identification paradigm in which a participant views a mock-crime being committed before usually making only a single identification decision, the mini-lineup procedure was chosen for an initial examination of the hypothesised effects for several reasons. Requiring the participants to make identification decisions for multiple stimuli economically addresses the sometimes problematic issue in eyewitness identification research that facial stimuli are highly idiosyncratic and, therefore, effects that are found for one target do not necessarily generalise to others (Brewer, Weber, & Semmler, 2005). Further, this procedure was also able to accommodate a within-subjects manipulation of whether or not the target appeared in the lineup. Moreover, the cue word presented with the each target face was able to be used to describe a crime and serve as the manipulation of the associated seriousness of the identification consequences. The initial purpose of pairing each stimulus with a cue word in the mini-lineup procedure was to encourage people to search the lineup for the face that was presented with that cue in the study phase, rather than just deciding if anyone in the array looks familiar. Previous research (Sauer et al., 2008; Weber & Brewer, 2006; Weber & Varga, 2012) has shown that people are able to make this distinction (i.e., they were able to correctly reject a face that was previously seen with a different cue). Participants in this experiment were not tested with cases where a

previously seen face was paired with the wrong cue. However, they were told that they would have to remember which face is paired with each cue.

Participants completed 40 mini-lineups (used in Sauer et al., 2008; Weber & Varga, 2012). Each lineup was comprised of a target face, three fillers and a target replacement for target-absent lineups. Following the procedure of Weber and colleagues the mini-lineup procedure was broken into two counterbalanced blocks of 20. In the study phase participants were shown the target face photographs in random order, each paired with a 1–2 word description of a crime (e.g., homicide, trespassing). The cue appeared first, presented on the screen for 2s, before the face appeared underneath for 1s. This was done to ensure that encoding of the cue word did not interfere with encoding of the face. The exposure duration of the face was chosen based on previous piloting with these stimuli to detect durations that would not produce ceiling performance. In the identification test phase participants were shown each crime cue (again in random order) above a lineup of four people. The face that was presented with that cue during the study phase was either present or absent in the lineup.[4] If participants thought the target was in the lineup they indicated this by clicking on the face of the person they thought it was. If they thought the target was not in the lineup they indicated this by clicking on a "not there" button centred below the lineup. Confidence in each identification decision was rated on an 11-point decile scale (0-100% confident, with anchors at each end of the scale).

Innocent suspect identification likelihood manipulation. Prior to making their own identification decisions participants were presented with (made-up) case information in which it was implied that a positive identification of the police's suspect had either been correct or incorrect. Preceding an account of what someone in the vicinity of a burglary witnessed, a

[4] Half of the lineups were target-present and half were target-absent. Within these parameters target presence was randomly determined for each participant.

short summary of case details contained the target information. The burglary suspect was described as having been identified by the witness and subsequently convicted of the crime. Information pertaining to the recovery of the stolen goods either suggested that the suspect had in fact been *guilty* or *innocent* (and, therefore, that either a rightful or wrongful conviction had occurred). That is, the stolen goods were either described to have been sold to a pawn store before or after the suspect was taken into custody. A full transcript of the information participants read can be found in Table 1.

As manipulation checks, participants were asked a series of questions to ascertain whether they had come to the intended conclusions and how well they had retained the key information from the scenario. They were asked how worried they were about implicating an innocent suspect and responded on a 7-point scale (0 = *not at all*; 6 = *extremely*). Then they selected whether the witness in the scenario had made a positive identification or had rejected the lineup. Finally they rated both the likelihood that the police's suspect had been the culprit and that justice had been served on 7-point scales (0 = *definitely not*; 6 = *definitely*).

Crime seriousness manipulation. As previously mentioned, the gravity of the identification decision consequences was manipulated by the seriousness of the crime described by the cue word(s) associated with each target face. To categorise crime seriousness a pilot was run where participants (N = 29) were given a list of 80 alphabetically ordered crimes and asked to rate each for how serious they felt it was (on an 11-point scale; 0 = *not at all serious*; 5 = *moderately serious*; 10 = *extremely serious*). People were also given the option of indicating that they did not know what the crime was, with crimes that more than one participant did not recognise (e.g., perjury, voyeurism) being excluded from consideration. Because a couple of respondents rated almost everything as "extremely serious", ratings more than two standard deviations from the mean were excluded to eliminate outliers. From the remaining data 20 high seriousness (Ms = 8–10) and 20 low

seriousness (*M*s = 3–5) crimes were selected. These were compiled into two (half high, half low seriousness) lists of crime cues to be counterbalanced across stimulus blocks. See Table 2 for a copy of the crime lists. The lists were staggered, so each face was paired once with a high seriousness and once with a low seriousness crime. Crime cues were not randomly paired with faces because I decided it was more important for the association to be somewhat realistic (e.g., sexual assault cues paired with male faces).

Table 1

Case Details (and Witnessing Event Information) Participants Read in Experiment 1

Condition	Case details
Both conditions	Police were investigating a series of highly publicised jewellery store robberies caught on CCTV camera. In each case the footage showed a person wearing a black coat and balaclava smashing open glass cabinets in the store and fleeing the premises with a bag of merchandise. Circumstantial evidence led to the arrest of a suspect. A witness, who saw the face of a person in the vicinity of the store on the night the latest robbery occurred, identified the police's suspect from a lineup. The person was convicted and sentenced to 10 years in prison.
Wrongful conviction condition	After this point, however, the robberies kept occurring and some of the jewellery from previous robberies was recovered from a pawn shop to which it had been sold <u>after</u> the suspect had been arrested.
Rightful conviction condition	After this point the robberies stopped and some of the jewellery was recovered from a pawn store to which it had been sold shortly <u>before</u> the suspect had been arrested.
	Witnessing event information
Both conditions	The witness had finished a shift at work around 11pm and was walking to their car parked in a well-lit lot across from a closed shopping precinct where an alarm was going off. A man came running out of the premises, carrying a duffle bag. He bumped into the witness as he rushed past and his coat fell open, showing a gun strapped to his belt. The man got into a white sedan parked in the car park and sped off. The witness called the police, who arrived at the scene minutes later, finding a jewellery store in the shopping precinct ransacked. The witness described the above events, also providing a detailed description of the appearance of the perpetrator. They described the man as Caucasian, medium height and build, having short dark hair and a scruffy beard. In particular, had had a piercing in his left ear and was wearing a belt with a big silver buckle.

Table 2

Crime Cues Used in Experiment 1, in Alphabetical Order by List Number and Crime Seriousness Category

	Crime seriousness category	
List number	Serious	Not serious
One	Attempted murder Bombing Child Abduction Hijacking Hit and run Homicide Mass shooting Poisoning Rape Torture	Cocaine Possession Fraud Graffiti Harbouring fugitives Illegal weapon possession Indecent exposure Pickpocketing Procuring Prostitution Speeding Trespassing
Two	Armed robbery Arson Assassination Child Abuse Hostage taking Human trafficking Manslaughter Serial killings Sexual assault Stabbing	Breaking and entering Bribery Disorderly conduct Forgery Handling stolen goods Illegal immigration Selling Ecstasy Serving minors Shoplifting Vandalism

Procedure

Participants in all the experiments were volunteers recruited from Flinders University.[5] They either signed up to participate (in a study on eyewitness memory) via an online research participation system, or were recruited in person on the university campus. Participation took place at the Flinders University eyewitness memory lab. After reading a letter of introduction and signing a consent form, each person was seated at a computer in an individual cubicle, told to enter their basic demographic information before following the

[5] The sample of participants in Pilot 3 also included some community members.

onscreen instructions through the experiment.[6] First, participants studied the target faces for one block of mini-lineups. As well as warning people to pay close attention to the faces, instructions stipulated that which face belonged with which cue word would need to be remembered. Next, the case information (comprising the previously outlined manipulation of choosing leading to a critical mistake) was read. This was framed as additional stimulus information about which people would be asked questions at a later time (which they were in the form of the manipulation check questions). Following this, the test phase was completed for the previously studied faces. Instructions briefly described the task and clarified (a) that the target may or may not be in the lineup, (b) how to indicate their response, (c) that no lineups would include a previously studied face with the wrong cue word, and (d) that confidence would be measured following each decision. Immediately prior to beginning the test phase participants were asked to try to imagine that each lineup for which they were making an identification decision related to a case in which the crime described had been committed. On completing the test block they were told that no more questions would refer to those faces. A second block of mini-lineups was then completed in the exact same manner as previously, with different faces and cues. The same case information (i.e., a repeat of the manipulation of innocent suspect identification likelihood) was presented as an opportunity to review the information before being questioned about it. Finally participants completed the manipulation check questions, were thanked for their participation and appropriately debriefed.

Statistical Approaches

As unconventional analysis techniques were used throughout this thesis, these are described and the rationale for implementing them is outlined in what follows. The

[6] They were also told they could ask for clarification of instructions at any point, which a small number of participants did.

manipulation check data in all of the experiments reported in this thesis had a tendency to not fall within normal distribution parameters. This is perhaps unsurprising as the manipulations were designed to cause ratings to diverge between experimental conditions and be skewed towards the upper end of the scales used in one condition and the lower end in the other. This characteristic of the data prevented standard practices of data transformation and removing outliers (see Field, Miles, & Field, 2012) from bringing the distributions within normal parameters. Therefore non-parametric tests—equivalents of 2 × 2 mixed ANOVAs, independent and paired samples *t*-tests tests (described in Field et al., 2012; Wilcox, 2012) using 20% trimmed means (recommended by Wilcox, 1995a, 1995b)—were run. These analyses were conducted using the WRS2 package (Mair, Wilcox, & Schoenbrodt, 2015) in the open-source program R (R Development Core Team, 2015). However, as the key results did not differ from what was indicated by standard versions of the tests, and the standard tests are much easier to interpret by readers, the inferential statistics from the parametric tests are reported in text. All of the non-parametric test results are tabulated in Appendix A. An alpha level of .05 was used as the cut-off for all inferential analyses. Further, means, standard deviations and measures of effect size (which were calculated from the raw data) are reported for all manipulation check results. The measure of effect size used for differences between two means was Cohen's *d*, for which guidelines for small, medium and large effects are .30, .50 and .80, respectively (Cohen, 1988, 1992). According to standard recommendations, confidence intervals are reported for *d* (Tabachnick & Fidell, 2013). For interactions Cohen's *f* was used as the effect size measure; guidelines for small, medium and large effects are .10, .25 and .40, respectively (Cohen, 1988, 1992).

All main[7] analyses in this thesis were conducted using mixed-effects modelling, which is an extension of regression (Baayen, Davidson, & Bates, 2008; Field et al., 2012).

[7] As well as one manipulation check test in Experiments 2 and 3

The reason for implementing this analysis technique was that in all of the experiments conducted participants made multiple identification decisions. Note that this type of analysis has been previously used in identification research with similar designs (e.g., Weber & Varga, 2012). While a common practice for examining data collected in multiple trials has been to aggregate people's responses (e.g., calculate the mean rating, or proportion correct) and analyse these scores using ANOVA (or *t*-tests when only comparing two groups), this approach has been found to be associated with an increased Type-I error rate (e.g., Jaeger, 2008; Murayama, Sakaki, Van, & Smith., 2014; Quené & van den Bergh, 2008). The increased propensity for error when analysing the data in this way has been attributed to the failure to account for random variability across items (Barr, Levy, Scheepers, & Tily, 2013). The significance of predictor variables are estimated in ANOVA with an error term that allows variance across subjects (Field et al., 2012). But the effects are also likely to vary depending on the characteristics of the stimuli (Clark, 1973), something that is not adequately accounted for when responses are collapsed across trials (Barr et al., 2013). Mixed-effects modelling is able to address this problem, as the data are analysed on a trial by trial basis and the regression parameters (i.e., intercepts and slopes) can be specified to allow random variation across both participants and stimuli (Baayen et al., 2008; Barr et al., 2013).

When constructing models for hypothesis testing, it has been stipulated that to appropriately account for random variability in subjects and items it is necessary to include a random intercept for both participant and stimulus, as well as random by unit (i.e., participant or stimulus) slopes for all manipulations in the experimental design occurring within that unit (Barr et al., 2013).[8] An exception to this is if there is only one observation at each level of the predictor for that unit. To clarify, consider the following example of an experiment using a 2

[8] Although the evidence for these guidelines is based on simulations with continuous outcome variables, the authors argued that logically their findings should extend to cases with dichotomous outcome variables (although this assertion is still technically an empirical issue).

× 2 mixed design to test the interaction between A and B on C. Imagine variable A was manipulated between-subjects, variable B was manipulated within-subjects and variable C represents the dependent variable ratings that were repeated for each participant across 10 stimulus faces (with half of the stimuli, faces 1–5, comprising the first level of B and the other half, faces 6–10, comprising the second level of B). Following the recommendations of Barr et al. (2013), the appropriate "random effects structure" for analysing the effects of A and B, as well as the A × B interaction, on C would be as follows. First, random intercepts would need to be entered for participant and stimuli. Because variable A was manipulated between subjects and, therefore, each participant was only exposed to one level of the manipulation, no random slope for A by participant would need to be added to the model. However, each stimulus was rated in association with both levels of A, across multiple occasions. Therefore, a random slope for stimulus by variable A would need to be added to the model. Further, each participant made several (5) ratings at each level of B, so a random slope for participant by B would be appropriate. In this example I have specified that stimuli 1–5 represented the first level of B and 6–10 represented the second level (imagine B was a manipulation of race and faces 1–5 were African-American and faces 6–10 were Asian). Therefore, by definition, the stimuli were not able to vary across variable B and it would not be sensible to add any random slope for stimuli by B. However, imagine that the stimuli were such that the within-subjects manipulation of B was able to be randomly assigned to each stimulus (e.g., the manipulation was whether or not a face was smiling and, therefore, both a smiling and not smiling photo was be taken of the same face and which version each participant saw was randomly determined). In this case a random slope for B by stimulus would be necessary. Further, in this alternative scenario a random slope for stimulus by the A × B interaction would also need to be included in the model.

However, there are practical issues associated with implementing the recommended random effects structure. Depending on the design of the experiment the appropriate random effect structure can cause the model to have too many parameters to fit the data (Baayen et al., 2008; Barr et al., 2013). In situations where the aim of modelling is to find the best explanation for the data it can be appropriate to simplify the random effect structure to include random intercepts only (e.g., Baayen et al., 2008). But when it comes to hypothesis testing, it has been argued that it may be preferable to leave the model with too many parameters (Barr et al., 2013). While there is no evidence to suggest that this causes any problems for accurately testing a hypothesis, models with only random intercepts can lead to an unacceptably large increase in Type-I error rate and loss of power (Barr et al., 2013). However, implementing a "maximal" random effect structure is not always possible. Often the added complexity it brings to the model will cause a failure to converge (i.e., the model breaks). Unfortunately this has been noted to be particularly common with categorical data (Barr et al., 2013) and occurred for almost all of the models reported in this thesis. Barr and colleagues (2013) demonstrated several simplifications that can be made to the random effects structure without compromising the power and error rate of the model. Through simulations they showed that models with the correlation parameter removed[9] performed comparably to models with a full random effects structure. Also, relevant to cases where an interaction is being tested, it has been demonstrated that as long as a random slope for the interaction was included, the lower order slopes (i.e., for the main effects) could be removed without compromising the error rate and power of the model (Barr, 2013). Even removing a (participant or stimulus) random intercept to preserve the inclusion of a critical random slope performed acceptably (Barr et al., 2013). Based on these recommendations, regarding which

[9] That is, correlation between random slopes and intercepts was not permitted. But importantly the random variances that could be confounded with the corresponding predictor variables (aka fixed effects) were still taken into consideration.

characteristics of a random effects structure can be removed without compromising the error rate and power of a model, I devised a step-wise system for simplifying the random effects structure in the models I constructed until the model converged. This procedure is reported in Appendix B, along with a description of the random effects structure for each model reported in this thesis. Note in a small number of cases the random effects structure had to be simplified to a random intercepts only model—one that is associated with an increased Type-1 error rate (Barr et al., 2013). These cases are clearly noted in text.

All mixed-effects models were built using the statistical software package lme4 (Bates, Maechler, Bolker, & Walker, 2016) in R (R Development Core Team, 2015). The main outcome variables were either whether people chose from a lineup (not choosing coded as 0, choosing coded as 1) or accuracy (incorrect responses coded as 0, correct responses coded as 1). In supplementary analyses for some experiments outcome variables were confidence (higher values equalling higher confidence) and willingness to testify (no coded as 0, yes coded as 1). Where the outcome variables were dichotomous, logistic mixed-effects models (i.e., an extension of logistic regression) were created using the logit link function. Conceptually this meant that the analysis predicted the probability of the outcome variable from the values of the predictor variable(s) (Field et al., 2012; Jaeger, 2008). Where the outcome variable was continuous a linear mixed-effects model was created, which is an extension of ordinary regression where the value of the outcome variable is estimated by the predictor variable (Field et al., 2012). In all cases, first a baseline model was fitted by defining the outcome variable and the random effects structure (i.e., the variables for which random variation across subjects and items should be allowed). Then the predictor variables of interest (i.e., fixed effects) were entered sequentially in order of importance. The improvement of model fit after the addition of each predictor is reported—which is essentially the comparison of the same model with and without that predictor. If this is

significant, it indicates that the more complex model provides a better fit to the data (Baayen et al., 2008; Barr et al., 2013; Field et al., 2012; Jaeger, 2008). However, the regression coefficients (*b*) were used as the primary means for interpreting the model as these indicate the significance of a predictor after controlling for all other predictors in the model (Gelman & Hill, 2006). Following Gelman and Hill (2007) confidence intervals were calculated for each regression coefficient—$b \pm (SE_b * 1.96)$—and where these did not include zero the corresponding variable was interpreted as a significant predictor of the outcome variable. Descriptive statistics are reported to illustrate all model results.

Results

Manipulation Checks

First, how well participants processed the information relevant to the manipulation was examined.[10] Most participants (90%) correctly indicated that the witness in the scenario had made a positive identification. Perceived guilt of the police suspect in the scenario was significantly higher in the rightful conviction (M = 3.71, SD = 0.85) than the wrongful conviction (M = 2.48, SD = 1.03) condition, $t(57) = -4.97$, $p < .001$, d = 1.30, d 95% CI [0.73, 1.86]. Perceived justice was also significantly higher in the rightful conviction (M = 3.71, SD = 1.15) than the wrongful conviction (M = 2.32, SD = 1.25) condition, $t(57) = -4.44$, $p < .001$, d = 1.15, d 95% CI [0.59, 1.70]. As correctly interpreting and remembering this information was essential to the manipulation of people's attitude towards choosing from a lineup, participants were coded for their accuracy on these three questions to identify any cases where they clearly did not process the relevant information. Correct answers included indicating that (a) the witness in the scenario had made a positive identification, (b) expressing the opinion that the suspect was "probably," "likely" or "definitely" guilty and

[10] One participant's data (in the rightful conviction condition) did not record correctly. Because this was not noticed until data collection was complete and analysis had commenced a replacement participant was not run.

justice had been served in the rightful conviction condition and that this was probably, likely or definitely not the case in the wrongful conviction condition. Note that "unsure" responses were omitted from this analysis. Three participants (two in the rightful conviction condition and one in the wrongful conviction condition) gave incorrect answers to all three questions. Their data were removed from all further analyses.

It was expected that those in the wrongful conviction condition would think that mistakenly implicating an innocent suspect was more likely than those in the rightful conviction condition and be more worried about this as a result. While worry about incorrectly picking an innocent suspect from a lineup was found to be slightly higher in the wrongful conviction condition (M = 4.69, SD = 1.17) than the rightful conviction condition (M = 4.26, SD = 1.43), this difference was not significant, $t(54) = 1.24$, $p = 0.22$, $d = 0.33$, d 95% CI [-0.20, 0.86]. Thus, the manipulation check results suggest that most participants processed the relevant information, but that this did not translate to a significant difference in reported apprehension towards implicating an innocent suspect.

Main Analyses

Choosing. The main prediction was that people in the wrongful conviction condition would be less likely to make positive identification decisions than those in the rightful conviction condition. Further, as evidence of consequence perceptions affecting their decisions, this effect was expected to be greater for cases associated with serious crimes. In order to test these predictions a logistic mixed-effects model was constructed with choosing as the outcome variable. Case outcome, followed by crime seriousness and their interaction were entered as predictors. No significant effects were observed. Aggregate choosing is summarised in Table 3. See Table 4 for model fit statistic and fixed effect coefficients.

Table 3

Percentage Choosing Rates Across the Case Outcome and Crime Seriousness Conditions

	Crime seriousness					
Case outcome	N trials	Serious	N trials	Not serious	N trials	Total
Rightful conviction	540	55.74	540	52.22	1080	53.98
Wrongful conviction	580	56.55	580	58.79	1160	57.67
Total	1120	56.16	1120	55.63	2240	55.90

Table 4

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Effect of Case Outcome and the Case Outcome by Crime Seriousness Interaction on Choosing

Fixed effect predictors	χ^2	*df*	*p*
Case outcome	1.26	1	.26
Crime seriousness	0.06	1	.81
Case Outcome × Crime Seriousness	1.76	1	.19
	b	SE_b	95% CI
Intercept	0.24	0.07	[0.11, 0.38]
Case outcome	-0.15	0.14	[-0.42, 0.11]
Crime seriousness	0.03	0.09	[-0.15, 0.21]
Case Outcome × Crime Seriousness	0.24	0.18	[-0.11, 0.60]

Accuracy. As no effect of case outcome on choosing was observed, no effects on accuracy were expected. Nevertheless, the prediction that those in the wrongful conviction condition would exhibit higher target-absent accuracy and lower target-present accuracy than those in the rightful conviction condition was tested.[11] A logistic mixed-effects model was built with accuracy as the outcome variable. Case outcome, target presence and the Case Outcome × Target Presence interaction were entered as predictors. Again, none of the predictors were found to be significant (see Table 5 for inferential and Table 6 for descriptive statistics).

Table 5

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Case Outcome and Target Presence on Accuracy

Fixed effect predictors	χ^2	*df*	*p*
Case outcome	1.87	1	.17
Target presence	0.43	1	.51
Case Outcome × Target Presence	0.31	1	.58
	b	SE_b	95% CI
Intercept	0.80	0.10	[0.61, 0.99]
Case outcome	0.26	0.19	[-0.11, 0.64]
Target presence	0.10	0.15	[-0.20, 0.39]
Case Outcome × Target Presence	-0.17	0.31	[-0.78, 0.43]

[11] Technically, the prediction that the shift in choosing would be greatest for decisions associated with serious crimes suggests that Case Outcome × Target Presence interaction on accuracy would also be greater for serious crimes. However, the results were collapsed across crime seriousness because (a) the descriptive statistics indicated that the pattern of results did not differ for serious and not serious crimes and (b) a model testing the Case Outcome × Target Presence × Crime Seriousness interaction was too complex to be fitted to the data with an appropriate random effects structure.

Table 6

Percentage Accuracy Rates Across the Case Outcome and Target Presence Conditions

	Target presence					
Case outcome	N trials	Present	N trials	Absent	N trials	Total
Rightful conviction	540	69.44	540	69.26	1080	69.35
Wrongful conviction	580	66.21	580	62.59	1160	64.39
Total	1120	67.77	1120	65.80	2240	66.80

Supplementary Analysis

Post-decision confidence. Although people's choosing was not found to be affected by the case outcome manipulation, I thought it was possible that, at a minimum, their post-decisional confidence was affected. Confidence in a decision made is qualifying information that might be more sensitive to changes in motivation than the decision itself. Therefore, it is possible that people were less confident in their positive identifications in the condition where they read about a wrongful conviction than the condition where they read about a rightful conviction. In order to explore this possibility confidence was entered as the outcome variable in a linear mixed-effects model, where case outcome and whether or not the lineup was chosen from, as well as the interaction between these variables, were sequentially entered as predictors. Target presence was controlled for as confidence was found to be significantly higher for target-present cases.[12] No other variables were found to be significant predictors of confidence (see Tables 7 and 8 for descriptive and inferential statistics, respectively).

[12] This is not a common finding in the literature (e.g., see Brewer & Wells, 2006), however, it was found fairly consistently in the experiments reported in this thesis. Tentatively perhaps getting people to think about identification consequences led to greater post-decisional ambivalence when no memory cue had been available in the lineup.

Table 7

Mean (and Standard Deviation) Post-Decision Confidence Across the Case Outcome and Lineup Decision Conditions

	Lineup decision					
Case outcome	N trials	Picked	N trials	Rejected	N trials	Total
Rightful conviction	669	63.86 (30.39)	497	57.61 (28.90)	1080	60.98 (29.86)
Wrongful conviction	583	62.75 (27.18)	491	55.21 (25.52)	1160	59.56 (26.73)
Total	1252	63.27 (28.71)	988	56.42 (27.28)	2240	60.25 (28.29)

Table 8

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Case Outcome and Lineup Decision on Post-Decision Confidence (After Controlling for Target Presence)

Fixed effect predictors	χ^2	*df*	*p*
Target presence	24.20	1	<.001
Case outcome	0.04	1	.84
Lineup decision	3.81	1	.05
Case Outcome × Lineup Decision	0.01	1	.93
	b	SE_b	95% CI
Intercept	6.01	0.22	[5.58, 6.44]
Target presence	0.73	0.15	[0.44, 1.02]
Case outcome	0.09	0.44	[-0.76, 0.95]
Lineup decision	0.34	0.18	[-0.002, 0.69]
Case Outcome × Lineup Decision	0.03	0.34	[-0.64, 0.71]

Discussion

Experiment 1 used a hypothetical consequences paradigm to investigate whether greater perceived likelihood of mistakenly implicating an innocent suspect would lead to decreased choosing from a lineup, especially for decisions when the consequences associated with making a mistake were more severe. While there was evidence indicating that participants processed the information relevant to the manipulation of likelihood perceptions, this did not translate to a significant difference in expressed concern about incorrectly implicating an innocent suspect (although the pattern was in the hypothesised direction). Two possible interpretations of this are (a) the manipulation was not strong enough and or (b) an insensitive manipulation check measure was used. The scenario participants read only implied that perhaps either a wrongful or rightful conviction had occurred. Therefore, it is conceivable that the manipulation did not provide a sufficiently decisive instance of a witness making a correct or incorrect suspect identification to influence people's attitudes regarding the probability that choosing from a lineup could result in a critical error. If the two conditions did not produce a difference in attitude towards the outcomes associated with choosing from a lineup, this would explain why choosing was not affected as predicted (and further why the crime seriousness manipulation had no effect, as this relied on the shift in choosing to operate). It is also possible that the measure of risk associated with choosing from a lineup was not sensitive to a difference in attitudes. Participants in the condition where choosing was presented positively had a relatively high rating of concern about mistakenly identifying an innocent suspect (M = 4.29 SD = 1.41, on a scale of 0–6). Thus, a ceiling effect may have constrained any increase in the condition where choosing was framed negatively. In hindsight, most people would be expected to express a high level of concern as this is arguably the socially acceptable attitude to have. In support of this interpretation, research

has shown that people have a tendency towards socially desirable responding in questionnaires (Paulhus, 2002).

In sum, whether or not the associated consequences of eyewitness identification decision influence choosing in this paradigm was not resolved by Experiment 1.

Experiment 2

To address the potential methodological limitations identified in Experiment 1, two major changes were made in Experiment 2. First, to strengthen the manipulation of mistaken identification likelihood, the scenario was altered so that the guilt or innocence of the suspect appeared to be irrefutably proven (i.e., with DNA evidence), making it clearer that either wrongful or rightful conviction had occurred. Second, the manipulation check measure was revised. As was previously discussed, it is possible that the manipulation used in Experiment 1 did sufficiently alter people's perception of the outcome associated with choosing from a lineup, but the difference in attitudes was not appropriately measured by asking people to express their concern about implicating an innocent suspect. This measure was intended to reflect participants' overall attitude towards choosing (and was expected to differ based on the manipulated likelihood that choosing would lead to positive vs. negative outcomes). However, in hindsight it may have been more appropriate to measure perceived likelihood of the negative outcome in question explicitly as this is what had been manipulated. Therefore, participants in Experiment 2 were asked to rate the likelihood that choosing from a lineup would lead to the error of picking an innocent suspect, with the expectation that this would be lower in the rightful conviction condition than in the wrongful condition.

Again I aimed to test the hypothesis that positive identification decisions would be lower in the wrongful conviction condition than the rightful conviction condition and that this effect would be greater for decisions associated with more serious crimes. Further, it was

predicted that this shift in choosing would lead to lower target-present and higher target-absent accuracy in the wrongful conviction condition than the rightful conviction condition.

