

Behavioral Confirmation in the Interrogation Room: On the Dangers of Presuming Guilt

Saul M. Kassir,^{1,2} Christine C. Goldstein,¹ and Kenneth Savitsky¹

A two-phased experiment tested the hypothesis that the presumption of guilt that underlies police interrogations activates a process of behavioral confirmation. In Phase I, 52 suspects guilty or innocent of a mock theft were questioned by 52 interrogators led to believe that most suspects were guilty or innocent. Interrogators armed with guilty as opposed to innocent expectations selected more guilt-presumptive questions, used more interrogation techniques, judged the suspect to be guilty, and exerted more pressure to get a confession—particularly when paired with innocent suspects. In Phase II, neutral observers listened to audiotapes of the suspect, interrogator, or both. They perceived suspects in the guilty expectations condition as more defensive—and as somewhat more guilty. Results indicate that a presumption of guilt sets in motion a process of behavioral confirmation by which expectations influence the interrogator's behavior, the suspect's behavior, and ultimately the judgments of neutral observers.

KEY WORDS: behavioral confirmation; interrogation; confessions.

In criminal law, confession evidence is the state's most potent weapon. Mock jury studies have shown that confessions are more persuasive than other forms of incriminating evidence, such as eyewitness identifications or character testimony (Kassin & Neumann, 1997). Indicating the operation of the fundamental attribution error, or correspondence bias, studies have also shown that the insertion of a confession into evidence increases the conviction rate even among mock jurors who believe that the confession was coerced and claim that it did not influence their decision making (see Kassin, 1997; Kassin & Sukel, 1997; Kassin & Wrightsman, 1980). Case studies of erroneous convictions corroborate the power of confessions. Leo and Ofshe (1998) found that 73% of defendants were convicted at trial in cases that contained evidence of confessions later proved to be false. Scheck, Neufeld, and Dwyer (2000) found that 23% of all DNA exoneration cases contained confessions, also apparently false.

¹Department of Psychology, Williams College, Williamstown, Massachusetts.

²To whom correspondence should be addressed at Department of Psychology, Williams College, Bronfman Science Center, Williamstown, Massachusetts 01267; e-mail: skassin@williams.edu.

Confessions have such far-reaching and rippling effects within the criminal justice system that researchers have sought to examine modern police interrogation practices and the social influences they produce. Training manuals are available that advise investigators in how to get confessions. The most popular is Inbau, Reid, Buckley, and Jayne's (2001) *Criminal Interrogation and Confessions*, which was first published in 1962, is now in its fourth edition, and forms the basis for the training of law enforcement professionals in the Reid Technique (www.reid.com). Specifically, Inbau et al. describe a nine-step procedure that is designed to overcome resistance by confronting the suspect with assertions of his or her guilt (a goal that often involves interrupting the suspect's denials and presenting false evidence), followed by the offering of sympathy, understanding, and an alternative "theme" that minimizes the moral seriousness of the act. Observational studies have revealed that use of these techniques in the United States is common (Leo, 1996; Wald, Ayres, Hess, Schantz, & Whitebread, 1967).

Modern police interrogations are so powerful that they have, at times, elicited coerced-compliant and coerced-internalized false confessions from innocent people (Gudjonsson, 2003; Kassin, 1997; Leo & Ofshe, 1998; Radelet, Bedau, & Putnam, 1992; Wrightsman & Kassin, 1993). Recent research highlights the risks associated with the use of certain techniques. In a laboratory experiment, for example, Kassin and Kiechel (1996) accused subjects of causing a desktop computer to crash by hitting a key they were specifically told not to press. All subjects were innocent. Yet among those who were forced to type quickly (a procedure that rendered them vulnerable to manipulation) and who were presented with false incriminating evidence (in the form of a confederate-eyewitness), 100% signed a confession, 65% internalized a belief in their own guilt, and 35% confabulated details to fit that newly created belief.

We believe that police interrogations are persuasive, and at times too persuasive, in part because they are theory-driven social interactions founded upon a presumption of guilt. Indeed, law enforcement professionals explicitly justify the use of heavy-handed tactics by claiming that they initiate interrogation only after an initial interview, during which time they analyze the suspect's verbal and nonverbal behavior to determine that he or she is deceptive and, hence, guilty (sometimes the judgment is bolstered by external evidence; often it is not). This is an important definitional point about how investigators are trained in the process of interrogation. As Inbau et al. (2001) put it, "An interrogation is conducted only when the investigator is reasonably certain of the suspect's guilt" (p. 8).

There are two potential problems with this two-step, interview-to-interrogation approach. First, there is no empirical evidence to suggest that police can distinguish between the denials made by guilty and innocent suspects at high levels of accuracy. Studies have shown that people are poor intuitive judges of truth and deception (Zuckerman, DePaulo, & Rosenthal, 1981). Some professionals do outperform the average person at this task (Ekman, O'Sullivan, & Frank, 1999). It is normatively clear, however, that police hold many false beliefs about the behavioral indicators of deception (Akehurst, Kohnken, Vrij, & Bull, 1996) and that groups of "experts" who make such judgments for a living—such as psychiatrists, police investigators, judges, customs inspectors, and polygraphers for the FBI, CIA, and military—are highly prone to error (DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991). In

an experiment specifically designed to examine the effects of training on people's ability to distinguish true and false denials, Kassin and Fong (1999) trained some college students but not others in common police methods of detecting deception, showed them videotaped mock interrogations of guilty and innocent suspects, and found that those who were trained were more confident in their judgments relative to naïve controls, but they were not more accurate. In a follow-up study, Meissner and Kassin (2002) found that even experienced detectives—many of whom were specially trained in interviewing and interrogation—also did not exceed chance-level performance. Compared to others, they also exhibited a deception response bias, leading them to commit an abundance of false positive errors. Thus, the pivotal decision to interrogate a suspect may well be based on prejudgments of guilt confidently made but frequently in error.

