The alibi-generation effect: Alibi-generation experience influences alibi evaluation

Elizabeth A. Olson\(^1\)* and Gary L. Wells\(^2\)
\(^1\)University of Wisconsin-Whitewater, Wisconsin, USA
\(^2\)Iowa State University, Iowa, USA

**Purpose.** Post-conviction DNA exonerations demonstrate a failure of alibis to protect innocent suspects. We contend one reason alibis are not believed is because evaluators underestimate how difficult it is for an innocent person to generate a convincing alibi. We hypothesized that asking evaluators to first generate an alibi of their own would lead evaluators to consider a suspect's alibi as more believable.

**Methods.** Participants (\(N = 147\)) were randomly assigned to either evaluate a suspect's alibi before generating their own alibi (evaluate-first) or generate their own alibi before evaluating the suspect's alibi (generate-first). Participants provided alibis from either 3 days previous or 30 days previous. In Experiment 2, participants (\(N = 255\)) were randomly assigned to either generate-first, evaluate-first, or read-experience (in which they read about alibi-generation difficulty) conditions. Half the participants were primed to think empathetically with Interpersonal Reactivity Index (IRI) subscales, and half were not. All participants evaluated the believability of the suspect's alibi as well as their own alibis.

**Results.** Across both experiments, participants who generated their own alibi first rated the suspect's alibi as more believable. This alibi-generation effect overshadowed alibi latency in Experiment 1 and the empathy manipulation in Experiment 2.

**Conclusions.** Alibi-generation experience seems to change the expectations evaluators have of alibis from criminal suspects. This effect likely emanates from increased awareness that alibis are difficult for innocent people to generate and from the emotional experience of having difficulty generating alibis.

Since 1989, forensic DNA testing has exonerated more than 250 people – after they wrongfully served a combined total of over 3,200 years in prison (Innocence Project, 2010). Whereas these convictions were largely due to mistaken eyewitness identification (Scheck, Neufeld, & Dwyer, 2000; Wells et al., 1998), a common question asked by the general public is, ‘If they were innocent, why couldn’t they just give their alibis and get out of it?’

\*Correspondence should be addressed to Elizabeth A. Olson, Department of Psychology, 800 W. Main St., Whitewater, WI 53190, USA (e-mail: olsone@uww.edu).

DOI:10.1111/j.2044-8333.2010.02003.x
Of course, innocent people may not have an alibi – they may not remember exactly where they were or what they had been doing. However, many of the wrongfully convicted bad alibis, suggesting a dramatic failure of alibi evidence to protect innocent people from prosecution and conviction. For example, Marvin Anderson offered four alibi witnesses to support his contention that he was washing his car at the time of the crime, and Stephen Avery presented 16 alibi witnesses to support his account of his whereabouts. Neither alibi defence was enough to prevent these men from being wrongfully convicted (Findley & Scott, 2006). The current work suggests one reason their alibis were not believed is that people (including jurors) underestimate the difficulty of an innocent person providing a convincing alibi even though the person is telling the truth. Moreover, we hypothesize that making people generate a truthful alibi for their own whereabouts from some time in the past can give them an appreciation for this problem that, in turn, makes them more accepting of someone else’s alibi.

There is some evidence in the literature that alibis are met with considerable scepticism and that it is difficult for innocent people to generate convincing, provable alibis for times in the past. We review literature that supports these two contentions and then describe our conceptualization of getting people to appreciate the difficulty of generating alibis.

