The Social Psychology of False Confessions

Saul M. Kassin
John Jay College of Criminal Justice

Inspired by DNA exoneration cases and other wrongful convictions of innocent people who had confessed to crimes they did not commit, and drawing from basic principles of social perception and social influence, a vast body of research has focused on the social psychology of confessions. In particular, this article describes laboratory and field studies on the “Milgramesque” processes of police interviewing an interrogation, the methods by which innocent people are judged deceptive and induced into confession, and the rippling effects of these confessions on judges, juries, lay and expert witnesses, and the truth-seeking process itself. This article concludes with a discussion of social and policy implications—including a call for the mandatory video recording of entire interrogations, blind testing in forensic science labs, and the admissibility of confession experts in court.

The 2012 film, The Central Park Five, tells a horrific tale about a profound, disturbing, and all-too-common manifestation of social influence. In 1989, a female jogger was raped, beaten, and left for dead in New York City’s Central Park. She managed to survive but could not remember anything about the attack—then or now. Within 72 hours, five African- and Hispanic-American boys, 14–16 years old, confessed to the assault. Solely on the basis of their oral confessions, four of which were videotaped, and all of which were vividly detailed, though often erroneous, the boys were convicted and sentenced to prison. Almost nobody questioned their guilt—even though there was no other evidence; even though DNA tests on sperm that was recovered from the victim and her clothing had excluded them all.

Thirteen years later, Matias Reyes, in prison for two rapes and a murder committed subsequent to the jogger attack, stepped forward to admit that he was the Central Park jogger rapist and that he acted alone. Reinvestigating the case, the Manhattan District Attorney questioned Reyes and discovered that he had accurate
and independently corroborated guilty knowledge of the crime and that the DNA samples originally recovered from the victim belonged to him. The DA issued a report that dismantled the confessions and other evidence. Shortly thereafter, the original convictions were overturned. Then in September of 2014, 25 years after the crime was committed, New York City awarded the defendants a $41 million settlement. The Central Park jogger case now stands as a shocking demonstration of five false confessions resulting from a single high-profile investigation (Kassin, 2002; for an extensive description of this case, see Burns, 2011).

An Historical Overview

It is hard to imagine any aspect of human behavior more counterintuitive than the proposition that an innocent person, as a function of social pressure, would knowingly confess to a heinous crime that he or she did not commit—an act that can cost the confessor liberty, and sometimes even his or her life. Yet false confessions occur with some measure of regularity throughout recorded history; in countries all over the world; and in criminal justice, military, and corporate settings (Drizin & Leo, 2004; Gudjonsson, 2003; Kassin, 1997; 2008; Kassin & Gudjonsson, 2004; Kassin et al., 2010).

More than one hundred years ago, Harvard psychology professor Hugo Munsterberg (1908) wrote about “untrue confessions” in his book, *On the Witness Stand*. Brilliant as they were, Munsterberg’s early insights did not inspire research within psychology, a yet-to-become applied science; nor did it inspire concern within the law. Sixty years later, when the United States Supreme Court in *Miranda v. Arizona* (1966) described American police interrogation practices as “inherently coercive,” there was still only a smattering of isolated articles on the subject. Bem (1966) published an empirical article in the *Journal of Personality and Social Psychology* entitled “Inducing belief in false confessions” in which he offered a self-perception analysis in the laboratory of how saying (induced confession) can lead to believing (feelings of guilt). The following year, Zimbardo (1967) published a social–psychological analysis of the police interrogation process in the inaugural issue of *Psychology Today*. At about the same time, occasional law review articles were published that offered “psychological” analyses of confessions—such as Driver’s (1968) “Confessions and the Social Psychology of Coercion,” which appeared in the *Harvard Law Review*, and Foster’s (1969) “Confessions and the Station House Syndrome,” which likened police interrogation to a trance-like state of hypnosis.

In 1985, Lawrence Wrightsman and I wrote a chapter on “Confession Evidence” in which we reviewed the law, described common social influence practices of police interrogation, reviewed the scant research literature, and introduced a taxonomy that is now widely used to distinguish three types of false
confessions—voluntary, coerced-compliant, and coerced-internalized (Kassin, 1997; Kassin & Wrightsman, 1985; Wrightsman & Kassin, 1993).