A second potential problem with the two-step approach is that it leads police to commence interrogations, after interviews, predisposed and reasonably certain of the suspect's guilt (Inbau et al., 2001)—a “frame of mind” that can influence an investigator's conduct while interviewing suspected offenders (Mortimer & Shepherd, 1999). Thus, an interrogation is a theory-driven social interaction led by an authority figure who holds a strong a priori belief about the target and who measures success by his or her ability to extract an admission from that target. For innocent suspects initially misjudged, one would hope that investigators would remain open-minded enough to monitor the situation and reevaluate their own beliefs. However, research suggests that once people form a belief, they tend unwittingly to seek, interpret, and create information in ways that verify that belief. This phenomenon, often referred to as a self-fulfilling prophecy, was first demonstrated by Rosenthal and Jacobson (1968) in their classic study of teacher expectation effects, with similar results later obtained in military, business, and other organizational settings (McNatt, 2000).

The process was first demonstrated in a controlled laboratory experiment by Snyder and Swann (1978), who brought together pairs of participants for a getting-acquainted interview. The interviewers were led to believe that their partners were introverted or extraverted, after which they selected interview questions from a prepared list. Two key results were obtained. First, interviewers adopted a confirmatory hypothesis-testing strategy, selecting introvert-oriented questions for the introvert (e.g., “Have you ever felt left out of a social group?”) and extravert-oriented questions for the extravert (“How do you liven up a party?”). Second, interviewers unwittingly manufactured support for their beliefs through the questions they asked, leading neutral observers to infer that the interviewees truly were introverted or extroverted, according to expectation. Other laboratory experiments have further shown that behavioral confirmation is the outcome of a three-step chain of events, by which (1) a perceiver forms a belief about a target person; (2) the perceiver unwittingly behaves toward that person in a manner that conforms to that belief; and (3) the target responds in turn, often behaving in ways that support the perceiver's belief (for reviews, see Darley & Fazio, 1980; Nickerson, 1998; Snyder, 1992; Snyder & Stukas, 1999).

In light of research on the behavioral confirmation bias, it is important to know if this process is at work in the interrogation room. Akehurst and Vrij (1999), for example, found that increased movement among police officers triggered increased

movement among their interviewees, behavior that is interpreted as suspicious. This study thus investigated whether the presumption of guilt influences the conduct of interrogators, the behavior of suspects, and the judgments made by neutral observers. This study was conducted in two phases. In Phase I, *suspects* stole \$100 or took part in a related but innocent act, after which they were interviewed via headphones by *interrogators* who were led to believe that most suspects were guilty or innocent. These sessions were taped for analysis and followed by postinterrogation questionnaires. In Phase II, *observers* listened to the taped interviews, judged the suspect as guilty or innocent, and rated their impressions of both participants.

Depending on whether interrogators adopt a theory-driven or data-driven approach to the task, two plausible hypotheses present themselves. On one hand, it is possible that interrogators mistaken in their initial judgments would rethink their beliefs during their interview of the suspect. On the other hand, their initial beliefs may set into motion a self-perpetuating process by which the interrogator's behavior constrains the suspect's behavior, thus leading to the apparent confirmation of that erroneous belief. In light of research conducted in other domains, we predicted the latter hypothesis, that interrogators armed with an expectation of guilt would ask more guilt-presumptive questions, conduct more aggressive interviews, cause suspects to behave defensively, and ultimately steer neutral observers into judging the suspects to be guilty.

PHASE I

Method

Participants and Design

Participating in pairs, 104 introductory psychology students received extra credit to serve as mock interrogators or suspects. Each pair was randomly assigned to one of four cells produced by a 2 (interrogator expectation, guilty vs. innocent) \times 2 (suspect status, guilty vs. innocent) factorial design ($N = 13$ pairs per cell).

Procedure

To prevent the members of a pair from volunteering together or meeting, separate sign-up sheets with time slots staggered at 15-min intervals were posted. Upon their arrival, all participants signed a consent form on which they were apprised of their role and assured, "I understand that the events are simulated, not real, and that all parties are aware that I am participating in a psychological experiment."

After each session, all participants were partially debriefed, thanked, credited for their time, and dismissed. The experimenter explained that the study was part of an ongoing research program examining police investigation practices and, more specifically, the behavior of interrogators and suspects. She then offered to fully explain the experiment via e-mail when all data collection and analyses were complete. As promised, these e-mails were sent after the last session was completed.

Interrogators. Interrogators were told that we were studying the processes of interviewing and interrogation and that they would enact the role of a detective trying

to solve a case. They were then given the details of a mock theft that was committed and a written incident report for their reference. The crime was described as follows: "In Room 100 of Miller House, someone took a key that was hidden behind a VCR on the fireplace. This key was then used to open a locked cabinet where \$100 was stolen from a basket. Whoever did this then put the key back and left with the money."

The experimenter then manipulated interrogator expectations through an instruction about base rates. The effect of this instruction manipulation on a priori expectations were established through pretesting. In the *guilty expectation* condition, she said that four out of every five suspects in the study (80%) actually commit the crime. In the *innocent expectation* condition, she stated that the number is one out of five (20%). Next the experimenter explained the purposes of interrogation. To model the incentives found in actual police work, two essential but sometimes incompatible goals were articulated: (1) to secure a confession and (2) to make an accurate determination of the suspect's guilt or innocence. Thus, while interrogators were directed to seek a confession, they were also offered an incentive (a gift certificate for a free lunch at a sandwich shop) if, ultimately, they correctly judge the suspect's guilt or innocence. Once the task was clear, the experimenter explained that the interrogator would be given a packet of materials and time to plan an interview strategy. To orient them to the task, the first page of the packet contained an introductory excerpt from Inbau et al.'s (2001) *Criminal Interrogation and Confessions*. This excerpt noted the importance of interviewing and interrogation in police work, the need to conduct sessions in private, and the need to use methods that are effective but not extreme.