**Believability of alibi evidence**

How believable an evaluator considers an alibi to be is dependent upon a number of factors - the context in which the evaluator sees the alibi, the characteristics of the alibi provider, the narrative details of the alibi itself, to name a few - and our understanding of how these factors interact is still rudimentary (see Burke, Turtle, & Olson, 2007; Sommers & Douglass, 2007). Olson and Wells (2004) attempted to capture a basic underlying component of alibi believability by examining one thing common to all alibis: the corroborating evidence. They divided corroborating evidence into two categories, person corroborators and physical evidence, and from there developed a taxonomy to classify evidence according to perceived believability. For example, a person corroborator who was perceived to have a motivation to lie for the alibi provider – such as a suspect’s close friend – would be perceived as less believable than a corroborator without such a relationship – such as a store clerk who was a complete stranger. Olson and Wells (2004) found that, as predicted by the taxonomy, alibis with higher levels of person and physical evidence were considered more believable; however, they noted that there remained a profound resistance to believe even the strongest alibis: the alibi with the suspect captured on a security video and a complete stranger corroborating the alibi garnered only a 7.4 on a 10-point belief rating scale (Olson & Wells, 2004).

One of the most common, and influential, components of alibi evidence is the person who corroborates an alibi – as social beings, we tend to spend our days around other people who could ostensibly testify to our whereabouts. Unfortunately, for innocent alibi providers, the people we spend time with, such as friends and relatives, are also considered the least believable. Culhane and Hosch (2004) manipulated the relationship of the alibi corroborator to their hypothetical defendant – the alibi witness was either a girlfriend or a neighbour who either expressed certainty or uncertainty in corroborating the alibi. Lindsay, Lim, Marando, and Cully (1986) manipulated the relationship between the defendant and the alibi witness. Both studies found that conviction rates dropped only
when the corroborator had no relationship to the defendant. Mock jurors seem especially disbelieving of a corroborator if they perceive that the relationship is a highly committed one (Jolly, Simonds, & Hosch, 2009). In the taxonomy research, Olson and Wells (2004) found that the alibis with a friend or relative corroborator were not considered more believable than alibis in which the alibi provider had no corroborating evidence at all.

The legal literature dealing with alibis suggests a healthy scepticism towards alibis – historically known as the ‘rogue’s defence’ because of a perceived ease of fabricating an alibi (Gooderson, 1977). Alibis as a defence in a trial are supposed to be considered evidence that would cast reasonable doubt on a suspect’s guilt (see *P v. Victor*, 1984); however, alibis often become something to be proved, in and of themselves, and mistakes or falsifications of alibi evidence may sometimes be used as an evidence of guilt for the crime, in spite of admonitions that they should not (Connelly, 1983; Wells *et al.*, 1998).

**Difficulty of alibi generation**

Research into the generation of alibis supports the idea that alibi generation can be a difficult process for innocent people (Culhane, Hosch, & Kehn, 2008; Olson & Charman, in press; Strange, Dysart, & Loftus, 2010). Olson and Charman (in press) found that innocent participants were largely willing to provide alibi stories from both a near-past date (3 days prior) and a distant-past date (6–14 weeks prior). However, after participants investigated their alibis, the majority of alibis offered fell into the weakest categories of Olson and Wells’ (2004) taxonomy: 81% of participants’ alibis had no or very weak physical or person evidence. Very few alibis were supported by compelling, convincing evidence: only 9% of alibis had strong physical evidence, and only 6% were corroborated by a non-motivated other (Olson & Charman, in press). Strange *et al.* (2010) similarly asked participants for a detailed memory account and sent them home to investigate their alibis. As in Olson and Charman (in press), very few had alibis supported by strong physical or person evidence, and fewer than half of the alibis were rated as consistent in details from the first account to the second (Strange *et al.*, 2010). Lastly, although Culhane *et al.* (2008) did not ask participants to investigate their alibis, most participants reported alibis that could only be corroborated by relatives or close friends.

Autobiographical memory research also suggests that the act of generating an alibi story can be a difficult task for an innocent person. People tend not to retain detailed memories for routine days or tasks (Brewer, 1988), and latency between an event and recall of that event can significantly degrade memory (see Schacter, 1999). Additionally, alibi providers might mistakenly report typical behaviour without recognizing that they may have done something atypical on that particular day (Olson & Charman, in press), or they might erroneously combine elements from different events (i.e., a memory conjunction error; see Lampinen, Faries, Neuschatz, & Toglia, 2000), leaving them vulnerable to a lack of alibi evidence when they discover their alibi is mistaken.

