

What Makes a Good Alibi? A Proposed Taxonomy¹

Elizabeth A. Olson^{2,3} and Gary L. Wells²

A taxonomy of alibis is proposed involving two forms of supporting proof: physical evidence and person evidence. Levels of physical evidence and person evidence were combined to create 12 cells in the taxonomy. Participants (n = 252), who were asked to assume the role of detectives, evaluated alibis representing these 12 cells. The believability of the alibis generally followed the taxonomy's predicted pattern, but physical evidence, when present, tended to overwhelm the person evidence more than had been expected. In addition, alibi evaluators seemed to not consider the possibility that a stranger who corroborated an alibi might be mistaken about the identity of the person. Trait inferences regarding the alibi providers tended to follow the believability data, even when the traits themselves were not relevant to believability of the alibi. We call for the development of a literature on the psychology of alibis, recommend the taxonomy as a framework, and suggest several avenues of inquiry.

KEY WORDS: alibi, alibi witness, corroborator, alibi provider, belief.

Alibi; n. A plea of having been at the time of the commission of an act elsewhere than at the place of commission.

Suppose you discovered, much to your surprise, that you were considered a suspect in a criminal investigation. You are innocent, so you simply have to account for your whereabouts for the time of the crime. If you remember where you were at the time of the crime, you have an alibi—an alibi is simply a claim that you were elsewhere. However, *having* an alibi is the easy part; *proving* an alibi is another matter altogether.

We first developed an interest in alibis in part through examination of cases in which innocent people were convicted by juries and their innocence was established later using forensic DNA testing that was unavailable at the time of their conviction (see Wells et al., 1998). Although the dominant evidence used against these innocent people was eyewitness identification testimony, we noticed that “weak alibis” were often exploited by prosecutors and used as incriminating evidence. Although the

¹Portions of these data were presented at the 2001 Biennial Meeting of the Society for Applied Research in Memory and Cognition.

²Department of Psychology, Iowa State University, Ames, Iowa.

³To whom correspondence should be addressed to Department of Psychology, Iowa State University, W112 Lagomarcino Hall, Ames, Iowa 50011; e-mail: vorndran@iastate.edu.

individuals were innocent of the offense (and, therefore, their alibis were likely true), their alibis were nevertheless ineffective.

These cases, in which innocent people's alibis did not prevent them from being erroneously convicted, led us to wonder, what is a weak alibi? Are there fundamental properties of alibis that can be identified as strong or weak? This seemed to us to be a natural interface between social psychology and law, and we expected to find an empirical literature addressing the issue. Instead, we have been surprised and a bit frustrated to discover that there is a dearth of empirical literature on alibis.

How should we approach the development of an empirical literature on the psychology of alibis? The legal literature provides us with little guidance because it focuses primarily on legal technicalities of the alibi defense. Two technicalities account for most of the legal writings on alibis, namely the prior notice rule and the issue of instructions to juries. The prior notice rule simply states that the defense must give pretrial notice that it intends to present an alibi defense. The legal reasoning behind this rule, which is applicable in most U.S. jurisdictions, is that an alibi can be easily manufactured in the final hours of the defense, thereby preventing a thorough investigation (Epstein, 1964). The instruction issue concerns the matter of whether or not to inform the jury that the defendant is not required to prove an alibi and that the prosecution has the burden to negate the alibi. The requirement of such instructions varies widely across U.S. jurisdictions (Friedman, 1998). Other legal writings allude to distinctions between interested alibi-corroborating witnesses (usually, family members) and disinterested alibi-corroborating witnesses, but fail to develop deeper understandings of this distinction (Gooderson, 1977; Sullivan, 1971).

Although an empirical literature on alibis has not yet been developed, there are a few empirical studies that have used alibis to test hypotheses about other issues, principally eyewitness identification issues. Leippe (1985) for example, used alibi testimony to examine mock juror judgments of various forms of eyewitness identification and nonidentification evidence, but the alibi information was held constant, not manipulated. McAllister and Bregman (1989) manipulated whether an alibi witness positively identified the defendant (thereby corroborating the alibi) or not and, as expected, failure of the alibi witness to corroborate the defendant's alibi led to more guilty verdicts from mock jurors. Lindsay, Lim, Marando, and Cully (1986) manipulated whether or not an alibi witness was a relative of the defendant and found that only the nonrelative alibi witness was able to reduce convictions when there was an eyewitness who had identified the defendant as the culprit. Although all three of these studies used alibi witnesses, the studies were focused primarily on issues of eyewitness identification, not alibis. Alibis were used merely as tools to find out how people think about eyewitness identification issues. The only empirical study that we have been able to find that was devoted to alibis per se was one conducted by Culhane and Hosch (in press). In their study, the alibi witness was either a neighbor or a girlfriend and the witness was either certain or not certain in making either an identification or nonidentification of the defendant as being at his home during the time of the crime. Their results, like those of Lindsay et al., showed that mock jurors were persuaded by the alibi only if the alibi corroborator had no relationship with the defendant.

Clearly, the relationship between the defendant and the alibi corroborator affects the believability of the alibi. But this represents a small start to what seems to us to be a potentially rich literature. It is our intent to facilitate the development of an empirical and conceptual literature on the psychology of alibis, and accordingly, this paper has two goals. First, we introduce some basic distinctions, concepts, and ideas that could serve as a framework for the development of an empirical literature on the psychology of alibis. In doing this, we hope to prime the interest of psychological researchers in the fascinating applied and theoretical possibilities that can accrue from the systematic study of alibis. Second, we propose and test a taxonomy of alibis. We argue that these goals are closely related—a taxonomy of alibis is needed to develop a coherent framework for hypotheses about alibis and to study alibis empirically.

Some Basic Concepts and Distinctions Regarding Alibis

We make a distinction between *alibi* in the colloquial sense and *alibi* as used in this paper. Generally, *alibi* is used to refer to the story, or claim, that is proffered by a suspect in a criminal investigation, and when people say “no alibi” they are referring to an alibi that has no accompanying proof (e.g., many people would consider a plea that the suspect was home alone to be “no alibi” at all). However, we use the definition of *alibi* from the legal system: “a defense that places the defendant at the relevant time of the crime in a different place than the scene involved and so removed therefrom as to render it impossible for one to be the guilty party” (Nolan, 1990). Hence, being home alone is in fact an alibi because it is a claim that places the person at a time and place that, if true, makes it impossible for that person to have committed the crime. From a legal defense perspective, however, the claim itself is only one component of the alibi because one must also have evidence, or proof, that is offered in support of the alibi claim.

Alibis can vary greatly in their surface attributes—the story line of the claim. By surface attributes, we mean the setting (e.g., home, work, public event), activities (e.g., sleeping, eating, playing, working), and other characteristics of the story. But, it is not the surface attributes themselves that make some alibis strong and others weak. Instead, it is the level of proof that can be produced in support of the story. An at-work alibi might be no more persuasive than an at-home alibi if there are no work records or coworker corroborators to support the alibi. What distinguishes one alibi from another is the level of proof that can be produced to support the alibi, not the surface attributes of the alibi itself. Although we acknowledge that there are key surface attributes to alibis (e.g., at-home setting vs. public event setting) that affect the likelihood or difficulty of having proof, we assert that the important dimension is proof, not the surface attributes themselves.