The second page contained an Interrogation Questions Checklist, a list of 13 questions about the mock theft. Interrogators were instructed to select six questions they might later want to ask. Following a neutral first question ("Where were you and what were you doing during the past hour?") were 12 that were constructed as pairs, with one question of each pair being guilt-presumptive (e.g., "How did you find the key that was hidden behind the VCR?" vs. "Do you know anything about the key that was hidden behind the VCR?"). The questions were randomly interspersed within a single list, not paired, so interrogators were free to select any six.

The third page of the packet contained an Interrogation Techniques Checklist. This list contained 13 techniques and an instruction to select six for possible use. The tactics were derived from Inbau et al.'s (2001) training manual as well as Leo's (1996) observations of actual police interrogations. Among the techniques included were making repeated accusations, exposing inconsistencies in the suspect's story, threatening to involve others, appealing to the suspect's self-interest, promising an end to questioning or leniency in consequences, appealing to the suspect's religious or moral character, offering sympathy and understanding, presenting false evidence, overplaying the seriousness of the offense, and minimizing the seriousness of the offense (e.g., by suggesting it was accidental, spontaneous, or provoked).

Participants were given 10 min and a notepad to formulate an interview plan. They were told that they were free to use or ignore items from the checklists and that their handwritten notes were theirs to keep and use in the interrogation. To ensure anonymity within the sessions, interrogators were asked to assume the name

"Detective Smith" and to refrain from asking for the suspect's name. The interrogators were then left alone to read through the packet, complete the checklists, and formulate a strategy.

Suspects. To ensure that participants did not meet, the suspects were scheduled 15 min after interrogators in a different location. Those assigned to the *guilty condition* were told that their first task was to commit a mock theft. The experimenter described, step-by-step, the actions required for commission of the crime (i.e., enter a designated room, find a key hidden behind a VCR, use the key to unlock a cabinet, take a \$100 bill from a basket, return the key, take the money, and leave). She then handed suspects a written copy of the instruction for their reference, and left. Those in the *innocent condition* were told merely to approach the targeted room, knock on the door, wait for an answer (which was not forthcoming), then meet the experimenter upstairs. The purpose of this latter script was to familiarize innocent suspects with the general area without disclosing details of the crime scene.

Modeling the incentives of people accused of real crimes, the experimenter explained to suspects that their goal was to deny involvement and convince the interrogator of their innocence. One guilty participant in pilot testing had broken down and confessed, so the following instruction was added: "No matter what happens, do not confess. If you are accused of taking money do not admit that you did or try to claim that you didn't really think it was stealing. Admitting having the stolen goods will be considered a confession . . . Imagine yourself in the role of a real suspect and consider how much could be lost by confessing."

All suspects were offered a gift certificate from a local sandwich shop if, after the session, the interrogator judged them to be innocent.

After suspects committed the crime or related innocent act, the experimenter escorted them to a cubicle and gave them time to formulate a defense. They were advised to say or do whatever they saw as necessary to convince the interrogator of their innocence (a truthful story for those who were innocent, a deceptive story for those who were guilty)—but not to cite their participation in a psychology experiment as part of that defense.

The Interrogation

For communication purposes, each participant was seated in a separate cubicle and fitted with headphones and a connected microphone. The two sets were connected through a tape recorder in a nearby control room, enabling the experimenter to record the session. The interrogator and suspect portions were recorded onto separate tracks so they could later be separated for presentation to observers. Interrogators were instructed to conduct interviews of up to 10 min in duration. Once they signaled the end of a session with a closing remark (e.g., "I have no more questions"), the experimenter administered questionnaires, individually, to both participants.

Postinterrogation Questionnaires

Interrogators were asked to judge whether the suspect was guilty or innocent and rate their confidence in that judgment on a 1–10 point scale. They also rated their

suspect's level of anxiety, defensiveness, and friendliness, and the forcefulness of his or her denials. Reflecting on their own behavior, interrogators then rated how hard they tried and how much pressure they put on the suspect to confess. Suspects completed a similar questionnaire framed from their perspective. First, they predicted whether their interrogator believed them to be guilty or innocent and rated their confidence in that prediction. Reflecting on themselves, the suspects rated how anxious, defensive, friendly, and forceful they were in their denials. Then they rated the interrogator's effort and amount of pressure exerted on them to confess, and how anxious, offensive, and friendly he or she was (all on 10-point scales).

Results

Interrogators

Three sources of data were collected from interrogators: preinterrogation selection of questions and techniques; interrogation behavior, as derived from a coding of the tapes; and postinterrogation judgments and self-reports.

Preinterrogation Plans. Interrogators selected six questions from a checklist in which pairs of neutral and guilt-presumptive questions were embedded. To check on the effectiveness of the pairings, we had 61 independent judges rate the degree to which each item was presumptive of guilt. As designed, mean ratings were higher for the guilt-presumptive questions than for the more neutral alternatives ($M_s = 8.17$ & 5.43 , respectively), $t(60) = 13.73$, $p < .0001$ ($d = 1.77$).³ For each interrogator, the number of guilt-presumptive questions chosen was summed, yielding a score that could range from 0 to 6. A 2 (suspect status) \times 2 (interrogator expectation) ANOVA on this score produced a significant and predicted main effect for expectation, $F(1, 48) = 6.24$, $p < .02$ ($d = .72$), as interrogators with guilty expectations chose more guilt-presumptive questions than did those with innocent expectations ($M_s = 3.62$ & 2.60 , respectively).