Method

Design and Participants

As in Experiment 1, a 2 (case outcome: rightful conviction, wrongful conviction) × 2 (crime seriousness: serious, not serious) × 2 (target presence: present, absent) mixed design was used where participants made multiple identification decisions from mini-lineups across which target presence and associated crime seriousness varied and were randomly assigned to the wrongful or rightful conviction condition. To account for the small possibility that the manipulation in Experiment 1 had been successful, but the effect was too small to detect given the sample size that was used, the previously used sample of 60 participants was doubled (and a few extra participants were run in each condition to account for the likelihood that—as in Experiment 1—a small number of people would be judged to have not adequately processed the case information within which the main manipulation was contained). Therefore, 128 participants were run (71 females and 57 males). Their ages ranges from 17 to 33 (M = 20.23, SD = 2.92).

Materials

Mostly the same materials as in Experiment 1 were used; however, the following changes and additions were made. In order to strengthen the manipulation of how likely it was perceived to be that choosing from a lineup could result in a serious error, the case example that participants read was modified to include post-conviction acquisition of DNA evidence that matched or did not match the defendant. Specific details of the changes that were made can be found in Table 9.

Table 9

Changes (Highlighted in Bold) to the Case Details Participants Read in Experiment 2

Condition	Case details
Both conditions	Police were investigating a series of highly publicised jewellery store robberies caught on CCTV camera. In each case the footage showed a person, wearing a black coat and balaclava smashing open glass cabinets in the store and fleeing the premises with a bag of merchandise. Circumstantial evidence led a **suspect to be brought in for questioning**. A witness who saw the face of a person in the vicinity of the store on the night the latest break occurred identified the police's suspect from a lineup. **Based on this the suspect was remanded in custody and brought to trial. The witness testified in court and the defendant** was convicted and sentenced to 10 years in prison.
Wrongful conviction condition	After this point, however, the robberies kept occurring, and some of the jewellery from previous robberies was recovered from a pawn shop to which it had been sold after the suspect's arrest. **Additionally, one of the jewellery stores which was vandalised and robbed began to undergo renovations. This led a shard of glass with the perpetrator's blood on it to be discovered under one of the smashed display cabinets. DNA testing on the blood sample showed that it was a non-match to the now convicted suspect.**
Rightful conviction condition	After this point the robberies stopped and some of the jewellery was recovered from a pawn store to which it had been sold shortly before the suspect's arrest. **Further, one of the jewellery stores which was vandalised and robbed began to undergo renovations. This led to a shard of glass with the perpetrator's blood on it being discovered under one of the smashed display cabinets. DNA testing on the blood sample confirmed that it was a match to the now convicted suspect.**

In addition to rating how confident they were, participants were asked to report after each lineup decision whether or not they were willing to testify to that decision. This is an additional measure of post-identification certainty that has been used in some previous identification decision research (e.g., Bradfield, Wells, & Olson, 2002) and may be useful in this particular paradigm to remind participants of the role their decision could have in the proceedings outcome of a case and better imagine the consequences that may ensue. Additional[13] manipulation check measures were also taken. That is, people were asked to

[13] Despite the potential problems associated with this measure, participants were again asked how worried they were about implicating an innocent suspect in order to compare to the results of Experiment 1.

indicate how likely they thought it was that choosing from would lead to an innocent suspect being identified (0 = *not at all likely*; 6 = *extremely likely*). To verify that the crime seriousness manipulation presented a different level of motivation for people avoid making a crucial error, participants were also asked to rate how important they thought it was for justice to be done for each crime cue (0 = *not at all important*; 6 = *extremely important*).

Procedure

The procedure was the same as Experiment 1, bar the necessary changes to accommodate the additional measures. Post-decisional confidence ratings were followed by indicating whether or not the participants would be willing to testify to that particular decision. The "likelihood' manipulation check measure preceded the "worry" measure, and lastly the crime cues were rated on importance.

Results

Manipulation Checks

As in Experiment 1, participants were asked to answer three questions assessing their recollection and comprehension of the scenario in which the perceived likelihood of implicating an innocent suspect was manipulated. Of the 127 participants who answered these questions,[14] 119 correctly indicated that the witness in the case example had chosen from the lineup. Ratings of guilt were significantly higher in the rightful conviction condition (M = 4.52, SD = 0.97) than the wrongful conviction condition (M = 2.33, SD = 1.26), $t(125) = -11.00$, $p < .001$, $d = 1.95$, d 95% CI [1.53, 2.37]. Justice was also perceived to be significantly higher in the rightful conviction condition (M = 4.42, SD = 1.27) than the wrongful conviction condition (M = 1.89, SD = 1.33), $t(125) = -10.96$, $p < .001$, $d = 1.95$, d 95% CI [1.53, 2.37]. As described in Experiment 1, each participant was coded for their accuracy on these questions. Three participants (one in the wrongful conviction condition and

[14] One participant had to leave early and did not complete the manipulation check ratings.

two in the rightful conviction condition) gave incorrect answers to all three questions. Their data were removed from all further analyses.

As in the previous experiment, participants were asked how *worried* they would be about mistakenly identifying an innocent suspect if they chose from the lineup. Again, people's rating of worry did not significantly differ between the choosing negative (M = 4.70, SD = 1.28) and the choosing positive condition (M = 4.43, SD = 1.42), $t(122) = 1.14$, $p = .26$, d = 0.20, d 95% CI [-0.15, 0.55]. Participants were also asked how *likely* they thought it was that choosing from a lineup would lead to an innocent suspect pick. As predicted, those in the wrongful conviction condition (M = 4.39, SD = 1.02) had significantly higher ratings than those in the rightful conviction condition (M = 3.97, SD = 0.93), $t(122) = 2.42$, $p = .02$, d = 0.43, d 95% CI [0.07, 0.79].

In order to determine whether people thought it was more important to achieve justice for the serious than the not serious crimes, a linear mixed-effects model was created with justice importance as the outcome variable and crime seriousness as the predictor. Results showed that crime seriousness was a significant predictor of justice importance, with importance being higher for more serious crimes, $\chi^2(1) = 69.67$, $p < .001$, b = 2.12, SEb = 0.17, b 95% CI [1.79, 2.45].[15] This is consistent with the aggregate statistics for the serious (M = 5.39, SD = 0.43) and not serious (M = 3.26, SD = 1.00) importance ratings.

Main Analyses

Choosing. As in the previous experiment, the main prediction was that people who were oriented towards negative consequences associated with choosing would be less willing to make positive identifications from lineups. Further, this effect was expected to be greater for lineups associated with serious versus not serious crimes. In order to test these predictions, a logistic mixed-effects model was create with choosing as the outcome variable

[15] Intercept b = 3.27, SEb = 0.14, b 95% CI [2.99, 3.54]

and case outcome, crime seriousness, as well as their interaction entered as predictors. No significant effects were obtained. Aggregate choosing rates appear in Table 10 and inferential statistics are summarised in Table 11.

Table 10

Percentage Choosing Rates Across the Case Outcome and Crime Seriousness Conditions

	Crime seriousness					
Case outcome	N trials	Serious	N trials	Not serious	N trials	Total
Rightful conviction	1240	59.76	1240	58.06	2480	58.91
Wrongful conviction	1260	55.56	1260	55.95	2520	55.75
Total	2500	57.64	2500	57.00	5000	57.30

Table 11

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Effect of Case Outcome and the Case Outcome by Crime Seriousness Interaction on Choosing

Fixed effect predictors	χ^2	*df*	*p*
Case outcome	1.63	1	.20
Crime seriousness	0.19	1	.66
Case Outcome × Crime Seriousness	0.46	1	.50
	b	SE_b	95% CI
Intercept	0.32	0.06	[0.21, 0.43]
Case outcome	0.15	0.11	[-0.08, 0.37]
Crime seriousness	0.03	0.06	[-0.10, 0.15]
Case Outcome × Crime Seriousness	0.09	0.14	[-0.18, 0.36]

Accuracy. As was done in Experiment 1, accuracy rates were examined despite the predicted effects on choosing not being found (which meant that no effects on accuracy were expected).[16] The predicted interaction between case outcome and target presence on accuracy was tested as follows. A logistic mixed-effects model was constructed, with accuracy as the outcome variable. Case outcome, target presence and the Case Outcome × Target Presence interaction were entered as predictors. None of these variables were found to significantly predict accuracy. The inferential statistics are shown in Table 12 and the descriptive statistics can be found in Table 13.

Table 12

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Case Outcome and Target Presence on Accuracy

Fixed effect predictors	χ^2	*df*	*p*
Case outcome	0.09	1	.76
Target presence	2.67	1	.10
Case Outcome × Target Presence	2.23	1	.14
	b	SE_b	95% CI
Intercept	0.69	0.07	[0.55, 0.82]
Case outcome	0.04	0.13	[-0.21, 0.29]
Target presence	0.19	0.11	[-0.03, 0.40]
Case Outcome × Target Presence	0.34	0.22	[-0.10, 0.78]

[16] Again, as in Experiment 2, the results were collapsed across crime seriousness, both because there was no indication in the descriptive statistics that results differed depending on this variable and because a model accounting for it was not able to be fitted without compromising the power and Type-I error risk in the results.

Table 13

Percentage Accuracy Rates Across the Case Outcome and Target Presence Conditions

	Target presence					
Case outcome	N trials	Present	N trials	Absent	N trials	Total
Rightful conviction	1240	68.55	1240	60.97	2480	64.76
Wrongful conviction	1260	64.60	1260	63.89	2520	64.24
Total	2500	66.56	2500	62.44	5000	64.50

Supplementary Analyses

Post-decision confidence. In line with what was done in Experiment 1, I examined whether those in the rightful conviction condition were more confident in their positive identifications than those in the wrongful conviction condition. To do so, a linear mixed-effects model was built with confidence as the outcome variable. As in Experiment 1, target presence was controlled for. Case outcome, lineup decision and their interaction were then entered into the model. Results showed that, in addition to target presence, both case outcome and lineup decision were significant predictors of confidence. That is, people were more confident in their positive identifications[17] and more confident in the rightful conviction condition (see Table 14 for descriptive statistics and Table 15 for inferential statistics).

[17] This finding was completely unexpected. I wondered whether perhaps it emerged because participants were not given the option to indicate that they were unsure whether the target was in the lineup or not. Clark's (2003) model of identification decision making stipulates that people first decide whether to choose from a lineup, then whether to reject it and finally, if neither of those options are pursued, people say they "don't know". Extrapolating from this, when a "don't know" option is not available, it could be the case that lineup rejections occur in all instances where people are not confident enough to make a positive identification decision. By default this would mean that the lineup rejections include the subset of cases where people would actually prefer to say they are not sure and, therefore, make these rejections with particularly low confidence (bringing down the average confidence for lineup rejections compared with positive identifications). However, while the same pattern of results occurred in the following experiment (i.e., Experiment 3) and a—just—non-significant trend in the same direction can be seen when referring back to Experiment 1, this effect disappears thereafter (in Experiments 4 and 5).

However, the interaction between case outcome and lineup decision did not significantly predict confidence.

Table 14

Mean (and Standard Deviation) Post-Decision Confidence Across the Case Outcome and Lineup Decision Conditions

	Lineup decision					
Case outcome	N trials	Picked	N trials	Rejected	N trials	Total
Rightful conviction	1461	69.43 (22.82)	1019	61.81 (24.31)	2480	66.30 (23.74)
Wrongful conviction	1405	64.77 (26.65)	1115	54.35 (27.84)	2520	60.16 (27.67)
Total	2866	67.15 (24.88)	2134	57.91 (26.47)	5000	63.20 (25.97)

Willingness to testify. In order to examine people's willingness to testify ratings in the same way as confidence, a logistic mixed-effects model was created with target presence (again to control for it), case outcome, lineup decision and the Case Outcome × Lineup Decision interaction entered as predictors. Note that only random intercepts were able to be fitted in the random effects structure for this model. As mentioned in the statistical approaches section of this chapter, this means that the results should be interpreted with caution because the risk of Type-I error is inflated. Apart from target presence, lineup decision was the only variable found to be a significant predictor of whether people were willing to testify to their decision, with this being more likely for positive identification decisions (see Table 16 for descriptive statistics and Table 17 for model output).

Table 15

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Case Outcome and Lineup Decision on Post-Decision Confidence (After Controlling for Target Presence)

Fixed effect predictors	χ^2	*df*	*p*
Target presence	61.18	1	<.001
Case outcome	6.26	1	.01
Lineup decision	43.31	1	<.001
Case Outcome × Lineup Decision	0.22	1	.64
	b	SE_b	95% CI
Intercept	6.28	0.13	[6.03, 6.53]
Target presence	0.79	0.09	[0.62, 0.96]
Case outcome	0.59	0.24	[0.13, 1.05]
Lineup decision	0.65	0.09	[0.48, 0.82]
Case Outcome × Lineup Decision	-0.08	0.16	[-0.40, 0.25]

Table 16

Percentage Willingness to Testify Rates Across the Case Outcome and Lineup Decision Conditions

	Lineup decision					
Case outcome	N trials	Picked	N trials	Rejected	N trials	Total
Rightful conviction	1461	59.75	1019	44.95	2480	53.67
Wrongful conviction	1405	55.23	1115	42.87	2520	49.76
Total	2866	57.54	2134	43.86	5000	51.70

Table 17

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Case Outcome and Lineup Decision on Willingness to Testify (After Controlling for Target Presence)

Fixed effect predictors	χ^2	*df*	*p*
Target presence	192.94	1	<.001
Case outcome	0.46	1	.50
Lineup decision	35.50	1	<.001
Case Outcome × Lineup Decision	2.93	1	.09
	b	SE_b	95% CI
Intercept	0.14	0.15	[-0.16, 0.43]
Target presence	0.75	0.07	[0.60, 0.89]
Case outcome	0.17	0.29	[-0.41, 0.75]
Lineup decision	0.47	0.08	[0.32, 0.62]
Case Outcome × Lineup Decision	0.25	0.14	[-0.03, 0.53]

Discussion

Interpretation of the results from Experiment 1 was clouded by queries about whether the main manipulation was strong enough and whether the manipulation check was appropriate. The manipulation check results for Experiment 2 suggest that both may have been a problem. The attempt to strengthen the manipulation in this follow-up appears to have been successful, in that there was a larger difference in perceptions of whether suspect was the culprit and justice had been served (i.e., the *ds* increased from ≈ 1.5 to ≈ 2). However, the main manipulation check question that was used in Experiment 1(i.e., asking participants how worried they were about implicating an innocent suspect) again failed to show a significant difference depending on the way in which the consequences of choosing were

framed. The alternate manipulation check question, measuring perceived likelihood that choosing from a lineup would lead to an innocent suspect pick, showed a difference in the hypothesised direction—but only a small one. Together these findings suggest that (a) the manipulation in Experiment 1 probably was not strong enough to affect likelihood perceptions and (b) the worry question was not an effective way of measuring the manipulation of perceptions regarding the chances that choosing from a lineup could lead to an innocent suspect identification.

Despite the improvements that were made, the predicted effects on choosing were still not observed. This could conceivably be attributed to the fact that the likelihood manipulation was not very strong (despite being strengthened from Experiment 1). Further, it is possible that the perceived likelihood that choosing could lead an innocent suspect to be implicated does not affect identification decisions (under hypothetical circumstances—or ever). The results of the supplementary analysis on confidence point to a different reason though. When no effect of the (innocent suspect identification likelihood) manipulation on choosing was observed, it was posited that it may have had an effect on confidence. Specifically, I suggested that those in the wrongful conviction condition might have been less confident in their positive identification decisions than those in the rightful conviction condition. Results showed that this indeed was the case. However, a similar decrease in confidence was also observed for lineup rejections (i.e., the interaction between lineup decision and case outcome on confidence was not significant). This finding suggests that possibly those in the wrongful conviction condition used their perception of how likely it was that choosing from a lineup would lead to a crucial mistake as a proxy for estimating how likely it was that rejecting a lineup would also lead to a serious error. In essence, the manipulation might have affected the perception of how likely it is that a serious error in general would occur, as opposed to as a result of specifically choosing from a lineup. Previous research has shown that increasing the

perceived difficulty of a general knowledge task led people to become less confident in their decisions (Arkes, Christensen, Lai, & Blumer, 1987). Perhaps in a similar manner the case outcome manipulation led people to have opposing views about how difficult an identification decision is and to adjust their confidence accordingly. If this is in fact what occurred, it would explain why choosing did not differ between the wrongful and rightful conviction conditions.

Whatever the cause, the failure to bias people's willingness to choose from a lineup meant that the question of whether the effect would be larger for decisions associated with more serious crimes could also not be answered. Therefore, further experiments are required in order to draw conclusions regarding whether consequence perceptions have the potential to affect identification decision making under hypothetical circumstances.

CHAPTER 3

The previous two experiments provided no indication that the hypothetical severity of consequences associated with identification decisions influences choosing behaviour. Rather than establishing that consequences do not affect identification decisions (in this paradigm at least), however, these results can be explained quite simply as follows. Observing an effect of consequence severity relied on choosing being influenced by manipulating the expected likelihood that an innocent suspect could be identified (as consequence severity, in the form of crime seriousness, was predicted to moderate this effect). Because the manipulation of likelihood perceptions failed to produce a difference in choosing, there was no scope for the consequence severity manipulation to influence the identification decisions that were made. Although there might have been a way to remedy this problem, I chose to instead change the focus of my research to manipulating whether mistakenly picking from or rejecting a lineup was seen as worse and testing whether this affected willingness to choose from a lineup. While I had previously judged that manipulating people's consequence perceptions so one mistake would be seen as worse than the other was not possible (because it seemed unfeasible to influence whether people saw causing a wrongful conviction or letting a guilty party get away as worse), I thought of a potential way to circumvent this obstacle—which is detailed below.

Experiment 3

As argued in Chapter 1, it might be expected that if people think that the consequences associated with mistakenly choosing from a lineup are worse than mistakenly rejecting a lineup then they would be less willing to choose from a lineup. Further, lower choosing rates would be expected to increase accuracy for cases when the target is not in the lineup, but to decrease accuracy in cases when the target is in the lineup. In order to test these propositions I sought to manipulate which mistake would be seen as worse by influencing

people's perceptions of how likely they thought it would be that the possible mistakes would actually lead to the associated worst case scenarios. While, as previously mentioned, people might have fairly fixed attitudes regarding how bad letting a guilty person get away would be compared with convicting someone for a crime they did not commit, the actual threat that these outcomes present could reasonably be expected to depend on how likely they are to occur. For example, a witness might think that wrongful conviction is worse than a guilty person walking free. But if they also believe[18] that the chances of a guilty party getting away—as a result of failing to identify them—are much higher than the chances of an innocent suspect being incarcerated—by identifying them—they might actually err on the side of thinking that mistakenly rejecting a lineup is worse (at least to a greater extent than someone who both thinks that a wrongful conviction is worse and fairly likely to occur as a result of a mistaken pick).

People's negative consequence likelihood perceptions were manipulated by telling them explicitly that in actual cases either mistakenly choosing (an innocent suspect[19]) or mistakenly rejecting a lineup would have a greater chance of ultimately leading to a negative outcome. This direct approach to manipulating perceptions of consequence likelihood was implemented in an attempt to avoid people making unintended inferences from the information they were provided with (i.e., similar to what occurred when attempting to manipulate perceptions of innocent-suspect-identification-likelihood in Experiments 1 and 2). The differing descriptions of consequence likelihood were presented to participants as part of background information about why identification research is conducted (in place of a filler task between the encoding and test phases of the experiment). Social persuasion theorists

[18] While it is not out of the question that actual witnesses consider the likelihood of negative consequences associated with the possible identification errors they could make, I am not making any claims about whether or to what extent witnesses do this.

[19] For convenience the phrases "mistakenly choosing" and "mistakenly choosing an innocent suspect" are used interchangeably here (because foil picks are never being referred to), but note that the latter phrase was always used in materials that participants read.

have argued that, when people do not have a strong motivation to critically evaluate the validity of an argument, they will judge its merit based on peripheral information, such as the expertise of the information source (Chaiken, 1987; Eagly & Chaiken, 1984; Petty & Cacioppo, 1986). Therefore, by presenting the manipulation of consequence perceptions in an offhand way (i.e., as mere background information), it was hoped that people would have minimal motivation to critically evaluate what they were told and, in turn, accept the perspective (provided by eyewitness memory experts) that the particular type of identification error they were presented with would be more likely to lead to negative consequences than any other error.

It was predicted that those who were told that mistakenly choosing would be more likely to lead to negative consequences (aka those in the "innocent-person-convicted-more-likely" condition) would think that mistakenly choosing from a lineup would be worse (compared with those who were told that mistakenly rejecting a lineup would be more likely to lead to negative consequences—aka those in the "guilty-person-released-more-likely" condition). As a result, those in the former condition were expected to be less likely to make positive identification decisions. Further, this difference in willingness to choose from a lineup was predicted to lead to lower accuracy for target-present lineups in the innocent-person-convicted-more-likely condition and lower accuracy for target-absent lineups in the guilty-person-released-more-likely condition. Similar to what was proposed in the first two experiments, the potential for these effects to be greater for decisions associated with serious (versus not serious) crimes was examined.

Pilot 1 and 2

Due the problems encountered with attempting to manipulate people's likelihood perceptions in Experiments 1 and 2—although these were of a different nature—the proposed manipulation of negative consequence likelihood was piloted prior to running Experiment 3.

Sixty volunteers were recruited to take part in the first pilot by approaching them on university campus and asking if they would like to participate in a survey on perceptions about eyewitness memory and the criminal justice system. They were given a page of information to read, including a description of why the police might ask an eyewitness to view a lineup, the errors that could result and a commentary on how likely it is that these could lead to serious consequences. Half of the participants were provided with information suggesting that mistakenly choosing an innocent suspect would be more likely to lead to a negative outcome and the other half were given information indicating that mistakenly rejecting a lineup would be more likely to lead to a negative outcome. In the condition where mistakenly choosing was framed as worse, mistakenly rejecting a lineup was described as not being that bad because, in the event that a suspect is in fact guilty, other evidence will be obtained to implicate them. However, mistakenly choosing an innocent suspect was stated to be extremely problematic because positive identifications are seen as a very strong piece of incriminating evidence and, alone, could form the basis for sending someone to jail. In the condition where mistakenly rejecting was framed as worse, people were told that mistakenly choosing from a lineup is ultimately not that bad because no other evidence will emerge to support the guilt of an innocent person, but that mistakenly rejecting a lineup with the guilty person in it can be very problematic because it prevents law enforcement from having probable cause to further investigate a suspect. A copy of the information both groups of participants were given to read can be found in Table 18.

Table 18

The Negative Consequence Likelihood Information Participants Read in Pilot 1

Condition	Identification decision background and consequence related information
Both conditions	If you were an eyewitness to an actual crime, you would be called in to make an identification decision if the police have a suspect and they want to know if you recognise them as the culprit or not. The decision you make could have important implications for the outcome of a criminal investigation. The suspect would usually be shown to you amongst several distracter faces (people who are known to be innocent). If you identified (i.e., thought you recognised) any of these distracter faces/people it would suggest to the police that your memory for the perpetrator isn't very good. The suspect themselves may or may not be the actual perpetrator (i.e., either the suspect would be guilty or innocent). So if you identified them there is a chance that you could be implicating an innocent person. On the other hand, if you rejected the lineup (said the guilty person wasn't there) you could be suggesting that the guilty person isn't the culprit. Research and real world cases have demonstrated that both mistakes occur.
Innocent-person-convicted-more-likely condition	If you picked the police's suspect out of the lineup, this is generally seen as really strong evidence in favour of their guilt. Depending on the nature of the evidence which led to them being a suspect in the first place, it may be enough to obtain a conviction and result in jail time. While this could mean that a guilty person has been brought to justice, if a mistake has been made an innocent person would be punished for something they didn't do (in this case the guilty party would also still be walking free, not being pursued and in the worst case scenario could reoffend). If you rejected the lineup, the police would probably not have enough evidence to charge the suspect at that point in the investigation. They may also consider other potential leads if they have them. However, as the investigation progressed it is likely that new evidence would emerge which either would point away from or back towards their initial suspect.
Guilty-person-released-more-likely condition	If you picked the police's suspect out of the lineup, this in and of itself wouldn't be enough evidence to obtain a conviction, but it would lead to this person being more thoroughly investigated. Throughout the course of the investigation then it is likely that new evidence would emerge which either would point away from or towards their initial suspect. If you rejected the lineup, the evidence which led to them being a suspect probably wouldn't be enough to hold them on and they would be released (for the time being at least). While this could mean that an innocent suspect isn't being further investigated, if a mistake has been made and the suspect is in fact guilty the police often waste time directing their investigation elsewhere. Because the guilty party knows the police have reason to suspect them, they will have the opportunity to destroy/dispose of evidence that hasn't been uncovered yet, flee and escape penalty, or in the worst-case scenario they might also reoffend.

On the following page participants were asked, amongst several filler questions (e.g., how many people they thought should be in a lineup), to rate how likely they thought it was that mistakenly picking an innocent suspect from a lineup would lead to a wrongful conviction and mistakenly rejecting a lineup would lead to a guilty party walking free (0 = *not at all likely*; 6 = *extremely likely*).

The effect of the negative consequence likelihood information on people's ratings of how likely mistakenly choosing from and mistakenly rejecting a lineup would be to lead to negative consequences was tested using 2 (negative consequence likelihood information: innocent-person-convicted-more-likely, guilty-person-released-more-likely) × 2 (type of mistake: incorrect suspect pick, incorrect lineup rejection) ANOVA. Descriptive statistics are reported in Table 19. The main effect of negative consequence likelihood information was not significant, $F(1, 58) = 2.96$, $p = .09$, $d = 0.37$, d 95% CI [-0.14, 0.88]. The main effect of type of mistake was significant, $F(1, 58) = 8.13$, $p = .01$, $d = 0.41$, d 95% CI [0.05, 0.77], with people thinking mistaken rejections would be more likely to lead to negative consequences than mistaken picks. The Negative Consequence Likelihood Information × Type of Mistake interaction was also significant, $F(1, 58) = 5.20$, $p = .03$, $f = 0.27$. Descriptive statistics suggest, and a simple effects analysis confirmed, that this interaction was characterised by mistakenly choosing being seen as less likely than mistakenly rejecting to lead to serious consequences in the condition where mistakenly rejecting was framed as being more likely to lead to negative consequences, $t(29) = -2.94$, $p = .01$, $d = 0.64$, d 95% CI [0.12, 1.16]. However, there was no difference in perceived likelihood of negative consequences eventuating as a result of mistakenly picking or rejecting a lineup in the condition where mistakenly choosing was framed as being more likely to lead to negative consequences, $t(29) = -0.58$, $p = .56$, $d = 0.11$, d 95% CI [-0.40, 0.62].

Table 19

Mean (and Standard Deviation) Perceived Likelihood of Negative Consequences as a Result of Incorrectly Picking from or Rejecting a Lineup, Across the Negative Consequence Likelihood Information Conditions

	Negative consequence likelihood information		
Type of mistake	Guilty-person-released-more-likely	Innocent-person-convicted-more-likely	Total
Incorrect lineup rejection	3.87 (1.57)	3.90 (1.18)	3.89 (1.38)
Incorrect suspect pick	2.67 (2.12)	3.77 (1.19)	3.22 (1.80)
Total	3.27 (1.50)	3.84 (1.01)	

While these results suggested that the information people were provided with in the innocent-person-convicted-more-likely and guilty-person-release-more-likely conditions led to a difference in negative consequence likelihood perceptions, there was considerable room for the difference between the two conditions to be increased. Therefore, another pilot was run in which several changes to the manipulation were made. The volume of the information provided to participants was reduced as this may have been confusing for people. Further, to emphasise the negative consequences that participants were told were more likely to occur, they were provided with a realistic-looking mock newspaper article that reported on a case in which these consequences (i.e., either wrongful conviction or a guilty part reoffending after being set free) had occurred.[20] A copy of the information people were given to read and the content of the newspaper articles can be found in Tables 20 and 21, respectively. Again, 60 volunteers, approached on university campus and asked to take part in a survey, participated. After reading the information sheet, as well as the accompanying newspaper article,

[20] Prior to reading the newspaper article, it was indicated to participants that the article had appeared in the Boston Globe on April 4th, 2009.

participants answered the same questions assessing negative consequence likelihood as in the previous pilot. Additionally they were asked to rate whether and to what degree they thought mistakenly choosing from or rejecting a lineup would be worse (0 = *mistakenly rejecting much worse*; 6 = *mistakenly choosing much worse*), as the ultimate aim of manipulating consequence likelihood perceptions was to lead people to differ in their perceptions of which mistake would be worse to make.