By drawing from the literature on normative and informational social influences (e.g., Asch, 1956; Kelman, 1958; Sherif, 1936), we distinguished, first, between the types of false confessions that arise when innocent people volunteer self-incriminating statements without pressure (often to high-profile crimes, as when 200 people volunteered false confessions to the 1932 kidnapping of Charles Lindbergh’s baby son) and those that come about through the interpersonal process of interrogation. Within the latter category, we then distinguished between cases in which innocent people are moved from denial to confession in an act of mere behavioral compliance, to escape a harsh interrogation or because they are led to perceive that confession serves their own self-interest (when it comes to stress, discomfort, and the deprivation of need states, everyone has a breaking point) and those rarer instances of internalization in which innocent people, subjected to highly misleading claims about the evidence, question their own innocence, come to infer their own guilt, and in some cases confabulate memories to support that inference. This taxonomy has provided a useful framework for the study of false confessions and has since been used, extended, and refined by others (Gudjonsson, 2003; Inbau et al., 2013; McCann, 1998; Ofshe & Leo, 1997).

Today, there is a substantial empirical literature on false confessions. Founded in 1992 by Barry Scheck and Peter Neufeld, the Innocence Project began to report on cases in which wrongfully convicted individuals were exonerated and set free through new forms of DNA testing of biological materials (e.g., blood, hair, semen, skin) previously collected and preserved. At present, the Innocence Project has reported on more than 300 such postconviction DNA exonerations, all involving rape and/or murder. In nearly 30% of these cases, false confessions were a contributing factor—and this sample represents only a fraction of all wrongful convictions (Garrett, 2011; www.innocenceproject.org/).

Contemporary research on false confessions has analyzed various aspects of the confession-taking process and has relied on a range of methodologies. One approach has involved a focus on actual case studies and aggregations of individual cases based on archived records. Other empirical methods have included naturalistic observations of live and videotaped police interrogations; self-report surveys and interviews that purport to describe normative practices and beliefs; correlational studies that link various personal suspect characteristics and the tendency to confess; and controlled experiments—in laboratory and field settings—designed to assess police judgments of truth and deception, the effects of certain interrogation tactics on confessions, and the impact that confessions have not only on judges and juries but, more recently, on lay witnesses and forensic examiners. This literature is now sufficiently mature and has served as the basis of an official White Paper of the American Psychology-Law Society, only the second in the history of this professional organization (Kassin et al., 2010; for a description of
the process by which this White Paper was vetted and produced, see Thompson, 2010).

This article reviews the current theoretical and empirical research literature on false confessions. In particular, this review is framed around four social psychologically loaded questions concerning the processes by which false confessions occur and wreak havoc on individuals and the criminal justice system as a whole: (1) why are innocent people often misidentified for suspicion during a preinterrogation interview? (2) What situational forces during the structure and tactics of interrogation lead innocent people confess to crimes they did not commit? (3) What adverse consequences follow from confession—both in the effects on judges and juries in the courtroom, and in the effects on witnesses, forensic examiners, and the truth-seeking process itself? Following a review of research on these questions, this article addresses the social policy implications—namely, (4) what can be done to prevent future miscarriages of justice based on false confessions?

**The Preinterrogation Interview: Judgments of Truth and Deception**

During an investigation, police identify one or more suspects for interrogation. Sometimes, this identification is based on witnesses, informants, a suspect’s past crimes, or other rational extrinsic evidence. Often, however, this identification is based on nothing more than a first impression formed during a preinterrogation interview. As described in *Criminal Interrogations and Confessions*, an influential manual on interrogation first published by Inbau and Reid (1962) and now in its fifth edition (Inbau, Reid, Buckley, & Jayne, 2013), police are trained in a two-step process called the Reid Technique by which the highly confrontational, accusatory process of interrogation is preceded by a neutral, information-gathering interview, the main purpose of which is to help determine if the suspect is truthful or deceptive; innocent or guilty.