Defining *alibi* in terms of the proof offered in support of the alibi allows us to discuss features of strong or weak alibis and compare various alibis with each other regardless of their surface attributes. The surface attributes of alibi stories can vary greatly, but the types of evidence offered to prove the alibi stories fall into one of two broad categories: physical evidence and person evidence. Physical evidence might include such items as receipts, videotape recordings, or other physical records that indicate the person was at a particular location at a given time. Person evidence

might involve testimony from friends, relatives, or strangers who can corroborate the claim that the alibi provider was at a particular place at a given time.⁴ An important quality of any corroborating evidence is that it must solve a time/space problem. The proof must speak to both the time of the crime and the space the alibi provider was occupying at that time. In short, a mere claim that one was at a particular location at a given time is likely to be insufficient; there must be some physical or person evidence to support the claim for an alibi to be considered at all believable.

In calling for a literature on the psychology of alibis, one distinction that we think should be made is between the alibi provider and the alibi itself. Undoubtedly, there are characteristics of alibi providers that make their alibi more or less believable than someone else who gives the same alibi with the same level of proof. The ex-convict who produces an ATM receipt from a distant location from the crime for the time the crime was committed is going to be less believable than the business owner with no prior record who produces an ATM receipt. It is tempting to think that varying the characteristics of alibi providers is the same as studying alibis, but the difference between the ex-convict and the businessman has nothing to do with alibis per se. Hence, it is important to make clear that an understanding of the believability of alibis requires us to study the alibis themselves, not the characteristics of the alibi provider.

The concept of alibi proof raises one of the most important questions that we can imagine regarding alibis. Even if innocent people can reconstruct where they were at a particular time, how often can they actually prove it? How often can we expect innocent people to have physical evidence to corroborate their alibis? How often are people alone, rendering them unable to provide person evidence to corroborate their alibi? How often are people around others in public who could not later recognize them so as to corroborate their alibi? How often are people with someone they know, but that person cannot absolutely attest to the time/space issue or remembers the time differently? These questions fall into a domain that we call “alibi generation.”

A Proposed Taxonomy of Alibis

We envision a research program on alibi generation in which innocent people would have to account for their whereabouts for specific days and times in the recent past. We are currently collecting such data, but the initial research reported in the current paper necessarily begins at a more rudimentary level. This research is concerned with the development of a taxonomy of alibis.

We argue that a taxonomy of alibis is needed to move on to other research questions about alibis. For instance, suppose a researcher wanted to “score” alibis from actual cases for descriptive purposes. How does one classify alibis for purposes of scoring them? Or, suppose a researcher hypothesized that people with certain demographic characteristics (e.g., unemployed, unmarried) are less likely to have good alibi proof than are people with other demographic characteristics (e.g., employed,

⁴The term *alibi provider* refers to the suspect or defendant who is being questioned regarding his or her whereabouts at the time of the crime. We call a person whose statements are put forward to support the alibi an *alibi corroborator*. Although an alibi corroborator is “providing an alibi” for the suspect in the colloquial sense, we reserve the term *alibi provider* for the suspect him or herself.

married). How does one score such alibis for “goodness” of proof or for qualitative differences in types of alibis? There must be some way to organize the alibis that are generated under varying conditions. Our proposed taxonomy is intended to function as an organizing structure that can be put to such purposes.

We think that the most useful taxonomy is one that sorts alibis according to the extent to which alibi evaluators will accept the alibi, that is, the extent to which evaluators *believe* that the alibi provider could not have committed the crime. Thus, the underlying psychological construct in our taxonomy is the “believability” of the alibi. Believability seems to be the basic psychological dimension along which all alibi evaluators, from detectives to prosecutors to judges to jurors, make their judgments about the alibi. Although there is no way to have a taxonomy of the “absolute level” of alibi believability (because alibis will interact with other details of a particular case to affect final judgments about the alibi provider), we can examine the *relative believability* of alibis within the taxonomy by holding nonalibi evidence constant. Believability is relatively easy to measure on a continuous scale; most people are accustomed to reporting such beliefs on a percentage scale. Accordingly, the believability construct not only drove our initial construction of the taxonomy but also served as the principal dependent measure for purposes of validating the taxonomy.

Whereas the underlying psychological dimension for our taxonomy is believability, the underlying structure of the taxonomy is based on the two forms of proof that can be used with alibis, namely physical evidence and person evidence. We can think of no other form of supporting evidence for an alibi other than a person who is brought forward to corroborate the alibi or physical evidence to support the alibi. Because an alibi provider can have physical evidence, person evidence, both physical and person evidence, or neither physical nor person evidence, it seemed natural for us to create a factorial taxonomy with physical evidence and person evidence as the factors.

Another underlying construct in our proposed taxonomy is “perceived ease of fabrication.” By perceived ease of fabrication, we mean the evaluator’s perception of how easily the alibi provider can manufacture or orchestrate the evidence supporting the alibi. For example, an alibi provider who claims he was watching a television program at the time of the crime might support his alibi by describing events in the television program. However, it would be easy to have taped the program and watched it later or downloaded a program summary from the Internet at a later time. On the other hand, it would be difficult for an alibi provider whose image appears on a dated and timed security video from a store across town to have fabricated the video. Of course, it is possible to have tampered with the store’s tapes, but the relative difficulty of fabricating the video should make the alibi proof very strong. We apply the same underlying construct to person corroborators. It would be easier to orchestrate alibi-corroborating testimony from a relative or friend than from a stranger, for instance, because their relationship to the alibi provider could motivate them to fabricate a story favoring the alibi provider. Sullivan (1971) foreshadows this distinction when he classes alibi corroborators into two general categories: those interested in the outcome of the case (such as relatives and friends) and those who are not. He goes on to assert that disinterested witnesses are rare and valuable to an alibi provider’s case.

On the basis of the ease-of-fabrication construct, we created multiple levels of person and physical evidence for the taxonomy. However, we felt that the ease-of-fabrication construct did not capture all of what needed to be captured for the person evidence dimension. A person who corroborates an alibi might be either lying or mistaken. A complete stranger runs the risk of misidentifying the alibi provider, but has no motive to lie. A person with a close relationship to the suspect (e.g., spouse), on the other hand, is unlikely to misidentify the alibi provider, but has a potential motive to lie for the suspect to protect him or her. Attribution theories show strong support for the principle of discounting in which the presence of one explanation leads observers to discount another explanation (Gilbert & Malone, 1995). Hence, attribution theory might predict that the focal explanation for the alibi corroborator's claim (i.e., the corroborator actually saw the suspect at the critical time) would be discounted by the possibility of mistaken identification in the stranger case and by lying in the close relationship case. Perhaps the most believable corroborator is one who is familiar with the alibi provider but not motivated to protect or lie for the alibi provider—a nonmotivated familiar other. Hence, we thought it would be important for the taxonomy to distinguish between three kinds of alibi corroborators: a motivated familiar other (who is not likely to be mistaken but might lie), a complete stranger (who is not motivated to lie but might be mistaken), and a nonmotivated familiar other (who is not motivated to lie for the person and also is not likely to have mistakenly identified the person).