Table 20

The Negative Consequence Likelihood Information Participants Read in Pilot 2

Condition	Identification decision background and consequence related information
Both conditions	In real life, you would be called in to make an identification decision if you witnessed an event for which the police have a suspect and they want to know if you recognise them as the culprit or not. The suspect would usually be shown to you amongst several distracter faces (people who are known to be innocent), which minimises the chance that you will mistakenly think you recognise an innocent suspect, but also doesn't make the task so confusing that a guilty party can't be recognised. Despite this precaution, errors still occur. However, the actual impact of mistakes depends a lot on the amount and nature of other evidence in the case.
Innocent-person-convicted-more-likely condition	While mistakenly rejecting a lineup with a guilty party in it can temporarily misguide an investigation, further evidence to implicate this person will likely emerge as the investigation progresses. However, it is proving to be a big problem when an innocent suspect is picked from a lineup, as this is seen as strong evidence of their guilt, which can have serious repercussions such as in the following case.
Guilty-person-released-more-likely condition	While mistakenly picking an innocent suspect out of lineup can temporarily misguide an investigation, no conclusive evidence will be found to implicate someone who hasn't actually committed the crime. However, it is proving to be a big problem when a lineup containing a guilty party is rejected, as the police then often don't have enough evidence to further pursue that person, which can have serious repercussions such as in the following case.

Table 21

Content of the Mock Newspaper Articles Participants Read in Pilot 2

Condition	Mock newspaper article content
Innocent-person-convicted-more-likely condition	Could This Have Been Prevented? By Shelley Murphy A week ago Mr. Timothy Richards was released after 5 years behind bars, following a surprise confession from someone else to having committed the infamous Charlestown jewellery heist. It has now been confirmed that the actual culprit had been taken into custody for a string of other robberies, one including a hostage situation, which were at first thought to be unrelated to the 2004 robbery in which several million dollars' worth of jewellery was stolen. As more information about the case emerges, the question of whether this grave miscarriage of justice could have been prevented is being asked. At the time of the first offense, Mr. Richards was known to police for several minor transgressions and brought in for questioning because he had purchased items described as 'suspicious' online and posted on social media that he was about to become rich. At this point a witness, who saw a man vandalise the security camera outside the premises that was robbed, identified Mr. Richards from a lineup. With this further reason for suspicion experts are saying that other circumstantial evidence was probably seen as more convincing than it would have on its own. The police have issued a statement saying that Mr. Richards' incarceration was the result of an unfortunate set of circumstances.
Guilty-person-released-more-likely condition	Could This Have Been Prevented? By Shelley Murphy A week ago, Mr. Timothy Richards was formally charged after 5 years on the run, following a surprise confession from him to having committed the infamous Charlestown jewellery heist. It has now been confirmed that the alleged culprit had been taken into custody for a string of other robberies, one including a hostage situation, which were at first thought to be unrelated to the 2004 robbery in which several million dollars' worth of jewellery was stolen. As more information about the case emerges, the question of whether these grave criminal offences could have been prevented is being asked. Prior to the first offense, Mr. Richards was known to police for several minor transgressions and questioned because he had purchased items described as 'suspicious' online and posted on social media that he was about to become rich. At this point a witness, who saw a man vandalise the security camera outside the premises that was eventually robbed, failed to identify Mr. Richards from a lineup. Without this further reason for suspicion, experts are saying that the police would have been hard pressed to obtain a warrant to search the suspect's home. The police have issued a statement saying that failing to apprehend Mr. Richards earlier was the result of an unfortunate set of circumstances.

The effect of the consequence perception manipulation on negative outcome likelihood ratings for mistaken picks and rejections was again tested using a 2 (negative

consequences likelihood information: innocent-person-convicted-more-likely, guilty-person-release-more-likely) × 2 (type of mistake: incorrect suspect pick, incorrect lineup rejection) ANOVA. Descriptive statistics are summarised in Table 22. The main effect of negative consequence likelihood information was not significant, $F(1, 58) = 3.36$, $p = .07$, $d = 0.47$, d 95% CI [-0.05, 0.98]. The main effect of mistake type was significant, $F(1, 58) = 16.01$, $p < .001$, $d = 0.60$, d 95% CI [0.23, 0.96], with people again indicating the belief that mistakenly rejecting a lineup is more likely to lead to serious consequences than mistakenly choosing from a lineup. The interaction was also significant, $F(1, 58) = 38.85$, $p < .001$, $f = 0.43$. Descriptive statistics suggest, and simple effects analyses confirmed, that people in the condition in which mistaken rejections were stipulated to be more likely to lead to negative outcomes gave higher consequence likelihood ratings for mistaken lineup rejections than for innocent suspect identifications, $t(29) = -6.09$, $p < .001$, $d = 1.49$, d 95% CI [0.91, 2.06]. However, in the condition where innocent suspect identifications were said to be more likely to lead to negative outcomes, participants only gave slightly higher ratings of consequence likelihood to innocent suspect identifications than mistaken lineup rejections, a mean difference that was not significant, $t(29) = 1.04$, $p = .31$, $d = 0.22$, d 95% CI [-0.29, 0.73]. Ratings of which mistake would be worse to make[21] were significantly different across the manipulation of negative consequence likelihood, $t(57) = -7.80$, $p < .001$, $d = 2.03$, d 95% CI [1.39, 2.66]. The average rating in the condition where incorrect identifications were framed as more likely to lead to negative consequences fell in between *mistakenly choosing* being seen as 'a little' or 'somewhat' worse ($M = 4.34$, $SD = 1.45$); in the condition where incorrect rejections were framed as being more likely to lead to negative consequences, people's average response fell between *mistakenly rejecting* being seen as "a little" or "somewhat" worse ($M = 1.63$, $SD = 1.22$). These results suggested that the two versions of negative

[21] One respondent in the wrongful conviction more likely condition did not answer this question.

consequence likelihood information led to a substantial difference in people's perception of which mistake would be worse to make. Therefore, this manipulation was used in Experiment 3 to test the potential for differing beliefs regarding which identification error is worse to affect people's willingness to choose from a lineup.

Table 22

Mean (and Standard Deviation) Perceived Likelihood of Negative Consequences as a Result of Incorrectly Picking from or Rejecting a Lineup, Across the Negative Consequence Likelihood Information Conditions

	Negative consequence likelihood information		
Type of mistake	Guilty-person-released-more-likely	Innocent-person-convicted-more-likely	Total
Incorrect lineup rejection	4.03 (1.16)	3.37 (1.43)	3.70 (1.33)
Incorrect suspect pick	2.04 (1.49)	3.65 (1.05)	2.85 (1.51)
Total	3.04 (.99)	3.51 (1.00)	

Method (Experiment 3)

Design and Participants

A 2 (negative consequences likelihood information: innocent-person-convicted-more-likely, guilty-person-released-more-likely) × 2 (crime seriousness: serious, not serious) × 2 (target presence: present, absent) mixed design was used. Participants were randomly assigned to a negative consequence likelihood information condition and made multiple identification decisions from mini-lineups, across which target presence and crime severity varied. As in the previous experiment 128 people participated. They were 74 females and 54 males and their ages ranged from 16 to 45 years (M = 21.30, SD = 3.94).

Materials and Procedure

The same materials were used as Experiment 2, apart from the new manipulation outlined in Pilot 2. The procedure was largely the same as the previous experiment, apart from a few necessary changes to accommodate the new manipulation. Instead of studying the face stimuli and then completing the mini-lineup decisions in two blocks of 20, as they had done previously, participants completed one block of 40 trials. This was done because the manipulation was presented to participants as background information on the role of identification decisions in the criminal justice system and, therefore, it would not have been sensible to repeat the manipulation in a second block. Due to concern that this would make the test block too long to maintain people's attention, the willingness to testify measure was dropped (i.e., people were only asked to indicate how confident they were following their decision). Inspection of the accuracy data from the first 20 participants indicated that presenting all of the trials in one block did not make the task too difficult, as average performance remained well above chance levels across both target-present and target-absent cases (i.e., 55%). The instructions participants were provided with also differed somewhat from the previous experiments. They were told that they would be participating in a study on eyewitness memory, but were given no further information about this before viewing the target faces other than the instructions provided in Experiments 1 and 2 regarding studying the faces and crime cues. After all the faces had been viewed participants were told that the motivation for studying eyewitness identification decisions is that the associated consequences can be very serious. Next the negative consequence likelihood information outlined in Pilot 2 was provided. Participants were then given the same instructions as the previous experiments for completing the identification tests, asked the manipulation check questions from Pilot 2 and asked to rate each crime cue on the importance of achieving justice as in Experiment 2.

Results

Manipulation Checks

The effect of the consequence perception manipulation on negative consequence likelihood ratings for mistaken lineup picks and rejections was testing using 2 (negative consequences likelihood information: innocent-person-convicted-more-likely, guilty-person-released-more-likely) × 2 (type of mistake: incorrect suspect pick, incorrect lineup rejection) ANOVA. Descriptive statistics are summarised in Table 23. The main effect of the negative consequence likelihood information was not significant, $F(1, 126) = 0.24$, $p = .63$, $d = 0.09$, d 95% CI [-0.26, 0.44]. Neither was the main effect of type of mistake, $F(1, 126) = 2.81$, $p = .10$, $d = 0.15$, d 95% CI [-0.01, 0.40]. The Negative Consequence Likelihood Information × Type of Mistake interaction was significant, $F(1, 126) = 4.29$, $p = .04$, $f = 0.10$. People in the condition where mistaken rejections were framed as more likely to lead to negative outcomes gave higher likelihood ratings to mistaken rejections than to mistaken picks, $t(63) = -2.46$, $p = .02$, $d = 0.34$, d 95% CI [-0.01, 0.69], whereas people in the condition where mistaken picks were framed as more likely to lead to negative outcomes did not differ in their consequence likelihood perceptions across lineup errors, $t(63) = 0.31$, $p = .76$, $d = 0.04$, d 95% CI [-0.31, 0.39]. Participants also recorded higher ratings on the measure of which mistake would be worse in the choosing more likely condition ($M = 3.86$, $SD = 2.19$) than those in the rejection more likely condition ($M = 2.78$, $SD = 2.16$), $t(125) = -2.80$, $p = .01$, $d = 0.50$, d 95% CI [0.15, 0.85].

Table 23

Mean (and Standard Deviation) Perceived Likelihood of Negative Consequences as a Result of Incorrectly Picking from or Rejecting a Lineup Across the Negative Consequence Likelihood Information Conditions

	Negative consequence likelihood information		
Type of mistake	Guilty-person-released-more-likely	Innocent-person-convicted-more-likely	Total
Incorrect lineup rejection	3.64 (1.57)	3.44 (1.55)	3.54 (1.56)
Incorrect suspect pick	3.05 (1.86)	3.50 (1.83)	3.27 (1.85)
Total	3.34 (1.42)	3.47 (1.49)	

As in Experiment 2, a linear mixed-effects model was created to determine whether the crimes coded as serious were rated as more important to achieve justice for than the not serious crimes. Crime seriousness code was found to be a significant predictor of justice importance ratings, $\chi^2(1) = 70.82$, $p < .001$, $b = 2.25$, $SE_b = 0.17$, b 95% CI [1.92, 2.59].[22] As predicted importance of achieving justice was higher for serious ($M = 5.43$, $SD = 0.43$) than not serious ($M = 3.16$, $SD = 0.87$) crimes.

Main Analyses

Choosing. The key focus of this experiment was on testing whether people who were told that the consequences associated with mistakenly choosing from a lineup were more likely to be serious than the consequences of mistakenly rejecting a lineup would be less willing to make positive identifications than those told the opposite. Tentatively, this effect was also expected to be greater for lineups associated with serious versus not serious crimes. To investigate these hypotheses, a logistic mixed-effects model was created with whether or not people chose from the lineup as the outcome variable. Negative consequence likelihood

[22] Intercept $b = 3.17$, $SEb = 0.14$, b 95% CI [1.92, 2.59]

information, crime seriousness and their interaction were added as fixed effect predictors. None of these variables were found to be significant predictors of choosing. Aggregate statistics are reported in Table 24 and the inferential statistics can be found in Table 26. Note that results are collapsed across crime seriousness for all further analyses, both because descriptive statistics indicated that the results did not vary across crime seriousness and the 3way interactions that would have needed to be examined to account for crime seriousness were too complex to be modelled properly (i.e., with appropriate random effects structures).

Table 24

Percentage Choosing Rates Across the Negative Consequence Likelihood Information and Crime Seriousness Conditions

	Crime seriousness					
Negative consequence likelihood Information	N trials	Serious	N trials	Not serious	N trials	Total
Innocent-person-convicted-more-likely	1280	58.52	1280	58.13	2560	58.33
Guilty-person-released-more-likely	1280	56.48	1280	60.23	2560	58.36
Total	2560	57.50	2560	59.18	5120	58.35

Accuracy. As an interaction between the consequence likelihood conditions and target presence on accuracy was predicted, another logistic mixed-effect model was constructed with accuracy as the outcome variable. Negative consequence likelihood information, target presence and the interaction between these two variables were entered as predictors. Accuracy was not significantly estimated by any of these predictors (see Table 25 for descriptive statistics and Table 27 for the model fit and fixed effect coefficients).

Table 25

Percentage Accuracy Rates Across the Negative Consequence Likelihood Information and Target Presence Conditions

	Target presence					
Negative consequence likelihood information	N trials	Present	N trials	Absent	N trials	Total
Innocent-person-convicted-more-likely	1280	54.45	1280	56.02	2560	55.24
Guilty-person-released-more-likely	1280	54.92	1280	54.38	2560	54.65
Total	2560	54.69	2560	55.20	5120	54.95

Table 26

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Effect of Case Outcome and the Case Outcome by Crime Seriousness Interaction on Choosing

Fixed effect predictors	χ^2	*df*	*p*
Negative consequence likelihood information	0.00	1	.98
Crime seriousness	1.49	1	.22
Negative consequence likelihood information × crime seriousness	2.26	1	.13
	b	SE_b	95% CI
Intercept	0.34	0.03	[0.28, 0.39]
Negative consequence likelihood information	-0.002	0.06	[-0.11, 0.11]
Crime seriousness	-0.07	0.06	[-0.18, 0.04]
Negative consequence likelihood information × crime seriousness	0.17	0.11	[-0.05, 0.39]

Table 27

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Negative Consequence Likelihood Information and Target Presence on Accuracy

Fixed effect predictors	χ^2	*df*	*p*
Negative consequence likelihood information	0.10	1	.75
Target presence	0.06	1	.82
Negative consequence likelihood information × target presence	0.23	1	.63
	b	SE_b	95% CI
Intercept	0.23	0.07	[0.10, 0.37]
Negative consequence likelihood information	0.04	0.13	[-0.21, 0.30]
Target presence	-0.03	0.12	[-0.26, 0.21]
Negative consequence likelihood information × target presence	-0.12	0.24	[-0.60, 0.36]

Supplementary Analysis

Post-decision confidence. While the negative consequence likelihood information was not found to influence the identification decisions that were made, it is possible that (similar to what was proposed in the earlier experiments) people's post-decisional confidence ratings could have been affected. Specifically, I thought it was possible that in the condition where the incorrect identification of an innocent suspect was framed as more likely to lead to a negative outcome than mistakenly rejecting a lineup with the guilty party present, people would be less confident in their positive identification decisions than their lineup rejections. Conversely, people in the condition where mistakenly rejecting a lineup was framed as being more likely than mistakenly choosing from a lineup to lead to a negative outcome, might have expressed lower confidence in their lineup rejections than their positive identifications.

In order to test whether there was an interaction between negative consequence likelihood information and lineup decision in predicting confidence, a linear mixed-effects model was constructed with post-identification confidence as the outcome variable. As in Experiment 1 and 2, target presence was controlled for by entering it into the model first; then negative consequence likelihood information, lineup decision and the Negative Consequence Likelihood Information × Lineup Decision interaction were entered as predictors. Besides target presence, only lineup decision was found to significantly predict confidence. As observed in Experiment 2, confidence in positive identification decisions was higher than lineup rejections. Descriptive and inferential statistics are shown in Table 28 and Table 29, respectively.

Table 28

Mean (and Standard Deviation) Post-Decision Confidence Across the Negative Consequence Likelihood Information and Lineup Decision Conditions

	Lineup decision					
Negative consequence likelihood information	N trials	Picked	N trials	Rejected	N trials	Total
Innocent-person-convicted-more-likely	1493	57.86 (26.51)	1067	52.15 (27.75)	2560	55.63 (28.32)
Guilty-person-released-more-likely	1494	56.36 (27.67)	1066	48.23 (28.55)	2560	52.97 (28.32)
Total	2987	57.11 (27.10)	2133	50.37 (28.23)	5120	

Table 29

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Negative Consequence Likelihood Information and Lineup Decision on Post-Decision Confidence (After Controlling for Target Presence)

Fixed effect predictors	χ^2	*df*	*p*
Target presence	41.26	1	<.001
Negative consequence likelihood information	0.98	1	.32
Lineup decision	20.67	1	<.001
Negative consequence likelihood information × lineup decision	0.85	1	.36
	b	SE_b	95% CI
Intercept	5.39	0.15	[5.11, 5.68]
Target presence	0.59	0.08	[0.43, 0.74]
Negative consequence likelihood information	0.28	0.28	[-0.26, 0.83]
Lineup decision	0.58	0.12	[0.34, 0.83]
Negative consequence likelihood information × lineup decision	-0.22	0.24	[-0.70, 0.25]

Discussion

The results of Experiment 3 did not show any evidence to suggest that lineup choosing behaviour depends on which mistake is perceived to be worse. Perhaps this is because diverging opinions regarding whether mistakenly choosing from or rejecting a lineup is worse do not lead to a difference in willingness to make a positive identification decision. However there are several, potentially modifiable, aspects of the experimental paradigm used that could have contributed to the null results obtained.

While telling half of the participants that mistakenly choosing from versus mistakenly rejecting a lineup is more likely to lead to negative consequences (in actual criminal

investigations) led to significant differences in both consequence likelihood perceptions and the perception of which mistake was worse, this difference was not as large as foreshadowed by the pilot test of the manipulation. Perhaps embedding the consequence likelihood information within the broader experiment somehow detracted from the attention it received by participants. Consequently, people may not have processed the information as well as they did in the pilot (where people read the relevant information in isolation from any other stimuli). There was also virtually no time delay between people reading the target information and providing their responses in the pilot. This could have meant that pilot participants more clearly remembered what they had been told when completing the measures of consequence likelihood and which mistake would be worse to make. Thus, strengthening the manipulation of consequence likelihood perceptions when embedded in the context of an experiment may be necessary. However, despite the fact that the key manipulation did not appear to operate as strongly as hoped, this is unlikely to account completely for the fact that people's identification decisions, as well as their post-decisional confidence ratings, were not affected as predicted. Although it is possible that the manipulation of consequence perceptions was simply not strong enough to influence choosing behaviour, there is an obvious alternative explanation to consider.

It may well be the case that the highly non-realistic paradigm used was unable to capture the impact of consequence perceptions on identification decisions. A key consideration here is whether the fact that the consequences associated with the decision were hypothetical prevented people from accounting for them in their decisions, or whether some other aspect of the method caused a problem. For example, people made a large number of identification decisions. While this approach had many practical benefits (e.g., economically allowing the predicted effects to be tested across multiple stimuli) it may have made the task too abstract to allow the mechanism being tested to operate. It has been argued that the

degree to which hypothetical versus real consequences affect people's behaviour in a similar way can be at least partially determined by the degree to which the actual task is realistically simulated in the hypothetical condition (e.g., Blascovich et al., 2002; FeldmanHall et al., 2012; Geller, 1978; Patil, Cogoni, Zangrando, Chittaro, & Silani, 2014; Rovira, Swapp, Spanlang, & Slater, 2009).

In one example of evidence obtained to support this argument, FeldmanHall et al., (2012) conducted a study in which they led participants to believe that they had to choose between the amount of monetary gain they would obtain and the strength of electric shocks another participant was receiving.[23] Across 20 trials they had to indicate on a sliding scale how much of £1 they would nominate for themselves, where taking no money in that trial meant that the other person received no shock and taking the full amount meant the other person would receive the highest level of shock possible.[24] In comparison to another group of participants who were asked to estimate (hypothetically) how much money they would take for themselves under the circumstances described above, those who thought they were completing the task for real kept a much higher amount of money for themselves (M = 12.52, SD = 4.80 in the "real" condition vs. M = 1.53, SD = 5.43 in the hypothetical condition, though note that the standard deviation in the hypothetical condition suggests considerable skew in the data). In a follow up study this difference was mitigated, however, by increasing the degree to which the hypothetical task simulated the real task. When people actually completed the task trial by trial (but knowing that they would not actually receive any money and no actual electric shocks would be administered) they kept an average of ≈ £8 for themselves. When participants also experienced the sample electric shock at the start of the

[23] The other participant, a confederate, did not actually receive any electric shocks. However, manipulation check measures suggested that people believed the ruse.

[24] The highest level of shock was not defined; however, prior to beginning the trials each participant was given a "sample shock" (2.5mA and 200ms at 400v) and told this was a low level shock. People rated this as M = 3.25, SD = 1.97 on a 10-point scale (1 = *no pain*; 10 = *extreme pain*).

hypothetical task, the amount of money they kept did not differ from those in the condition who thought they were making decisions involving actual monetary gain and electric shocks (i.e., ≈ £12). These results provide at least some basis for thinking that the manipulation of consequence likelihood perceptions did not affect the choices people made in Experiment 3 because the identification task people completed was not similar enough to an actual identification procedure. In particular, making a large number of choices might have made it harder for people to treat each one as if it was an important decision.

Another issue to note was that the participants in the present experiment may not have felt socially accountable for the decisions they made, and that this may be an important factor influencing whether the predicted effects are realised. Vieider (2011) proposed that differences observed between real and hypothetically made decisions are often confounded with accountability (e.g., Burke, Carter, Gominiak, & Ohl, 1996; Epley & Gilovich, 2005; Wright & Anderson, 1989). When people make decisions with real consequences (e.g., payoffs), their decisions are usually made known to others. In such cases, people may feel additional pressure to make what are perceived to be good decisions because their misjudgements are subject to scrutiny from others. It has been found that the mere presence of an experimenter (Guerin, 1993; Zajonc & Sales, 1966), or feeling that their decisions can be traced back to them (Price, 1987; Reicher & Levine, 1994; Schoppler et al., 1995; Williams, Harkins, & Latane, 1981), leads people to feel more socially accountable for the responses they provide. In contrast, when people make hypothetical decisions, these tend to be anonymous and this anonymity would presumably negate any social pressure. Along these lines, participation in the present experiment was completed in an individual cubicle fitted with a door. Further, the letter of introduction people were provided with detailed that their data would be stored anonymously. While in my experience few participants appeared to read this information closely, it seems reasonable to assume that people would not have felt that

they were particularly subject to scrutiny. Therefore, this confound may well have inadvertently occurred here. Importantly, a lack of accountability could have contributed to the null results rather than the hypothetical consequences.

Vieider (2011) conducted an experiment that demonstrated an instance where accountability was a motivating force in altering decision making but realism of the consequences was not. Participants were asked to indicate their preferred option in two variations of the same scenario. Both times they had to decide between a sure option or a prospect gamble (i.e., ending up with $25 for sure, or having a 25% chance of ending up with $40 and a 75% chance of ending up with $20). However, in one instance the choices were framed positively as gains (i.e., having $20 to begin with and choosing the option of an extra $5 for sure, or having a 25% chance of gaining $20 and a 75% chance of gaining nothing) and in the other the choices were framed negatively as losses (i.e., having $40 to begin with and choosing the option of losing $15 for sure, or having a 75% chance of losing $20 and 25% chance of losing nothing). While people should be consistent in their preferences across frames, it has been typically found that people prefer the sure amount when the options are framed positively and the prospect gamble when the options are framed negatively (Tversky & Kahneman, 1981). Vieider (2011) compared the incidence of preference reversals for participants who were assigned to think they were playing for real money or that they would be making hypothetical decisions. Additionally, they were assigned to one of two accountability conditions. Either they were told that they would discuss their answers with the experimenter or that their responses would be anonymous.[25] He found that the number of preference reversals did not differ depending on whether people made hypothetical or real decisions. However, they did differ depending on accountability, with participants who were

[25] An elaborate set-up was used so that participants in the real consequences/low accountability condition believed that they would receive their winnings in a way that preserved their anonymity.

made to feel accountable for their decisions showing fewer preference reversals (38.27%) than participants who were not made to feel accountable (54.12%). This shift towards consistency was attributed to the fact that accountability leads people to try and make decisions that are most easily defensible (Lerner & Tetlock, 1999). These results suggest that accountability can play a key role in shaping people's decision making.

It is possible, therefore, that potential consequences of an identification decision may, at least in part, bias choosing behaviour because people want to make a justifiable decision. For example, if someone believes that it is highly likely that a wrongful rejection of a lineup is going to lead a guilty party walking free, then it would be difficult for them to rationalise, defend or excuse why they rejected the lineup, unless they were absolutely positive that the target was not present. However, if no one is privy to their decision, as was the case in Experiment 3, they will not feel like they might have to defend the decision and might just make it without any consideration of possible consequences. Social accountability might be particularly relevant in this context, because the consequences of incorrect identification decisions are most likely to affect people other than the decision maker (e.g., an innocent person going to jail). Therefore, public scrutiny for a mistake invokes the consequences directly relevant to the witness and, therefore, may be the most motivating.

Only one study (to my knowledge), has examined the effect of accountability on eyewitness identification decisions. Robinson and Johnson (1998) compared identification performance for people who were told they would have to justify their decision to others with that of people who were not, but found no difference in performance across these conditions. Given it has been stipulated that, if greater cognitive effort can improve performance, then accountability will motivate this (Simonson & Nye, 1992), perhaps these results indicate that thinking more carefully about identification decisions cannot improve the outcome. However, as the perceived consequences associated with choosing from or rejecting the lineup were not

manipulated (or measured), Robinson and Johnson's study (1998) study does not resolve whether accountability is crucial for consequences to bias willingness to choose from a lineup.

The possibility that, in order for hypothetical consequence perceptions to affect identification decision, it is necessary for the identification task to more closely simulate an actual identification procedure and people to be feel socially accountable for their decisions, was addressed in the following experiment.

CHAPTER 4

Experiment 4

Experiment 3 produced no evidence that people's lineup choosing behaviour is affected by considering the hypothetical consequences that could result from a mistaken eyewitness identification decision. Perhaps the consequences of identification decisions do not affect people's decisions. Or, maybe a hypothetical consequences paradigm is not a viable way to research the relationship between consequences and identification decisions. A third possibility is that the hypothesised shift in choosing behaviour was not observed because of other[26] methodological problems, discussed in Chapter 3. In particular the fact that, unlike a real investigation, people made a large number of essentially anonymous identification decisions may have prevented people from accounting for negative consequence likelihood in their decisions. Evidence that hypothetical decisions may not match real decisions when (a) social accountability is absent, which can be caused by making anonymous decisions, (Vieider, 2011) and (b) the task people would complete in reality is not adequately simulated in an experimental paradigm (FeldmanHall et al., 2012), suggests that perhaps the specific form of the hypothetical paradigm used in Experiment 3 was not conducive to people weighing up the consequences associated with their choices as part of the identification decision making process. Therefore, several procedural amendments were made to address these potential methodological problems in Experiment 4.

First, in order to emulate an actual identification decision more closely than was done in the previous experiment, Experiment 4 used a more traditional identification paradigm with participants being shown video footage of two targets,[27] whom they would later try to

[26] While the potential methodological problems being referred to are related to the fact that the identification decision consequences were only hypothetical, the term "other" is used because these problems have the potential to be addressed whilst still using a hypothetical consequences paradigm.

[27] I chose to have two targets in favour of only one, because it is not uncommon for actual criminal investigations to involve several perpetrators and this way the predicted effects were at least able to be tested across two different sets of stimuli.

identify. Further, the participant's role as an imaginary witness in a criminal investigation was emphasised by giving them detailed information regarding the way in which the case unfolded leading up to their identification decision and asking them to provide a description of the targets they saw.[28] Second, given that accountability may be necessary for people to take identification consequences into account, Experiment 4 examined the effects of hypothetical consequences under circumstances designed to make witnesses feel responsible for their decisions. One way that people can be led to feel accountable for their decisions is by leading people to expect that another person will become aware of them (Lerner & Tetlock, 1999). Therefore, participants were told that they would be informed at the end of the study (by the experimenter) as to what the likely outcome of the case would be had they been making a real identification decision.