From a second checklist, interrogators selected six techniques they would consider using. To determine if the expectations manipulation influenced these selections, we had 61 judges evaluate the coerciveness of each technique. The ratings ranged from 4.48 ("Appeal to the suspect's conscience") up to 8.39 ("Tell the suspect that there is independent evidence of his or her guilt"), with an overall mean of 6.63. Based on a median split of ratings, we classified each technique as being high or low in coerciveness. A 2 (suspect status) \times 2 (interrogator expectation) ANOVA revealed no significant effects on the total number of high-coercion techniques selected (overall $M = 3.23$) or on mean ratings of the techniques selected (overall $M = 6.82$), all $F_s < 1$.

Interrogation Behavior. The taped sessions were transcribed and analyzed within a coding scheme that emerged from a presampling of interrogations. The scheme we used involved counting discrete statements that indicated use of specific techniques (rapport, assertions of guilt or disbelief, use of guilt-presumptive questions, appeals to self-interest, appeals to conscience, threats of punishment, promises

³Cohen (1988) has established guidelines for interpreting effect sizes as small ($d = .2$), medium ($d = .5$), and large ($d = .8$).

of leniency, minimization, maximization, and the presentation of false evidence). All transcripts were coded by the second author. To assess interrater reliability, 10 interrogation transcripts were randomly selected for independent coding by the first author. On the simple count of how many discrete tactics were used in a given session, the agreement rate was 100% (on the assignment of tactics to specific categories, the results of which are not reported in this study, the agreement rate was 94%).

The interrogation tapes were coded in their entirety. We also separately analyzed the first minute to examine what techniques were used at the outset—before the influence of the suspect's behavior on the interaction. For both time frames, we totaled the number of techniques used. On average, interrogators used 1.37 techniques in the first minute of interrogation. A two-way ANOVA further revealed that more techniques were used at the outset by interrogators with guilty expectations than by those with innocent expectations ($M_s = 1.92$ & $.81$, respectively), $F(1, 48) = 10.26$, $p < .01$ ($d = .92$). There was no effect for actual guilt and no interaction (both $F_s < 1$). Across entire sessions, which lasted 6.15 min, interrogators used 9.29 techniques. While expectations did not significantly affect this total as it did in the opening minute ($M_s = 10.35$ & 8.23 in the guilty and innocent expectations conditions; $F < 1$), interrogators used more techniques overall when they interviewed suspects who were actually innocent ($M = 11.42$) than when interviewing those who were guilty ($M = 7.15$), $F(1, 48) = 3.80$, $p < .06$ ($d = .56$).

Postinterrogation Judgments and Self-Reports. On the postinterrogation questionnaires, only 16 out of 54 interrogators (30%) judged their suspects guilty. These data were also marginally consistent with the hypothesis that preinterrogation expectations would bias judgments, as 42% of those with guilty expectations judged the suspect guilty, compared to only 19% with innocent expectations, $\chi^2(1, N = 52) = 3.25$, $p < .07$. Indicating that the interrogators could not distinguish truth and deception, actual guilt did not similarly affect these ultimate judgments, $\chi^2(1, N = 52) < 1$. The results are shown in Figure 1.

Our main hypothesis is that the presumption of guilt would set into motion a more pressure-filled interrogation. As such, we had interrogators rate how hard they tried to get a confession and how much pressure they put on the suspect in the process. Similar, surprising, and significant effects were obtained on both measures. Interrogators saw themselves as trying harder to get a confession when the suspect was innocent than when he or she was guilty ($M_s = 6.12$ & 4.73 , respectively), $F(1, 48) = 4.54$, $p < .05$ ($d = .62$). They also said they had exerted more pressure on the suspect who was innocent than guilty ($M_s = 5.34$ & 4.08 , respectively), $F(1, 48) = 5.11$, $p < .05$ ($d = .65$). In neither instance was there a significant effect for interrogator expectations or an interaction. In short, interrogators saw themselves as the most aggressive when they interviewed suspects who—unbeknownst to them—were truly innocent.

Suspects

Suspects filled out a questionnaire that was the mirror image of that given to interrogators. First, they predicted whether they were judged guilty or innocent and rated themselves and their interrogator on the dimensions previously described.

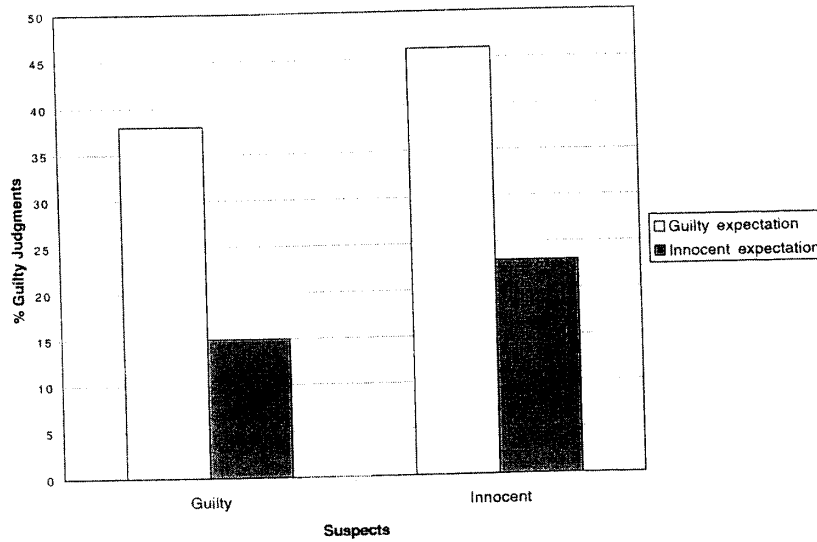


Fig. 1. Interrogator judgments as a function of their expectations and suspects' guilt or innocence.