Kassam, Gilbert, Swencionis, and Wilson (2009) demonstrated that evaluators of someone else’s memory expect that motivation to remember an event should improve memory. Unfortunately for innocent alibi providers, motivation at the time of retrieval does not improve memory for information that was not well-encoded in the first place.
This could leave alibi providers in a psychologically difficult position, feeling both motivated to remember and unable to do so.

Alibi-generation experience and judgments about a suspect’s alibi

If evaluators expect that alibis should be easy to generate, they would likewise have high expectations about how strong an alibi should be in order to declare a suspect innocent. Similar to the actor-observer effect, in which observers rely on different types of information to make judgments than do actors (Jones & Nisbett, 1972; Malle, Knobe, & Nelson, 2007), evaluators may believe that the process of alibi generation seems simple – in an abstract, heuristic sense – and may have little appreciation for how rarely an innocent person is able to generate a convincing, provable alibi. Actor-observer differences in judgments depend largely on the history of the person making the judgment (increased interaction between actors and observers changes the effect; see Robins, Spranca, & Mendelsohn, 1996). Changes to a person’s experience (e.g., by watching a video from the visual direction of the other person; see Storms, 1973) or to a person’s perspective (through empathy instructions; see Galper, 1976) can eliminate or even reverse the actor-observer effect. Thus, if we can alter expectations about how easy alibi generation should be, we should also be able to alter evaluators’ considerations of alibi believability: if evaluators expect that innocent people have difficulty generating strong alibis, they may be more willing to consider a weaker alibi as more believable. Altering these expectations may occur via a cognitive appreciation of the difficulty of alibi generation – a change in knowledge, created by an experience of memory failure or a realization that their own evidence is poor – or via an emotional appreciation of generation difficulty – the phenomenological feeling-of-difficulty when generating the alibi results in an increase in empathy towards the suspect and a change in expectations.

One method of altering expectations about the process of alibi generation is to ask evaluators to undergo that very same difficult process. In this way, evaluators can use their experience of attempting to generate alibis as a metacognitive cue – the ease with which they can recall autobiographical information for their own alibis becomes an anchor for their judgments about a suspect’s experience, especially when their expectation differs from the actual experience (Raghubir & Menon, 2005). To the extent that alibi evaluators find the alibi-generation process difficult, we would expect that generating an alibi would result in greater belief in a suspect’s alibi.

EXPERIMENT 1

Perceptions of alibi-generation difficulty were manipulated in two ways: Participants were asked to either evaluate a criminal suspect’s alibi first or to generate an alibi of their own before evaluating the suspect’s alibi, and participants were asked to generate alibis for a time either 3 days previous or 30 days previous. All participants in the generate-first conditions were expected to rate the suspect’s alibi as more believable than those in the evaluate-first conditions. However, it was hypothesized that latency and experience would interact. Specifically, because it is more difficult to generate an alibi for 30 days ago than for 3 days ago, the impact of generating an alibi before evaluating the suspect’s alibi should be stronger in the 30-day condition than in the 3-day condition. But, of course, this should only occur in the generate-first conditions.
Method

Participants
Participants were 147 undergraduate students (51 men, 96 women) at a large Midwestern university, recruited for a study entitled ‘The Psychology of Belief’. They earned extra credit in their psychology classes for their participation.