Overview

We combined three levels of the factor physical evidence (none, easy to fabricate, and difficult to fabricate) and four levels of the factor person evidence (none, motivated familiar other, nonmotivated familiar other, nonmotivated stranger) into the taxonomy in Fig. 1. The taxonomy is organized from the lowest level of predicted believability (upper left of the factorial taxonomy) to the highest level of believability (lower right of the factorial taxonomy).

We created several exemplar alibis for each cell of the taxonomy. Using multiple exemplars for each cell of the taxonomy (e.g., mother, friend, brother representing the motivated familiar other category) was important so as to not confuse the effect of a category-level cell (e.g., motivated familiar other) with a particular instance (e.g., mother) of that category cell (see Wells & Windschitl, 1999). Participants read a brief description of a crime and then read the alibis of three suspects. Participants were asked to evaluate each alibi as though they were detectives investigating this criminal case. Each alibi was evaluated sequentially before reading the next alibi. Our primary measure was the believability of the alibi, but participants also made judgments of the likelihood of guilt. Participants also evaluated the alibi providers on a number of personality trait dimensions, some of which were relevant to alibi believability and some of which were not.

We expected each of the two factors in the taxonomy, that is, physical evidence and person evidence, to have main effects on believability. We also expected an interaction. In particular, although we expected to find differences across levels of person evidence when physical evidence was nonexistent or easy to fabricate, we

	Physical Evidence		
Person Evidence	None	Easy to Fabricate	Difficult to Fabricate
None	Least believable		
Motivated Familiar Other (easy to fabricate, not likely mistaken)			
Non-Motivated Stranger (difficult to fabricate, but possibly mistaken)			
Non-Motivated Familiar Other (difficult to fabricate, not likely mistaken)			Most believable

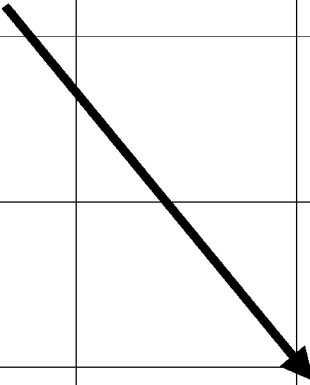


Fig. 1. A taxonomy of alibi believability.

expected levels of person evidence to not have an impact at the highest level of physical evidence. We expected the likelihood-of-guilt and relevant trait measures to follow the same general pattern as the believability data, but we did not expect these measures to be as sensitive to the taxonomy levels as the believability measure. We did not expect the traits that were irrelevant to believability to be affected by levels of the taxonomy.

METHOD

Participants

Participants for the main study were 252 students from a large Midwestern university recruited for an experiment titled “Police Detective Reasoning Skills.” A separate sample of 46 participants were used to conduct a manipulation check on the scenarios. Participants earned extra credit in psychology classes for their participation. All participants were treated in accordance with the “Ethical Principles of Psychologists and Code of Conduct” (American Psychological Association, 1992).

Design

This experiment was a 4 (Person evidence) \times 3 (Physical evidence) mixed factorial design. Presentation of physical evidence was manipulated within participants, and presentation of person evidence was manipulated between participants. Because participants viewed three alibis within one level of person evidence, it was necessary to create three exemplars for each category of person evidence to reduce suspicion and to alleviate stimulus sampling concerns (Wells & Windschitl, 1999). For example, the exemplars from the motivated familiar other category were mother, brother, and best friend, whereas the nonmotivated familiar other category included a grocery store cashier, a bookstore clerk, and a taxi driver. This multiple-exemplar approach ensured that a given participant would not encounter the same exemplar of a person corroborator in more than one alibi.

The order of the within-subjects physical evidence was counterbalanced so that each level of physical evidence occurred equally often in the first, second, and third alibis. For instance, one participant might see alibis with no physical evidence, easy-to-fabricate physical evidence, and difficult-to-fabricate physical evidence, in that order. Another participant might see alibis with difficult-to-fabricate physical evidence, no physical evidence, and easy-to-fabricate physical evidence, in that order. Likewise, the order of exemplars was counterbalanced so that each exemplar appeared equally often in the first, second, and third alibis, as well as equally often with each level of physical evidence.

Materials

Stimulus Materials

The crime scenario was in the form of a police report consisting of two parts. The initial report contained an account of a police officer's investigation of an armed robbery at a convenience store. A follow-up report (shown on the same page) explained that the investigation had reached a dead end with multiple suspects, and it had become necessary to evaluate the suspects on the basis of their alibis. There were 36 short alibi reports used in the experiment, three exemplars for each cell of the 4 \times 3 taxonomy.

Table 1 gives four examples of the 36 alibis. Example A is from the no person evidence and no physical evidence cell, example B is from the motivated familiar-other and easy-to-fabricate physical evidence cell, example C is from the nonmotivated familiar other and difficult-to-fabricate physical evidence cell, and example D is from the nonmotivated stranger and difficult-to-fabricate physical evidence cell. Some readers might be concerned that the differences between one alibi and another alibi in Table 1 involve more than just the evidence offered because they also differ in the basic story line (e.g., one was eating pizza, one was taking a walk, and so on). However, it is crucial to note that a given story line (e.g., was in a grocery store at the time) was used as the setting for no physical evidence, easy-to-fabricate physical evidence (e.g., cash receipt), and difficult-to-fabricate physical evidence (security video) *equally often across participants*. Hence, although no individual participant evaluated more than one alibi with the same alibi story line, alibi story lines were not

Table 1. Alibi Examples

Type of alibi	Example
No person evidence No physical evidence	Suspect A. M. initially could not remember where he was between 7:30 and 8:30 on the evening in question. Later in the interview he claimed that he had been out for a walk in his neighborhood on the east side of the city. A. M. has no record of gun ownership.
Motivated familiar other Easy-to-fabricate	Suspect B. L. said he was entertaining his brother in his home on the west side of the city on the evening in question. He said they had ordered pizza and provided a pizza delivery receipt, paid in cash, timed 8:07 pm. A statement from the brother was taken; the brother claimed they had been in the home for the entire time between 7:30 and 8:30 pm. Suspect does not own a gun.
Nonmotivated familiar other Difficult-to-fabricate	Suspect C. Z. claimed he had been at a Check-Into-Cash store in the central city between 8:00 and 8:20 pm on the evening in question. He said that he regularly goes to that store to get cash. The teller at the store recognized a picture of the suspect and agreed that he is a regular customer. The teller also indicated that he was there that night. Security camera video from the store showed the suspect in the store between 8:03 and 8:18. Suspect has no history of gun ownership.
Nonmotivated stranger Difficult-to-fabricate	Suspect C. Z. claimed that he had been at a grocery store in the city between 7:45 and 8:30 pm on the evening in question. He said that he had been grocery shopping, and then stopped at the ATM in the grocery store to withdraw cash. After seeing a picture of the suspect, a cashier at the customer service desk said that she remembered seeing the suspect at the store that night at approximately 8 pm. Camera video from the ATM showed the suspect withdrawing money at 8:26 pm. Suspect has no history of gun ownership.