Overall, only 15 suspects (29%) thought their interrogator would judge them guilty. This low rate was similar to that obtained for the actual judgments, which is not surprising, as interrogators were given no extrinsic evidence of culpability. These predictions were not affected by the suspect's actual guilt or by interrogator expectations (though suspects were somewhat more likely to anticipate a guilty judgment from interrogators who had expectations of guilt rather than innocence, 35% vs. 23%).

Paralleling the results of interrogator self-ratings, there were significant effects on suspects' perceptions of the interrogator's behavior. Specifically, those who were innocent as opposed to guilty saw their interrogators as trying harder to get a confession ($M_s = 5.50$ & 4.12 , respectively), $F(1, 48) = 4.09$, $p < .05$ ($d = .58$). Those who were innocent as opposed to guilty also saw the interrogator as exerting more pressure on them ($M_s = 5.38$ & 3.88), $F(1, 48) = 3.85$, $p < .06$ ($d = .57$). There were no other differences on these measures. Across conditions, suspects also did not differ in self-ratings of how anxious, defensive, or friendly they were during interrogation, or in how forcefully they denied their involvement in the crime.

PHASE II

In Phase I of this study, interrogators' expectations influenced their perceptions and behavior before, during, and after interrogation. Prior to interrogation, those armed with guilty expectations chose more guilt-presumptive questions, used more techniques at the outset of interrogation and, in the end, judged the suspects guilty. While actual guilt and innocence did not affect the judgments that interrogators

ultimately made, it did affect their behavior. Specifically, both sets of participants said that interrogators exerted more pressure to confess on suspects who, unbeknownst to them, were truly innocent than on those who were guilty.

To determine if the suspects themselves were influenced by the manipulated expectations, we had independent observers listen to the taped sessions and rate the participants. On the basis of past studies of behavioral confirmation biases, we predicted that interrogators with guilty expectations would unwittingly elicit behavior in suspects that others see as consistent with that expectation.

Method

Seventy-eight introductory psychology students participated for extra credit. Scheduled in pairs, and randomly assigned, these observers listened to taped sessions of (1) interrogators only, (2) suspects only, or (3) both tracks of four interrogations (one from each cell of the 2×2 design).

Each tape recording contained four interrogations that were randomly grouped except for the constraint that each cell of the design be represented (i.e., guilty suspect-guilty expectation, guilty suspect-innocent expectation, innocent suspect-guilty expectation, innocent suspect-innocent expectation). Across tapes, the ordering of conditions was partially counterbalanced (1234, 2341, 3412, 4123) to ensure that each group appeared in each position an equal number of times.

The experimenter told participants that the purpose of the study was to examine the processes of interviewing and interrogation and that they would be asked to evaluate tapes of criminal interrogations. Observers in the dual-tracks condition were told only that they would answer various questions following each interrogation. Those in the suspects-only condition were further told, "You will only hear the suspects' responses, not the interrogators' questions, so there will be moments of silence when the interrogators would have been speaking. . . . Although you will not actually hear the interrogator, you may be asked to infer things about his or her behavior."

Observers in the interrogators-only condition also received the latter instruction, except the words "suspect" and "interrogator" were interchanged. Importantly, observers were not told about the variations in suspect status or interrogator expectation. The experimenter then started the tape, left the room, and returned after each segment to pause the tape and administer a questionnaire.

First and foremost, observers judged whether each suspect was guilty or innocent and rated their confidence in that judgment. They were also asked to indicate whether the interrogator had judged the suspect guilty or innocent and their confidence in that judgment. To see if observers could discern interrogators' initial expectations, we also had them rate the degree to which the interrogator presumed the suspect's guilt at the outset. Then they rated how hard the interrogator tried to get a confession and how much pressure he or she put on the suspect. Turning to the suspect, observers rated how anxious and defensive he or she was, how firmly he or she denied the accusation, and how plausible his or her alibi seemed. All ratings were made on 10-point scales.

Results

Although observers were blind to the variations in suspect status and interrogator expectations (in fact, they were unaware that such variations existed), these factors still influenced them. Overall, they judged as guilty 42% of truly guilty suspects, compared to only 28% of those who were innocent, $\chi^2(1, N = 306) = 6.28, p < .02$ ($r = .14$). While the difference is not quite significant, they also judged as guilty 40% of suspects in the guilty expectations condition, compared to only 30% in the innocent expectations condition, $\chi^2(1, N = 306) = 3.19, p < .08$ ($r = .10$).

Perceptions of Interrogators' Beliefs

Phase I results revealed that interrogators were affected by the manipulated expectations. Were these resulting behavioral differences detectable by observers unaware that such expectations were even in place? To answer this question, we asked observers to predict the judgments that interrogators were likely to make. The result was clear and important: 76% of interrogators with guilty expectations were seen as guilt-biased, compared to only 61% of those with innocent expectations, $\chi^2(1, N = 306) = 7.98, p < .005$ ($r = .16$). There were no significant differences as a function of actual guilt, $\chi^2(1, N = 306) < 1$.