Procedure
Upon entering the lab, participants were escorted to a small cubicle where they filled out the questionnaire packet. The design was a 2 (experience order: evaluate-first vs. generate-first) × 2 (latency: alibi for 3 days ago vs. 30 days ago) factorial. Participants who were randomly assigned to the evaluate-first condition read an alibi from a criminal suspect and rated the believability of the alibi. Participants in the generate-first condition were asked to provide an alibi for some time in their own past. For their second task, these roles were reversed: participants in the evaluate-first condition were asked to provide an alibi, and participants in the generate-first condition evaluated the believability of the criminal suspect’s alibi. Participants who were randomly assigned to the 3-day condition were asked to generate an alibi for their whereabouts for a time frame that was 3 days in the past and those randomly assigned to the 30-day condition were asked to generate an alibi for their whereabouts for a time frame that was 30 days in the past. Finally, as an exploratory measure, participants were asked to estimate how much time during a typical day they would be without alibi corroboration. After they had completed the questionnaires, participants were thanked, debriefed, and dismissed.

Dependent measures
When participants were asked to provide their own alibi, they also filled out the generation measures. First, they were asked to write an alibi in narrative form and answer several questions about the supporting evidence they thought they could produce for their alibi, including the type of physical evidence and type of person corroboration. The confidence question asked participants how confident they were that they could produce the supporting evidence on a 0–10 scale (0 = not at all certain, 10 = totally certain). Finally, the judgment question asked participants to estimate how believable they thought a police detective or juror would find their alibi on a 0–10 scale (0 = not at all believable, 10 = completely believable).

The criminal suspect’s alibi (given to all participants; see below) was adapted from a moderately believable alibi, in which there was no physical evidence and the person evidence was a corroborating friend. This alibi was used in Olson and Wells (2004) and this type of alibi (no physical evidence; only a friend to corroborate) was the most common type of alibi proffered by innocent participants in Olson and Charman (in press). Participants read the following:

A detective was investigating an armed robbery at a convenience store. The robbery occurred at 12:30 pm. Her best lead disintegrated, so the detective visited a new suspect one month after the crime occurred. Following is her report of his alibi story: “B. L. claimed that he had been at a friend’s house on the day of the robbery. He insisted he had nothing to do with the crime. He said that he and his friend had been planning on watching a football game that day, and he was going to bring snacks. So he left his house on foot at around 11:30, stopped at a nearby grocery store and bought snacks and beer, and continued on to his friend’s house, intending to be there by kickoff at noon. I contacted the friend, who
confirmed that B. L. was at his house for the football game that day, and they watched the entire game.

The evaluation measures examined the participants’ belief in the alibi: first, the belief question asked participants how much they believed the suspect’s alibi story on a 0–10 scale (0 = not at all, 10 = completely). Then, the qualitative questions were exploratory questions asking participants to explain what about the alibi made them believe it and what about the alibi made them disbelieve it. Finally, the detective belief question asked participants to imagine themselves as a police detective, and in this context, estimate how much they would believe the suspect’s alibi on a 0–10 scale (0 = not at all, 10 = completely).

### Results

**Evaluation measures**

Because the belief question and the detective belief question both measured participants’ perceptions of the believability of the suspect’s alibi and the measures were moderately correlated ($r = .64$), we averaged them to form a composite believability rating; the means and standard deviations can be found in Table 1. As predicted, participants who generated an alibi before evaluating the suspect’s alibi rated the suspect’s alibi as significantly more believable than did participants who simply evaluated the suspect’s alibi first $F(1, 143) = 11.55, p < .001$, partial $\eta^2 = .075$. The difficulty of generating alibis due to latency did not affect participants’ belief in the suspect’s alibi $F(1, 143) = 1.16, p = .28$, and there was no interaction between latency and experience order for participants’ believability ratings for their own alibis $F(1, 143) = 0.07, p = .79$ or for their confidence in their ability to produce evidence $F(1, 143) = 0.05, p = .83$.