To help investigators at distinguishing truth telling from deception, Inbau et al. (2013) provide investigators with a list of “Behavioral Analysis Interview” (BAI) questions (e.g., What do you think should happen to the person who did this?) and instruct them to detect lies by observing changes in the suspect’s verbal and nonverbal behavior (e.g., eye contact, pauses, posture, fidgety movements). Based on one study, Inbau et al. (2013) claim that training in the Reid technique produces an exceedingly high level of accuracy. Yet the claim is based on data from a single flawed study in which Horvath, Jayne, and Buckley (1994) selected 60 interview tapes from the Reid collection, the ground truths of which could not be established with certainty. Then they edited the tapes in a manner that was not specified, showed these edited tapes to four experienced in-house staff employees of their training, and concluded from their judgments that the Reid technique produced high levels of accuracy (no comparison group of untrained or lay evaluators was included).
Equally as important, the claim of high levels of accuracy as a function of training is grossly out of step with the bulk of basic research, which has consistently shown that the demeanor cues touted by the Reid technique do not significantly discriminate between truth-telling and deception (DePaulo et al., 2003). This research shows that on average laypeople are only 54% accurate; that training seldom produces appreciable improvement compared to naive control groups; and that police, judges, psychiatrists, customs inspectors, and other so-called experts perform only slightly better, if at all (for recent reviews, see Bond & DePaulo, 2006; Hartwig & Bond, 2011; Vrij, 2008).

Experiments specifically designed to test the BAI have also failed to support the efficacy of the approach. One study showed that the verbal and nonverbal demeanor cues that investigators are instructed to use do not increase judgment accuracy. Kassin and Fong (1999) randomly trained some lay participants but not others in the use of the “behavioral symptoms” cited by the Reid technique. All participants then watched videotaped interviews of mock suspects.

By random assignment, half of these taped suspects actually committed one of four mock crimes, while seeking to evade detection. Following instructions from the experimenter, some guilty suspects had shoplifted jewelry or a stuffed animal from a local gift store during store hours; others broke into a campus building to steal the answer key to an exam, thereby setting off an alarm; others vandalized a public wall by chalking obscenities on it; still others logged onto a college computer and broke into another student’s private email account with her username and password. In contrast, innocent suspects were instructed to merely report to one of these four sites without actually committing a mock crime. In all cases, guilty and innocent suspects alike were then apprehended by a young man posing as a security officer and brought into the laboratory for questioning. Before questioning, suspects were incentivized to be judged innocent by a threat of a brief detention; all denied their involvement.

As in the typical laboratory experiment, Kassin and Fong’s (1999) participant observers could not reliably differentiate between suspects who denied involvement truthfully and those who lied. Moreover, those participants who underwent training in the Reid Technique were less accurate, more confident, and more biased toward seeing deception. In a follow-up study using these same taped interviews, Meissner and Kassin (2002) tested experienced police investigators and found that they too exhibited these erroneous and biased tendencies.

In another study, Vrij, Mann, and Fisher (2006) had some participants commit a mock crime while others did not. All participants were then interviewed about the crime using the BAI questions. Overall, the results showed that verbal and behavioral responses to the questions did not significantly distinguish between truth tellers and liars in the Reid-predicted manner. Of the significant differences that were found, innocent participants were more likely than those who were guilty to exhibit behaviors supposedly associated with deception (such as crossing legs
and changing posture). In short, the questions provided by the Reid technique failed to produce diagnostic responses.

Reasonably, proponents of the Reid technique and others have suggested that laboratory experiments lack external validity because they often involve college student participants lying or telling the truth in a low stakes situation (Buckley, 2012; O’Sullivan, Frank, Hurley, & Tiwana, 2009). In a meta-analysis of studies spanning over 40 years, however, Hartwig and Bond (2014) found that the detectability of deception did not differ as a function of whether the speaker was a college student or nonstudent, whether the speaker’s motivation level was high or low, or whether the speaker lied in a monologue or in a question-and-answer interview.

Recent research suggests a possible explanation for the empirical failures of lie-detection training: By focusing on such cues as gaze avoidance, fidgeting, and changes in posture, the Reid Technique merely formalizes the folk wisdom that laypeople already use without much success (Masip, Barba, & Herrero, 2012; Masip, Herrero, Garrido, & Barba, 2011). In contrast, social psychologists have identified other ways to improve police lie detection performance. In one important line of research, for example, Vrij, Fisher, Mann, and Leal (2006) theorized that because lying is more effortful than telling the truth, interviewers should tax a suspect’s cognitive load and attend to cues that betray cognitive effort. Thus, when interviewers challenge truth tellers and liars—for example, by having them recount their stories in reverse chronological order—observers become more accurate in their ability to distinguish between truthful and deceptive accounts (Vrij & Granhag, 2012; Vrij, Granhag, & Porter, 2011).