On all scalar measures, the data were submitted to a 3 (observer condition) \times 2 (suspect status) \times 2 (interrogator expectation) mixed ANOVA.⁴ On ratings of the interrogator's initial presumption of guilt, observers correctly detected a difference, seeing those armed with guilty expectations as more presumptive of guilt ($M = 6.78$ vs. 6.00), $F(1, 294) = 7.18, p < .01$ ($d = .31$). Consistent with the paradoxical data reported earlier, a significant main effect for suspect status also indicated that interrogators were seen as more presumptive of guilt when paired with suspects who were actually innocent rather than guilty ($M_s = 6.77$ vs. 6.00), $F(1, 294) = 7.15, p < .01$ ($d = .31$).

Perceptions of Interrogators' Behavior

As with the suspects and interrogators, observers rated the amount of pressure and effort the interrogator exerted to get a confession. Paralleling the results obtained from the interactants, a main effect on perceived pressure indicated that interrogators with guilty expectations were seen as exerting more pressure than those with innocent expectations ($M_s = 6.40$ & 5.72, respectively), $F(1, 294) = 4.73, p < .05$ ($d = .25$). Importantly, too, interrogators were seen as exerting far more pressure on suspects who were truly innocent than on those who were guilty ($M_s = 6.52$ & 5.60, respectively), $F(1, 294) = 8.61, p < .005$ ($d = .34$).

Similar results were obtained on the related question of how hard the interrogator tried to get a confession. Those with guilty expectations were seen as trying harder than those with innocent expectations ($M_s = 6.40$ & 5.71), $F(1, 294) = 4.96, p < .05$ ($d = .26$). Again, interrogators were seen as trying harder when paired with suspects who were truly innocent than guilty ($M_s = 6.51$, & 5.60), $F(1, 294) = 8.33, p < .005$ ($d = .34$). A significant two-way interaction was also obtained on this measure,

⁴Effects for the observer condition are not reported in this study but are available upon request.

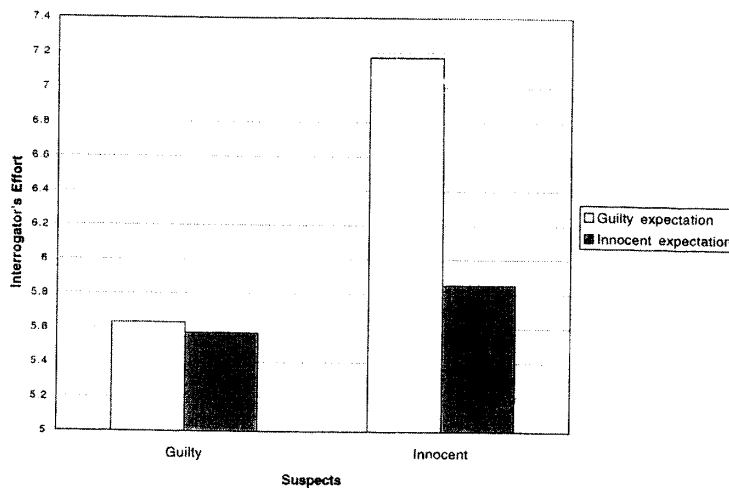


Fig. 2. Observer ratings of how hard interrogators tried to get a confession as a function of interrogators' expectations and suspects' guilt or innocence.

$F(1, 294) = 3.93, p < .05 (d = .23)$. Figure 2 shows that interrogators were seen as trying hardest to get a confession when armed with guilty expectations and interviewing a truly innocent suspect ($M = 7.17$). At $p < .05$, this cell was significantly higher than all the others which, in turn, did not differ from each other (combined $M = 5.68$).

Perceptions of Suspects

Did the added interrogative pressure in the guilty expectations condition, particularly against innocent suspects, cause these suspects to alter their behavior in noticeable ways? Of particular interest was whether suspects were forced to behave defensively. On this measure, a significant main effect confirmed that suspects in the guilty expectations condition were seen as more defensive than those in the innocent expectations condition ($M_s = 5.74$ & 5.11 , respectively), $F(1, 294) = 4.94, p < .05 (d = .26)$. No differences emerged as a function of suspect's actual guilt or innocence. Observers also rated how plausible they found the suspects' alibis and denials. Interestingly, the stories told by innocent suspects were seen as more plausible than those told by guilty suspects ($M_s = 6.62$ & 6.03 , respectively), $F(1, 294) = 4.37, p < .05 (d = .24)$.

DISCUSSION

Modern American police interrogations are theory-driven social influence events, founded upon a presumption of guilt, that sometimes draw false confessions from innocent people. Police justify their use of powerful techniques by claiming that they can make accurate initial judgments of truth and deception—judgments that determine whether to proceed (Inbau et al., 2001; Vessel, 1998). The problem is,

trained law enforcement investigators are not reliably more accurate than the average person—and often fail to exceed chance level performance despite high levels of confidence (DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991)—a pattern of results also found when judgments were made from mock crime interrogations (Kassin & Fong, 1999; Meissner & Kassin, 2002). Thus, the pivotal decision to interrogate a suspect is based on prejudgments of guilt confidently made but frequently in error.

In light of the profound risks associated with modern police interrogations, we developed a laboratory paradigm to examine the process. By independently varying interrogators' beliefs and suspects' actual guilt or innocence, we tested the hypothesis that interrogator expectations would trigger a range of behavioral confirmation effects, ultimately biasing perceptions of guilt. The results strongly supported this hypothesis. Interrogators with guilty expectations chose more guilt-presumptive questions, used more techniques in their interrogation (including the presentation of false evidence and promises of leniency), and were more likely to see suspects in incriminating terms, exhibiting a 23% increase in guilty judgments relative to those with innocent expectations. The suspect's actual guilt or innocence had a paradoxical and particularly disturbing effect on interrogators, leading them to exert the most pressure on innocent suspects. In short, a presumption of guilt triggered aggressive interrogations, which constrained the behavior of suspects and led others to infer their guilt—thus confirming the initial presumption.