**Generation measures**

Means and standard deviations for the generation measures can be found in Table 2. The alibi-generation effect also influenced how participants assessed their own alibis – participants who generated their alibi first rated their own alibis as significantly more believable than did participants who evaluated the suspect’s alibi first $F(1, 143) = 6.95, p = .009$, partial $\eta^2 = .05$. There were no significant interactions between latency and experience order on either participants’ believability ratings for their own alibis $F(1, 143) = 0.07, p = .79$ or for their confidence in their ability to produce evidence $F(1, 143) = 0.05, p = .83$.
The alibi-generation effect

Table 2. Experiment 1: Participants’ evaluation of their own alibis as a function of experience and latency

<table>
<thead>
<tr>
<th>Latency</th>
<th>Generate-first</th>
<th>Evaluate-first</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>How confident are you that you could produce some evidence supporting your alibi?</td>
<td></td>
</tr>
<tr>
<td>3 Days</td>
<td>90.00 (17.07)</td>
<td>92.43 (18.32)</td>
</tr>
<tr>
<td>30 Days</td>
<td>78.68 (26.83)</td>
<td>82.78 (26.68)</td>
</tr>
<tr>
<td></td>
<td>How believable do you think a detective would find your alibi?</td>
<td></td>
</tr>
<tr>
<td>3 Days</td>
<td>81.39 (16.06)</td>
<td>71.76 (15.37)</td>
</tr>
<tr>
<td>30 Days</td>
<td>68.95 (25.23)</td>
<td>61.11 (21.75)</td>
</tr>
</tbody>
</table>

Note. Standard deviations are in parentheses.

There was a main effect of latency, supporting the contention that participants found temporally distant alibis more difficult to generate. Participants who were asked for alibis from 30 days previous were significantly less confident in their ability to produce evidence \( F(1, 143) = 7.82, p = .006, \text{ partial } \eta^2 = .05 \) and rated their own alibis as significantly less believable \( F(1, 143) = 12.14, p = .001, \text{ partial } \eta^2 = .08 \) than did those who were asked for alibis from 3 days previous.

Participants estimated that they would be without alibi corroboration for approximately 15 hours in a typical day \( (M = 14.91, SD = 5.15) \). There was no significant interaction of latency and experience on their estimates, nor were there significant main effects (all \( F_s < 2.2, ps > .13 \)).

Discussion

Participants who generated their own alibis rated a criminal suspect’s alibi as more believable than those participants who simply evaluated the alibi. This alibi-generation effect emerged across multiple measures, influencing participants’ evaluations of the suspect’s alibi, their estimations of a detective’s belief in the suspect’s alibi, and even participants’ perceptions of their own alibis. The alibi-generation effect was robust enough to appear whether participants were asked to generate alibis from a mere 3 days ago or 30 days ago. Participants did seem to recognize that latency should result in greater alibi-generation difficulty – those generating alibis from 30 days previous were less confident in their ability to produce evidence and felt their own alibis were less believable. Interestingly, latency in generating their own alibi did not affect how participants evaluated the suspect’s alibi. Perhaps the 3-day latency experience allowed them to extrapolate and recognize that it would be even more difficult at 30 days, so they did not actually have to experience the longer latency in order to appreciate the 30-day situation of the suspect.

The alibi-generation effect is consistent with earlier research demonstrating that people use their own knowledge and emotional states to estimate what others may know or feel (e.g., Nickerson, 1999; Van Boven & Loewenstein, 2003): If evaluators extrapolate from their own experience (that their alibis are weak or difficult to prove) to the experience of a criminal suspect (if he were innocent, his alibi would probably be weak, too), they would be more believing of a relatively weak alibi. The experience of alibi generation thus lowers evaluators’ expectations about what a suspect should be able
to produce for an alibi. It is unclear from Experiment 1, however, if their expectations were altered as a result of a cognitive change or a more emotional change, that is, an increase in empathy towards the suspect.