confounded with evidence levels. All 36 alibis can be viewed at the following internet address: <http://www.psychology.iastate.edu/faculty/glwells/alibiscenarios.htm>

Manipulation Checks on the Scenarios

It was intended that the three person corroborator conditions would differ along two dimensions: the willingness of the corroborator to lie for the suspect and the chances that the corroborator would be honestly mistaken. More specifically, it was intended that the main difference between the motivated-other corroborator and the nonmotivated familiar-other corroborator would primarily be one of willingness to lie whereas the main difference between the nonmotivated familiar-other corroborator and the stranger corroborator would primarily be in the chances of an honest mistake. With regard to physical evidence, it was intended that the two physical evidence scenarios differ in terms of perceived ease of fabricating the evidence. A separate sample of 46 participants were asked to evaluate the person corroborator scenarios for willingness of the corroborator to lie for the suspect and for the chances that the corroborator could have been honestly mistaken. These same 46 participants also evaluated the easy and difficult-to-fabricate physical evidence scenarios for ease of fabrication. As expected, the motivated other was perceived as much more likely to lie for the suspect than was the nonmotivated familiar other, $M = 7.11$ and 3.63 , respectively, $t(45) = 7.94$, $p < .001$. Also as expected, the stranger corroborator was perceived

as more likely to be mistaken than was the nonmotivated familiar other, $M = 6.24$ and 5.34 , respectively, $t(45) = 2.09$, $p < .05$. Finally, the difficult-to-fabricate physical evidence was perceived as more difficult to fabricate than was the easy-to-fabricate physical evidence, $M = 7.50$ and 4.76 , respectively, $t(45) = 6.00$, $p < .001$.

Dependent Measures

For the primary dependent measure, participants rated each alibi on an 11-point Likert-type scale of believability ($0 = I$ do not believe him at all, $10 = I$ believe him completely) immediately after reading each alibi (hereafter called the *belief* measure). Participants also rated on an 11-point Likert-type scale the probability that the suspect is the one who committed the crime ($0 =$ totally unlikely, $10 =$ he is certainly the gunman) after reading each alibi (hereafter called the *likelihood* measure). This likelihood measure was not considered a primary measure of believability because it does not focus on the alibi itself; rather, such a judgment could evoke a broad range of considerations, such as the number of other suspects or the absence of other evidence. Although not the primary measure, the likelihood-of-guilt judgment was included to see if it followed the same pattern as the primary measure (belief).

Participants also rated each alibi provider on 20 traits. Some of these traits we considered to be relevant to the believability of an alibi provider (e.g., *suspicious*, *cunning*, *scheming*, *deceitful*) and others we considered to be irrelevant to the believability of an alibi provider (e.g., *shy*, *curious*, *funny*, *content*). We used the resulting correlations between traits and alibi believability to sort traits into the relevant and irrelevant categories for purposes of analysis. Traits were rated on a 7-point Likert-type scale ($1 =$ does not describe this suspect at all, $7 =$ describes this suspect perfectly). We expected the relevant traits to be affected by our manipulations of alibis and we expected the irrelevant traits to not be affected by our manipulations.

After participants had read and rated each alibi individually, they were asked an open-ended, exploratory question regarding what made them believe and disbelieve each alibi. The open-ended question was not subjected to formal analyses, but was included merely to assist the current authors' development of future hypotheses about how people evaluate alibis.

Procedure

The crime scenario, consisting of an initial crime report and a follow-up report, was given to all participants upon arrival at the study. Participants were asked to assume the role of a detective and evaluate the alibis of six suspects. Although participants were told that they would evaluate six alibis, participants were given only three alibis to evaluate; this was done to help prevent order effects. If participants believed they were approaching the last of the possible suspects, they may have injected some strategy (e.g., "I didn't think it was any of the others, so it must be this one") into their decision, rather than evaluating the alibi itself. After participants finished evaluating their three alibis, they were fully debriefed, thanked, and dismissed.

Table 2. Mean Alibi Belief as a Function of Evidence

Person evidence	Physical evidence			Overall mean
	None	Easy-to-fabricate	Difficult-to-fabricate	
None	4.79 (1.88) _a	6.44 (2.18) _{bd}	6.97 (2.22) _{bd}	6.07 (2.28)
Motivated				
Familiar other	5.40 (2.20) _{ab}	6.83 (2.28) _{cd}	7.19 (2.21) _{cd}	6.47 (2.35)
Nonmotivated				
Familiar other	5.83 (2.29) _{ab}	6.46 (2.39) _{bd}	7.41 (2.23) _{cd}	6.57 (2.38)
Nonmotivated				
Stranger	6.63 (2.06) _b	6.68 (2.43) _{bd}	7.11 (2.66) _{bd}	6.81 (2.39)
Overall mean	5.66 (2.20)	6.60 (2.31)	7.17 (2.33)	

Note. Standard deviations are given in parentheses. Scale range is from 0 to 10. Higher numbers indicate greater belief in the alibi. Means not sharing a common subscript within row or column differ at $p < .05$ (a Bonferroni correction was used to hold overall alpha at $p < .05$).

RESULTS

We expected that participants would differentiate among types of physical evidence and types of person evidence, rating some alibis as more believable than others in ways reflected in the taxonomy as shown in Fig. 1. We also hypothesized that difficult-to-fabricate physical evidence would be powerful evidence and would overwhelm all person evidence with which it was combined, resulting in no effect across levels of person evidence for difficult-to-fabricate physical evidence.