When coupled with the finding that police investigators do not judge truth and deception at high levels of accuracy, despite confidence, the results also suggest that cognitive confirmation biases may exacerbate the problem—that erroneous prejudgments of guilt color the information gathering process such that plausible denials are discounted or misinterpreted. One would hope that erroneous prior expectations would not blind interrogators to contradictory evidence in the form of forceful and plausible assertions of innocence. However, our study suggests otherwise, as does other research showing that suspicion impedes the detection of truths and lies (Burgoon, Buller, Ebesu, & Rockwell, 1994).

Turning from the perceivers to their targets, results showed that suspects played an all-too-predictable role in the behavioral confirmation process. Paralleling Snyder and Swann's observation that a confirmatory approach to questioning constrains a target's response options, suspects in the guilty expectations condition became noticeably more defensive. It is not clear what aspects of their behavior gave rise to this impression. But it is not hard to imagine, as our results suggest, that people trapped in coercive interrogations may well look away, slouch, sigh in despair, or exhibit other cues that trained police regard as indicators of guilt (Inbau et al., 2001).

Remarkably, whatever suspects did to be perceived as defensive, those who were presumed guilty by interrogators later tended to be judged as such by neutral observers. By neglecting to account sufficiently for the way in which the suspect's behavior was shaped by the interrogative situation, observers thus committed the fundamental attribution error, or correspondence bias (Gilbert & Malone, 1995; Jones, 1990; Ross, 1977). Even the interrogators failed to account for the impact of their own behavior, a result that resembles the Gilbert and Jones (1986) finding that perceivers inferred an actor's true attitude from the position espoused in a low-choice speech even when they had personally assigned the actor to assert that position.

Particularly disturbing was the effect of innocence on the perceiver-target interaction. According to neutral observers, innocent suspects told plausible denial stories. Yet these suspects brought out the worst in the guilt-presumptive interrogators (consistently, by everyone's account, this cell produced the most pressure-filled interviews). Interrogators who approached the task with a guilty base-rate expectation never stopped to reevaluate this belief—even when paired with innocent suspects who issued plausible denials. Rather, it appears that they interpreted the denials as proof of a guilty person's resistance—and redoubled their efforts to elicit a confession.

While it is significant that interrogator expectations biased the selection of questions, the use of techniques, the pressure exerted, and the reactive behavior of suspects, the most devastating possible effect is on the judgments of observers who, in the criminal justice setting, represent trial judges, lawyers, and jurors. The impressions formed by observers—who were not privy to the variations in suspect status or interrogator expectations—supported the hypothesis that beliefs can transform important outcomes in the interrogation room. Two interesting results were obtained from the observers. First, they were able to distinguish interrogators with guilty and innocent expectations and detect differences in their behavior. Listening to interrogators who—unbeknownst to them—had been armed with guilty expectations, observers “heard” a presumption of guilt and perceived greater effort and more pressure to get a confession (they were able to make these distinctions even when listening to tapes that contained only the suspects). Second, observers saw the innocent suspects as having told the more plausible denial stories, suggesting that there were detectable narrative cues indicative of guilt and innocence—cues apparently missed by the interrogators themselves.

Most importantly, the observer data indicated that suspects had behaviorally confirmed the prior beliefs of their interrogators. Observers rated suspects in the guilty expectations condition as more defensive than those in the innocent expectations condition. By a 10% margin, they also tended to judge more of them guilty. Although not quite significant, this latter tendency suggests the possibility of a far-reaching influence of interrogator expectations on the final link in the behavioral confirmation process. As noted earlier, this result is also consistent with the fundamental attribution error, as observers did not sufficiently adjust their judgments of suspects for the social pressures they knew to be present in the situation.

Overall, this study suggests that behavioral confirmation is a risk that is incurred when the police presume guilt as a basis for interrogation. Questions may be raised about the extent to which our laboratory paradigm adequately captured the dynamics of real police-suspect interactions. All participants were college students, they were fully aware that the crime was simulated, and their performance incentives were small relative to the high stakes of a real criminal investigation. However, these aspects of the study were constant across variations in interrogator expectations and suspect status—and still, the results for both factors were consistent and in the predicted directions.

One might also argue that participants discerned the experimental hypothesis and responded to the demand characteristics accordingly. There was, however, no evidence in pretesting or in the actual experiment to support any such speculation.

All participants were queried after each session. Suspects did not report knowing that interrogators were preset with an expectation or that other suspects were assigned to a different condition; interrogators did not question the veracity of the base rate information or even know that other interrogators received a different base rate. Furthermore, all suspects were offered the same incentive to be judged innocent; all interrogators had the same incentive for getting a confession and for making a correct judgment.

We believe that our results may underestimate the risks incurred in criminal justice settings in which police are trained to have confidence, unfounded as it is, in their ability to divine guilt from a suspect's interview behavior; in which police are trained in the use of psychological interrogation techniques; in which police are motivated by career aspirations to solve cases; in which police pressure suspects over the course of hours, not minutes, of interrogation; and in which the presumption of guilt is a self-generated hypothesis—a condition that exacerbates confirmation biases (Haverkamp, 1993).

Questions may also be raised about the observers, who listened to the tapes without a context of other evidence and who rendered their judgments individually rather than conferring in groups, as juries do. These are notable possible limitations. At the same time, the observer results were highly consistent with research in other domains, as in showing that people are not highly accurate in making judgments of truth and deception and that people tend to underweight the impact of situational forces that constrain others. With regard to group deliberation, mock jury studies have shown that the extralegal biases that plague individual jurors persist in groups despite deliberation (Carretta & Moreland, 1983; Kramer, Kerr, & Carroll, 1990). More to the point, Kassin and Fong (1999) found that observers were not more accurate at judging true and false denials when they made the decisions in groups than alone.