Empathy is an ability to put oneself in another’s shoes, to see the world from another’s point of view, and to feel an emotional connection to another’s situation (Duan & Hill, 1996; Smith, 2006). There is some empirical evidence showing that similar experiences increase feelings of empathic concern towards a target (Hodges, Kiel, Kramer, Veach, & Villanueva, 2010), and similar experience may increase perspective-taking, an important antecedent to empathetic responses (Batson, Eklund, Chermok, Hoyt, & Ortiz, 2007). Perspective-taking instructions are also enough to eliminate actor-observer differences in explanations of behaviour (Galper, 1976). Davis et al. (2004) demonstrated that asking participants to engage in perspective-taking increases self-related cognitions – which could explain how evaluators can make the inferential step from ‘criminal suspect’ to ‘if he were innocent, his alibi could be weak’; after all, evaluators know they are innocent. If the empathy interpretation is correct, then priming evaluators with empathetic, perspective-taking thoughts should increase the alibi-generation effect, as greater empathy should be associated with more belief in the suspect’s alibi.

EXPERIMENT 2

Experiment 2 was designed to replicate the alibi-generation effect by asking participants to either generate an alibi of their own or evaluate the criminal suspect’s alibi first. We also sought to separate the influence of the knowledge of alibi-generation difficulty from the phenomenological experience of that difficulty, so we added an additional condition in which participants read about the difficulty of generating alibis without actually generating one of their own. We refer to this as the read-only condition and it represents a manipulation of passive knowledge about alibi-generation difficulty without actually experiencing that difficulty first hand. Also, we primed some participants to think empathetic thoughts by administering two subscales of the Interpersonal Reactivity Index (IRI), an empathy measure that includes items such as, ‘Before criticizing somebody, I try to imagine how I would feel if I were in their place’ (Davis, 1980). If the empathy interpretation is correct, then the empathy prime should interact with alibi-generation experience, so that the generation effect would be larger when empathy is primed than when it is not.

Method

Participants

A total of 255 undergraduate student participants (73 men, 119 women, 63 did not record gender) at a small Midwestern comprehensive university were recruited for a study entitled ‘The Psychology of Belief’. They earned extra credit in psychology and general-education social science classes for their participation. The study was conducted in a classroom setting with between one and five participants per session.

Procedure

Upon entering the classroom, participants were randomly assigned to one of six experimental conditions in a 2 (empathy) × 3 (experience) design. Participants in the empathy-prime conditions were given the Perspective-Taking and Empathic-Concern
The alibi-generation effect

159

subscales of the IRI, and participants in the no-empathy-prime conditions were given a questionnaire about music preferences. Then, participants either generated their own alibi (generation-experience condition), evaluated a criminal suspect’s alibi (evaluation-experience condition), or read a statement asserting that alibis are difficult to generate (read-only condition). The statement consisted of two paragraphs summarizing the reasons for alibi-generation difficulty mentioned earlier: that people may find it difficult to remember their whereabouts, that alibis require evidence that can be hard to recover, and that family and friends may not be considered believable corroborators.

Next, participants in the generation-experience and the read-only conditions evaluated the criminal suspect’s alibi, and participants in the evaluation-experience condition generated their own alibis. Lastly, all participants completed the Perspective-Taking and Empathic-Concern subscales of the IRI before being fully debriefed, thanked, and dismissed.

The generation measures, the criminal suspect’s alibi, and the evaluation measures were the same as used for Experiment 1. Participants in the evaluation-experience and the generation-experience conditions were asked for alibis for a time frame 3 days in the past.

Results

Evaluation measures

As in Experiment 1, we combined the belief measure and the detective belief measure ($r = .77$) into a composite believability rating; means and standard deviations can be found in Table 3. As predicted, the alibi-generation effect was replicated: participants who generated their own alibis first rated the suspect’s alibi as significantly more believable than did those who evaluated the suspect’s alibi first, with those in the read-only condition in-between $F(2, 249) = 3.97, p = .02$, partial $\eta^2 = .031$. Surprisingly, participants in the empathy-prime conditions did not rate the criminal suspect’s alibi as significantly more believable than did those in the non-empathy-prime conditions $F(1, 249) = 2.60, p = .09$, partial $\eta^2 = .012$. There was not a significant interaction between empathy and experience order on participants’ belief in the suspect’s alibi story $F(2, 249) = 0.02, p = .98$.