**Inside Interrogation: Police-Induced False Confessions**

When the United States Supreme Court in *Miranda v. Arizona* (1966) sought to understand what transpires during an in-custody police interrogation, a process that the Court ultimately described as “inherently coercive,” it turned to Inbau and Reid’s (1962) manual, *Criminal Interrogations and Confessions* (see Inbau et al., 2013). In this approach, investigators are advised to isolate the suspect in a small, private, windowless room, which increases anxiety and, hence, the incentive to escape. A nine-step process then ensues involving the interplay of negative and positive incentives. On the one hand, the interrogator confronts the suspect with accusations of guilt, assertions that may be bolstered by evidence, real or manufactured, and refuses to accept objections and denials. On the other hand, the interrogator offers sympathy and moral justification, introducing “themes” that minimize the crime and lead suspects to see confession as an expedient means of escape. The use of these techniques has been documented in naturalistic observational studies (Feld, 2013; King & Snook, 2009; Leo, 1996a) and in surveys of police (Kassin et al., 2007; also see Meyer & Reppucci, 2007; for critiques, see...
Kassin, 1997, 2006; Starr, 2013; for an historical account of interrogation in the United States, see Leo, 2008).

**Structural Aspects of Interrogation**

It goes without saying that individuals differ in the extent to which they comply or can resist figures of authority pressing for a confession. Over the years, individual differences research has focused on suspect characteristics that are associated with compliance, suggestibility, and other forms of social influence. In particular, research has shown that juveniles—adolescents who exhibit immaturity of judgment across a range of domains (Owen-Kostelnik, Reppucci, & Meyer, 2006)—as well as adults with intellectual disabilities and various psychological disorders (for a review, see Gudjonsson, 2003) are at risk in this situation.

Individual differences notwithstanding, there are two structural aspects of a typical police interrogation that are striking to this social psychologist. The first concerns the fact that interrogation is, by definition, a guilt-presumptive process—a theory-driven social interaction led by an authority figure who has formed a strong belief about the suspect, sometimes through a pre-interrogation interview, and who single-mindedly measures success by whether he or she is able to extract a confession. The guilt-presumption that accompanies the start of interrogation thus provides fertile ground for the operation of cognitive and behavioral confirmation biases.

In a study that demonstrates the point, Kassin, Goldstein, and Savitsky (2003) had some participants but not others commit a mock crime, after which all were questioned by participant interrogators who by random assignment were led to presume guilt or innocence. In a study that was modeled after Snyder and Swann’s (1978) classic confirmatory hypothesis-testing experiment in which participants were led to believe they were interviewing people who were introverted or extroverted, interrogators who presumed guilt chose to ask more incriminating questions, conducted more coercive interrogations, and tried harder to get the suspect to confess. In turn, this more aggressive style made the suspects sound defensive and led observers who later listened to the tapes to judge them as guilty, even when they were innocent. Follow-up research has confirmed this chain of events in suspect interviews (Hill, Memon, & McGeorge, 2008; Narchet, Meissner, & Russano, 2011).

A second striking feature of interrogation concerns the “Milgramesque” nature of the process itself. Fifty-one years ago, Milgram (1963) published his classic, the first of 18, obedience experiment in which 65% of participants obeyed an experimenter’s commands to deliver increasingly painful electric shocks to a confederate—in their view, up to 450 V. Milgram (1974) described his elegant method, summarized his findings, and theorized about the implications in his book *Obedience to Authority* (for reviews, see Blass, 2004; Miller, 1986; Perry, 2013).
The parallels between police interrogations and the protocol that Milgram established to elicit obedience are striking. In both venues, the subject is isolated—without access to friends, family, or other means of social support—in a specially designed space, whether the laboratory or an interrogation room. In both venues, the subject is confronted by a figure of authority—a psychology experimenter or a detective; the subject then engages a contractual agreement with that authority figure to proceed—volunteering and receiving payment in advance of participation in Milgram’s paradigm; signing a waiver of Miranda rights to silence and to counsel in the interrogation setting.