In addition to these key methodological changes, the manipulation of people's negative consequence likelihood perceptions was also altered. Although the information that was provided to participants in Experiment 3 appeared to have produced a difference in opinion about whether mistakenly choosing from or rejecting a lineup would be more likely to lead to negative consequences (and which mistake would be worse to make), this effect was not very strong—only just reaching the cut-off for a medium effect, with a wide confidence interval. In an attempt to strengthen the manipulation, Experiment 4 participants were not only told which mistake would be more likely to lead to negative consequences but were also asked to generate their own reasons for why this might be the case. I thought that altering the manipulation in this manner would increase the likelihood of people adopting the attitudes about consequences that they were presented with, as people who are asked to come up with convincing arguments themselves—as opposed to merely reading an argument

[28] Although there was scope for the identification procedure to be more realistically simulated, it was hoped that these steps towards making the task seem more realistic (i.e., contextualising the decision within a criminal investigation scenario and only requiring each person to make a small number of identification decisions) would be sufficient.

provided by somebody else—have been found to be more likely to genuinely change their attitudes towards target perceptions (e.g., Elms, 1966; Watts, 1967).

In summary, in Experiment 4 a number of changes to the method by which the potential influence of hypothetical consequences on identification decision making was evaluated were introduced, while largely testing the same hypotheses as in Experiment 3. It was predicted that those who were told that mistakenly choosing an innocent suspect from a lineup would be more likely to lead to negative consequences than mistakenly rejecting a lineup would (a) think that mistakenly choosing from a lineup would be worse than mistakenly rejecting a lineup, (b) make fewer positive identification decisions and (c) exhibit lower target-present, but higher target-absent accuracy than those who were told that incorrectly rejecting a lineup would be more likely to lead to negative consequences than wrongfully identifying an innocent suspect. In Experiment 3 consequence severity was also manipulated (by varying crime seriousness) and the effect of the negative consequence likelihood perceptions was hypothesised to be greater for more serious crimes; however, in the interest of maximising the power of the present experiment, the main hypothesis in this study was tested for decisions associated with a serious crime only.

Method

Participants and Design

A 2 (negative consequence likelihood information: guilty-person-released-more-likely, innocent-person-convicted-more-likely) × 2 (target: male, female) × 2 (target presence: present, absent) design was used. Participants were randomly assigned to a negative consequence likelihood information condition and made identification decisions for two targets across which target presence varied.[29] Given the fact that the number of lineups each participant completed was greatly reduced from the previous experiment, the outcome

[29] Each participant saw one target-present and one target-absent lineup

variable is binary and the predicted effect could potentially be small I aimed to recruit the largest sample possible in the timeframe allowed. Three hundred and fifty-two participants were obtained. They were 78 males and 274 females and their ages ranged from 17 to 52 (M = 21.66, SD = 6.10).

Materials

Stimulus video. As previously mentioned, a short video was used as the stimulus event in this experiment. This video was previously used in research conducted in the Flinders University eyewitness memory laboratory (e.g., Brewer, Weber, Wootton, & Lindsay, 2012). It shows a young dark haired man sitting on a bench reading a newspaper, before looking up, then getting up, tucking the paper under his arm and crossing the street towards a branch of the ANZ bank. He walks out of view as the camera focusses on a young blonde woman making a withdrawal from an ATM outside of the same branch. She then takes a pamphlet from a dispenser next to the ATM, tucks it into a small handbag and proceeds to walk in the same direction the young man. The video lasts for 1 minute and 40 seconds, with each of the targets being in view for about 40s and clear close-up look at their faces being available for approximately 8s.

Negative consequence likelihood manipulation. Participants were provided with the same information to manipulate their consequence perceptions as in Pilot 2 and Experiment 3. However, in addition to being given a description of why a particular mistake would be more likely to lead to negative consequences, people were asked to come up with their own reasons why this might be the case. They were first told that either mistakenly choosing from or mistakenly rejecting a lineup is more problematic. Then they were prompted to think about and write down why they think this might be the case. Finally, they read the brief account of why this is thought to be the case and the accompanying newspaper article (describing a case

where the most likely negative outcome, either a wrongful conviction or a guilty party being released and reoffending, had occurred).[30]

In addition to the manipulation check measures used in Experiment 3, participants were asked to rate whether they thought mistakenly choosing from or mistakenly rejecting a lineup would be more likely to lead to the worst case scenario (i.e., either a wrongful conviction or a guilty party walking free) (0 = *guilty walking free much more likely*; 6 = *innocent conviction much more likely*). I thought that obtaining this rating may be distinct from measuring the likelihood that each mistake would lead to negative consequences, because the *relative* perceived chance of negative consequences is being explicitly probed.

Lineups and identification procedure. Participants were shown an eight-person lineup that either contained the target or a target replacement and seven fillers (filler selection strategy documented in Brewer et al., 2012). The position of all the lineup members was randomly determined. All lineup photographs showed the person from the chest up, and were in colour (4 cm × 5.75 cm) on a white background. The lineups were shown on a 21 inch computer monitor (resolution 1,024 × 768 pixels) in a simultaneous presentation procedure. That is, the lineup members were shown in two rows of four with a number from 1–8 underneath each. Below the lineup there was a centred "not there" button. An identification decision was made by clicking on a lineup member or on the "not there" button. Confidence and willingness to testify were measured as in previous experiments (confidence described in Experiment 1 and willingness to testify described in Experiment 2).

[30] The software used to conduct Experiment 4 prevented the inclusion of a realistic-looking version of the newspaper article, therefore, presenting the content of the article to participants was prefaced by telling people that they were about to read a transcript of a newspaper article (and, as was done in the previous experiment, where it had been published).

Procedure

Participants viewed the stimulus video under instructions to pay close attention. Then, as in Experiment 3, they were told that the motivation for studying eyewitness identification decisions is that the associated consequences can be very serious. Next the negative consequence likelihood manipulation was administered, as described in the materials section above. Participants were then instructed that they would be taking part in a simulated crime investigation in which they would be playing the role of a witness providing evidence. Further, to lead people to feel that they would be held accountable for their decisions, they were advised that they would discuss the outcome of the case with the experimenter at the end of the mock-investigation. To encourage participants to think about what it would be like to give evidence in a real criminal investigation, instructions prompted them to think back to the video they had watched and imagine that they had seen this in real life. Next they were told to envisage that shortly after this event there was a breaking news story that there had been an armed hold-up and shooting in Tanunda (where the video had been filmed) and police were looking to talk to people who saw anyone in the vicinity of the ANZ bank at the time they did. Participants were then asked to provide a description of the people they saw (i.e., what they looked like and what they were doing). Following this they were told to picture that after giving their statement they heard on the news that it had been confirmed that a number of people were shot in a botched robbery and the police were questioning both a female and male suspect in connection with the crime. After this the police had contacted them and asked them to come in to try and identify the people they had seen from a lineup. At this point participants were shown the lineups for the male and female targets in a counterbalanced order. Before viewing each lineup they were given unbiased lineup instructions (i.e., reminded that the target may or may not be in lineup). They were also reminded of the crime that had been committed and that their decision would have a bearing

on the outcome of the simulated crime investigation. After making each identification decision participants were asked to rate how confident they were in their decision and indicate whether or not they were willing to testify to their decision. Finally they were asked the manipulation check questions, thanked for their participation and appropriately debriefed.

Results

Manipulation Checks

A 2 (negative consequence likelihood information: guilty-person-released-more-likely, innocent-person-convicted-more-likely) × 2 (type of mistake: incorrect suspect pick, incorrect lineup rejection) mixed ANOVA was run to test whether consequence likelihood ratings differed across the conditions where people were given opposing information regarding the probability of negative outcomes. There was a main effect of identification error type, $F(1, 350) = 12.61$, $p < .001$, $d = 0.23$, d 95% CI [0.08, 0.38], with the average likelihood rating that mistakenly rejecting a lineup would lead to negative consequences ($M = 3.78$, $SD = 1.52$) being slightly higher than the rating of whether mistakenly choosing an innocent suspect from a lineup would lead to negative consequences ($M = 3.44$, $SD = 1.50$). The effect of negative consequence likelihood information was not significant, $F(1, 350) = 0.88$, $p = .35$, $d = 0.07$, d 95% CI [-0.14, 0.28]. The interaction between the negative consequence likelihood information and identification error type conditions was significant, $F(1, 350) = 55.21$, $p < .001$, $f = 0.24$. Simple effects analyses showed that (a) those in the guilty-person-released-more-likely condition thought that mistakenly rejecting a lineup would be *more* likely to lead to negative consequences than mistakenly choosing an innocent suspect from a lineup, $t(175) = -7.37$, $p < .001$, $d = 0.71$, d 95% CI [0.49, 0.92] and (b) those in the innocent-person-convicted-more-likely condition thought that mistakenly rejecting a lineup would be (slightly) *less* likely than mistakenly choosing an innocent suspect from a

lineup to lead to negative consequences, $t(175) = 2.91$, $p = .004$, $d = 0.25$, d 95% CI [0.04, 0.46]. The descriptive statistics are show in Table 30.

Table 30

Mean (and Standard Deviation) Ratings of Negative Consequence Likelihood as a Result of Incorrectly Picking from or Rejecting a Lineup, Across the Negative Consequence Likelihood Information Conditions

	Negative consequence likelihood information		
Identification error type	Guilty-person-released-more-likely	Innocent-person-convicted-more-likely	Total
Incorrect lineup rejection	4.19 (1.36)	3.37 (1.55)	3.78 (1.52)
Incorrect suspect pick	3.14 (1.57)	3.74 (1.38)	3.44 (1.50)
Total	3.67 (1.47)	3.56 (1.47)	

Ratings of which mistake was seen as *more* likely to lead to negative outcomes were higher in the innocent-person-convicted-more-likely condition ($M = 2.83$, $SD = 1.40$) than the guilty-person-released-more-likely condition ($M = 2.00$, $SD = 1.68$), $t(339.28) = -5.03$, $p < .001$, $d = 0.53$, d 95% CI [0.32, 0.74]. Thus, this measure of consequence likelihood did not indicate a divergence in perceptions (as a result of the consequence likelihood manipulation) that differed in size from what was indicated by the interaction between the negative consequence-likelihood conditions and identification error type on negative consequence likelihood above. Ratings of which mistake would be worse to make were also higher in the innocent-person-convicted-more-likely condition ($M = 4.15$, $SD = 1.46$) than the guilty-person-released-more-likely condition ($M = 3.20$, $SD = 1.90$), $t(328.59) = -5.22$, $p < .001$, $d = 0.56$, d 95% CI [0.35, 0.77].

Main Analyses

Choosing. The overall choosing rate was 44.74%. In order to test whether the negative consequence likelihood information affected people's willingness to choose from a lineup, a logistic mixed-effects model was constructed with choosing as the outcome variable and negative consequence likelihood information as a fixed factor. The inferential results indicated that, although aggregate choosing was in the direction predicted, being slightly lower in the innocent-person-convicted-more-likely condition (41.48%) than the guilty-person-released-more-likely (48.01%) condition, negative consequence likelihood information was not a significant predictor of choosing, $\chi^2(1) = 2.10$, $p = .15$, $b = -0.31$, $SE_b = 0.16$, b 95% CI [-0.63, 0.003].[31]

Accuracy. To test whether the effect of the negative consequence likelihood information on identification decision accuracy depended on target presence, a logistic mixed-effects model was created with accuracy as the outcome variable. Negative consequence likelihood information, followed by target presence and their interaction, were entered as fixed effects. Negative consequence likelihood information did not predict accuracy; however, target presence did, with accuracy being higher for target-absent, than target-present cases.[32] The Negative Consequence Likelihood Information × Target Presence interaction was not a significant predictor of accuracy. Descriptive statistics are shown in Table 31 and model fit statistics, as well as the fixed effect coefficients, appear in Table 32.

[31] Intercept $b = -0.25$, $SEb = 0.08$, b 95% CI [-0.41, -0.09]

[32] This effect was unexpected, however, as it was not replicated in the following experiment it may have been a Type-I error.

Table 31

Percentage Accuracy Rates Across the Negative Consequence Likelihood Information and Target Presence Conditions

	Target presence					
Negative consequence likelihood information	N trials	Present	N trials	Absent	N trials	Total
Innocent-person-convicted-more-likely	176	45.45	176	76.14	352	60.80
Guilty-person-released-more-likely	176	43.18	176	68.18	352	55.68
Total	352	44.32	352	72.16	704	58.24

Table 32

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Negative Consequence Likelihood Information and Target Presence on Accuracy

Fixed effect predictors	χ^2	*df*	*p*
Negative consequence likelihood information	2.05	1	.15
Target presence	60.02	1	<.001
Negative Consequence Likelihood Information × Target Presence	1.06	1	.30
	b	SE_b	95% CI
Intercept	0.40	0.10	[0.21, 0.59]
Negative consequence likelihood information	0.29	0.19	[-0.08, 0.65]
Target presence	-1.32	0.19	[-1.69, -0.95]
Negative Consequence Likelihood Information × Target Presence	-0.36	0.34	[-1.03, 0.30]

Supplementary Analyses

Post-decision confidence. As was done in Experiment 3, I examined whether perceived negative consequence likelihood influenced post-identification decision confidence. Specifically, I tested whether (a) those who were told mistakenly choosing would be more likely to lead to negative consequences were less confident in their positive identifications than their lineup rejections and (b) those who were told mistakenly rejecting would be more likely to lead to negative consequences expressed less confidence in their lineup rejections than their positive identifications. To do so, a linear mixed-effects model was constructed with post-identification decision confidence as the outcome variable. Target presence was controlled for, as was done in previous experiments, by entering it into the model first.[33] Next, negative consequence likelihood information, lineup decision and finally the Negative Consequence Likelihood Information × Lineup Decision interaction were entered as fixed-effects. The model fit statistics and fixed effect coefficients appear in Table 34. Negative consequence likelihood information and lineup decision were not significant predictors of confidence. However, the interaction between these two predictors was significant. Aggregate statistics, summarised in Table 33, suggest that overall confidence was higher for lineup rejections, but that this difference was greater in the innocent-person-conviction-more-likely condition versus the guilty-person-released-more-likely condition.

[33] Note that in this Experiment target presence was not actually found to be a significant predictor of confidence (or willingness to testify), but was included as a control variable in the analysis for consistency across experiments.

Table 33

Mean (and Standard Deviation) Post-Decision Confidence Across the Negative Consequence Likelihood Information and Lineup Decision Conditions

	Lineup decision					
Negative consequence likelihood information	N trials	Picked	N trials	Rejected	N trials	Total
Innocent-person-convicted-more-likely	145	59.79 (19.20)	207	68.16 (20.70)	352	64.72 (20.49)
Guilty-person-released-more-likely	169	61.36 (20.32)	183	64.86 (21.78)	352	63.18 (21.34)
Total	314	60.64 (19.80)	390	66.62 (21.25)	704	63.95 (20.92)

Table 34

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Negative Consequence Likelihood Information and Lineup Decision on Post-Decision Confidence (After Controlling for Target Presence)

Fixed effect predictors	χ^2	*df*	*p*
Target presence	1.48	1	.22
Negative consequence likelihood information	0.20	1	.66
Lineup decision	3.63	1	.06
Negative Consequence Likelihood Information × Lineup Decision	3.44	1	.06
	b	SE_b	95% CI
Intercept	63.51	0.89	[61.76, 65.23]
Target presence	1.95	1.39	[-0.78, 4.67]
Negative consequence likelihood information	0.64	1.79	[-2.86, 4.15]
Lineup decision	-6.65	2.99	[-12.51, -0.78]
Negative Consequence Likelihood Information × Lineup Decision	-6.01	2.85	[-11.60, -0.41]

The data were then split by the negative consequence likelihood conditions and two separate linear mixed-effects models were constructed to test whether lineup decision predicted post-identification decision confidence (after controlling for target presence). Lineup decision was only found to be a significant predictor of confidence in the innocent-person-convicted-more-likely condition, with confidence being higher for lineup rejections than positive identification decisions. Model fit statistics and fixed effect coefficients for both analyses can be found in Table 35 (refer back to Table 33 for descriptive statistics).

Table 35

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Models Testing the Relation Between Lineup Decision and Post-Decision Confidence (After Controlling for Target Presence), Separately for Each Level of Negative Consequence Likelihood Information

	Negative consequence likelihood information					
	Innocent-person-convicted-more-likely			Guilty-person-released-more-likely		
Fixed effect predictors	χ^2	*df*	*p*	χ^2	*df*	*p*
Target presence	0.01	1	.94	2.06	1	.15
Lineup decision	5.39	1	.02	1.93	1	.16
	b	SE_b	95% CI	*b*	SE_b	95% CI
Intercept	63.92	1.23	[61.52, 66.32]	63.17	1.30	[60.62, 65.72]
Target presence	0.49	1.98	[-3.38, 4.37]	3.38	1.97	[-0.48, 7.24]
Lineup decision	-8.97	2.39	[-13.66, -4.28]	-4.22	3.42	[-10.93, 2.49]

Willingness to testify. Similar to the effect on post-decision confidence, alleged negative consequence likelihood may have affected people's stated willingness to testify to their decisions. It was expected that those who were told that mistakenly choosing from a

lineup is more likely to lead to negative consequences would be less likely to indicate that they would be willing to testify to their positive identifications than their lineup rejections. Conversely, those who were told that mistakenly rejecting a lineup is more likely to result in negative consequences were expected to be less likely to indicate that they would be willing to testify to their lineup rejections than their positive identifications. To test this, a logistic mixed-effects model was created in which willingness to testify was entered as the outcome variable. Target presence was controlled for, by entering it into the model first. Then negative consequence likelihood information and lineup decision, followed by the interaction between those two variables, were entered as fixed factors. Negative consequence likelihood information did not significantly predict willingness to testify, whereas lineup decision did. Willingness to testify was higher for lineup rejections than positive identification decisions. Further, the interaction between negative consequence likelihood information and lineup decision was significant. The increased willingness to testify to lineup rejections than positive identifications appeared to be greater in the innocent-person-convicted-more-likely than the guilty-person-released-more-likely condition. See Table 36 for descriptive statistics and Table 37 for the relevant model output.

Table 36

Percentage Willingness to Testify Rates Across the Negative Consequence Likelihood Information and Lineup Decision Conditions

	Lineup decision					
Negative consequence likelihood information	N trials	Picked	N trials	Rejected	N trials	Total
Innocent-person-convicted-more-likely	145	48.26	207	68.12	352	59.94
Guilty-person-released-more-likely	169	55.62	183	61.20	352	58.52
Total	314	52.22	390	64.87	704	59.23

Table 37

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Negative Consequence Likelihood Information and Lineup Decision on Willingness to Testify (After Controlling for Target Presence)

Fixed effect predictors	χ^2	*df*	*p*
Target presence	0.87	1	.35
Negative consequence likelihood information	0.03	1	.87
Lineup decision	16.80	1	<.001
Negative Consequence Likelihood Information × Lineup Decision	2.49	1	.11
	b	SE_b	95% CI
Intercept	0.43	0.11	[0.21, 0.66]
Target presence	0.47	0.20	[0.09, 0.85]
Negative consequence likelihood information	-0.02	0.22	[-0.44, 0.41]
Lineup decision	-0.84	0.21	[-1.26, -0.43]
Negative Consequence Likelihood Information × Lineup Decision	-0.90	0.39	[-1.67, -0.14]

Again the data were split by the negative consequence likelihood conditions and separate models were constructed to test whether lineup decision predicted willingness to testify (after controlling for target presence). As was found with post-decisional confidence, lineup decision only predicted willingness to testify in the innocent-person-convicted-more-likely condition. As shown in Table 36, people in this condition indicated that they would be willing to testify to their lineup rejections with greater frequency than their positive identifications. Model fit statistics and fixed effect coefficients appear in Table 38.

Table 38

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Models Testing the Relation Between Lineup Decision and Willingness to Testify (After Controlling for Target Presence), Separately for Each Level of Negative Consequence Likelihood Information

	Negative consequence likelihood information					
	Innocent-person-convicted-more-likely			Guilty-person-released-more-likely		
Fixed effect predictors	χ^2	*df*	*p*	χ^2	*df*	*p*
Target presence	0.28	1	.60	4.77	1	.03
Lineup decision	5.82	1	.06	1.65	1	.20
	b	SE_b	95% CI	*b*	SE_b	95% CI
Intercept	0.41	0.15	[0.12, 0.71]	0.47	0.17	[0.14, 0.80]
Target presence	0.27	0.31	[-0.35, 0.88]	0.68	0.28	[0.12, 1.23]
Lineup decision	-1.16	0.30	[-1.75, -0.57]	-0.47	0.30	[-1.06, 0.11]

Discussion

The negative consequence likelihood information provided to participants in this experiment did not significantly affect their willingness to choose from a lineup (or associated accuracy). However, people's post-identification confidence ratings and their stated willingness to testify to their decisions were affected by the manipulation of consequence perceptions in a way that was consistent with the predicted effects on choosing. People who were told that mistakenly identifying an innocent suspect would be more likely to lead to negative consequences than mistakenly rejecting a lineup were both more confident in and more likely to be willing to testify about their lineup rejections than their positive identification decisions. Those told the opposite did not differ significantly in their

confidence or willingness to testify depending on what decision they made. As post-identification confidence ratings in Experiment 3 were not in any way affected by the negative consequence likelihood manipulation (willingness to testify was not measured), the results obtained in Experiment 4 suggest that the changes to the method that were made resulted in some bias attributable to consequence perceptions arising. While it is not possible to determine which specific aspect(s) of the changes were responsible for the observed effects of consequence perceptions on confidence and willingness to testify emerging, the attempt to strengthen the difference between people's perceptions of which mistake would be worse to make did not appear to have been successful. The effect size was again just above the cut-off for a medium effect.[34] Therefore, it is likely that some combination of the other changes that were made—leading people to feel more accountable for their decisions and, or, more accurately simulating an identification procedure in the current experiment—caused people's post-decision ratings to be affected by the consequence likelihood information.

Although choosing behaviour did not appear to be affected as hypothesised in this experiment, I considered whether perhaps the number of positive identification decisions had been affected by the consequence information in another way. That is, I thought it might be possible that prompting people to think about the negative consequences of identification decisions constrained overall choosing. The worst case scenario of mistakenly choosing from a lineup is that an innocent person goes to jail and, by default, the guilty party is not apprehended, whereas a mistaken rejection can only result in the guilty person getting away (Wells, Steblay, & Dysart, 2012). Therefore, people might have been apprehensive about choosing from a lineup across the board because the overall severity of consequences associated with implicating an innocent suspect were seen (correctly, see Wells, Steblay, &

[34] The confidence intervals for the effect size in this experiment were smaller; however, presumably this reflected the larger sample of participants. From this we can be more certain that the manipulation led to a small–medium difference in people's judgements about which identification error would be worse to make.

Dysart, 2012) as worse than those of failing to implicate a guilty party. This could have created a floor effect that prevented a difference in choosing as a result of differences in perceived negative consequence likelihood to arise. This possibility was unable to be verified, however, because no control condition (where participants were not prompted to think about the consequences associated with identification decisions) was included in the design of the experiment. Archival data using the same stimuli in our laboratory do indicate that perhaps choosing was lower (between 10-20%) than usual in this experiment. However, making this direct comparison is problematic because other experiments conducted with the same stimuli varied on a number of levels (e.g., using additional stimuli, including a "don't know" option and using fewer lineup members). For this reason an ad-hoc pilot study was conducted, in which the choosing rates for the stimuli used in Experiment 4, without having people explicitly considering consequences, could be determined.

Pilot 3

One hundred volunteers from the wider community participated in the pilot. They were 48 males and 52 females. Their ages ranged from 16 to 66 (M = 28.90, SD = 12.94). Participants followed a similar procedure to the one implemented in Experiment 4, except they did not receive the identification decision-related consequence information. From the outset they were told they would be taking part in a short memory test, to be completed on a laptop. After entering their basic demographic information they were shown the stimulus video used in Experiment 4, followed by a distractor video to replace the consequence perception manipulation. The distracter video was 30 seconds long, featuring an excerpt from the popular children's movie *Finding Nemo*, in which a clown fish messes up telling a joke. They were then asked to provide (appearance and behaviour) descriptions of the male and female targets in the stimulus video. Next, participants made identification decisions and gave a post-decisional confidence rating for each target (one of which was present and one of

which was absent in the lineup) in a counterbalanced order. People were not asked how willing they were to testify to their decisions as they were not supposed to be thinking about the implications of identification decisions.

As the purpose of this pilot was to informally obtain baseline identification decision data for the stimuli used in Experiment 4 in order to provide a point of comparison and help guide further research, inferential analyses were not conducted. The choosing rate was 47%, compared with 41.48% and 48.01%, for the innocent-person-convicted-more-likely and guilty-person-released-more-likely conditions in Experiment 4, respectively. The overall accuracy rate was 61.5%, compared with 60.8% in the innocent-person-convicted-more-likely condition and 55.68% in the guilty-person-released-more-likely condition. Thus identification decision patterns did not obviously differ in Experiment 4 from the baseline data. For the pilot data, confidence in positive identification decisions was 63.21% and confidence in lineup rejections was 64.08%, indicating that there was no difference between these conditions. This lack of a difference in confidence for positive identifications versus lineup rejections is consistent with what has been found in other research (e.g., Brewer & Wells, 2006). This pattern of results was matched in the guilty-person-released-more-likely condition, but not the innocent-person-convicted-more-likely condition in Experiment 4. Those in the former condition showed no significant difference in identification confidence for lineup rejections 64.86% versus positive identifications 61.36%; however, those in the latter condition had significantly lower confidence in positive identifications 59.79% than lineup rejections 68.96% (see results section of Experiment 4 for inferential analysis). Therefore, tentatively, the innocent-person-convicted-more-likely, but not the guilty-person-released-more-likely, information led confidence to differ from baseline. Importantly, the pilot results did not provide any striking evidence to suggest that the manipulations in Experiment 4 altered choosing behaviour from baseline, thereby presenting no basis for

further considering the possibility that thinking about identification consequences lowers overall choosing rates.

Discussion Cont. (Experiment 4)

After ruling out the possibility that leading people to think about the consequences associated with actual identification decisions affected choosing in a different way to what was predicted, I turned to interpreting the fact that the only significant effects of the consequence perception manipulation were on people's post-decisional confidence and willingness to testify. As these measures represent qualifying information about the identification decisions people made, they arguably represent more sensitive indicators of people's attitude towards choosing from a lineup.[35] Therefore, the difference in people's consequence perceptions in the current experiment appears to have been not strong enough to affect people's decision making, but strong enough to affect the degree to which people were willing to stand behind their decisions.

One explanation for why only more sensitive measures of people's identification decision approach were affected is that the steps taken to improve the hypothetical consequences paradigm from Experiment 3 were only moderately successful. People were guided through imagining what it would be like to be a witness in an actual criminal investigation and completed related tasks, such as providing a description of what they witnessed, before being asked to view and make identification decisions from a small number of lineups. Further, they were led to feel socially accountable for their decisions by being told that they would learn the likely outcome of a real criminal investigation based on the decisions they made. Yet the experience of being a witness in an actual criminal investigation could have been simulated much more realistically. For example, people could have attended

[35] This point was already made in Chapter 2 as a basis for conducting supplementary analyses on people's post-decision ratings of confidence.

several sessions to more accurately emulate the timeline of an investigation and interacted with someone dressed as a police officer who treated them as an actual witness to a crime. Perhaps the study could have also been run at a police station, instead of in a laboratory on university campus. However, conducting the experiment in this way would have required a significant use of resources and, despite the potential importance of facilitating people's ability to accurately imagine being an actual eyewitness in a criminal investigation, the benefit of going to the lengths described above to simulate an eyewitness identification procedure is somewhat unclear. Indicating that realistically simulating a criminal investigation is perhaps unnecessary, a paper (Kantner, et al., 2015) published shortly after Experiment 4 was run showed evidence of hypothetical consequences shifting people's willingness to make positive recognition decisions in a paradigm that was not particularly realistic.