The Belief Measure

Table 2 shows means and standard deviations for the *belief* measure according to condition, and Fig. 2 shows the overall pattern of means. For ease of viewing, we

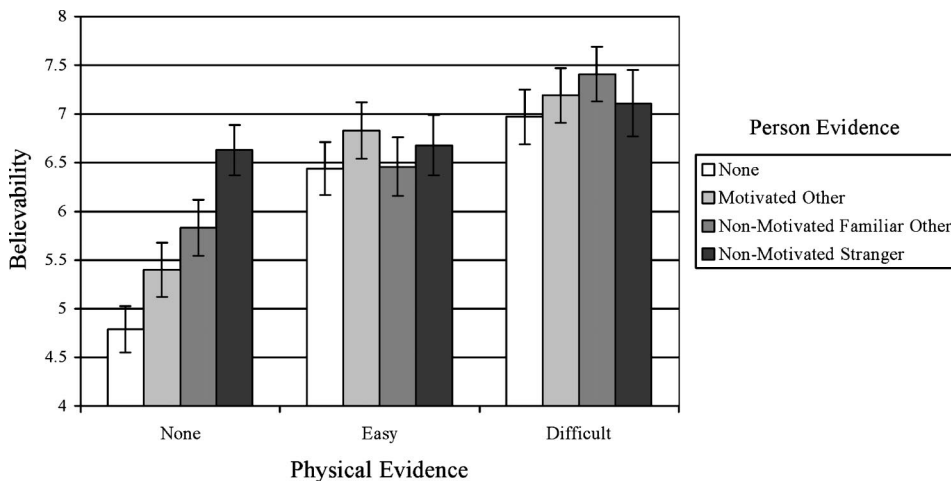


Fig. 2. Mean believability rating as a function of corroborating evidence. Bars represent one standard error.

switched the order of the nonmotivated familiar other and nonmotivated stranger conditions to reflect the order of the obtained means rather than the order in the taxonomy. A general 4×3 mixed ANOVA revealed a significant interaction between physical evidence and person evidence on how believable evaluators rated alibis, $F(6, 494) = 2.49$, $MSE = 4.53$, $p < .05$, Cohen's $f = 0.17$. There was also a significant main effect for person evidence, $F(3, 247) = 2.76$, $MSE = 6.21$, $p < .05$, Cohen's $f = 0.18$, as well as a significant main effect for physical evidence, $F(2, 494) = 32.68$, $MSE = 4.53$, $p < .05$, Cohen's $f = 0.36$.

Planned comparisons were then performed to clarify the relationships among person and physical evidence. The first three simple effects were tests of whether person evidence affected believability within each level of physical evidence. Results indicated that type of person evidence only affected believability when there was no physical evidence. There was no significant simple effect of person evidence at the level of easy-to-fabricate evidence $F(3, 247) = .39$, $MSE = 5.38$, *ns*, Cohen's $f = 0.07$, or at the level of difficult-to-fabricate physical evidence $F(3, 247) = .40$, $MSE = 5.45$, *ns*, Cohen's $f = 0.07$. There was, however, a significant simple effect for person evidence at the level of no physical evidence $F(3, 248) = 8.48$, $MSE = 4.46$, $p < .05$, Cohen's $f = 0.32$. Six single-degree-of-freedom contrasts were conducted to examine differences among the four levels of person evidence under conditions in which there was no physical evidence. We used the Bonferroni correction to hold the overall alpha level at $p < .05$ by dividing the alpha level by 6, thereby requiring that any single contrast be significant at $p < .008$. Three of these contrasts were significant. The nonmotivated stranger was considered more believable than no person evidence, $t(251) = 4.90$, $p < .008$, $d = 0.62$, and was also considered more believable than the motivated other, $t(251) = 3.29$, $p < .008$, $d = 0.42$. Also, the nonmotivated familiar other was considered more believable than no person evidence $t(251) = 2.77$, $p < .008$, $d = 0.35$.

The next four simple effects were tests of whether physical evidence affected believability within each level of person evidence. Results indicated that physical evidence affected believability at all levels of person evidence except the nonmotivated stranger level. When the alibis included no person evidence, physical evidence made a significant difference in the ratings of alibis, with no physical evidence the weakest and difficult-to-fabricate evidence the strongest, $F(2, 124) = 19.92$, $p < .05$, Cohen's $f = 0.57$. Likewise, there was a simple effect of physical evidence at the level of motivated familiar-other person evidence, $F(2, 124) = 11.72$, $p < .05$, Cohen's $f = 0.44$, as well as at the level of nonmotivated familiar-other person evidence, $F(2, 124) = 8.66$, $p < .05$, Cohen's $f = 0.38$. However, there was no significant simple effect for physical evidence when a stranger was a corroborator, $F(2, 122) = 1.08$, *ns*, Cohen's $f = 0.13$.

Three single-degree of-freedom contrasts were conducted to examine differences among the three levels of physical evidence in the absence of person evidence. We used the Bonferroni correction to hold the overall alpha level at $p < .05$ by dividing the alpha level by 3, thereby requiring that each single contrast be significant at $p < .017$. Two of these three contrasts were significant. Easily fabricated physical evidence was rated significantly more believable than no physical evidence, $t(251) = 5.06$, $p < .017$, $d = 0.80$. Likewise, difficult-to-fabricate physical evidence

Table 3. Mean Likelihood of Guilt Rating as a Function of Evidence

Person evidence	Physical evidence			Overall mean
	None	Easy-to-fabricate	Difficult-to-fabricate	
None	5.41 (1.77) _a	3.98 (2.37) _b	3.35 (2.20) _{bd}	4.25 (2.29)
Motivated				
Familiar other	4.59 (2.04) _{ab}	3.32 (2.35) _{bc}	3.22 (1.98) _{cd}	3.71 (2.21)
Nonmotivated				
Familiar other	4.49 (2.19) _{ab}	3.60 (2.41) _{bc}	2.66 (2.10) _{cd}	3.59 (2.35)
Stranger	3.98 (2.16) _b	3.37 (2.23) _b	2.81 (2.05) _{bd}	3.38 (2.19)
Overall mean	4.62 (2.10)	3.56 (2.34)	3.01 (2.09)	

Note. Standard deviations are given in parentheses. Scale range is from 0 to 10. Higher numbers indicate greater likelihood that the alibi provider is the gunman. Means not sharing a common subscript within row or column differ at $p < .05$ (a Bonferroni correction was used to hold overall alpha at $p < .05$).

was significantly more believable than no physical evidence $t(251) = 5.60, p < .017, d = 0.71$.

The Likelihood Measure

Table 3 shows the means and standard deviations for the *likelihood* question according to condition, and Fig. 3 shows the overall pattern of means for the *likelihood* question according to condition. A general 4×3 mixed ANOVA revealed no significant interaction between physical evidence and person evidence on judgments of likelihood that the alibi provider was the gunman, $F(6, 486) = 0.98, MSE = 3.99, ns$, Cohen’s $f = 0.11$. However, there was a significant main effect for physical evidence $F(2, 486) = 40.84, MSE = 3.99, p < .05$, Cohen’s $f = 0.41$, as well as a significant