Once the structure of these situations is in place, the authority figure uses deception to reframe the purposes and consequences of the subject’s actions. In Milgram’s experiments, subjects were led to believe that the objective was to test the effects of punishment on a learner through the administration of shocks that may be painful but do not cause harm. In an interrogation, suspects are led to believe that confession serves their personal self-interest better than denial. In both venues, the authority figure then proceeds to make a series of unwavering and relentless demands. Milgram used four scripted prompts and prods (ranging from “Please continue” to “You have no other choice, you must go on”); the Reid technique offers a series of nine steps (beginning with the “positive confrontation” and culminating in “converting the oral admission into a written confession”). In both cases, full obedience is achieved through the elicitation of gradually escalating acts of compliance, culminating in 450 V in Milgram—and, of course, a full confession in the interrogation room.

Two additional similarities are worth noting. One concerns questions that are often raised about ethics. In social psychology, controversy erupted shortly after the publication of Milgram’s first article (Baumrind, 1964; Milgram, 1964) and continues to exert influence over current day Institutional Review Boards (IRBs) in the behavioral sciences. In law, similar questions are typically framed within a rubric of concerns for the “voluntariness” of a suspect’s confession and, hence, its admissibility as evidence at trial (Kamisar, 1963; McCormick, 1972).

The second additional point of similarity concerns the value of recording. When Milgram (1965) released his classic film, Obedience, all of us were able to observe the structure, protocol, and power of the situation that elicited his earlier published results. Although I am aware of no data that surgically address the impact of the film on people’s attributions for the behavior of obedient subjects, Safer (1980) reported that students who saw the film compared to those who did not later overestimated the amount of shock that subjects would administer in a no-command control version of Milgram’s paradigm, suggesting an increased appreciation for the power of the situation (see Reeder, Monroe, & Pryor, 2008). Hence, it is certainly reasonable to suggest that seeing the process sheds light on the background forces and is necessary for others to render judgment as to the outcome. As will be discussed later in this article, the proposal that all interrogations be video
recorded in their entirety represents an important recommendation for reform, in part so that judges and juries can evaluate the voluntariness of the process and the credibility of the resulting statements.

**Confession as a Decision-Making Dilemma**

As to why anyone would confess to police, research on human decision-making has shown that people make choices that they think will maximize their well-being given the constraints they face (Herrnstein, Rachlin, & Laibson, 1997). In addition, studies on temporal discounting show that people tend to be impulsive in their orientation, preferring outcomes that are immediate rather than delayed, with delayed consequences depreciating over time in their subjective value (Rachlin, 2000). In this context, it is easy to appreciate the power of a psychological approach to interrogation—which is explicitly designed to increase the anxiety associated with denial and to decrease the anxiety associated with confession, thereby making it easier for the rational suspect to make the decision to confess (Ofshe & Leo, 1997).

In the context of how people respond to interrogation, recent research illustrates the point. Madon, Guyll, Scherr, Greathouse, and Wells (2012) asked participants to report on whether they had ever committed 20 criminal and unethical acts of misconduct. In one condition, participants faced a short-term negative consequence for each denial of misconduct (having to answer repetitive questions) but risked a larger long-term consequence for admissions of misconduct (having to discuss their responses with a police officer at a later date). Given the choice, participants exhibited a tendency to make admissions of misconduct to avoid the short-term consequence of denial even though it increased the risk of the larger long-term consequence.

This tendency toward short-sighted decision making has been used to characterize what suspects face in a police interrogation setting and can be exacerbated by a number of factors, such as the expected length of an interrogation (Madon, Yang, Smalarz, Guyll, & Scherr, 2013) and being sleep deprived, or questioned during “off-peak” periods of alertness (Scherr, Miller, & Kassin, 2014). As suggested by a self-regulation perspective, even suspects who vigorously refuse to confess at first will become exhausted over time and lose their will to resist (Davis & Leo, 2012). Hence, whereas the typical interrogation lasts for 1 or 2 hours, in proven false confession cases in which interrogation time was recorded, 34% lasted 6–12 hours and 39% lasted 12–24 hours (Drizin & Leo, 2004).

**Analyses of Police Interrogation Tactics in the Laboratory**

As noted earlier, current research on false confessions has analyzed various aspects of the confession-taking process and outcomes using a broad range of methodologies—including case studies, naturalistic observations, self-report
surveys and interviews, and laboratory and field experiments. These approaches have been used to examine the effects of personal and situational factors on the elicitation of false confessions.