Participants in Kanter et al. (2015) were told they were taking part in a simulated security patrol and studied a list of suspicious individuals before "embarking on the patrol"—they were then shown a series of still images on the backdrop of an inner city or suburban outskirts. For each trial they had to indicate whether they wanted to stop and search that person or not. In one condition instructions specified that people should operate as if it was okay to miss some suspicious people, but that it was extremely important to not attempt to search innocent people (because they would attempt to flee and probably be harmed in the pursuit). In the other condition people were instructed to act as if it was extremely important to catch all of the suspicious individuals, but that searching innocent people was only a minor inconvenience to them. Results showed that people in the condition where mistakenly pursuing an innocent person was described as a mistake that should be avoided were less likely to indicate that they recognised individuals than people in the condition where mistakenly failing to pursue a guilty person was described as important to avoid. The degree

to which the hypothetical consequences manipulation affected people's willingness to search people was similar to what was found for a base rate manipulation (e.g., 30% or 70% target-present cases, Experiments 1, 2 & 3) and a real payoff manipulation[36] (i.e., +10c for a correct decision in both conditions, -20c for a miss and ±0c for a false alarm in one condition and the opposite in the other condition, Experiments 2 & 3). Therefore, these results suggest that hypothetical consequences in a low-realism simulation can form a basis for biasing people's recognition decisions.[37]

However, I noted two differences between the current experiment and those conducted by Kantner and colleagues (2015) that could potentially account for why the hypothetical consequence perceptions were found to bias people's decisions in their experiments, but not in mine. First, their consequences manipulation struck me as a lot stronger (and therefore more likely to affect people's choosing behaviour) than the one I used; however, as no manipulation check measures were implemented this could not be empirically verified. The reason I judged their manipulation to be stronger was that participants in Kantner et al. (2015) were explicitly told that one type of mistake was okay to make, whereas the consequence likelihood manipulation I implemented at best implied that one mistake might end up being less bad to make. Further, people in the security patrol experiments were clearly instructed to account for the consequences in their decisions, something that was not done in my experiments (although the attempt to lead people to feel

[36] The authors argued that observing a similar effect for the payoff manipulation as the hypothetical consequences manipulation showed evidence that hypothetical decisions match real ones; however, the size of the effect appeared to be confounded (i.e., small real monetary penalties were being compared to large hypothetical consequences). While the degree to which people shifted their decision criterion may well be the maximum people will shift their criterion, it is nevertheless possible that large real consequences would have a greater effect and (or) small hypothetical consequences would have a smaller effect.

[37] Note that participants in Kantner et al. (2015) made a large number or recognition decisions, but this made sense in terms of the security patrol scenario in which the decisions were being made. Therefore it is not my argument that it was unimportant to modify the procedure from Experiment 3 so people were not making a large number of identification decisions. It may well be the case that it is important for people to make a realistic number of identification decisions, but it may not be important for the context in which the decisions are made to be highly realistic.

accountable for their decisions in the current experiment—by telling people they would receive feedback on the likely outcome of an actual case based on their decision—aimed to implicitly cause people to be motivated to account for their consequence perceptions in their decisions). The strength of the manipulation in my experiments was something I was already concerned about, but had not found a way to improve. However, the second key difference I noted between the security patrol experiments and my own highlighted a potential issue with my method that could be quite easily addressed.

Unlike what I had done, which was to use stimuli that led to mid-range performance,[38] Kantner et al. (2015) used highly ambiguous stimuli,[39] which led accuracy to only just exceed chance levels. In doing so they arguably created circumstances where evidence of a shift in willingness to make a positive identification would be expected to be particularly apparent. To clarify, from a signal detection perspective, old and new stimuli that are hard to correctly classify (as such) would be expected to form two highly overlapping distributions on the continuum of signal strength. This in turn would mean that the peaks of the two distributions are close to one another and a shift in people's response criterion (somewhere between the peaks of the distributions) would affect a larger number of cases (compared with a circumstances where the distributions are far apart—i.e., discriminability is high—and a shift in the response criterion would only affect cases in the upper tail end of the "new" distribution and those in the lower tail end of the "old" distribution). Similar to this, it has been argued in the identification literature that under circumstances where people's memory for a target is impaired, factors that have the potential to influence people's willingness to

[38] Although the pilot of baseline performance for the stimuli used in Experiment 4 indicated that accuracy was substantially below ceiling (≈ 60%), people's performance was still well above chance. Note that in an 8-person lineup with a "not present" option chance performance would be 1 in 9 (although this might be adjusted considerably depending on the functional size of the lineup—i.e., how many of the lineup members present plausible options (Wells, Leippe, & Ostrom, 1979)—which is not known for these lineups).

[39] Each target had a unique face (generated using FaceGen software); however, they had 1 of 6 hairstyles and wore 1 of 6 outfits (in a variety of 13 colours).

choose from a lineup are more likely to do so, or should do so to a greater extent (Brewer & Wells, 2011). In the current experiment (and the earlier experiments reported) I expected that the decisions people were less confident in making would be the ones more likely to be affected by the manipulated shift in willingness to make a positive identification and, therefore, that these cases would responsible for driving the effects that were predicted. Along these lines I suggested in Chapter 1 that intuitively it seems plausible for people to be most likely to consider and allow their decisions to be affected by their negative consequence perceptions if they feel that an error could be made. However, I had assumed (perhaps naively it seems) that enough cases would be affected to bear this out in a main effect. However, it might be that a combination of stimuli that people could readily discriminate and a manipulation of consequence perceptions that was small-medium in strength resulted in only a few cases meeting the conditions for people to be motivated to allow their consequence perceptions to guide their decisions. Further, people made their identification decisions under circumstances that were not only conducive to people clearly exceeding chance performance but may have also been subjectively perceived as close to ideal. The encoding conditions were good (i.e., people were warned to pay attention, the target's faces were in view from a close distance for several seconds) and the retention interval was very short (5–10min). Therefore, distinct from their actual ability to correctly choose from or reject the lineups they were shown, many people may have also thought that their memory for the target should be pretty good. As a result they may have felt reasonably confident in making a definitive decision from the lineup and been unlikely to allow their consequence perceptions to affect their decisions.

The above argument suggests that while there might not have been many of these cases, a subset of decisions in this experiment where people felt unsure about whether to choose from or reject the lineup, may have been affected by the consequence perception

manipulation. Consistent with this speculation, the descriptive statistics showed a trend in choosing that was in the hypothesised direction. That is, choosing in the guilty-person-released-more-likely condition was around 6.5% higher than in the innocent-person-convicted-more-likely condition. While this difference could have been a chance occurrence, it could also be indicative of a slight overall shift in willingness to make a positive identification decision driven by a small number of cases where people were not sure which decision to make and allowed their consequence perceptions to guide their decisions. For the following reasons pursuing this possibility seemed worthwhile. If the hypothesised effect could be shown in the current paradigm for a subset of unconfident decisions, it would suggest that there is scope for using an improved version of a hypothetical consequences method to research the effect of consequence perceptions on identification decisions (and value in further developing the paradigm). In contrast, if even in cases where people were quite uncertain about which decision to make their consequences perceptions had no bearing on their lineup choices, then it would be time to conclude that a hypothetical paradigm is not an appropriate means for investigating the relationship between consequence perceptions and identification decisions, or that consequence perceptions do not actually affect people's identification decisions.

CHAPTER 5

Experiment 5

The predicted relationship between people's identification decisions and the consequences associated with their choice was not found in the previous experiment. However, it was argued (in Chapter 4) that perhaps decisions people were not confident in making had been affected by the manipulation of people's consequence perceptions, but there were simply not enough cases where this occurred for the effect to be reflected in overall choosing rates. In light of this possibility, Experiment 5 was designed to replicate Experiment 4 and allow the exploration of whether a subset of low confidence decisions was significantly affected by the negative consequence likelihood manipulation. Several studies have measured people's pre-decision confidence (e.g., Brigham & Cairns, 1988; Cutler & Penrod, 1988; Fleet, Brigham, & Bothwell, 1987; Murray & Wells, 1982) by asking them to indicate how confident they were in making an identification decision prior to seeing the lineup. However, I was interested in how confident people felt in making a particular identification decision once they had seen the lineup: this construct will henceforth be referred to as "initial confidence". Therefore, in an unorthodox procedure, people were shown each lineup requiring an identification decision twice. The first time they were told that they were not required to make a final decision about whether to choose from or reject the lineup, but that they should indicate how confident they were in making a decision. They were not asked which decision they were leaning towards in an attempt to avoid causing people to stick with their initial inclination because they felt they had already committed to that decision (see the following for investigations of commitment effects in the identification literature, Deffenbacher, Bornstein, & Penrod, 2006; Gorenstein & Ellsworth, 1980; Palmer, Brewer, &

Weber, 2012). The second time they were shown the lineup, following the manipulation of people's consequence likelihood perceptions,[40] people were asked to make a final decision.

It was predicted that at lower levels of initial confidence (a precise point was not defined) people's decisions would be more likely to be affected by whether they were told that mistakenly choosing from or mistakenly rejecting a lineup would be more likely to lead to negative consequences (with choosing being lower in the condition where mistakenly choosing was framed as being more likely to lead to negative consequences).

Method

Participants and Design

A total of 512 participants completed the study (176 male and 336 female). Their ages ranged from 17 to 56 (M = 21.26, SD = 5.76). As I aimed to explore the difference in choosing between the consequence likelihood conditions across 11 levels of pre-decision confidence (i.e., 0-100%), the largest sample I could obtain (given time and funding considerations) was targeted. I estimated that 500 subjects would be feasible to recruit within my timeframe and 512 was the closest number to this target that ensured an equal number of participants completed the various experimental conditions. Following the design of Experiment 4, a 2 (negative consequence likelihood information: guilty-person-released-more-likely, innocent-person-convicted-more-likely) × 2 (target: male, female) × 2 (target presence: present, absent) design was used, where participants were randomly assigned to a negative consequence likelihood information condition and made identification decisions for two targets across which target presence varied.

[40] While people would not necessarily be expected to form their opinion about negative consequence likelihood until after seeing a lineup, it was important to measure confidence in making a choice before manipulating consequence likelihood information to avoid people's confidence ratings being affected by the manipulation.

Materials

Negative consequence likelihood manipulation. People's consequence likelihood perceptions were manipulated as they were in Experiment 4; however, different mock-newspaper articles (describing the worst case scenario that people were told would be more likely to occur) were shown to people. The articles were adapted from an actual opinion piece published in the New York Times (Thompson, 2000), where the witness from a high profile wrongful conviction case outlined her experience of making a positive identification from a lineup and finding out 10 years later that the wrong man had been sent to jail. In the innocent-person-convicted-more-likely condition the article closely matched the real witness's account of what had occurred. In the guilty-person-released-more-likely condition the narrative was significantly altered to describe a scenario where the lineup was rejected and later the witness learned that the offender had been released as a result and gone on to kidnap someone and hold them hostage for the next 10 years. The exact content of the articles is shown in Table 39. These new articles were used because providing people with a first person of account being a witness who made a serious identification error could potentially have several benefits. Firstly, they might be more interesting reading that better engages people and improves the chances that they will take on board the consequence likelihood information presented to them. Secondly, it might help people to imagine and relate to what it would be like to make a crucial error and to motivate people to account for their consequence perceptions in their identification decisions despite making them in a non-consequential context.

Table 39

Content of the Mock Newspaper Articles Participants Read in Experiment 5

Condition	Mock newspaper article content
Innocent-person-convicted-more-likely condition	"I was certain but I was wrong", by Jennifer Thompson In 1984 I was a 22-year-old college student with a grade point average of 4.0, and I really wanted to do something with my life. One night someone broke into my apartment, put a knife to my throat and raped me. During my ordeal, some of my determination took an urgent new direction. I studied every single detail on the perpetrator's face. I looked at his hairline; I looked for scars, for tattoos, for anything that would help me identify him. When and if I survived, I was going to make sure that he was put in prison and he was going to rot. When I went to the police department later that day, I worked on a composite sketch to the very best of my ability. I looked through hundreds of noses and eyes and eyebrows and hairlines and nostrils and lips. Several days later, looking at a series of police photos, I identified my attacker. I knew this was the man. I was completely confident. I was sure. I had picked the right guy, and he was going to go to jail. If there was the possibility of a death sentence, I wanted him to die. I wanted to flip the switch. Based on my testimony, Ronald Cotton was sentenced to prison for life. I was filled with grim satisfaction because justice had been done. But in 1995 the case was reopened. I learned that **another** man than the one I had identified from that lineup all those years ago had supposedly claimed to be my attacker and was bragging about it on arrival in the same prison wing where Ronald Cotton had been held for the last 11 years. This man, Bobby Poole, was brought into court, and I was asked, "Ms. Thompson, have you ever seen this man?" I answered: "I have never seen him in my life. I have no idea who he is." I was asked to provide a blood sample so that DNA tests could be run on evidence from the rape. I agreed because I knew that Ronald Cotton had raped me and DNA was only going to confirm that. I will never forget the day I learned about the DNA results. I was standing in my kitchen when the detective and the district attorney visited. They were good and decent people who were trying to do their jobs -- as I had done mine, as anyone would try to do the right thing. They told me: "Bobby Poole is the man who raped you." The man I was so sure I had never seen in my life was the man who was inches from my throat, who raped me, who hurt me, who took my spirit away, who robbed me of my soul. And the man I had identified so emphatically, which had led to his conviction and subsequent incarceration for 11 years, was absolutely innocent. Ronald Cotton and I are the same age, so I knew what he had missed during the last 11 years of being imprisoned. My life had gone on. I had gotten married. I had graduated from college. I worked. I was a parent. Ronald Cotton hadn't gotten to do any of that. Although he is now moving on with his own life, I live with constant anguish that my profound mistake cost him so dearly. If anything good can come out of what Ronald Cotton suffered because of my limitations as a human being, let it be an awareness of the fact that eyewitnesses can and do make mistakes.

Condition	Mock newspaper article content
Guilty-person-released-more-likely condition	"I was certain but I was wrong", by Jennifer Thompson In 1984 I was a 22-year-old college student with a grade point average of 4.0, and I really wanted to do something with my life. One night someone broke into my apartment, put a knife to my throat and raped me. During this ordeal, some of my determination took an urgent new direction. I studied every single detail on the perpetrators face. I looked at his hairline; I looked for scars, for tattoos, for anything that would help me identify him. When and if I survived, I was going to make sure that he was put in prison and he was going to rot. When I went to the police department later that day, I worked on a composite sketch to the very best of my ability. I looked through hundreds of noses and eyes and eyebrows and hairlines and nostrils and lips. Several days later, looking at a series of police photos, I did not identify my attacker. I knew none of these was the man. I was completely confident. I was sure. Based on my testimony, Bobby Poole was released from facing a life sentence in prison. I was filled with grim satisfaction because justice had been done. I had steered the police away from the wrong guy and when they found the actual rapist he was going to go to jail. If there was the possibility of a death sentence, I wanted him to die. I wanted to flip the switch. It wasn't until 1995 the case was reopened. I learned that the **same** man that had been in the lineup I rejected all those years ago had supposedly claimed to be my attacker and bragged about it to a kidnapping victim, Connie Cotton, that he had held hostage for the past 11 years. This man, Bobby Poole, was brought into court, and I was asked, "Ms. Thompson, have you ever seen this man?" I answered: "I have never seen him in my life. I have no idea who he is." I was asked to provide a blood sample so that DNA tests could be run on evidence from the rape. I agreed because I knew that this was not the person who had raped me and DNA was only going to confirm that. I will never forget the day I learned about the DNA results. I was standing in my kitchen when the detective and the district attorney visited. They were good and decent people who were trying to do their jobs -- as I had done mine, as anyone would try to do the right thing. They told me: "Bobby Poole is the man who raped you." The man I was so sure I had never seen in my life was the man who was inches from my throat, who raped me, who hurt me, who took my spirit away, who robbed me of my soul. And the fact that I had not identified him had led to his release and subsequent kidnapping of someone absolutely innocent for 11 years. Connie Cotton and I are the same age, so I knew what she had missed during the last 11 years of being imprisoned. My life had gone on. I had gotten married. I had graduated from college. I worked. I was a parent. Connie Cotton hadn't gotten to do any of that. Although she is now moving on with her own life, I live with constant anguish that my profound mistake cost her so dearly. If anything good can come out of what Connie Cotton suffered because of my limitations as a human being, let it be an awareness of the fact that eyewitnesses can and do make mistakes.

Initial confidence. People's initial confidence in making an identification decision was measured by showing them each lineup from which they would be required to choose using the procedure outlined in Experiment 4. They were instructed to indicate as quickly as possible how confident they felt about making a decision. Asking people to make this judgement rapidly was done to limit the chances that a definitive decision would be made at that stage. They received these instructions prior to being shown the lineup with the confidence scale (0–100% confident in making a decision) underneath. The lineup was presented in a different configuration (i.e., the position of the lineup members was different) to the second time it would be shown the participant,[41] in an attempt to avoid people basing their decision on something other than their recognition memory or the consequence information (e.g., remembering that they thought number 1 looked the most like the target the first time).

Apart from the changes and additions outlined above, the materials used in this experiment were the same as Experiment 4.

Procedure

Participants first viewed the stimulus video, before being told that they were taking part in a simulated criminal investigation and guided through imagining the case unfolding (as outlined in Experiment 4). After being asked to describe the appearance and behaviour of the targets, people's initial confidence was measured as delineated in the materials section above. Next, the consequence likelihood manipulation was administered (as in Experiment

[41] The lineups for obtaining initial confidence were presented with the target (or the target replacement for target-absent lineups) inserted in one of the 8 positions (with the fillers in a fixed order for each version). Due to a software issue that arose from having the confidence scale on the same page as the lineup, the position of the lineup members could not be randomly determined. The second time people were shown that lineup (to make their final decision) the position of the lineup members was randomly generated. This meant that there was no guarantee that none of the lineup members would appear in the same position in both presentations of the lineup, however, it was extremely unlikely that both lineups would be configured in the exact same way.

4)[42] before people were asked to make their final identification decisions and indicate their post-decisional confidence level and willingness to testify. Finally the manipulation check measures were obtained as in Experiment 4.

Results

Manipulation Checks

A 2 (negative consequence likelihood information: guilty-person-released-more-likely, innocent-person-convicted-more-likely) × 2 (identification error type: incorrect suspect pick, incorrect lineup rejection) mixed ANOVA showed that overall consequence likelihood perceptions were not affected by the negative consequence likelihood information, $F(1, 510) = 2.25$, $p = .13$, $d = 0.10$, d 95% CI [-0.07, 0.27]. However, a significant main effect for identification error type was detected, with lineup rejections rated as being more likely to lead to negative consequences than positive identifications, $F(1, 510) = 34.45$, $p < .001$, $d = 0.31$, d 95% CI [0.19, 0.43]. The Negative Consequence Likelihood Information × Identification Error Type interaction was also significant, $F(1, 150) = 72.77$, $p < .001$, $f = 0.23$, with descriptive statistics (see Table 40) showing that, in the guilty-person-released-more-likely condition, mistaken lineup rejections were rated as more likely to lead to negative consequences. In contrast, in the innocent-person-convicted-more-likely condition there was a weak trend towards mistakenly choosing being rated as more likely to lead to negative consequences. Simple effects analyses showed that the difference in the guilty-person-released-more-likely condition was significant, $t(255) = 9.97$, $p < .001$, $d = 0.76$, d 95% CI [0.59, 0.95], but the difference in the innocent-person-convicted-more-likely condition was not, $t(255) = 1.92$, $p = .06$, $d = 0.15$, d 95% CI [-0.02, 0.32].

[42] Although note that people were told the newspaper article content they were going to read was published On June 8th 2000 in the New York Times (instead of April 4th 2009 in the Boston Globe, which was what people were told in Experiments 3 and 4).

On the relative consequence likelihood scale, where the low end of the scale (i.e., 0) indicates the belief that a mistaken rejection would be much more likely to lead to negative consequences and the high end of the scale (i.e., 6) reflects the position that mistakenly choosing (an innocent suspect) is far more likely to have negative repercussions, those in the innocent-person-convicted-more-likely condition (M = 2.96, SD = 1.64) gave higher ratings than those in the guilty-person-released-more-likely condition (M = 1.86, SD = 1.62), $t(510) = -7.59$, $p < .001$, $d = 0.67$, d 95% CI [0.49, 0.85]. On the equivalent scale measuring which mistake would be worse to make, those in the innocent-person-convicted-more-likely condition (M = 4.09, SD = 1.64) again gave higher ratings than those in the guilty-person-released-more-likely condition (M = 2.60, SD = 1.77), $t(510) = -9.85$, $p < .001$, $d = 0.87$, d 95% CI [0.69, 1.05].

Table 40

Mean (and Standard Deviation) Ratings of Negative Consequence Likelihood as a Result of Incorrectly Picking from or Rejecting a Lineup, Across the Negative Consequence Likelihood Information Conditions

	Negative consequence likelihood information		
Identification error type	Guilty-person-released-more-likely	Innocent-person-convicted-more-likely	Total
Incorrect lineup rejection	4.01 (1.46)	3.46 (1.53)	3.74 (1.52)
Incorrect suspect pick	2.82 (1.62)	3.68 (1.44)	3.25 (1.59)
Total	3.42 (1.65)	3.57 (1.48)	3.50 (1.57)

Main Analyses

The present study was conducted to investigate whether people's consequence perceptions only affected their identification decisions when people were ambivalent about whether the target appeared in the lineup or not. First, however, the data were analysed across

all cases (to provide a direct comparison of results to Experiment 4). Then the results were examined at each level of people's first impression of how confident they were in making a decision from the lineups they were shown.

Choosing. The overall choosing rate was 61.33%. To test whether the consequence perception manipulation affected choosing behaviour, a logistic mixed-effects model was constructed with choosing as the outcome variable and negative consequence likelihood information as a fixed factor. Negative consequence likelihood information was found to significantly estimate choosing, $\chi^2(1) = 4.49$, $p = .03$, $b = -0.38$, $SE_b = 0.13$, b 95% CI [-0.64, -0.11],[43] characterised by a lower rate in the innocent-person-convicted-more-likely condition (57.23%) than the guilty-person-released-more-likely condition (65.43%).

Accuracy. To test whether the effect of consequence perceptions on accuracy depended on target presence (with target-present accuracy expected to be higher in the guilty-person-released-more-likely condition and target-absent accuracy expected to be higher in the innocent-person-convicted-more-likely condition), a logistic mixed-effects model was created with accuracy as the outcome variable. Negative consequence likelihood information, followed by target presence and the Negative Consequence Likelihood Information × Target Presence interaction were entered as fixed effects. Negative consequence likelihood information and target presence were not related to accuracy. However, the Negative Consequence Likelihood Information × Target Presence interaction significantly estimated accuracy, with aggregate statistics showing the predicted pattern. That is, as shown by the descriptive statistics in Table 41, target-present accuracy was higher in the guilty-person-released-more-likely condition and target-absent accuracy was higher in the innocent-person-convicted-more-likely condition. Model fit statistics and fixed effect coefficients are shown in Table 42.

[43] Intercept $b = 0.52$, $SEb = 0.21$, b 95% CI [0.11, 0.93]

Table 41

Percentage Accuracy Rates Across the Negative Consequence Likelihood Information and Target Presence Conditions

	Target presence					
Negative consequence likelihood information	N trials	Present	N trials	Absent	N trials	Total
Innocent-person-convicted-more-likely	256	41.41	256	56.64	512	49.02
Guilty-person-released-more-likely	256	55.86	256	48.05	512	51.95
Total	512	48.63	512	52.34	1024	50.49

Table 42

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Negative Consequence Likelihood Information and Target Presence on Accuracy

Fixed effect predictors	χ^2	*df*	*p*
Negative consequence likelihood information	0.87	1	.35
Target presence	1.70	1	.19
Negative Consequence Likelihood Information × Target Presence	5.77	1	.02
	b	SE_b	95% CI
Intercept	0.02	0.29	[-0.55, 0.59]
Negative consequence likelihood information	-0.14	0.15	[-0.45, 0.16]
Target presence	-0.18	0.14	[-0.45, 0.09]
Negative Consequence Likelihood Information × Target Presence	-1.11	0.28	[-1.66, -0.56]

Models testing the effect of negative consequence likelihood information on accuracy separately for target-present and target-absent cases (see Table 43 for inferential statistics)

showed that only target-present accuracy was significantly different, being 14.45% lower in the innocent-person-convicted-more-likely than the guilty-person-released-more-likely condition. In summary, the results of this experiment showed a modest but significant shift in choosing behaviour as a result of contrasting information about the consequences associated with identification decisions. This shift further translated into the predicted opposite trends in accuracy for target-present and target-absent cases.

Table 43

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Models Testing the Effect of Negative Consequence Likelihood Information on Accuracy Separately for Target-Present and Target-Absent Cases

	Target presence					
	Present			Absent		
Fixed effect predictors	χ^2	*df*	*p*	χ^2	*df*	*p*
Negative consequence likelihood information	3.36	1	.07	3.04	1	.08
	b	SE_b	95% CI	*b*	SE_b	95% CI
Intercept	-0.07	0.38	[-0.81, 0.68]	0.09	0.13	[-0.17, 0.36]
Negative consequence likelihood information	-0.62	0.23	[-1.08, -0.17]	0.35	0.18	[-0.003, 0.70]

Initial confidence. To recap, the main focus of this experiment was on distinguishing between people who were and were not confident in making a particular identification decision and testing whether those who tended towards the latter were more likely to let their decision be guided by their consequence perceptions. In order to measure how confident people were in making each identification decision people were shown each lineup twice. The first time they were asked to indicate only their confidence in making a decision from that lineup (i.e., they were not actually asked to make a final decision or indicate what

decision they were leaning towards). Then, after being provided with the negative consequence likelihood information, people were asked to make their final identification decision. While the general hypothesis was that those who were less certain about which decision to make would be more likely to be affected by the potential consequences associated with their decision, I had no inkling of the level of initial confidence that might be expected to mark the point of increase. Therefore, the data were examined by comparing the choosing rate for the two negative consequence likelihood conditions after removing all of the participants who were 100% confident in making a decision, then all who were 90% confident, 80% confident and so on. The descriptive data from this procedure are summarised in Table 44 and the pattern is also shown in Figure 1. Choosing remained a little (≈ 10%) higher in the guilty-person-released-more-likely condition than in the innocent-person-convicted-more-likely condition until the point where initial confidence was 20% or lower. At this point the difference doubled in size, suggesting that perhaps these very unsure participants were more likely to have allowed their perceptions about consequences to influence their decisions. The apparent pattern could have been confounded by target presence, a variable that strongly influences choosing (being lower for target-absent cases) and was not necessarily equivalent across conditions. Therefore, a logistic mixed-effects model testing the effect of the negative consequence likelihood manipulation on choosing (and controlling for target presence) was constructed after each level of initial confidence was removed. Significant differences are indicated by an asterisk in Table 44. Notably, the 23.85% difference in choosing across consequence conditions for people who expressed initial confidence that was 20% or lower (i.e., < 30%), was significant.