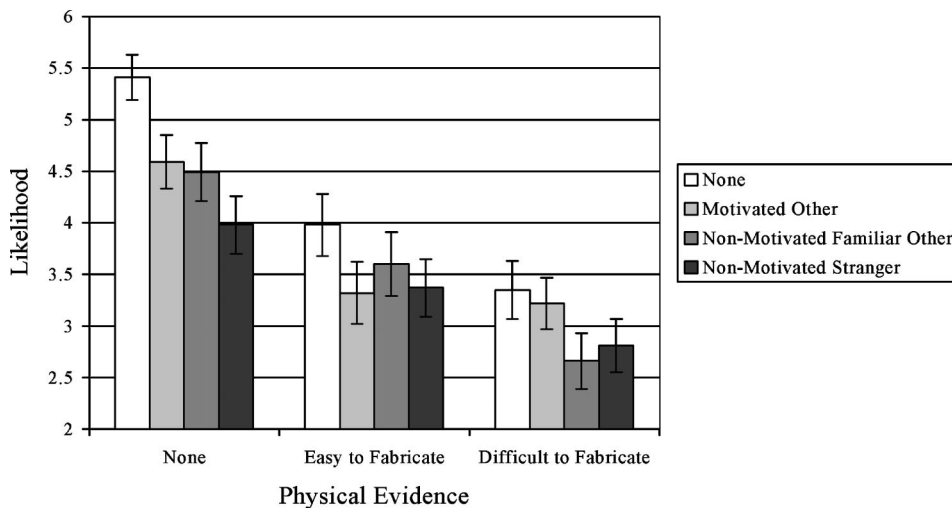


Fig. 3. Mean likelihood rating as a function of corroborating evidence. *Note.* Bars represent one standard error.

main effect for person evidence $F(3, 243) = 4.29$, $MSE = 5.97$, $p < .05$, Cohen's $f = 0.23$.

Planned comparisons were then performed to clarify the relationships among person and physical evidence. The first three simple effects were tests of whether person evidence affected likelihood judgments within each level of physical evidence. Results indicated that type of person evidence affected likelihood judgments only when there was no physical evidence. There was no significant simple effect of person evidence at the level of easy-to-fabricate physical evidence $F(3, 246) = 1.05$, $MSE = 5.48$, *ns*, Cohen's $f = 0.12$, or at the level of difficult-to-fabricate evidence $F(3, 247) = 1.55$, $MSE = 4.34$, *ns*, Cohen's $f = 0.14$. There was a significant simple effect for person evidence at the level of no physical evidence $F(3, 246) = 5.21$, $MSE = 4.18$, $p < .05$, Cohen's $f = 0.25$. Six single-degree-of-freedom contrasts were conducted to examine differences among the four levels of person evidence under conditions in which there was no physical evidence. We used the Bonferroni correction to hold the overall alpha level at $p < .05$ by dividing the alpha level by 6, thereby requiring that any single contrast be significant at $p < .008$. One of these contrasts was significant: Alibi providers with a nonmotivated stranger corroborator were considered less likely to be the gunman than were alibi providers with no person corroborator $t(250) = 3.87$, $p < .008$, $d = 0.49$.

The next four simple effects were tests of whether physical evidence affected likelihood judgments within each level of person evidence. Results indicated that type of physical evidence affected likelihood judgments at all levels of person evidence. When the alibi provider offered no person evidence, physical evidence made a significant difference in the judgments of likelihood that the alibi provider was the gunman, with no physical evidence the most likely and difficult-to-fabricate evidence the least likely, $F(2, 122) = 17.19$, $p < .05$, Cohen's $f = 0.53$. Likewise, there was a similar simple effect of physical evidence when a motivated familiar other was a corroborator, $F(2, 124) = 8.19$, $p < .05$, Cohen's $f = 0.36$, when a nonmotivated familiar other was a corroborator, $F(2, 120) = 12.02$, $p < .05$, Cohen's $f = 0.45$, and when a stranger was a corroborator, $F(2, 120) = 6.14$, $p < .05$, Cohen's $f = 0.32$.

Three single-degree of-freedom contrasts were conducted to examine differences among the three levels of physical evidence in the absence of person evidence. We used the Bonferroni correction to hold the overall alpha level at $p < .05$ by dividing the alpha level by 3, thereby requiring that each single contrast be significant at $p < .017$. Two of these three contrasts were significant. Alibi providers with easily fabricated physical evidence were rated significantly less likely to be the gunman than those with no physical evidence, $t(251) = 4.65$, $p < .017$, $d = 0.59$. Likewise, alibi providers with difficult-to-fabricate physical evidence were rated significantly less likely to be the gunman than those with no physical evidence $t(251) = 5.52$, $p < .017$, $d = 0.70$.

The Trait Measures

All trait ratings were scored so that higher scores bore a negative connotation. Correlations between traits and alibi believability were then calculated to determine which traits were relevant and which were irrelevant. The traits *honest*, *deceitful*,

Table 4. Mean Relevant Traits Rating as a Function of Evidence

Person evidence	Physical evidence			Overall mean
	None	Easy-to-fabricate	Difficult-to-fabricate	
None	4.15 (0.71) _a	3.68 (0.90) _{ab}	3.58 (0.90) _{bd}	3.80 (1.11)
Motivated				
Familiar other	3.90 (0.82) _{ab}	3.37 (0.95) _{bc}	3.30 (0.90) _{cd}	3.52 (1.10)
Nonmotivated				
Familiar other	3.89 (0.74) _{ab}	3.61 (0.90) _{bc}	3.41 (0.74) _{cd}	3.64 (1.10)
Nonmotivated				
Stranger	3.60 (0.75) _b	3.47 (0.79) _b	3.40 (0.86) _{bd}	3.49 (1.10)
Overall mean	3.89 (0.78)	3.53 (0.89)	3.42 (0.85)	

Note. Standard deviations are given in parentheses. Scale range is from 0 to 10. Higher numbers indicate greater ratings of the alibi provider as dishonest, suspicious, deceitful, untrustworthy, etc. Means not sharing a common subscript within row or column differ at $p < .05$ (a Bonferroni correction was used to hold overall alpha at $p < .05$).

suspicious, conniving, open, cunning, sincere, scheming, trustworthy, calculating, loyal, and intelligent were significantly correlated with the *belief* measure ($.13 < r > .35$), and the traits *content, shrewd, friendly, caring, ambitious, curious, funny, shy* were not significantly correlated with the *belief* measure ($.01 < r > .09$). Thus, the former 12 traits were considered relevant and the latter eight traits were considered irrelevant for purposes of the main analyses.⁵ Coefficient alpha for the 12 relevant traits was .79 and for the eight irrelevant traits was .43.

Relevant Traits

The relevant trait ratings were averaged into a single score for each participant. Table 4 shows means and standard deviations for the overall relevant traits rating according to condition. A general 4×3 mixed ANOVA revealed no significant interaction between physical evidence and person evidence on the relevant trait ratings of alibi providers $F(6, 494) = 1.07$, $MSE = 0.67$, *ns*, Cohen's $f = 0.12$. However, there was a significant main effect for physical evidence, $F(2, 494) = 23.17$, $MSE = 0.67$, $p < .05$, Cohen's $f = 0.31$, as well as a significant main effect for person evidence, $F(3, 247) = 2.93$, $MSE = 1.08$, $p < .05$, Cohen's $f = 0.19$.