The false evidence effect. Looking to study the “Milgramesque” interrogation tactics sanctioned by the Reid technique, my colleagues and I sought to develop an ethical laboratory paradigm that would both meet with IRB approval and confront innocent participants with a personally meaningful decision to confess. It was clear that entrapping people to cheat, steal, or otherwise commit an act that would cast them in a negative light would not be permitted. With these limits in mind, Kassin and Kiechel (1996) devised an experimental paradigm now variously referred to as the computer crash or ALT key experiment in which the experimenter accused participants typing on a desktop computer of causing the hard drive to crash by inadvertently pressing the ALT key he had explicitly instructed them to avoid. Despite their actual innocence and initial denials, participants were asked to sign a confession. The purpose of the study was to test the hypothesis that police lies about evidence can lead innocent people both to confess and to internalize a belief in their own guilt. In some sessions but not others, therefore, a confederate said she witnessed the participant hit the forbidden key. This false evidence manipulation nearly doubled the number of students who signed a written confession, from 48% to 94%. Many of those who signed also internalized the erroneous belief in their own culpability and confabulated false memories of how it happened.

Follow-up studies went on to replicate this effect to the extent that the accusation was plausible (Horselenberg et al., 2006; Klaver, Lee, & Rose, 2008), even when the confession was said to bear a financial or other consequence (Horselenberg, Merckelbach, & Josephs, 2003; Redlich & Goodman, 2003), and even among informants who are pressured to report on a confession allegedly made by another person (Swanner, Beike, & Cole, 2010). The effect is particularly evident among children and juveniles who tend to be both more compliant and suggestible than adults (Candel, Merckelbach, Loyen, & Reyskens, 2005; Redlich & Goodman, 2003).

Using a completely different paradigm, Nash and Wade (2009) then used digital editing software to fabricate video evidence of participants in a computerized gambling experiment “stealing” money from the “bank” during a losing round. Presented with this false evidence, all participants confessed—and most internalized the belief in their own guilt. Together, these studies serve as a basis for a critical analysis of police-induced false confessions—and, in particular, the coercive effects of the false evidence ploy, which American police are permitted to use (see also Wright, Wade, & Watson, 2013; for a discussion of the implications, see Kassin, 2007b).

Minimization effects. A second potentially problematic police tactic concerns the use of minimization. Among suspects feeling trapped by the highly
confrontational stages of interrogation, interrogators are trained to minimize the crime through “theme development”—a process of providing moral justification or face-saving excuses, making confession seem like an expedient means of escape. Interrogators may suggest to suspects that their actions were spontaneous, accidental, provoked, peer pressured, or otherwise justifiable by external factors. Over the years, the U.S. courts had ruled that confessions extracted by promises of leniency and threats of harm or punishment were not voluntary and, hence, not admissible in court. But what about the use of subtler, lawful tactics that produce the same net effects on suspects’ expectations?

In a series of paper-and-pencil studies, Kassin and McNall (1991) had participants read transcripts of suspect interrogations. In each case, three versions were produced in which the detective: (1) made a conditional promise of leniency, (2) used minimization by blaming the victim, or (3) used neither technique. Participants read one version and estimated the sentence that they thought would be imposed on the suspect upon confession. The result: Minimization tactics led people to infer by pragmatic implication that leniency in sentencing will follow from confession—as if an explicit promise had been made.

If people infer leniency from minimization remarks, it stands to reason that minimization would encourage false confessions from innocent suspects who feel trapped and unable to extricate themselves. To test this hypothesis, Russano, Meissner, Kassin, and Narchet (2005) devised a “cheating paradigm” that enabled the manipulation of guilt and innocence for a willful act and a test of the behavioral effects of minimization on the diagnosticity of the resulting confession (as measured by the ratio of true to false confessions). In this study, participants were paired with a confederate for a problem-solving study and instructed to work alone on some problems and jointly on others. In the guilty condition, the confederate sought help on a problem that was supposed to be solved alone, inducing a violation of the experimental prohibition. In the innocent condition, the confederate did not make this request to induce the crime. The experimenter soon “discovered” a similarity in their solutions, separated the participant and confederate, and accused the participant of cheating. The experimenter tried to get the participant to sign an admission by overtly promising leniency, making minimizing remarks, using both tactics, or using no tactics.