Table 44

Descriptive and Inferential Statistics for the Effect of Negative Consequence Likelihood Information on Choosing After Removing Each Descending Level of Initial Confidence in Making a Decision

Initial confidence	All	<100	<90	<80	<70	<60	<50	<40	<30	<20	<10
Number of trials											
Innocent-person-convicted-more-likely	512	473	416	327	243	182	137	94	60	42	31
Guilty-person-released-more-likely	512	449	395	317	237	170	112	78	58	39	30
Number of picks											
Innocent-person-convicted-more-likely	293	279	254	207	157	114	83	50	25	18	13
Guilty-person-released-more-likely	335	312	281	229	171	126	80	52	38	24	18
Percentage of picks											
Innocent-person-convicted-more-likely	57.23	58.99	61.06	63.30	64.61	62.64	60.58	53.19	41.67	42.86	41.94
Guilty-person-released-more-likely	65.43	69.49	71.14	72.24	72.15	74.12	71.43	66.67	65.52	61.54	60.00
Percent difference in choosing	8.20*	10.50*	10.08*	8.94*	7.54	11.48*	10.84	13.48	23.85*	18.68	18.06

*$p < .05$

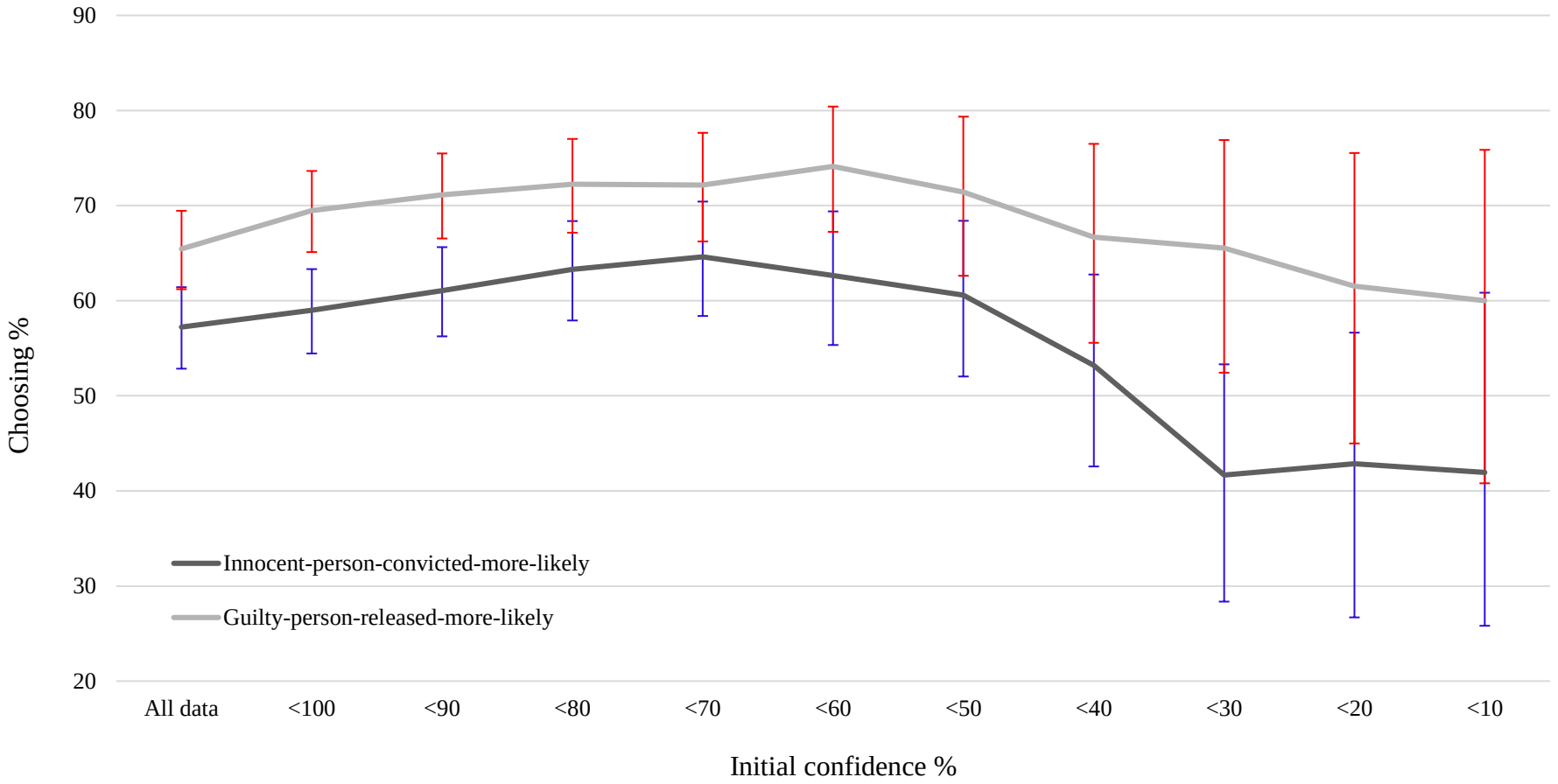

Figure 1. Percentage choosing rates (and 95% confidence intervals) for the innocent-person-convicted-more-likely and guilty-person-released-more-likely conditions at each step of removing descending levels of initial confidence

It seemed unlikely that the pattern of results just described would lead to a significant interaction between negative consequence likelihood information and initial confidence on choosing. Nevertheless, a mixed-effects model was constructed with choosing as the outcome variable. As target presence was not necessarily the same across levels of initial confidence, its effects on choosing were controlled for by entering it into the model as the first fixed predictor. Negative consequence likelihood information, initial confidence[44] and their interaction were then entered. Model fit statistics and fixed effect coefficients are shown in Table 45. As already established earlier, negative consequence likelihood information was a significant predictor of choosing, with choosing being lower in the innocent-person-convicted-more-likely condition than the guilty-person-released-more-likely condition. Initial confidence was also a significant predictor of choosing, with the fixed effect coefficient suggesting that, as initial confidence decreased, choosing increased. The pattern is shown in Figure 2. The interaction between the negative consequence likelihood information and initial confidence was not a significant predictor of choosing.

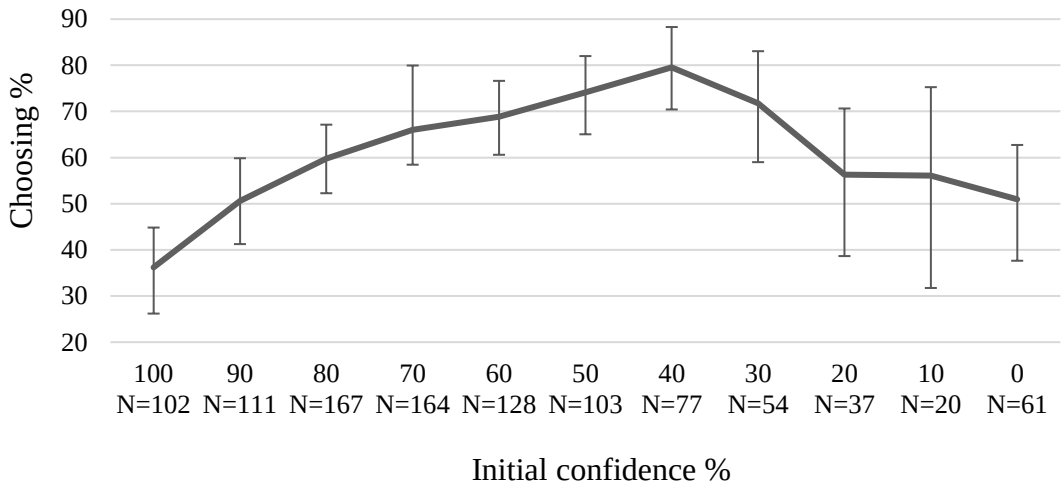

Figure 2. Choosing (with 95% confidence intervals) at each level of initial confidence

[44] As per recommendations for continuous predictor variables in regression, the mean was centred prior to being included in the analysis (Cohen, Cohen, West, & Aiken, 2003; Dalal & Zickar, 2012).

Table 45

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Negative Consequence Likelihood Information and Initial Confidence on Choosing (After Controlling for Target Presence)

Fixed effect predictors	χ^2	*df*	*p*
Target presence	82.36	1	<.001
Negative consequence likelihood information	4.51	1	.03
Initial confidence	4.96	1	.03
Negative Consequence Likelihood Information × Initial Confidence	0.98	1	.32
	b	SE_b	95% CI
Intercept	0.52	0.07	[0.38, 0.65]
Target presence	1.20	0.14	[0.93, 1.47]
Negative consequence likelihood information	-0.41	0.14	[-0.68, -0.14]
Initial confidence	-0.09	0.03	[-0.14, -0.04]
Negative Consequence Likelihood Information × Initial Confidence	0.05	0.05	[-0.05, 0.15]

Whether the pattern of results for accuracy depended on initial confidence was not formally investigated (i.e., with inferential tests). The relevant comparisons would have involved examining an interaction at each level of initial confidence and, therefore, been hopelessly underpowered. However, graphs were constructed to show the relevant interactions after successively removing cases where initial confidence was 100%, 90% and so forth. Figure 3 shows that for target-absent cases the difference between the innocent-person-convicted-more-likely and guilty-person-released-more-likely conditions is much larger at the lower end of the scale (i.e., when only cases with very low initial confidence are considered).

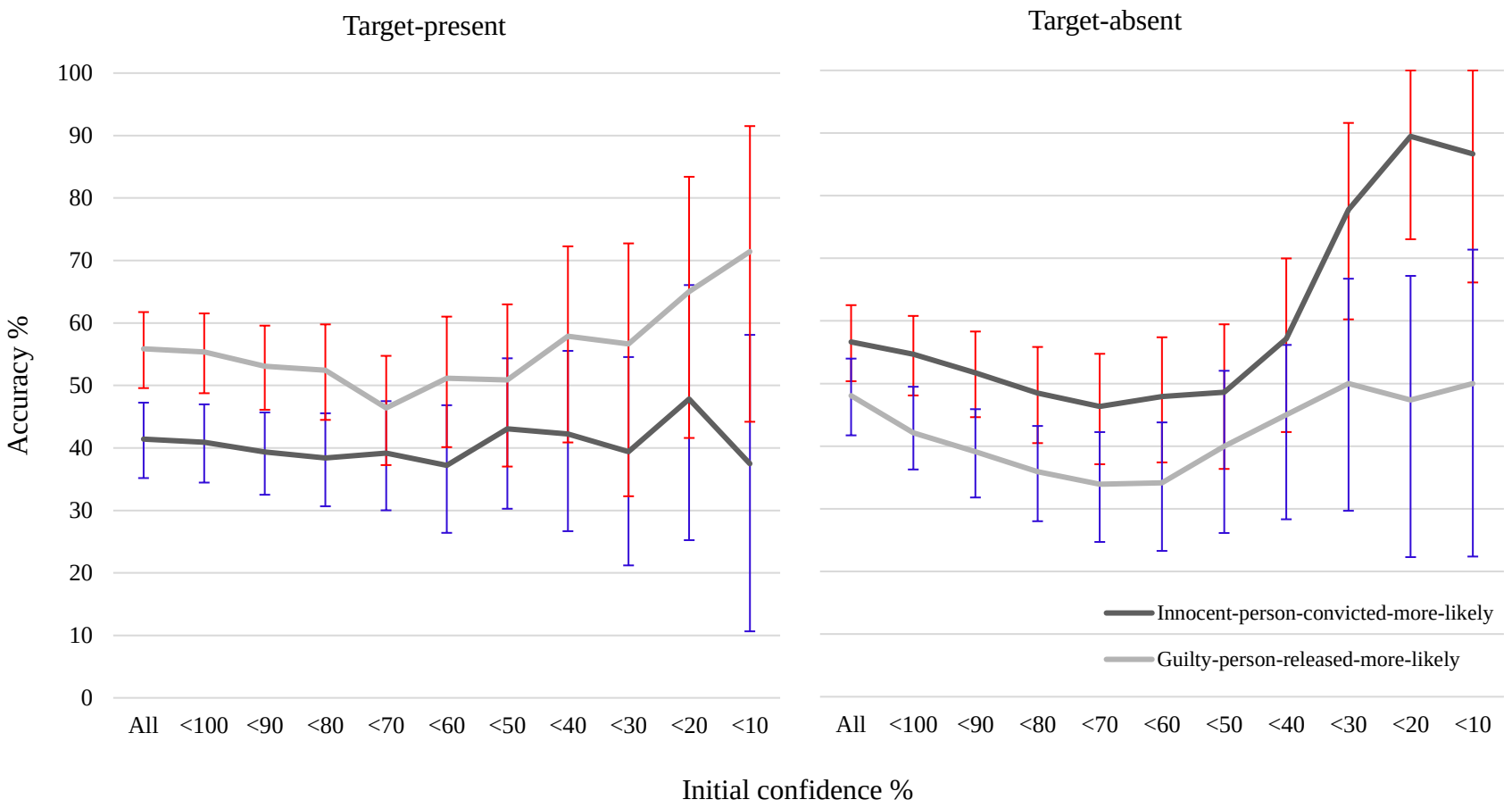

Figure 3. Percentage accuracy rates (with 95% confidence intervals) for the innocent-person-convicted-more-likely and guilty-person-released-more-likely conditions at each step of removing descending levels of initial confidence, with separate panels for target-present and target-absent cases

Supplementary Analyses

Post-decision confidence. As in all previous experiments, I examined whether confidence differed depending on the consequence related information with which people were provided and the lineup decision they made. In Experiment 4, people in the innocent-person-convicted-more-likely condition were found to be more confident in their lineup rejections than their positive identifications, whereas people in the guilty-person-released-more-likely condition were not found to differ. It was therefore expected that this effect would be replicated here. A linear mixed-effects model was constructed with post-identification decision confidence as the outcome variable. Target presence was entered into the model first to control for it. Then negative consequence likelihood information, lineup decision and finally the interaction between those two variables were entered as fixed effects. The model fit statistics and fixed effect coefficients can be found in Table 47. Target presence was a significant predictor of confidence, with the fixed effect coefficient indicating that confidence was higher for target-present cases. Negative consequence likelihood information and lineup decision were not significant predictors of confidence. However, there was a significant interaction between the two. Aggregate statistics (see Table 46) suggest that, in line what was found in Experiment 4, those in the innocent-person-convicted-more-likely condition were more confident in their lineup rejections than their positive identifications, whereas those in the guilty-person-released-more-likely condition recorded similar mean levels of confidence across the type of lineup choice they made. Models testing the relationship between lineup decision and confidence separately for the innocent-person-convicted-more-likely and guilty-person-released-more-likely conditions (again controlling for target presence) confirmed that people's decisions only predicted their confidence in the condition where people were told that choosing would be more likely to lead to negative consequences (see Table 48 for inferential statistics).

Table 46

Mean (and Standard Deviation) Post-Decision Confidence Across the Negative Consequence Likelihood Information and Lineup Decision Conditions

	Lineup decision					
Negative consequence likelihood information	N trials	Picked	N trials	Rejected	N trials	Total
Innocent-person-convicted-more-likely	293	56.38 (21.37)	219	61.74 (20.15)	512	58.67 (21.01)
Guilty-person-released-more-likely	335	60.45 (21.53)	177	60.23 (21.87)	512	60.37 (21.63)
Total	628	58.55 (21.54)	396	61.06 (20.93)	1024	59.52 (21.33)

Table 47

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Negative Consequence Likelihood Information and Lineup Decision on Post-Decision Confidence (After Controlling for Target Presence)

Fixed effect predictors	χ^2	*df*	*p*
Target presence	6.13	1	.01
Negative consequence likelihood information	0.68	1	.41
Lineup decision	3.55	1	.06
Negative Consequence Likelihood Information × Lineup Decision	4.16	1	.04
	b	SE_b	95% CI
Intercept	59.73	0.78	[58.20, 61.26]
Target presence	5.55	1.13	[3.34, 7.76]
Negative consequence likelihood information	-1.19	1.56	[-4.24, 1.87]
Lineup decision	-3.03	1.31	[-5.61, -0.45]
Negative Consequence Likelihood Information × Lineup Decision	-6.73	2.49	[-11.61, -1.84]

Table 48

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Models Testing the Relation Between Lineup Decision and Post-Decision Confidence (After Controlling for Target Presence), Separately for Each Level of Negative Consequence Likelihood Information

	Negative consequence likelihood information					
	Innocent-person-convicted-more-likely			Guilty-person-released-more-likely		
Fixed effect predictors	χ^2	*df*	*p*	χ^2	*df*	*p*
Target presence	4.72	1	.03	5.37	1	.02
Lineup decision	4.51	1	.03	0.01	1	.90
	b	SE_b	95% CI	*b*	SE_b	95% CI
Intercept	59.13	1.06	[57.04, 61.21]	60.34	1.14	[58.10, 62.58]
Target presence	5.16	1.59	[2.05, 8.27]	5.99	1.61	[2.84, 9.14]
Lineup decision	-6.28	2.18	[-10.55, -2.00]	0.23	1.91	[-3.52, 3.98]

Willingness to testify. To examine whether willingness to testify differed depending on consequence condition and lineup decision (in the same way as confidence), a logistic mixed-effects model was created in which willingness to testify was entered as the dependent variable. Target presence was entered as the first fixed effect to control for it, followed by negative consequence likelihood information, lineup decision and the Negative Consequence Likelihood Information × Lineup Decision interaction. Descriptive statistics are summarised in Table 49. Model fit statistics and fixed effect coefficients are shown in Table 50. Target presence was again a significant predictor of willingness to testify, with people being more likely to be willing to testify when the target was present in the lineup. Negative consequence likelihood information did not predict people's willingness to testify. However, lineup

decision did, with people overall being more willing to testify to their rejections than their positive identifications. The interaction between the two key predictors was also significant and the descriptive data indicates that, similar to what was found for confidence ratings, the main effect was driven by people being more willing to testify to their lineup rejections in the innocent-person-convicted-more-likely condition, while those in the guilty-person-released-more-likely condition showed virtually no difference.

Table 49

Percentage Willingness to Testify Rates Across the Negative Consequence Likelihood Information and Lineup Decision Conditions

	Lineup decision					
Negative consequence likelihood information	N trials	Picked	N trials	Rejected	N trials	Total
Innocent-person-convicted-more-likely	293	48.46	219	62.10	512	54.30
Guilty-person-released-more-likely	335	59.10	177	61.02	512	59.77
Total	628	54.14	396	61.62	1024	57.03

Again, models testing the relationship between the decision people made and their stated willingness to testify, for each level of the consequence likelihood manipulation, were constructed. These indicated that lineup decisions only predicted willingness to testify for those who were told that choosing would be more likely to lead to lead to negative consequences[45] (see Table 51 for inferential statistics).

As was the case with the accuracy data, whether the pattern of results for post-decisional confidence and willingness to testify differed depending on initial confidence was not formally tested. However, the descriptive data are shown below. Figure 5, summarising

[45] But note that only a random intercepts model was able to be fitted, so these results should be treated with caution.

post-decisional confidence ratings, indicates that in the guilty-person-released-more-likely condition, the difference between picks and rejections (i.e., higher confidence for rejections) appears to become larger as cases of higher initial confidence are excluded. The same pattern is evident for willingness to testify, shown in Figure 6. The confidence intervals clearly show that these trends are by no means definitive. However, these patterns warrant comment because the effect of consequence perceptions on identification decisions—and particularly the possibility that the effects might be larger for people who are unsure—has received so little attention.

Table 50

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Model Testing the Interaction Between Negative Consequence Likelihood Information and Lineup Decision on Willingness to Testify (After Controlling for Target Presence)

Fixed effect predictors	χ^2	*df*	*p*
Target presence	2.79	1	.10
Negative consequence likelihood information	1.58	1	.21
Lineup decision	9.48	1	.002
Negative Consequence Likelihood Information × Lineup Decision	3.88	1	.05
	b	SE_b	95% CI
Intercept	0.44	0.10	[0.24, 0.64]
Target presence	0.40	0.16	[0.09, 0.71]
Negative consequence likelihood information	-0.26	0.20	[-0.64, 0.13]
Lineup decision	-0.54	0.18	[-0.90, -0.19]
Negative Consequence Likelihood Information × Lineup Decision	-0.83	0.35	[-1.51, -0.14]

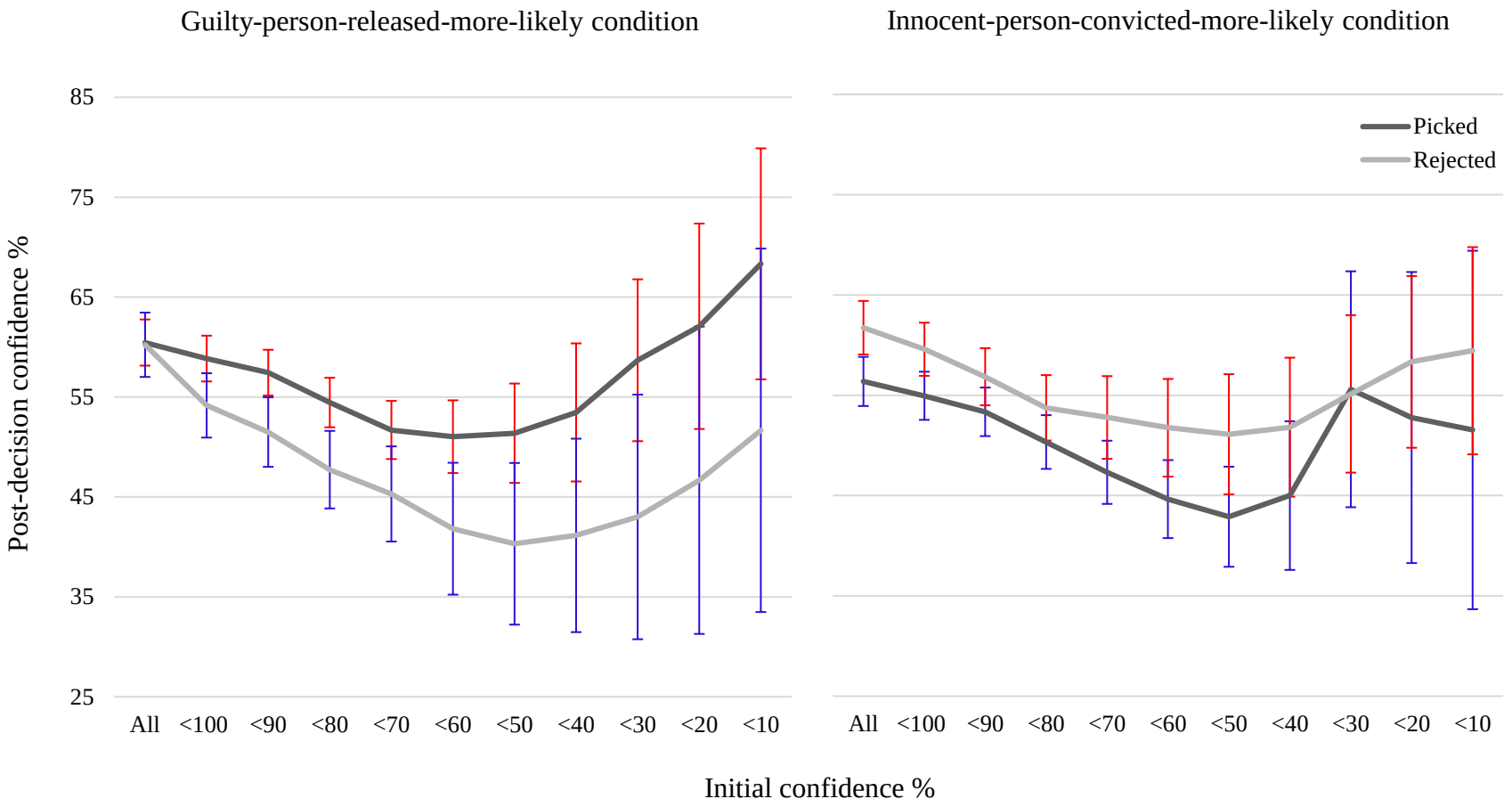

Figure 4. Mean post-decision confidence (with 95% confidence intervals) for lineup picks and rejections at each step of removing descending levels of initial confidence, with separate panels for the innocent-person-convicted-more-likely and guilty-person-released-more-likely conditions

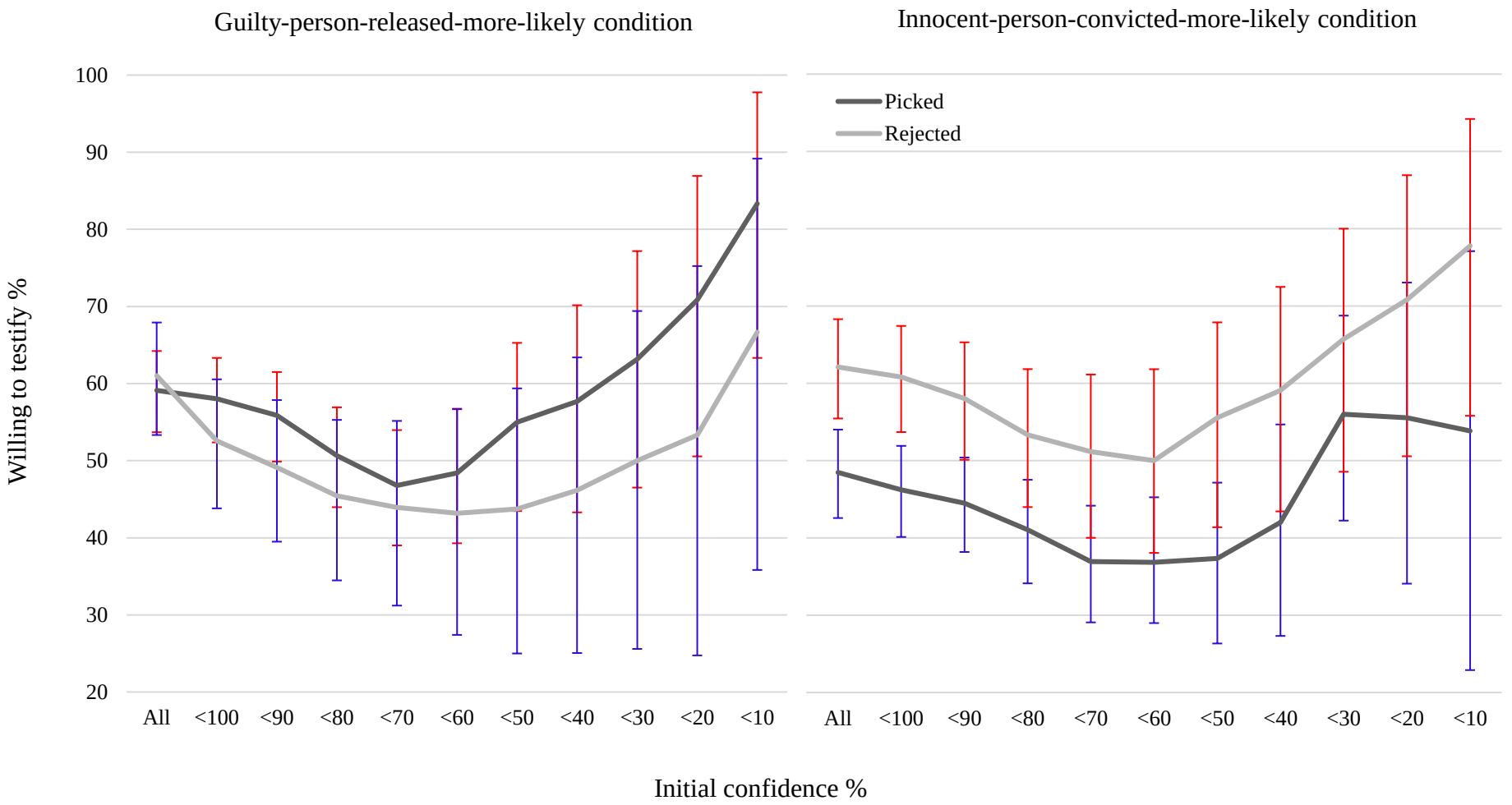

Figure 5. Percentage willingness to testify rates (with 95% confidence intervals) for lineup picks and rejections at each step of removing descending levels of initial confidence, with separate panels for the innocent-person-convicted-more-likely and guilty-person-released-more-likely conditions

Table 51

Improvement of Model Fit Statistics and Fixed Effect Coefficients for the Mixed-Effects Models Testing the Relation Between Lineup Decision and Willingness to Testify (After Controlling for Target Presence), Separately for Each Level of Negative Consequence Likelihood Information

	Negative consequence likelihood information					
	Innocent-person-convicted-more-likely			Guilty-person-released-more-likely		
Fixed effect predictors	χ^2	*df*	*p*	χ^2	*df*	*p*
Target presence	2.13	1	.14	3.66	1	.06
Lineup decision	6.05	1	.01	0.31	1	.58
	b	SE_b	95% CI	*b*	SE_b	95% CI
Intercept	0.31	0.14	[0.04, 0.58]	0.58	0.15	[0.28, 0.87]
Target presence	0.35	0.23	[-0.09, 0.80]	0.45	0.22	[0.002, 0.89]
Lineup decision	-0.93	0.27	[-1.46, -0.41]	-0.14	0.26	[-0.65, 0.36]

Discussion

The main focus of the present experiment was to explore the choosing data broken down by people's initial confidence in making a decision (expressed when first viewing the lineup), to determine whether the consequence likelihood information people were provided with appeared to influence their choosing behaviour at low levels of initial confidence. In line with what was predicted, results showed descriptive evidence to indicate that people who expressed being very unconfident about which identification decision to make were more likely to be affected by the consequence likelihood information they were provided with. That is, while overall choosing was only around 8% higher in the guilty-person-released-more-likely condition than the innocent-person-convicted-more-likely condition, people who

indicated that they were less than 30% confident in making a decision showed a difference in choosing between the two negative consequence likelihood conditions that was around 23%. This finding supports the view that the non-significant trend in choosing that was observed in Experiment 4 was evidence that decisions people were not very confident in making had been affected by their consequence perceptions.

The overall difference in choosing between the negative consequence likelihood information conditions in the present experiment was found to be significant. Further, this difference in choosing translated to a significant difference in accuracy depending on whether or not the target was in the lineup. Target-present accuracy was higher in the guilty-person-released-more-likely (vs. the innocent-person-convicted more likely) condition and target-absent accuracy was higher in the innocent-person-convicted-more-likely (vs. the guilty-person-convicted-more-likely) condition.[46] On reflection, it might have been the case that the larger sample of participants run in this experiment led to a sufficient increase in power for the non-significant trend in choosing that was observed in Experiment 4 to reach significance. Additionally, it appears that tweaking the manipulation of people's consequence perceptions with new mock-newspaper articles may have strengthened it. While the difference in people's ratings of which mistake would be more likely to lead to negative consequences across the two levels of the manipulation in the present experiment was not notably higher than the previous one (*d*s 0.67 vs. 0.53, respectively), the difference in people's ratings of which mistake would be worse to make was large in Experiment 5 ($d = 0.87$) compared with medium ($d = 0.56$) in Experiment 4. It has been argued that evoking emotion can give rise to an increased feeling of risk and be a strong motivator of behaviour (Lowenstein, Weber, Hsee, & Welch, 2001). Therefore, it is possible that reading a first person account of

[46]Note that while the overall interaction was significant, the difference in target-absent accuracy between the two negative consequence likelihood conditions did not quite reach significance when analysed separately.

someone making an incorrect identification decision that had serious consequences increased people's feeling of that particular mistake being worse to make (despite not increasing their perceived likelihood of that mistake leading to negative consequences) and contributed to the overall effect of consequence perceptions on choosing behaviour reaching significance in this experiment.