Generation measures

Unlike in Experiment 1, participants did not alter their confidence in their ability to find supporting evidence in their alibis ($M_s = 88.84–99.56, SDs = 12.46–23.82$) as a result.

Table 3. Experiment 2: Participants’ composite believability rating of the criminal suspect’s alibi as a function of experience and empathy

<table>
<thead>
<tr>
<th>Experience</th>
<th>Generation</th>
<th>Evaluation</th>
<th>Read-only</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empathy prime</td>
<td>6.20 (1.49)</td>
<td>5.49 (1.84)</td>
<td>5.67 (1.72)</td>
<td>5.78 (1.71)</td>
</tr>
<tr>
<td>No empathy prime</td>
<td>6.60 (1.59)</td>
<td>5.89 (1.69)</td>
<td>5.99 (1.96)</td>
<td>6.17 (1.78)</td>
</tr>
<tr>
<td>Total</td>
<td>6.41 (1.58)$^*$</td>
<td>5.69 (1.76)$^*$</td>
<td>5.83 (1.84)</td>
<td></td>
</tr>
</tbody>
</table>

Note. Standard deviations are in parentheses. Means with an asterisk (‘$^*$’) are significantly different at $p < .05$. 
of experience $F(1, 166) = 0.30, p = .59$ or empathy prime $F(1, 166) = 0.08, p = .78$. Likewise, participants did not change their estimation of a detective’s belief in their alibis ($Ms = 79.50–85.33, SDs = 16.32–25.02$) as a result of experience $F(1, 166) = 0.005, p = .95$ or of empathy prime $F(1, 166) = 0.09, p = .77$.

**Empathy scale**

Using a median split of the initial IRI measures taken by participants in the empathy-prime conditions, participants who scored high on the empathy measures were not significantly different from participants who scored low in their composite believability rating of the suspect’s alibi $F(1, 116) = 1.61, p = .21$, suggesting the main effect of the empathy manipulation was not the result of participants’ existing empathetic disposition. Using the final IRI measures taken by all participants after the primary measures, we found that the composite believability rating was not correlated with either the Empathic Concern subscale ($r = .08$) or the Perspective-Taking subscale ($r = .12$).

**Discussion**

The alibi-generation effect was replicated in Experiment 2: As expected, those who generated their own alibi first rated the criminal suspect’s alibi as more believable than did those who evaluated the criminal suspect’s alibi first. The read-only condition yielded believability ratings in-between the generation-experience and the evaluation-experience conditions – suggesting that the alibi-generation effect is due in part to a change in knowledge about alibi-generation difficulty. However, mere knowledge is apparently not enough to produce the alibi-generation effect, as participants in the read-only condition did not change their belief ratings as much as those in the generation-experience condition. It appears that the phenomenological experience of attempting to provide an alibi is necessary to obtain appreciable change in participants’ beliefs about alibis and their evaluation of the alibi provided by the suspect.

It seems, however, that the phenomenological experience of providing an alibi may not change participants’ evaluations due to empathy, as the empathy prime did not affect participants’ belief in the suspect’s alibi. It is important to note, however, that our empathy prime was unrelated to the actual task of generating an alibi, and a more pointed empathy prime may produce the expected effect. For example, asking participants to specifically imagine how they would feel if wrongfully accused of a crime may result in a heightened alibi-generation effect by asking participants to more explicitly place themselves in the perspective of the criminal suspect. On the other hand, empathy scores were uncorrelated to belief in the suspect’s alibi, suggesting that general empathic concern is not involved in alibi belief at all.