Overall, the confession rate was higher among guilty participants than innocent, when leniency was promised than when it was not, and when minimization was used than when it was not. Importantly, diagnosticity was highest when no tactics were used (46% of guilty suspects confessed vs. only 6% of innocents). Paralleling the effects of an explicit promise of leniency, minimization reduced diagnosticity by increasing not only the rate of true confessions (from 46% to 81%) but even more so the rate of false confessions (from 6% to 18%).
Effects of actual innocence. On the basis of anecdotal evidence suggesting that innocent people think and behave differently from guilty suspects in an interrogation setting, Kassin (2005) proposed that innocence itself could be a risk factor for confession. Noting that innocent people believe that the truth will prevail, Kassin and Norwick (2004) found, in a mock crime experiment, that innocent suspects are more likely to waive their Miranda rights to silence and to counsel even when in the presence of an officer who appears guilt-presumptive, hostile, and closed-minded (Kassin & Norwick, 2004; also see Moore & Gagnier, 2008).

Other research shows that innocent people do not use self-presentation “strategies” in their narratives when interviewed by police (Hartwig, Granhag, Strömwall, & Vrij, 2005; Hartwig, Granhag, & Strömwall, 2007); they offer up alibis freely, without regard for the fact that police would view minor inaccuracies with suspicion (Olson & Charman, 2012); and they become less physiologically aroused in response to the stress of an accusatory interrogation (Guyll et al., 2013). In the plea bargaining domain, experiments have shown that most participants who are accused of a transgression they did not commit—compared to those who are guilty—refuse to accept a plea offer, often to their own detriment, because they are confident of acquittal (Gregory, Mowen, & Linder, 1978; Tor, Gazal-Ayal, & Garcia, 2010). The sense of reassurance that accompanies innocence may reflect a generalized and perhaps motivated belief in a just world in which human beings get what they deserve and deserve what they get (Lerner, 1980). It may also occur because of an “illusion of transparency,” a tendency for people to overestimate the extent to which their true thoughts, emotions, and other inner states can be seen by others (Gilovich, Savitsky, & Medvec, 1998; Miller & McFarland, 1987). Whatever the reason, a good deal of research now supports the hypothesis.

Innocence as a mental state can have nonintuitive effects on a suspect’s response to various interrogation tactics. In a series of experiments, Perillo and Kassin (2011) examined the relatively benign bluff technique by which interrogators pretend to have evidence without further claiming that it implicates the suspect (e.g., stating that biological materials were collected and sent for testing). The theory underlying the bluff is simple: Fearing the evidence to be processed, perpetrators will succumb to pressure and confess; not fearing that alleged evidence, innocents would not succumb and confess. Yet in two experiments, Perillo and Kassin found that innocent participants were substantially more likely to confess to pressing a forbidden key, causing a computer to crash, when told that their keystrokes had been recorded for later review. In a third experiment, innocent participants were more likely to confess to willful cheating when told that a surveillance camera had taped their session. Afterward, these participants consistently stated that the bluffed camera offered an assurance of future exoneration, which paradoxically made it easier to confess.
The Consequences of Confession

A good deal of research has focused on the factors that lead people to confess to police, where agreement to sign a confession has served as the dependent measure. An important second direction has been to consider the consequences of confession in studies in which confession has served as an independent variable. In particular, research has focused on two criminal justice venues in which confession evidence is potent: in trial and appellate courts, and during the processes of criminal investigation and prosecution.

How Juries and Judges Perceive Confessions

When a suspect retracts a confession, pleads not guilty, and goes to trial, a judge determines at a pretrial suppression hearing whether the confession was voluntary and hence admissible as evidence. There are no simple criteria for making this judgment, but over the years the courts have ruled that whereas various forms of trickery and deception are permissible, confessions cannot be produced by physical violence, threats, or harm or punishment, explicit promises of leniency, or interrogations conducted in violation of a suspect’s Miranda rights.