Irrelevant Traits

The irrelevant trait ratings were averaged into a single score for each participant. Table 5 shows means and standard deviations for the overall irrelevant traits rating according to condition. A general 4×3 mixed ANOVA revealed no significant interaction between physical and person evidence $F(6, 492) = 1.40$, $MSE = 0.21$, *ns*, Cohen's $f = 0.13$. Likewise, there was no main effect for person evidence, $F(3, 246) = 2.29$, $MSE = 0.78$, *ns*, Cohen's $f = 0.17$. Interestingly, there was a significant main effect for physical evidence $F(2, 492) = 21.55$, $MSE = 0.21$, $p < .05$,

⁵To avoid overanalysis of the trait data, we did not perform tests of the simple effects for the relevant and irrelevant trait measures.

Table 5. Mean Irrelevant Traits Rating as a Function of Evidence

Person evidence	Physical evidence			Overall mean
	None	Easy-to-fabricate	Difficult-to-fabricate	
None	4.39 (0.53) _a	4.21 (0.57) _a	4.24 (0.61) _a	4.28 (0.96)
Motivated				
Familiar other	4.29 (0.62) _a	3.94 (0.60) _a	4.00 (0.65) _a	4.08 (0.96)
Nonmotivated				
Familiar other	4.26 (0.62) _a	4.23 (0.64) _a	4.15 (0.54) _a	4.21 (0.96)
Nonmotivated				
Stranger	4.29 (0.60) _a	4.10 (0.58) _a	4.13 (0.59) _a	4.18 (0.96)
Overall mean	4.31 (0.60)	4.12 (0.60)	4.13 (0.60)	

Note. Standard deviations are given in parentheses. Scale range is from 0 to 10. Higher numbers indicate greater ratings of the alibi provider as shy, uncaring, unfriendly, unfunny, etc. Means not sharing a common subscript within row or column differ at $p < .05$ (a Bonferroni correction was used to hold overall alpha at $p < .05$).

Cohen's $f = 0.30$. Thus, participants rated those alibi providers with stronger physical evidence as funnier, more caring, friendlier, etc.

DISCUSSION

Our proposed taxonomy of alibi believability received some support from participants' alibi evaluations. Main effects for person evidence and for physical evidence indicate that participants in the role of detectives tended to make distinctions among these alibis along the lines we had expected. In addition, the presence of an interaction followed the general idea that difficult-to-fabricate physical evidence overwhelms person evidence. Nevertheless, there were some aspects to the results that were not expected. Our interaction prediction was that the difficult-to-fabricate physical evidence would trump the person evidence, but we had not expected that even the easy-to-fabricate physical evidence was sufficient to render the person evidence irrelevant. Some caution is required in generalizing this conclusion because the impact of a manipulated variable is sensitive to the particular ways in which it is operationalized in a given study. Nevertheless, we were surprised that even easy-to-fabricate physical evidence (e.g., a cash receipt) wiped out the effects of having a person to corroborate the alibi even when that person had no apparent motivation to lie for the suspect.

Another mild surprise was that participants did not find the alibi corroborators to be more credible when they were nonmotivated familiar others than when they were nonmotivated strangers. In both cases there was no motive for the alibi corroborator to lie, but the stranger should be more likely than the familiar other to be honestly mistaken. If anything, alibi corroboration from the nonmotivated stranger was seen as more credible than corroboration from the nonmotivated familiar other (although we remind the reader that this difference was not statistically significant). Recall that our manipulation check participants perceived the stranger to be more likely than the nonmotivated familiar other to have been honestly mistaken. However, the manipulation check participants were explicitly asked to consider the chances of an

honest mistake by the corroborator and this might have made that dimension salient for the manipulation check participants in a way that was not salient for the participants who were asked to evaluate the believability of the alibi. This suggests to us that our results might have shown the expected differences between the nonmotivated familiar-other corroborator and the stranger corroborator if this had been a trial setting in which a prosecutor cross-examined the corroborator along lines suggesting that the stranger corroborator could have been mistaken. For this reason, we advise maintaining the taxonomy distinction between stranger corroborators and nonmotivated familiar-other corroborators.

As expected, the pattern of responses for the likelihood of guilt measure followed the (inverse) pattern of the believability measure. Although the interaction pattern between physical and person evidence was similar for the likelihood measure and the believability measure, the interaction did not reach statistical significance for the likelihood of guilt measure. We believe that this is because likelihood of guilt measures are not particularly sensitive measures of alibi believability. There are several reasons why we believe that believability of the alibi rather than likelihood of guilt is a more sensitive measure for studying alibis. First, questions about guilt are likely to lead at least some participants to invoke a variety of other considerations such as the presumption of innocence, the totality of evidence, the presentation of evidence by both sides, and the need for affirmative evidence of guilt. These other considerations that people use to make judgments of guilt can introduce noise into a paradigm that is specifically attempting to focus on the believability of the alibi itself. In this sense, guilt measures are, at best, indirect measures of alibi believability. Second, the absence of proof to support one's alibi might undermine the believability of the alibi, but it is not actually evidence of guilt. Being alone at home at the time that an offense was committed, but having no way to prove it, should not count as affirmative evidence of guilt. Third, even if an alibi is proven wrong, it does not automatically mean the suspect is guilty. For instance, false alibis can be given to cover up social indiscretions (e.g., an affair). Our general point is that we were not surprised that the likelihood of guilt measure was not as sensitive as the believability measure in response to our manipulations. Our taxonomy is a taxonomy of the believability of alibis, not a taxonomy of the perceived likelihood of guilt of those using these alibis. Although the believability of an alibi and the likelihood of guilt are related, guilt judgments invoke a much broader range of considerations and a broader range of evidence than just the alibi.

We found an interesting pattern of responses for the personality trait ratings of the alibi provider. Although we did not have a strong prediction for the trait ratings, we assumed that some traits, such as deceitful and conniving, would be related to alibi believability. We expected that other traits, however, such as caring and funny, should not be affected by alibi proof. A correlation analysis indicated that the 12 relevant traits were each significantly correlated with alibi believability, whereas none of the eight irrelevant traits were significantly correlated with alibi believability. The results showed that both relevant and irrelevant traits were affected by the taxonomy manipulations. The effects on irrelevant traits were not as strong as the effects on relevant traits, but we were surprised to find irrelevant traits to be affected at all by alibi proof. For instance, why should a person with a weak alibi be rated as less funny

than would a person with a strong alibi? In hindsight, this seems not so surprising. All traits can be sorted along a good–bad dimension, even those traits that seem not related to alibis. A suspect whose alibi is evaluated negatively is perceived as someone who is likely to have committed the crime in question and is, therefore, likely to suffer from a number of negative inferences about his or her character. Serendipitously, perhaps we have discovered an implicit, or at least indirect, method for assessing alibi believability. As a proxy for directly asking people to evaluate the believability of an alibi, they could be asked to make trait inferences about the alibi provider.