Finally, the effect that the consequence perception manipulation had on people's post-decision ratings of confidence and willingness to testify in the previous experiment was replicated in this experiment. Namely, as in Experiment 4, it was found that confidence was higher and stated willingness to testify was more likely for lineup rejections than positive identifications in the condition where people were told that mistakenly choosing from a lineup was more likely to lead to negative consequences than mistakenly rejecting a lineup. In the condition where people were told the opposite, people's confidence and willingness to testify ratings were not found to differ depending on the decision they made. Although the influence of people's consequence perceptions on their evaluation of the decisions they made was not a key focus of this research, repeating the same finding across two experiments is nevertheless noteworthy. These results provide consistent evidence that perhaps consequence perceptions can affect people's approach towards their identification decisions (even when the decisions themselves are not necessarily affected).

In summary, the results from Experiment 5 provide tentative evidence that people's consequence perceptions can affect their willingness to make a positive identification decision from a lineup. The implications of these results are addressed in the following chapter.

CHAPTER 6

General Discussion

The experiments reported in this thesis comprise a number of attempts to investigate a potentially important and previously under-researched issue in the eyewitness memory literature: whether the consequences associated with identification decisions have the potential to bias people's willingness to choose from a lineup. As discussed in Chapter 1, a methodological framework that was both practical and sufficiently ecologically valid had not previously been established to investigate questions of this nature. Hence, the experiments carried out involved both testing basic theoretically driven (Green & Swets, 1966) predictions about how consequence perceptions might be expected to affect identification decisions and trialling different versions of a hypothetical consequences paradigm as a means for showing the predicted relationship. Unexpected difficulties were encountered with manipulating people's consequence perceptions in a way that would be expected to lead to a shift in decision making behaviour—necessitating several revisions to the initial methodology. Therefore, although the collective results from these experiments did not provide strong support for the main hypotheses put forward, various methodological issues provide plausible explanations for this and caution against the conclusion that consequence perceptions do not have the potential to affect identification decisions. Moreover, after addressing these methodological issues, tentative support for the expected relationship between consequence perceptions and identification decisions emerged. Nevertheless it is still largely unclear whether or to what extent consequence perceptions affect identification decisions as there are several possible interpretations of the results. In what follows, what can be gleaned from the experiments that were conducted is discussed, with a focus on potential future directions that research in this area could take.

The experiments reported collectively suggest that developing a good methodology for this area of research involves more considerations than were first envisaged. When first considering the viability of a hypothetical consequences paradigm, it seemed plausible (as argued in Chapter 1) that having people (a) focus on the consequences of real identification decisions and (b) imagine making their decisions as part of an actual criminal investigation might create a context where people would be compelled to shift their choosing behaviour to account for their consequence perceptions. However, when null results were obtained (in Experiment 3—the first experiment where the manipulation could have been reasonably expected to cause a shift in choosing behaviour) several potential problems with the hypothetical consequences method that was used were identified. In particular, it was proposed that in order for hypothetical consequences to affect identification decisions (a) people needed to be led to feel socially accountable for their decisions, instead of making essentially anonymous decisions, and (b) the task needed to be more similar to an actual identification procedure than the mini-lineup procedure that was used. Steps taken to fulfil these possible requirements for using a hypothetical consequences paradigm effectively—that is, a traditional crime video–photo lineup paradigm was used and people were told that at the end of the experiment they would learn what the likely outcome of an actual criminal investigation would have been based on their decisions—led to more promising indications that the manipulation of people's consequence perceptions was affecting their decision making behaviour. That is, in Experiment 4 people's post-decision confidence and willingness to testify ratings were affected by the manipulation. Further, a non-significant trend in choosing was evident, leading to the supposition that perhaps a small subset of decisions—namely those people were not confident in making—had been influenced by the consequence likelihood information. This interpretation of the results was supported by Experiment 5, where people's confidence in making a decision was measured and descriptive

results indicated that the difference in choosing between the two consequence perception conditions was greater for decisions associated with very low confidence. Therefore, the results from Experiments 4 and 5 (when compared with the results from Experiment 3) suggest that, in a context where identification decision consequences are only hypothetical, using a traditional video–lineup identification paradigm and (or[47]) leading people to feel socially accountable for their decisions is necessary for detecting a relationship between consequence perceptions and identification decisions.

However, what these results indicate in terms of the effect that consequence perceptions have on real identification decisions is unclear. One possibility is that the effect observed in Experiment 5 was weaker than what would be observed for identification decisions made in the context of a real criminal investigation and further modifications to the method would reflect this. For example, as mentioned in Chapter 4, it is possible that simulating a criminal investigation in a highly realistic manner might increase the likelihood that consequence perceptions would affect people's identification decisions. Under these circumstances they might be able to better imagine what it would be like being a witness to a crime and, therefore, role-play making an identification decision associated with real consequences more accurately. However, the potential benefits of using such a paradigm would need to be carefully considered and weighed against the decreased practicality associated with implementing it. A more convenient, and equally effective, option might be to increase people's feelings of accountability. Participants in Experiments 4 and 5 were told that at the end of the experiment they would learn what the likely outcome of an actual investigation would have been based on the identification decisions they made. People were told this so they would focus on the fact that the researcher would be aware of the decisions

[47] It could be the case that only one of these changes would have made a difference to the results, but as both the use of a traditional identification procedure and leading people to feel accountable for their decisions were introduced together this possibility cannot be verified.

that they made—something that has been linked with increased feelings of accountability (Lerner & Tetlock, 1999). As a result of this, people were expected to feel like they had to make justifiable decisions and be more motivated to account for their consequence perceptions in their decisions. Another way of leading people to feel accountable for their decisions is to explicitly tell them that they will have to justify their decisions to other people (Lerner & Tetlock, 1999). Intuitively, this approach seems like it might provide a stronger incentive for people to account for their consequence perceptions in their decisions, because the aim of making a justifiable decision appears to be more salient. Unfortunately, the level of accountability that was experienced by participants was not measured in the current experiments. It was not measured In Experiments 1–3 because the potential importance of leading people to feel accountable for their decisions was not (yet) under consideration. In Experiments 4 and 5 steps were taken to lead people to feel more responsible for their decisions but, perhaps because accountability was not being manipulated, it did not occur to me to measure it. In hindsight, failing to measure the level of accountability people experienced means that it is unclear whether there is scope for manipulating this construct more strongly. Further there is no point of comparison for determining whether different manipulations of accountability that might be used in future experiments are stronger than what was used in the current research. Thus, the potential benefits of tweaking people's perceptions of accountability are difficult to gauge, but nevertheless might be fruitful to investigate as they might increase the effect that people's consequence perceptions have on their identification decisions.

It is of course also entirely possible that to see any increase in the strength of the consequence perception–identification decision relationship (from what was observed in Experiment 5) people would need to think that their decisions could actually have the consequences being considered. Unfortunately, gaining a sense of what would occur in real

cases presents a significant challenge for future research as the problems associated with running an experiment where people think they are making real identification decisions are substantial (these are discussed at length in Chapter 1). But perhaps whether the degree to which consequence perceptions affect identification decisions is likely to differ depending on the realness of the consequences could be investigated in the following manner. An experiment could test whether artificial consequences (e.g., monetary) have a different effect on people's identification decisions depending on whether they are real or hypothetical (and maybe also large or small). If a difference in effect size across real and hypothetical payoffs was observed, then it could be tentatively inferred that actual identification decisions would be likely to be more strongly affected by consequence perceptions than mock identification decisions. Finding evidence to suggest that a hypothetical consequences paradigm leads to weaker effects than would be observed if the consequences were real would not necessarily mean that there is no value in using a hypothetical consequences paradigm to conduct research (particularly because of the difficulties associated with a "real consequences" paradigm). But it is clearly important to be aware of how the results are likely to generalise.

Another aspect of the methodology used that warrants scrutiny is the manipulation of people's consequence perceptions. Following two failed attempts to use a crime seriousness manipulation as a means for magnifying the effect of other biasing variables on people's choosing behaviour, influencing people's perceptions about whether mistakenly choosing from or rejecting a lineup would be more likely to lead to negative consequences emerged as a promising manipulation of people's consequence perceptions. Across two pilots and three experiments several revisions were made to strengthen this manipulation, including (a) the addition of a mock newspaper article to emphasise the consequences associated with the error portrayed as more serious (b) having people generate their own reasons for explaining why one mistake was more problematic to make than the other and (c) using a revised, more

emotionally evocative, version of the mock newspaper article. Overall, these revisions did appear to lead to a larger difference in people's perceptions of whether mistakenly picking from or rejecting a lineup would be worse. However, the effect on choosing behaviour that resulted from the tweaked manipulation was still not very impressive. Although it is clear from the preceding discussion that other factors could have contributed to these results, it is also worth considering the possibility that, despite the encouraging manipulation check results, the manipulation did not meaningfully influence people's consequence perceptions. As pointed out in the preamble to Experiment 1, whether mistakenly choosing from or rejecting a lineup is seen as worse is likely to be strongly tied to ideological beliefs (Carroll, Perkowitz, & Lurigio, 1987; Miller, 1973), that are rigid by definition (Gerring, 1997; Miller, 1973). Manipulating people's perception of which mistake was more likely to lead to a negative outcome was identified as a possible way to influence people's attitude towards which mistake would be more serious to make without attempting to change their belief about whether a wrongful conviction or a guilty person getting away is worse. However, this approach may have fallen short of influencing whether people were ultimately more concerned about mistakenly choosing from or rejecting a lineup. It is possible that people responded to the manipulation check question in line with what they had been explicitly told to think, without truly believing their response. Such demand effects are commonly observed in research when it is clear to people what behaviour is expected of them (Orne, 1962). If explicit consequence information is not actually influencing people's attitudes and a way to subtly manipulate people's consequence perceptions cannot be found,[48] then perhaps simply measuring, rather than trying to manipulate, people's consequence perceptions would be the

[48] The attempts to do this in Experiments 1 and 2 were unsuccessful, but that does not mean that it would be impossible to do so.

best way of examining the relation that they have to people's lineup choosing behaviour, despite the associated drawbacks with determining causality.

However, the idea that there is a stronger relationship between consequence perceptions and identification decisions to demonstrate is only one possible interpretation of the results. It is entirely possible that most of the experiments conducted did not show the hypothesised effects because consequence perceptions do not tend to affect identification decisions. The sole experiment showing the hypothesised effect used a manipulation that strongly suggested to participants which mistake they ought to avoid and still the difference in choosing rates was not very large between the two consequence information conditions. These results may indicate that even when the consequences of identification errors are explicitly considered (which they might not be unless prompted) they are relatively unlikely to affect a person's identification decision. Most people would presumably agree that making the error associated with the most serious consequences should be avoided, yet this consideration may not actually end up factoring into their identification decisions. A witness must compare all of the lineup members to their memory for the perpetrator, as well as to each other and work out whether this evidence warrants a positive identification or a lineup rejection (Clark, 2003). They have also been argued to come into the task with preconceptions about how likely it is that the target is actually present (Wells, 1993) and may also be attending to a number of other considerations. Therefore, it is possible that the effort invested in making the right choice exhausts the available mental resources the witness has and ends up preventing consequence perceptions associated with potential errors from factoring into the decision. After all, while often serious, the consequences associated with an incorrect identification decision eventuate in the distant future. They do not present a pressing matter and may not be prioritised for this reason. The degree to which the consideration of future consequences shape people's current behaviour (such as smoking, exercising or

studying) is something that has been associated with large individual differences (Strathman, Gleicher, Boninger, & Edwards, 1994). Similarly, it seems likely that only a subset of people are predisposed to consequence perceptions influencing their identification decisions. Further, those whose identification decision making is susceptible to being influenced by consequence perceptions would in many cases find no reason to factor this consideration into their decision because they are certain enough that the culprit either is or is not in the lineup.

Irrespective of which interpretation(s) of the results is correct, further research in this area would appear to benefit from making the following methodological change. Engineering the encoding, storage and (or) test conditions to make the identification decisions people are asked to make more difficult— similar to Kantner et al. (2015), described in Chapter 4— would presumably cause a larger number of people to be uncertain enough about whether to choose from or reject the lineup that their consequence perceptions would be more likely to affect their choosing behaviour. Therefore, a more targeted sample would be obtained, perhaps decreasing the number of participants that would need to be recruited to show the hypothesised effects. Further, these circumstances might arguably be more reflective of the conditions under which an actual witness would make an identification decision. For example, the retention interval in the present experiments was very short (5–10 minutes), a timeframe within which witnesses in a real life case would never make an identification test decision. Therefore, it is likely that given similar encoding and test conditions (which could also be much worse in a real case) actual witnesses would have a much harder time making an identification decision purely due to their memory having degraded over time. This would presumably lead them to feel more susceptible to making an error and result in a greater reliance on consequence perceptions in making their decisions. Within this framework other aspects of the methodology (discussed above) may be able to be improved.

Turning to broader implications of the current research, the fact that there are various possible interpretations of the results means that it would be premature to draw any firm conclusions from them. But the possibility that consequences can affect lineup choosing behaviour was somewhat encouraged and, therefore the potential implications of a relationship between consequence perceptions and identification decisions seems to be worth discussing. From a theoretical perspective, the pattern of people's responses in Experiment 5 was broadly consistent with what SDT (Green & Swets, 1966) suggested would occur. SDT holds that if the penalties associated with false positive responses were worse than those associated with false negative responses then people would require stronger evidence to make a positive identification decision than if those penalties were reversed. However, the results obtained do not actually tell us whether the consequence perception manipulation shifted the amount of evidence people required to make a positive identification, or whether perhaps some other mechanism was in operation. For example, it has been shown that people sometimes engage in what has been termed "biased pre-decision processing" (Brownstein, 2003), where forming an early inclination towards making a particular decision colours the interpretation of later evidence (in favour of making that decision). Therefore, it could be the case that people's consequence perceptions actually led people to distort their perception of the degree to which the lineup members were plausible matches to their memory for the target they were attempting to identify. It seems important for future research to clarify the mechanism that was operating. Nevertheless, finding some evidence that consequence perceptions can affect lineup choosing behaviour indicates that there is potentially an extra layer of complexity in the process by which eyewitnesses arrive at their (often incorrect) identification decisions. This finding supports the argument that the potential for people's cognitive and metacognitive judgements to affect their lineup choosing behaviour needs to be incorporated into existing theoretical accounts of identification decision making and be a part

of future research seeking to understand this decision process (Brewer et al., 2007). In particular, it has been argued that the interaction between different variables on identification decisions should be a focus of future research, to form a better understanding of when and how erroneous decisions occur (Brewer & Wells, 2011). In the current research, the potential for consequence perceptions to affect identification decisions appeared to be more pronounced for decisions people were not very confident in making. This possibility, as well as the potential for other factors to exacerbate the relationship, appears to be worth investigating further.

From an applied perspective, the possibility that people's negative consequence perceptions can bias their identification decisions highlights the potential importance of implementing procedures that are able to prevent this from occurring. For example, perhaps including a "don't know" option in the identification procedure would inhibit the effect that viewing a mistaken pick or rejection as worse could have on people's decisions. The results of Experiment 5 suggest that decisions people are not very confident making are most likely to be affected by biasing consequence perceptions. This is the same subset of cases for which people are likely to select a "don't know" option if it is available to them (Weber & Perfect, 2012). Providing people with an explicit "don't know" option (as distinct from advising them that they do not have to make a positive decision) has also been found to increase the accuracy of the positive identifications and lineup rejections that are made and is easy to include in an identification procedure (Weber & Perfect, 2012). Hence, there appear to be no drawbacks involved with implementing it to also protect against people's consequence perceptions affecting their decisions. The potential for consequence perceptions to bias people's propensity to choose from or reject a lineup also potentially endorses the use of alternate identification procedures, where information pertaining to whether people recognise anyone in a lineup is collected without requiring them to make a yes or no identification

decision. For example, Sauer et al. (2008) asked people to rate their confidence in each lineup member being the target and used an algorithm to classify the response pattern as a positive identification or a lineup rejection (see also, Brewer et al., 2012; Sauer, Weber, & Brewer, 2012). This method was found to be more accurate than a conventional (simultaneous or sequential) lineup procedure—particularly when the target was not present. If people's consequence perceptions affect their threshold for making a positive identification, then testing their recognition of the suspect in the manner just described would presumably prevent the consequence perceptions from affecting the evidence obtained. The potential for this procedure, as well as the potential for a "don't know" option to mitigate the influence of biasing consequence perceptions on identification decisions, points to research that may be fruitful to pursue following the clearer establishment of (a) whether and to what extent consequence perceptions affect identification decisions and (b) a suitable paradigm within which to conduct research on this topic.

References

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting—With and without satisficing knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1224-1245. doi: 10.1037/a0012938

Arkes, H. R., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, *39*, 133-144. doi: 10.1016/0749-5978(87)90049-5

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412. doi: 10.1016/j.jml.2007.12.005

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*, 1-2. doi: 10.3389/fpsyg.2013.00328

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi: 10.1016/j.jml.2012.11.001

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2016). Lme4: Linear mixed-effects models using Eigen and S4 classes. Retrieved from https://CRAN.R-project.org/package=lme4

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, *25*, 475-491. doi: 10.1023/A:1012840831846

Benjamin, A. S., Diaz, M., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychological Review*, *116*, 84-115. doi: 10.1037/a0014351

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13*, 103-124. doi: 10.1207/S15327965PLI1302_01

Bradfield, A. L., Wells, G. L., & Olson, E. A. (2002). The damaging effect of confirming feedback on the relation between eyewitness certainty and identification accuracy. *Journal of Applied Psychology, 87*, 112-120. doi: 10.1037/0021 9010.87.1.112

Brewer, N., & Palmer, M. A. (2010). Eyewitness identification tests. *Legal and Criminological Psychology, 15*, 77-96. doi: 10.1348/135532509X414765

Brewer, N., Weber, N., & Semmler, C. (2005). Eyewitness identification. In N. Brewer & K. D. Williams (Eds.), *Psychology and law: An empirical perspective* (pp. 177-221). New York: The Guilford Press.

Brewer, N., Weber, N., & Semmler, C. (2007). A role for theory in eyewitness identification research. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology: Memory for people* (Vol. 2, pp. 201-218). Mahwah, NJ: Erlbaum.

Brewer, N., Weber, N., Wootton, D., & Lindsay, D. S. (2012). Identifying the bad guy in a lineup using confidence judgments under deadline pressure. *Psychological Science, 23*, 1208-1214. doi: 10.1177/0956797612441217

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied, 12*, 11-30. doi: 10.1037/1076 898X.12.1.11

Brewer, N., & Wells, G. L. (2011). Eyewitness identification. *Current Directions in Psychological Science, 20,* 24-27. doi: 10.1021/nl200722z

Brigham, J. C., & Cairns, D. L. (1988). The effect of mugshot inspections on eyewitness identification accuracy. *Journal of Applied Social Psychology, 18,* 1394-1410. doi: 10.1111/j.1559-1816.1988.tb01214.x

Brigham, J. C., Maass, A., Martinez, D., & Wittenberger, G. (1983). The effect of arousal on facial recognition. *Basic and Applied Social Psychology, 4,* 279-293. doi: 10.1207/s15324834basp0403_6

Brownstein, A. L. (2003). Biased predecision processing. *Psychological Bulletin, 129,* 545-568. doi: 10.1037/0033-2909.129.4.545

Burke, M. S., Carter, J. R., Gominiak, R. D., & Ohl, D. F. (1996). An experimental note on the allais paradox and monetary incentives. *Empirical Economics, 21,* 617-632. doi: 10.1007/BF01180705

Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology, 14,* 88-96. doi: 10.1016/0022-1031(78)90062-8

Carroll, J. S., Perkowitz, W. T., Lurigio, A. J., & Weaver, F. M. (1987). Sentencing goals, causal attributions, ideology, and personality. *Journal of Personality and Social Psychology, 52,* 107-118. doi: 10.1037/0022-3514.52.1.107

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (Vol. 5, pp. 3-39). Hillsdale, NJ: Erlbaum.

Charman, S., & Wells, G. L. (2007). Applied lineup theory. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Handbook of eyewitness psychology: Memory for people* (Vol. 2, pp. 219-256). Mahwah, NJ: Erlbaum.

Christianson, S. Å. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin, 112,* 284-309. doi: 10.1037/0033-2909.112.2.284

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior. 12,* 335-359. doi: 10.1016/S0022-5371(73)80014-3

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

Clark, S. E. (2005). A re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior, 29,* 575-604. doi: 10.1007/s10979-005-7121-1

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Mahwah, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159. doi: 10.1037/0033-2909.112.1.155

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cohen-Hatton, S. R., & Honey, R. C. (2015). Goal-oriented training affects decision-making processes in virtual and simulated fire and rescue environments. *Journal of Experimental Psychology: Applied, 21*, 395-406. doi: 10.1037/xap0000061

Curran, T., DeBuse, C., & Leynes, P. A. (2007). Conflict and criterion setting in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 2-17. doi: 10.1037/0278-7393.33.1.2

Cutler, B. L., & Penrod, S. D. (1988). Improving the reliability of eyewitness identification: Lineup construction and presentation. *Journal of Applied Psychology, 73,* 281-290. doi: 10.1037/0021-9010.73.2.281

Dalal, D. K., & Zickar, M. J. (2012). Some common myths about centering predictor variables in moderated multiple regression and polynomial regression. *Organizational Research Methods, 15,* 339-362. doi: 10.1177/1094428111430540

Davis, M. H., Mitchell, K. V., Hall, J. A., Lothert, J., Snapp, T., & Meyer, M. (1999). Empathy, expectations, and situational preferences: Personality influences on the decision to participate in volunteer helping behaviors. *Journal of Personality, 67,* 469-503. doi: 10.1111/1467-6494.00062

Deffenbacher, K. A., Bornstein, B. H., & Penrod, S. D. (2006). Mugshot exposure effects: Retroactive interference, mugshot commitment, source confusion, and unconscious transference. *Law and Human Behavior, 30,* 287-307. doi: 10.1007/s10979-006 9008-1

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28,* 687-706. doi: 10.1007/s10979-004-0565-x

Dunning, D., & Stern, L. B. (1994). Distinguishing accurate from inaccurate eyewitness identifications via inquiries about decision processes. *Journal of Personality and Social Psychology, 67,* 818-835. doi: 10.1037/0022-3514.67.5.818

Eagly, A. H., & Chaiken, S. (1984). Cognitive theories of persuasion. *Advances in Experimental Social Psychology, 17,* 267-359. doi: 10.1016/S0065-2601(08)60122-7

Eastwick, P. W., Hunt, L. L., & Neff, L. A. (2013). External validity, why art thou externally valid? Recent studies of attraction provide three theoretical answers. *Social and Personality Psychology Compass, 7,* 275-288. doi: 10.1111/spc3.12026

Ebbesen, E. B., & Flowe, H. D. (2002). *Simultaneous v. sequential lineups: What do we really know?* Retrieved from https://dspace.lboro.ac.uk/2134/20167

Elms, A. C. (1966). Influence of fantasy ability on attitude change through role playing. *Journal of Personality and Social Psychology*, *4*, 36-43. doi: 10.1037/h0023509

Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision Making*, *18*, 199-212. doi: 10.1002/bdm.495

Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: A meta-analytic review of the 'weapon focus' literature. *Psychology, Crime & Law*, *19*, 35-66. doi: 10.1080/1068316X.2011.599325

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition*, *123*, 434-441. doi: 10.1016/j.cognition.2012.02.001

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: Sage Publication.

Fleet, M. L., Brigham, J. C., & Bothwell, R. K. (1987). The confidence-accuracy relationship: The effects of confidence assessment and choosing. *Journal of Applied Social Psychology*, *17*, 171-187. doi: 10.1111/j.1559-1816.1987.tb00308.x

Foster, R. A., Libkuman, T. M., Schooler, J. W., & Loftus, E. F. (1994). Consequentiality and eyewitness person identification. *Applied Cognitive Psychology*, *8*, 107-121. doi: 10.1002/acp.2350080203

Geller, D. M. (1978). Involvement in role-playing simulations: A demonstration with studies on obedience. *Journal of Personality and Social Psychology, 36,* 219-235. doi: 10.1037/0022-3514.36.3.219

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, MA: Cambridge University Press.

Gerring, J. (1997). Ideology: A definitional analysis. *Political Research Quarterly, 50,* 957-994. doi: 10.1177/106591299705000412

Gorenstein, G. W., & Ellsworth, P. C. (1980). Effect of choosing an incorrect photograph on a later identification by an eyewitness. *Journal of Applied Psychology, 65,* 616-622. doi: 10.1037/0021-9010.65.5.616

Greathouse, S. M., & Kovera, M. B. (2009). Instruction bias and lineup presentation moderate the effects of administrator knowledge on eyewitness identification. *Law and Human Behavior, 33,* 70-82. doi: 10.1007/s10979-008-9136-x

Green, D. S., & Swets, S. (1966). *Signal detection theory and psychophysics.* New York: J Wiley.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293,* 2105-2108. doi: 10.1126/science.1062872

Guerin, B. (1989). Reducing evaluation effects in mere presence. *The Journal of Social Psychology, 129,* 183-190. doi: 10.1080/00224545.1989.9711719

Hall, D. F., & Ostrom, T. M. (1976). Accuracy of eyewitness identification after biased or unbiased instructions. *Unpublished Manuscript,* University of Florida. (Paper presented at the meeting of the American Psychological Association, Chicago, August 1975).

Healy, A. F., & Kubovy, M. (1978). The effects of payoffs and prior probabilities on indices of performance and cutoff location in recognition memory. *Memory & Cognition, 6*, 544-553. doi: 10.3758/BF03198243

Healy, A. F., & Kubovy, M. (1981). Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 344-354. doi: 10.1037/0278 7393.7.5.344

Hope, L., Lewinski, W., Dixon, J., Blocksidge, D., & Gabbert, F. (2012). Witnesses in action: The effect of physical exertion on recall and recognition. *Psychological Science, 23*, 386-390. doi: 10.1177/0956797611431463

Horry, R., Halford, P., Brewer, N., Milne, R., & Bull, R. (2014). Archival analyses of eyewitness identification test outcomes: What can they tell us about eyewitness memory? *Law and Human Behavior, 38*, 94-108. doi: 10.1037/lhb0000060

Ihlebæk, C., Løve, T., Eilertsen, D. E., & Magnussen, S. (2003). Memory for a staged criminal event witnessed live and on video. *Memory, 11*, 319-327. doi: 10.1080/09658210244000018

Innocence Project (2016). *Innocence Project.* Retrieved November 2016, from http://www.innocenceproject.org

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446. doi: 10.1016/j.jml.2007.11.007

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior, 77*, 129-146. doi: 10.1901/jeab.2002.77-129

Kang, M. J., Rangel, A., Camus, M., & Camerer, C. F. (2011). Hypothetical and real choice differentially activate common valuation areas. *Journal of Neuroscience, 31*, 461-468. doi: 10.1523/JNEUROSCI.1583-10.2011

Kantner, J., Vettel, J. M., & Miller, M. B. (2015). Dubious decision evidence and criterion flexibility in recognition memory. *Frontiers in Psychology*, *6*, 1-18. doi: 10.3389/fpsyg.2015.01320

Kekessie, S. (2013). *The impact of the "not sure" option on lineup identification decisions* (Master's thesis, Florida International University, Florida, USA). Retrieved from http://digitalcommons.fiu.edu/etd/850.