This study suggests that police interrogations founded upon an interview-based presumption of guilt can set into motion a biased chain of events. In light of the potentially tragic implications, as seen in recent accounts of false confessions, more research is needed to test various aspects of this process. Studies have shown that initial judgments of guilt are not made at high levels of accuracy, even by trained professionals, but more needs to be known about factors that affect these judgments, such as a suspect's sex, race, age, intelligence, physical appearance, or even social anxiety and interaction style. Research is also needed to examine the motives of interrogators, for example, by systematically varying the relative emphases on speed vs. accuracy and error, or on the elicitation of confessions vs. the due process rights of the accused. Finally, from a remedial perspective, it remains to be seen whether the processes uncovered in this study can be prevented inside the interrogation room by educating police about the power of expectations and the associated risks.

REFERENCES

- Akehurst, L., Kohnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology, 10*, 461-471.
- Akehurst, L., & Vrij, A. (1999). Creating suspects in police interviews. *Journal of Applied Social Psychology, 29*, 192-210.

- Burgoon, J. K., Buller, D. B., Ebesu, A. S., & Rockwell, P. (1994). Interpersonal deception V. Accuracy in deception detection. *Communication Monographs*, *61*, 303-325.
- Carretta, T. R., & Moreland, R. L. (1983). The direct and indirect effects of inadmissible evidence. *Journal of Applied Social Psychology*, *13*, 291-309.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Darley, J. M., & Fazio, R. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist*, *35*, 867-881.
- DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology*, *16*, 249-267.
- Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, *46*, 913-920.
- Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science*, *10*, 263-266.
- Gilbert, D. T., & Jones, E. E. (1986). Perceiver-induced constraint: Interpretations of self-generated reality. *Journal of Personality and Social Psychology*, *50*, 269-280.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21-38.
- Gudjonsson, G. (2003). *The psychology of interrogations and confessions: A handbook*. New York: John Wiley & Sons.
- Haverkamp, B. E. (1993). Confirmatory bias in hypothesis testing for client-identified and counselor self-generated hypotheses. *Journal of Counseling Psychology*, *40*, 303-315.
- Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). *Criminal Interrogation and Confessions* (4th ed.). Gaithersburg, MD: Aspen.
- Jones, E. E. (1990). *Interpersonal perception*. New York: Freeman.
- Kassin, S. M. (1997). The psychology of confession evidence. *American Psychologist*, *52*, 221-233.
- Kassin, S. M., & Fong, C. T. (1999). "I'm innocent!": Effects of training on judgments of truth and deception in the interrogation room. *Law and Human Behavior*, *23*, 499-516.
- Kassin, S. M., & Kiechel, K. L. (1996). The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science*, *7*, 125-128.
- Kassin, S. M., & Neumann, K. (1997). On the power of confession evidence: An experimental test of the "fundamental difference" hypothesis. *Law and Human Behavior*, *21*, 469-484.
- Kassin, S. M., & Sukel, H. (1997). Coerced confessions and the jury: An experimental test of the "harmless error" rule. *Law and Human Behavior*, *21*, 27-46.
- Kassin, S. M., & Wrightsman, L. S. (1980). Prior confessions and mock juror verdicts. *Journal of Applied Social Psychology*, *10*, 133-146.
- Kramer, G. P., Kerr, N. L., & Carroll, J. S. (1990). Pretrial publicity, judicial remedies, and jury bias. *Law and Human Behavior*, *14*, 409-438.
- Leo, R. A. (1996). Inside the interrogation room. *Journal of Criminal Law and Criminology*, *86*, 266-303.
- Leo, R. A., & Ofshe, R. J. (1998). The consequences of false confessions: Deprivations of liberty and miscarriages of justice in the age of psychological interrogation. *Journal of Law and Criminology*, *88*, 429-496.
- McNatt, D. B. (2000). Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology*, *85*, 314-322.
- Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior*, *26*, 469-480.
- Mortimer, A., & Shepherd, E. (1999). Frames of mind: Schemata guiding cognition and conduct in the interviewing of suspected offenders. In A. Memon & R. Bull (Eds.), *Handbook of the psychology of interviewing*. Chichester, England: Wiley.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175-220.
- Radelet, M. L., Bedau, H. A., & Putnam, C. E. (1992). *In spite of innocence: Erroneous convictions in capital cases*. Boston: Northeastern University Press.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart, & Winston.
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, *10*, 174-221.
- Scheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual innocence: Five days to execution and other dispatches from the wrongly convicted*. New York: Doubleday.
- Snyder, M. (1992). Motivational foundations of behavioral confirmation. *Advances in Experimental Social Psychology*, *25*, 67-114.
- Snyder, M., & Stukas, A. A. (1999). Interpersonal processes: The interplay of cognitive, motivational, and behavioral activities in social interaction. *Annual Review of Psychology*, *50*, 273-303.

- Snyder, M., & Swann, W. B. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, *36*, 1202-1212.
- Vessel, D. (1998, October). Conducting successful interrogations. *FBI Law Enforcement Bulletin*, 1-6.
- Wald, M., Ayres, R., Hess, D. W., Schantz, M., & Whitebread, C. H. (1967). Interrogations in New Haven: Impact of *Miranda*. *Yale Law Journal*, *76*, 1519-1648.
- Wrightsmann, L. S., & Kassin, S. M. (1993). *Confessions in the courtroom*. Newbury Park, CA: Sage.
- Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology*, *14*, 1-59.