We find it curious that the condition representing the strongest alibi, which included not only a neutral person corroborating the alibi but also a dated, timed security video capturing a clear image of the suspect, nevertheless yielded a mean believability score of only 7.4 (out of a possible 10.0). What kind of proof would it take to get a believability score of an 8, 9, or 10? In an earlier draft of this paper, we suggested that this speaks to the incredible skepticism with which people seem to greet alibis. We suggested that the mere labeling of a statement as being an alibi evokes a sense of disbelief and challenges people to create imaginative scenarios worthy of a great crime novel as to how the person could nevertheless have committed the crime. Reviewers of the earlier version of this paper objected to our interpretation, noting that there are many reasons why participants might have avoided the high end of the scale. We agree with the reviewers' cautions, and so we offer the alibi-skepticism idea as only a hypothesis for future research. We think it is quite plausible that the term alibi is an extremely loaded term. Consider someone responding to the question "Where were you last night?" versus responding to the question "What is your alibi for where you were last night?" Asking for an alibi is more than just being inquisitive; it is a form of accusation regarding a bad act for which there seems reason to suspect the person being asked. In effect, giving an alibi is a form of denial of criminality and research on innuendo shows that denials of criminality can themselves be incriminating. For example, research participants who read the headline "Andrew Winters Not Connected to Bank Embezzlement" produced assessments of Andrew Winters that were as negative as those coming from the headline "Andrew Winters Connected to Bank Embezzlement" (Wegner, Wenzlaff, Kerker, & Beattie, 1981). Research on anchoring and adjustment processes in human judgment illustrate the power of "starting points" (Koehler, 1991). If alibi evaluators start with an assumption that the alibi is likely to be false (an anchor), then adjustments toward believing the alibi based on proffered proof are likely to be insufficient.

Our primary motivation for developing this taxonomy was to help launch a systematic literature on the psychology of alibis. We think that this taxonomy represents a reasonable starting point for thinking about alibis. In developing this taxonomy, we tried to maintain a degree of simplicity that would permit easy comprehension while serving a heuristic role for generating new hypotheses. We invite others to refine this taxonomy or to propose alternative taxonomies, but we think any useful taxonomy will need to be based on the underlying concept of proof to support the alibi rather than on the surface attributes of the alibi. This is not to say that the surface attributes of the alibi story cannot affect alibi evaluations at all. Consider two men who claim to

have been alone at home at the time of the crime but one says he was smoking crack at the time and the other says he was reading his Bible. Clearly, inferences about the character of the two men based on the surface attributes of their alibi stories are likely to lead to different evaluations. Still, we remind readers that these differences are more at the level of differences in alibi providers, not differences in the alibis themselves.

It would be misleading to suggest that our alibi taxonomy, or any taxonomy, can capture all the dynamics that determine the believability of an alibi. Clearly, the taxonomy assumes that other variables are held constant. These other variables include such things as background characteristics of the alibi provider (e.g., prior record, no prior record) and behavioral cues of the person giving the alibi story (e.g., nervousness, pauses). We also think that alibi evaluators might be especially suspicious of someone who is able to readily provide an alibi for a distant past event when there seems to have been no reason for an innocent person to be able to recall what she or he was doing that day. And, what if the person initially errs in recall (“I was playing golf with friends that afternoon”) only to discover later that she recalled the wrong day and has to change the alibi story? These dynamic features are not captured in the taxonomy and we do not think that the taxonomy itself can be easily modified to capture them. Nevertheless, we think that the taxonomy represents a useful structure from which to study these other variables. For example, we suspect that these other variables (e.g., prior record, nervousness, change of alibi story) will affect alibi credibility primarily when the level of proof is weak (upper left cells of the taxonomy) and will matter less when levels of proof are strong (lower right cells of the taxonomy). Hence, the taxonomy should be useful for researchers even when their hypotheses concern alibi variables that are not part of the taxonomy itself.

A big door that we hope we have opened on the study of alibis concerns the ability of innocent people to generate and prove alibis. Clearly, some innocent people simply will not be able to reconstruct where they were at a particular time in the past. Others might mistakenly recall where they were and then later have to change their alibi story, a behavior that is likely to be met with considerable incredulity. But, even when innocent people can accurately recall where they were, how often could they provide proof at one of the higher levels of our taxonomy? How often would people be stuck in the upper-left cell of the taxonomy (i.e., no person or physical evidence to support the alibi)? We suspect that physical evidence is rather rare and that person evidence, when it exists, will usually come from a motivated familiar other (friend or relative), which is the second-lowest level of believability in the taxonomy. At the same time, we suspect that people overestimate their ability to prove their alibi. For example, they might think that their girlfriend or cousin will somehow be a powerful form of corroboration or that their corroborator will always recall the time frame the same way that they do. In theory, it seems that the alibi defense ought to work with some regularity for innocent people; in practice we worry that it will rarely work because of the difficulty of proving the alibi and the incredible skepticism that seems to accompany alibi evaluations. We challenge psychological researchers to take up some of these important questions.

ACKNOWLEDGMENT

This research is based on the first author's Master's thesis.

REFERENCES

- American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47*, 1597–1611.
- Culhane, S. E., & Hosch, H. M. (in press). An alibi witness's influence on jurors' decision making. *Journal of Applied Social Psychology*.
- Epstein, D. M. (1964). Advance notice of alibi. *The Journal of Criminal Law, Criminology, and Police Science, 55*, 29–38.
- Friedman, J. P. (1998). Alibi instructions and due process of law. *Western New England Law Review, 20*, 343–386.
- Gilbert, D. T., & Malone, D. S. (1995). The correspondence bias. *Psychological Bulletin, 117*, 21–38.
- Gooderson, R. N. (1977). *Alibi*. London, UK: Heinemann Educational.
- Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin, 110*, 499–519.
- Leippe, M. R. (1985). The influence of eyewitness nonidentifications on mock jurors' judgments of a court case. *Journal of Applied Social Psychology, 15*, 656–672.
- Lindsay, R. C. L., Lim, R., Marando, L., & Cully, D. (1986). Mock-juror evaluations of eyewitness testimony: A test of metamemory hypotheses. *Journal of Applied Social Psychology, 15*, 447–459.
- Nolan, J. R. (1990). *Black's law dictionary* (6th ed.). St. Paul, MN: West Publishing.
- McAllister, H. A., & Bregman, N. J. (1989). Juror underutilization of eyewitness nonidentifications: A test of the disconfirmed expectancy explanation. *Journal of Applied Social Psychology, 19*, 20–29.
- Sullivan, T. P. (1971). Presentation of the defense. *The Journal of Criminal Law, Criminology, and Police Science, 62*, 153–172.
- Wegner, D. M., Wenzlaff, R., Kerker, R. M., & Beattie, A. E. (1981). Incrimination through innuendo: Can media questions become public answers? *Journal of Personality and Social Psychology, 40*, 822–832.
- Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22*, 603–647.
- Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling in social psychological experimentation. *Personality and Social Psychology Bulletin, 25*, 1115–1125.