Köhnken, G., & Maass, A. (1988). Eyewitness testimony: False alarms on biased instructions? *Journal of Applied Psychology*, *73*, 363-370. doi: 10.1037/0021 9010.73.3.363

Kühberger, A., Schulte-Mecklenbeck, M., & Perner, J. (2002). Framing decisions: Hypothetical and real. *Organizational Behavior and Human Decision Processes*, *89*, 1162-1175. doi: 10.1016/S0749-5978(02)00021-3

Leippe, M. R., Wells, G. L., & Ostrom, T. M. (1978). Crime seriousness as a determinant of accuracy in eyewitness identification. *Journal of Applied Psychology*, *63*, 345-351. doi: 10.1037/0021-9010.63.3.345

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125*, 255-275. doi: 10.1037/0033 2909.125.2.255

Levine, M., Cassidy, C., Brazier, G., & Reicher, S. (2002). Self-categorization and bystander non-intervention: Two experimental studies. *Journal of Applied Social Psychology*, *32*, 1452-1463. doi: 10.1111/j.1559-1816.2002.tb01446.x

Lindsay, R. C., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, *70*, 556-564. doi: 10.1037/0021-9010.70.3.556

Locey, M. L., Jones, B. A., & Rachlin, H. (2011). Real and hypothetical rewards in self control and social discounting. *Judgment and Decision Making*, *6*, 552-564.

Loewenstein, G. (2005). Hot-cold empathy gaps and medical decision making. *Health Psychology*, *24*, S49-S56. doi: 10.1037/0278-6133.24.4.S49

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267-286. doi: 10.1037/0033 2909.127.2.267

Loftus, E. F., & Burns, T. E. (1982). Mental shock can produce retrograde amnesia. *Memory & Cognition*, *10*, 318-323. doi: 10.3758/BF03202423

Maass, A., & Köhnken, G. (1989). Eyewitness identification: Simulating the "weapon effect". *Law and Human Behavior*, *13*, 397-408. doi: 10.1007/BF01056411

Maddox, W. T., & Bohil, C. J. (1998). Base-rate and payoff effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1459-1482. doi: 10.1037/0278-7393.24.6.1459

Mair, P., Wilcox, R., & Schoenbrodt, F. (2015). WRS2: A collection of robust statistical methods. Retrieved from https://CRAN.R-project.org/package=WRS2

Malpass, R. S., & Devine, P. G. (1980). Realism and eyewitness identification research. *Law and Human Behavior*, *4*, 347-358. doi: 10.1007/BF01040626

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

Malpass, R. S., Tredoux, C. G., & McQuiston-Surrett, D. (2007). Lineup construction and lineup fairness. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P.

Toglia (Eds.), *The handbook of eyewitness psychology: Memory for people* (Vol 2, pp. 155-178). Mahwah, NJ: Erlbaum.

Memon, A., Havard, C., Clifford, B., Gabbert, F., & Watt, M. (2011). A field evaluation of the VIPER system: A new technique for eliciting eyewitness evidence. *Psychology, Crime & Law, 17,* 711-729. doi: 10.1080/10683160903524333

Miller, W. B. (1973). Ideology and criminal justice policy: Some current issues. *The Journal of Criminal Law and Criminology*, *64*, 141-162. doi: 10.2307/1142984

Morgan, C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., ... & Southwick, S. M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry*, *27*, 265-279. doi: 10.1016/j.ijlp.2004.03.004

Murayama, K., Sakaki, M., Yan, V. X., & Smith, G. M. (2014). Type I error inflation in the traditional by-participant analysis to metamemory accuracy: A generalized mixed effects model perspective. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *40*, 1287- 1306. doi: 10.1037/a0036914

Murphy, J. J., Allen, P. G., Stevens, T. H., & Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics*, *30*, 313-325. doi: 10.1007/s10640-004-3332-z

Murray, D. M., & Wells, G. L. (1982). Does knowledge that a crime was staged affect eyewitness performance? *Journal of Applied Social Psychology*, *12*, 42-53. doi: 10.1111/j.1559-1816.1982.tb00847.x

National Research Council (2014). *Identifying the culprit: Assessing identification evidence.* Washington D.C.: National Academies Press. Retrieved from: http://www.nap.edu/catalog/18891/identifying-the-culprit-assessing-eyewitness-identification

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17,* 776-783. doi: 10.1037/h0043424

Palmer, M. A., & Brewer, N. (2012). Sequential lineup presentation promotes less-biased criterion setting but does not improve discriminability. *Law and Human Behavior, 36,* 247-255. doi: 10.1037/h0093923

Palmer, M. A., Brewer, N., & Weber, N. (2012). The information gained from witnesses' responses to an initial "blank" lineup. *Law and Human Behavior, 36,* 439-447. doi: 10.1037/h0093939

Patil, I., Cogoni, C., Zangrando, N., Chittaro, L., & Silani, G. (2014). Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. *Social Neuroscience, 9,* 94-107. doi: 10.1080/17470919.2013.870091

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49-69). Mahwah, NJ: Erlbaum.

Perfect, T. J., & Weber, N. (2012). How should witnesses regulate the accuracy of their identification decisions: One step forward, two steps back? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1810-1818. doi: 10.1037/a0028461

Petrova, D., Garcia-Retamero, R., Catena, A., & Van der Pligt, J. (2016). To screen or not to screen: What factors influence complex screening decisions? *Journal of Experimental Psychology: Applied, 22,* 247-260. doi: 10.1037/xap0000086

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion*: *Central and peripheral routes to attitude change.* New York: Springer-Verlag.

Pickel, K. L. (1998). Unusualness and threat as possible causes of "weapon focus". *Memory, 6,* 277-295. doi: 10.1080/741942361

Pickel, K. L. (1999). The influence of context on the "weapon focus" effect. *Law and Human Behavior, 23,* 299-311. doi: 10.1023/A:1022356431375

Pike, G., Brace, N., & Kynan, S. (2002). *The visual identification of suspects: Procedures and practice*. London, UK: Policing and Reducing Crime Unit, Home Office Research, Development and Statistics Directorate. doi: 10.1037/e665432007-001

Pozzulo, J. D., & Dempsey, J. (2006). Biased lineup instructions: Examining the effect of pressure on children's and adults' eyewitness identification accuracy. *Journal of Applied Social Psychology, 36,* 1381-1394. doi: 10.1111/j.0021 9029.2006.00064.x

Price, K. H. (1987). Decision responsibility, task responsibility, identifiability, and social loafing. *Organizational Behavior and Human Decision Processes, 40,* 330-345. doi: 10.1016/0749-5978(87)90020-3

Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language, 59,* 413–425. doi: 10.1016/j.jml.2008.02.002

R Development Core Team. (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reicher, S., & Levine, M. (1994). Deindividuation, power relations between groups and the expression of social identity: The effects of visibility to the out-group. *British Journal of Social Psychology, 33,* 145-163. doi: 10.1111/j.2044-8309.1994.tb01015.x

Robinson, M. D., & Johnson, J. T. (1998). How not to enhance the confidence–accuracy relation: The detrimental effects of attention to the identification process. *Law and Human Behavior, 22,* 409-428. doi: 10.1023/A:1025770926338

Rovira, A., Swapp, D., Spanlang, B., & Slater, M. (2009). The use of virtual reality in the study of people's responses to violent incidents. *Frontiers in Behavioral Neuroscience*, *3*, 1-10. doi: 10.3389/neuro.08.059.2009

Sanders, G. S., & Warnick, D. H. (1981). Truth and consequences: The effect of responsibility on eyewitness behavior. *Basic and Applied Social Psychology*, *2*, 67-79. doi: 10.1207/s15324834basp0201_6

Sauer, J. D., Brewer, N., & Weber, N. (2008). Multiple confidence estimates as indices of eyewitness memory. *Journal of Experimental Psychology: General*, *137*, 528-547. doi: 10.1037/a0012712

Sauer, J. D., Weber, N., & Brewer, N. (2012). Using ecphoric confidence ratings to discriminate seen from unseen faces: The effects of retention interval and distinctiveness. *Psychonomic Bulletin & Review*, *19*, 490-498. doi: 10.3758/s13423 012-0239-5

Schopler, J., Insko, C. A., Drigotas, S. M., Wieselquist, J., Pemberton, M. B., & Cox, C. (1995). The role of identifiability in the reduction of interindividual-intergroup discontinuity. *Journal of Experimental Social Psychology*, *31*, 553-574. doi: 10.1006/jesp.1995.1025

Schwartz, D., Bruine de Bruin, W., Fischhoff, B., & Lave, L. (2015). Advertising energy saving programs: The potential environmental cost of emphasizing monetary savings. *Journal of Experimental Psychology: Applied, 21*, 158-166. doi: 10.1037/xap0000042

Simonson, I., & Nye, P. (1992). The effect of accountability on susceptibility to decision errors. *Organizational Behavior and Human Decision Processes, 51,* 416-446. doi: 10.1016/0749-5978(92)90020-8

Skinner, B. F. (1953). *Science and human behaviour.* New York: Macmillan.

Stanny, C. J., & Johnson, T. C. (1999). Effects of stress induced by a simulated shooting on recall by police and citizen witnesses. *The American Journal of Psychology, 113,* 359-386. doi: 10.2307/1423364

Steblay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior, 16,* 413-424. doi: 10.1007/BF02352267

Steblay, N. M. (1997). Social influence in eyewitness recall: A meta-analytic review of lineup instruction effects. *Law and Human Behavior, 21,* 283-297. doi: 10.1023/A:1024890732059

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 25,* 459-473. doi: 10.1023/A:1012888715007

Steblay, N. K., Dysart, J. E., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychology, Public Policy, and Law, 17,* 99-139. doi: 10.1037/a0021650

Stevens, T., Brevers, D., Chambers, C. D., Lavric, A., McLaren, I. P., Mertens, M., ... & Verbruggen, F. (2015). How does response inhibition influence decision making when gambling? *Journal of Experimental Psychology: Applied, 21,* 15-36. doi: 10.1037/xap0000039

Strathman, A., Gleicher, F., Boninger, D. S., & Edwards, C. S. (1994). The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology, 66,* 742-752. doi: 10.1037/0022-3514.66.4.742

Sussman, A. B., Sharma, E., & Alter, A. L. (2015). Framing charitable donations as exceptional expenses increases giving. *Journal of Experimental Psychology: Applied, 21,* 130-139. doi: 10.1037/xap0000047

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). New Jersey: Pearson Education.

Taylor, M. P. (2013). Bias and brains: Risk aversion and cognitive ability across real and hypothetical settings. *Journal of Risk and Uncertainty*, *46*, 299-320. doi: 10.1007/s11166-013-9166-8

Thompson, J. (2000, June 18). 'I was certain, but I was wrong'. *The New York Times.* Retrieved from http://www.nytimes.com/2000/06/18/opinion/i-was-certain-but-i-was wrong.html

Thompson, W. B., & Johnson, J. (2008). Biased lineup instructions and face identification from video images. *The Journal of General Psychology*, *135*, 23-36. doi: 10.3200/GENP.135.1.23-36

Tollestrup, P. A., Turtle, J. W., & Yuille, J. C. (1994). Actual victims and witnesses to robbery and fraud: An archival analysis. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 144–162). Cambridge, UK: Cambridge University.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207-232. doi: 10.1016/0010-0285(73)90033 9

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453–458. doi: 10.1126/science.7455683

Ulehla, Z. J. (1966). Optimality of perceptual decision criteria. *Journal of Experimental Psychology, 71*, 564-569. doi: 10.1037/h0023007

Vieider, F. M. (2011). Separating real incentives and accountability. *Experimental Economics*, *14*, 507-518. doi: 10.1007/s10683-011-9279-3

Vlaev, I. (2012). How different are real and hypothetical decisions? Overestimation, contrast and assimilation in social interaction. *Journal of Economic Psychology*, *33*, 963-972. doi: 10.1016/j.joep.2012.05.005

Von Winterfeldt, D., & Edwards, W. (1982). Costs and payoffs in perceptual research. *Psychological Bulletin*, *91*, 609-622. doi: 10.1037/0033-2909.91.3.609

Warnick, D. H., & Sanders, G. S. (1980). Why do eyewitnesses make so many mistakes? *Journal of Applied Social Psychology*, *10*, 362-366. doi: 10.1111/j.1559 1816.1980.tb00716.x

Watts, W. A. (1967). Relative persistence of opinion change induced by active compared to passive participation. *Journal of Personality and Social Psychology*, *5*, 4-15. doi: 10.1037/h0021196

Weber, N., & Brewer, N. (2006). Positive versus negative face recognition decisions: Confidence, accuracy, and response latency. *Applied Cognitive Psychology*, *20*, 17-31. doi: 10.1002/acp.1166

Weber, N., & Perfect, T. J. (2012). Improving eyewitness identification accuracy by screening out those who say they don't know. *Law and Human Behavior*, *36*, 28-36. doi: 10.1037/h0093976

Weber, N., & Varga, M. (2012). Can a modified lineup procedure improve the usefulness of confidence? *Journal of Applied Research in Memory and Cognition*, *1*, 152-157. doi: 10.1016/j.jarmac.2012.06.007

Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology*, *14*, 89-103. doi: 10.1111/j.1559-1816.1984.tb02223.x

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist*, *48*, 553-571. doi: 10.1037/0003-066X.48.5.553

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

Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist, 55,* 581-598. doi: 10.1037/0003-066X.55.6.581

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology, 54,* 277-295. doi: 10.1146/annurev.psych.54.101601.145028

Wells, G. L., & Quinlivan, D. S. (2009). Suggestive eyewitness identification procedures and the Supreme Court's reliability test in light of eyewitness science: 30 years later. *Law and Human Behavior, 33,* 1-24. doi: 10.1007/s10979-008-9130-3

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22,* 603-647. doi: 10.1023/A:1025750605807

Wells, G. L., Steblay, N. K., & Dysart, J. E. (2012). Eyewitness identification reforms: Are suggestiveness-induced hits and guesses true hits? *Perspectives on Psychological Science, 7,* 264-271. doi: 10.1177/1745691612443368

Wells, G. L., & Turtle, J. W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin, 99,* 320-329. doi: 10.1037/0033-2909.99.3.320

Wilcox, R. R. (1995a). ANOVA: A paradigm for low power and misleading measures of effect size? *Review of Educational Research, 65,* 51-77. doi: 10.3102/00346543065001051

Wilcox, R. R. (1995b). Three multiple comparison procedures for trimmed means. *Biometrical Journal, 37,* 643-656. doi: 10.1002/bimj.4710370602

Wilcox, R. R. (2012). *Introduction to robust estimation and hypothesis testing* (3rd ed.). Waltham, MA: Academic Press.

Williams, K., Harkins, S. G., & Latané, B. (1981). Identifiability as a deterrant to social loafing: Two cheering experiments. *Journal of Personality and Social Psychology*, *40*, 303-311. doi: 10.1037/0022-3514.40.2.303

Wilson VanVoorhis, C. R., & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology*, *3*, 43-50. doi: 10.20982/tqmp.03.2.p043

Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review*, *121*, 262-276. doi: 10.1037/a0035940

Wright, W. F., & Anderson, U. (1989). Effects of situation familiarity and financial incentives on use of the anchoring and adjustment heuristic for probability assessment. *Organizational Behavior and Human Decision Processes*, *44*, 68-82. doi: 10.1016/0749-5978(89)90035-6

Wright, D. B., & McDaid, A. T. (1996). Comparing system and estimator variables using data from real line-ups. *Applied Cognitive Psychology*, *10*, 75-84. doi: 10.1002/(SICI)1099-0720(199602)10:175::AID-ACP3643.3.CO;2-5

Zajonc, R. B., & Sales, S. M. (1966). Social facilitation of dominant and subordinate responses. *Journal of Experimental Social Psychology*, *2*, 160-168. doi: 10.1016/0022-1031(66)90077-1

Appendix A

Inferential Statistics for the Non-Parametric Manipulation Check Tests Conducted

Table A1

Inferential Statistics for the Non-Parametric Manipulation Check Tests Run in Experiment 1

Analyses testing the effect of case outcome on ratings of:	Test statistic	Significance
Likelihood that justice was served	$T_y(34.65) = 4.32$	$p < .001$
Likely suspect guilt	$T_y(34.96) = 6.32$	$p < .001$
Worry about implicating an innocent suspect	$T_y(32.57) = 1.43$	$p = .16$

Table A2

Inferential Statistics for the Non-Parametric Manipulation Check Tests Run in Experiment 2

Analyses testing the effect of case outcome on ratings of:	Test statistic	Significance
Likelihood that justice was served	$T_y(71.74) = 9.21$	$p < .001$
Likely suspect guilt	$T_y(64.76) = 11.98$	$p < .001$
Worry about implicating an innocent suspect	$T_y(32.57) = 1.43$	$p = .16$
Likelihood that choosing leads to an innocent suspect conviction	$T_y(68.12) = 2.66$	$p = .01$

Table A3

Inferential Statistics for the Non-Parametric Manipulation Check Tests Run in Experiment 3 Pilot 1

Effect being tested	Test statistic	Significance
The interaction between negative consequence likelihood information × type of mistake on ratings of negative consequence likelihood		
Main effect of negative consequence likelihood information	*Bwtrim* = 2.97	$p = .10$
Main effect of mistake type	*Bwtrim* = 8.13	$p = .01$
Interaction	*Bwtrim* = 4.57	$p = .04$
The effect of mistake type on ratings of negative consequence likelihood in the:		
Innocent-person-convicted-more-likely condition	$T_y(17) = -0.80$	$p = .43$
Guilty-person-released-more-likely condition	$T_y(17) = -2.79$	$p = .01$

Table A4

Inferential Statistics for the Non-Parametric Manipulation Check Tests Run in Experiment 3 Pilot 2

Effect being tested	Test statistic	Significance
The interaction between negative consequence likelihood information × type of mistake on ratings of negative consequence likelihood		
Main effect of negative consequence likelihood information	*Bwtrim* = 4.21	$p = .06$
Main effect of mistake type	*Bwtrim* = 14.90	$p < .001$
Interaction	*Bwtrim* = 23.77	$p < .001$
The effect of mistake type on ratings of negative consequence likelihood in the:		
Innocent-person-convicted-more-likely condition	$T_y(17) = 0.75$	$p = .46$
Guilty-person-release-more-likely condition	$T_y(17) = -5.95$	$p < .001$
The effect of negative consequence likelihood information on ratings of which mistake would be worse to make	$T_y(29.9) = 6.21$	$p < .001$

Table A5

Inferential Statistics for the Non-Parametric Manipulation Check Tests Run in Experiment 3

Effect being tested	Test statistic	Significance
The interaction between negative consequence likelihood information × type of mistake on ratings of negative consequence likelihood		
Main effect of negative consequence likelihood information	*Bwtrim* = 0.03	$p = .87$
Main effect of mistake type	*Bwtrim* = 2.75	$p = .10$
Interaction	*Bwtrim* = 5.60	$p = .02$
The effect of mistake type on ratings of negative consequence likelihood in the:		
Innocent-person-convicted-more-likely condition	$T_y(39) = -2.55$	$p = .01$
Guilty-person-release-more-likely condition	$T_y(39) = 0.58$	$p = .57$
The effect of negative consequence likelihood information on ratings of which mistake would be worse to make	$T_y(75.42) = 2.72$	$p = .01$

Table A6

Inferential Statistics for the Non-Parametric Manipulation Check Tests Run in Experiment 4

Effect being tested	Test statistic	Significance
The interaction between negative consequence likelihood information × type of mistake on ratings of negative consequence likelihood		
Main effect of negative consequence likelihood information	*Bwtrim* = 1.09	$p = .30$
Main effect of mistake type	*Bwtrim* = 7.98	$p = .01$
Interaction	*Bwtrim* = 41.14	$p < .001$
The effect of mistake type on ratings of negative consequence likelihood in the:		
Innocent-person-convicted-more-likely condition	$T_y(105) = 2.85$	$p = .01$
Guilty-person-release-more-likely condition	$T_y(105) = -5.95$	$p < .001$
The effect of negative consequence likelihood information on ratings of which mistake is more likely to lead to negative consequences	$T_y(184.1) = 5.65$	$p < .001$
The effect of negative consequence likelihood information on ratings of which mistake would be worse to make	$T_y(167.37) = 4.19$	$p < .001$

Table A7

Inferential Statistics for the Non-Parametric Manipulation Check Tests Run in Experiment 5

Effect being tested	Test statistic	Significance
The interaction between negative consequence likelihood information × type of mistake on ratings of negative consequence likelihood		
Main effect of negative consequence likelihood information	*Bwtrim* = 2.93	$p = .09$
Main effect of mistake type	*Bwtrim* = 30.13	$p < .001$
Interaction	*Bwtrim* = 66.25	$p < .001$
The effect of mistake type on ratings of negative consequence likelihood in the:		
Innocent-person-convicted-more-likely condition	$T_y(153) = 1.81$	$p = .07$
Guilty-person-release-more-likely condition	$T_y(153) = -10.00$	$p < .001$
The effect of negative consequence likelihood information on ratings of which mistake is more likely to lead to negative consequences	$T_y(298.43) = 6.14$	$p < .001$
The effect of negative consequence likelihood information on ratings of which mistake would be worse to make	$T_y(305.93) = 10.00$	$p < .001$

Appendix B

Random Effects (RE) Structures Used in all Mixed-Effects Models Reported in the Main Body of This Thesis

Table B1

Code Number Assigned to and Description of the Systematic Random Effects Structure Simplifications Employed When Constructing the Mixed-Effects Models Used to Test the Main Hypotheses in the Experiments Described in This Thesis

Code	Description of the random effects structure
1	Fully maximal model with random intercepts for participant and stimulus, as well as appropriate random slopes for all variables justified by the design (i.e., including variables for which fixed effects were not being tested—note that design variables were not included if they led to more than a 3way interaction in the RE structure but control variables were always attempted to be included when relevant)
2	Both random intercepts and all critical random slopes (i.e., when interaction by unit slopes were appropriate the associated main effect by unit slopes were removed)
3a & b	Both random intercepts and random slopes with correlations removed (CR) (a) with all applicable random slopes (b) with critical random slopes only
4a & b	As above but applicable control variable removed (a) with correlations removed (b) without correlations removed
5a & b	Both random intercepts and critical interaction random slope only (a) without correlations removed (b) with correlations removed
6	Critical interaction random slope (with correlations removed) and one or both random intercepts (i.e., those that have within-unit slopes) removed
7	Random intercepts only.

Table B2

Random Effects Structures for Mixed-Effects Models Run to Test Hypotheses in Experiment 1

Mixed-effects model	Random effects structure	Code
Case outcome × crime seriousness on choosing	Random intercepts: Participant and stimulus Random slopes (CR): Crime seriousness | participant; crime seriousness × case outcome | stimulus	3b
Case outcome × target presence on accuracy	Random intercepts: Participant and stimulus Random slopes (CR): Target presence | participant; target presence × case outcome | stimulus	3b
Case outcome × lineup decision on post-decision confidence	Random intercepts: Participant and stimulus Random slopes (CR): target presence | participant; lineup decision | participant; target presence | stimulus; case outcome | stimulus; lineup decision | stimulus; case outcome × lineup decision | stimulus	3a

Table B3

Random Effects Structures for Mixed-Effects Models Run to Test Hypotheses in Experiment 2

Mixed-effects model	Random effects structure	Code
Crime seriousness code on justice importance	Random intercepts: Participant and stimulus Random slopes (CR): Crime seriousness code | participant	1
Case outcome × crime seriousness on choosing	Random intercepts: Participant and stimulus Random slopes (CR): Crime seriousness | participant; case outcome × crime seriousness | stimulus	3b
Case outcome × target presence on accuracy	Random intercepts: Participant and stimulus Random slopes (CR): target presence | participant; case outcome × target presence | stimulus	3b
Case outcome × lineup decision on post-decision confidence	Random intercepts: Participant and stimulus Random slopes (CR): target presence | participant; lineup decision | participant; target presence | stimulus; lineup decision × case outcome | stimulus	3b
Case outcome × lineup decision on willingness to testify	Random intercepts: Participant and stimulus	7

Table B4

Random Effects Structures for Mixed-Effects Models Run to Test Hypotheses in Experiment 3

Mixed-effects model	Random effects structure	Code
Crime seriousness code on justice importance	Random intercepts: Participant and stimulus Random slopes: Crime seriousness code | participant	1
Negative consequence likelihood information × crime seriousness on choosing	Random intercepts: None Random slopes(CR): crime seriousness | participant; negative consequence likelihood information × crime seriousness | stimulus	6
Negative consequence likelihood information × target presence on accuracy	Random intercepts: Participant and stimulus Random slopes (CR): target presence | participant; negative consequence likelihood information × target presence | stimulus	3b
Negative consequence likelihood information × lineup decision on post-decision confidence	Random intercepts: Participant and stimulus Random slopes (CR): target presence | participant; lineup decision | participant; target presence | stimulus; negative consequence likelihood information | stimulus; lineup decision | stimulus; negative consequence likelihood information × lineup decision | stimulus	3a

Table B5

Random Effects Structures for Mixed-Effects Models Run to Test Hypotheses in Experiment 4

Mixed-effects model	Random effects structure	Code
Negative consequence likelihood information on choosing	Random intercepts: Participant and stimulus Random slopes (CR): Negative consequence likelihood information | stimulus; target presence | stimulus; negative consequence likelihood information × target presence | stimulus	3a
Negative consequence likelihood information × target presence on accuracy	Random intercepts: Participant Random slopes (CR): Negative consequence likelihood information × target presence | stimulus	6
Negative consequence likelihood information × lineup decision on post-decision confidence	Random intercepts: Participant and stimulus Random slopes (CR): Target presence | stimulus; negative consequence likelihood information | stimulus; lineup decision | stimulus; negative consequence likelihood information × lineup decision | stimulus	3a
Lineup decision on post-decision confidence in the innocent-person-convicted-more-likely condition	Random intercepts: Participant and stimulus Random slopes (CR): Target presence | stimulus; lineup decision | stimulus	3a
Lineup decision on post-decision confidence in the guilty-person-released-more-likely condition	Random intercepts: Participant and stimulus Random slopes (CR): Target presence | stimulus, lineup decision | stimulus	3a
Negative consequence likelihood information × lineup decision on willingness to testify	Random intercepts: Participant and stimulus Random slopes (CR): Negative consequence likelihood information | stimulus; lineup decision | stimulus; negative consequence likelihood information × lineup decision | stimulus	4a
Lineup decision on willingness to testify in the innocent-person-convicted-more-likely condition	Random intercepts: Participant and stimulus Random slopes (CR): lineup decision | stimulus	4a
Lineup decision on willingness to testify in the guilty-person-released-more-likely condition	Random intercepts: Participant Random slopes (CR): lineup decision | stimulus	4a

Table B6

Random Effects Structures for Mixed-Effects Models Run to Test Hypotheses in Experiment 5

Mixed-effects model	Random effects structure	Code
Negative consequence likelihood information on choosing	Random intercepts: Participant and stimulus Random slopes: Negative consequence likelihood information | stimulus; target presence | stimulus; negative consequence likelihood information × target presence | stimulus	1
Negative consequence likelihood information × target presence on accuracy	Random intercepts: Participant and stimulus Random slopes (CR): Negative consequence likelihood information × target presence | stimulus	3b
Negative consequence likelihood information on accuracy for target present cases	Random intercepts: Participant and stimulus Random slopes: Negative consequence likelihood information | stimulus	1
Negative consequence likelihood information on accuracy for target absent cases	Random intercepts: Participant and stimulus Random slopes: Negative consequence likelihood information | stimulus	1
Negative consequence likelihood information on choosing after removing each descending level of initial confidence	Random intercepts: Participant (all) and stimulus (except for when initial confidence < 10%) Random slopes: Target presence | stimulus; negative consequence likelihood information | stimulus (all)	-
Negative consequence likelihood information × initial confidence on choosing	Random intercepts: Participant and stimulus Random slopes: target presence | stimulus; negative consequence likelihood information | stimulus; initial confidence | stimulus; negative consequence likelihood information × initial confidence | stimulus	1
Negative consequence likelihood information × lineup decision on post-decision confidence	Random intercepts: Participant and stimulus Random slopes (CR): Target presence | stimulus; negative consequence likelihood information | stimulus; lineup decision | stimulus; negative consequence likelihood information × lineup decision | stimulus	3a
Lineup decision on post-decision confidence for the innocent-person-convicted-more-likely condition	Random intercepts: Participant and stimulus Random slopes (CR): Target presence | stimulus; lineup decision | stimulus	3a
Lineup decision on post-decision confidence for the guilty-person-released-more-likely condition	Random intercepts: Participant and stimulus Random slopes (CR): Target presence | stimulus; lineup decision | stimulus	3a

Negative consequence likelihood information × lineup decision on willingness to testify	Random intercepts: Participant and stimulus Random slopes: Negative consequence likelihood information × lineup decision | stimulus	3b
Lineup decision on willingness to testify for the innocent-person-convicted-more-likely condition	Random intercepts: Participant and stimulus	7
Lineup decision on willingness to testify for the guilty-person-released-more-likely condition	Random intercepts: Participant and stimulus Random slopes (CR): Lineup decision | stimulus	4a