The generation measures were not responsive to experience; we suspect that this is largely due to a ceiling effect, as participants uniformly reported that their alibis were very believable and they were highly confident in their ability to find evidence, perhaps because all participants were asked for alibis from 3 days prior. This ceiling effect may also be due to overconfidence in the quality of their evidence: Strange *et al.* (2010) found no relationship between participants’ confidence in their alibi evidence and the actual believability of the alibis, as rated by independent observers. Future research could investigate the effects of manipulations to reduce overconfidence – for example, if evaluators were presented an argument that certain types of evidence were poor corroboration, would they become equally suspicious of their own alibi and the suspect’s alibi?
GENERAL DISCUSSION

Evaluation of alibis occurs at several points in the criminal justice process: Investigators might use an alibi to determine whether to continue investigating a suspect, a prosecutor may decide an alibi is not strong enough to prevent the case from going forward, a juror must determine whether an alibi defence provides a reasonable doubt about the guilt of the accused (see Burke et al., 2007). Each of these types of evaluators are likely to have differing expectations about how strong an alibi needs to be considered believable. These different expectations may arise, in part, from different experiences with the process of alibi generation; for example, investigators may be very familiar with everyday memory errors, and they may thus have a less stringent criterion than do jurors, who might operate under more conventional assumptions about memory. The current research is the first to examine how expectations of alibi evaluators influence their judgments of alibi believability. The results of the current studies suggest that such expectations are malleable by experience.

Experience with alibi generation can be influential on evaluators’ judgments, but it is unclear exactly how the influence operates, as we used a single alibi for both experiments. It is possible that the alibi-generation effect would be more influential when evaluators are judging believability of weak alibis, as those are more common (Olson & Charman, in press). Alternatively, the alibi-generation effect may be more influential when evaluators are judging strong alibis, perhaps via a contrast effect (see Wedell, Hicklin, & Smarandescu, 2007): evaluators may lower their expectations about how strong an alibi should be, and when confronted with strong physical and person evidence, they may judge that alibi as even more believable than they otherwise would have. This study offers a first look into the influence of expectations regarding the difficulty an innocent person should have when generating an alibi - and changes in those expectations - on the judgments of those who hold the expectations, unlike much expectancy-effect research, which addresses the effect of expectations on the targets of those expectations (see Jussim, 1986, for a review of the self-fulfilling prophecy literature).

The alibi-generation effect found in the current research is likely due to a combination of a phenomenological feeling-of-difficulty in generating alibis and a cognitive knowledge that alibi generation is difficult. Manipulations to these elements should also change the strength of the alibi-generation effect; for example, if we were to ask participants to evaluate suspects’ alibis after they had attempted to investigate their own alibis (and many of them would likely discover they had been mistaken on at least some details; see Olson & Charman, in press), we might expect a stronger alibi-generation effect. Additionally, the alibis we asked participants to generate were relatively artificial – our participants were not accused of a crime and were likely less emotionally invested in their alibis than they would have been with a more impactful experience. That we found a robust alibi-generation effect, even under such artificial conditions, suggests that a more ecologically accurate protocol would produce a stronger effect.

The alibi-generation protocol we used most closely resembles a tactic a defence attorney might use with a jury - by asking jurors to engage in a thought experiment of attempting to generate an alibi, attorneys may wish to increase jurors’ understanding and empathy towards the defendant. The current research indicates, however, that perhaps general empathy-raising is not the correct path; perhaps jurors need a more explicit, transferable empathy. Also, we do not know how increased alibi-generation experience - such as a criminal investigator who regularly sees suspects who cannot recall their whereabouts - would affect how evaluators approach alibis. As Sommers and Douglass (2007) demonstrated, alibi evaluation is influenced by the context in which an
alibi is presented. The current research demonstrates that the experience – or lack of it – that an evaluator brings into the evaluation can be equally important in how they rate alibi believability.

Acknowledgement

We would like to thank Mary Amon, Yvonne Anwanwan, Abigail Blackmun, Jason Cliffgard, Erin Dahl, Angela Rehling, and Sarah Theis for their help in data collection. We would also like to extend our deep appreciation to Amy Bradfield Douglass for her helpful comments on earlier drafts.

References


Received 2 July 2010; revised version received 27 September 2010