Whatever the criteria, confessions ruled voluntary are admitted at trial. Hearing the admissible confession, the jury then determines whether the defendant is guilty beyond a reasonable doubt. But are people accurate and discriminating judges of confessions? The wrongful convictions of innocent confessors suggest a negative answer to this question. Other research too indicates that confession evidence is devastating when presented in court. In fact, it is the power of confessions to influence social perceptions that sparked my current day interest in the processes of interrogation and the validity of the confessions that are produced (Kassin & Wrightsman, 1980, 1981; for a review, see Kassin & Wrightsman, 1985).

To test whether police can distinguish between true and false confessions to actual crimes, Kassin, Meissner, and Norwick (2005) recruited male prison inmates to take part in a pair of videotaped interviews. Each inmate gave both a true narrative confession to the crime for which he was incarcerated and a false confession to a crime he did not commit. Using this procedure, Kassin et al. compiled a videotape of 10 confessions known to be true or false. College students and police investigators judged these statements, and the results showed that neither group exhibited significant accuracy but that police were more confident in their judgments. People’s inability to distinguish between true and false confessions was recently replicated in a study involving the confessions of juvenile offenders (Honts, Kassin, & Craig, 2014).

Over the years, mock jury studies have shown that confessions have a great impact on jury verdicts—a greater impact, for example, than eyewitness and
character testimony (Kassin & Neumann, 1997). Research has also shown that people do not adequately discount confession evidence even when the confessions are perceived to have been coerced by police (Kassin & Sukel, 1997); even when jurors are told that the defendant suffers from a mental illness or interrogation-induced stress (Henkel, 2008); even when the defendant is a juvenile (Redlich, Ghetti, & Quas, 2008; Redlich, Quas, & Ghetti, 2008); even when the confession was given not by the defendant but by a second-hand informant who was motivated to lie (Neuschatz, Lawson, Swanner, Meissner, & Neuschatz, 2008; Neuschatz et al., 2012); and even, at times, when the confession is contradicted by exculpatory DNA (Appleby & Kassin, 2011).

In a study that well illustrates the potency of confession evidence, Kassin and Sukel (1997) presented participants with one of three versions of a murder trial transcript. In a low-pressure version, the defendant was said to have confessed to police immediately upon questioning. In a high-pressure version, participants read that the suspect was in a state of physical discomfort and interrogated aggressively for a long period of time. A control version contained no confession in evidence. In some ways, participants presented with the high-pressure confession responded in a legally appropriate manner. They judged the statement involuntary and said it did not influence their decisions. Yet this confession significantly boosted the conviction rate.

This same pattern of results was recently replicated in a study involving judges. Wallace and Kassin (2012) presented 132 experienced judges with a case summary with strong or weak evidence and a confession elicited by either high- or low-pressure interrogation tactics, plus a no confession control group. As expected, judges were less likely to see the confession as voluntary when it resulted from a high-pressure than a low-pressure interrogation (29% vs. 84%, respectively). However, even the high-pressure confession significantly increased the percentage of guilty verdicts. In the weak evidence condition, which produced a mere 17% conviction rate without a confession, a significant increase was produced not only by the low-pressure confession (96%) but by the high-pressure confession as well (69%). As with lay juries, it appears that judges are so influenced by confession evidence that they do not discount it when it is coerced and hence they are legally required to do so.

In actual cases, there are two reasons why confessions are highly persuasive. The first reason is that the common sense of attribution leads us to trust other people’s statements against self-interest. Hence, research shows that people are more likely to believe peoples’ admissions of guilt than denials (Levine, Kim, & Blair, 2010). Surveys show that most people believe that they would never confess to a crime they did not commit—and that they evaluate others accordingly (Henkel, Coffman, & Dailey, 2008; Leo & Liu, 2009). Part of the problem is that people are ignorant of the interrogation tactics that are used by police and the dispositional and situational risk factors that would lead someone to make a false
There is a second important reason why confessions carry so much decision-making weight, even when they are false. Analyzing 38 proven false confessions from the Innocence Project’s data base of DNA exonerations, Garrett (2010) found that a striking 95% contained accurate and often vivid details about the crime that were not in the public domain. Often the prosecution featured these details at trial, suggesting that they could have only been known by the perpetrator. The confessors in these cases were innocent, so they could not have possessed firsthand guilty knowledge. Thus, it appears that police had communicated these details during the process of interrogation. To further complicate matters, Appleby, Hasel, and Kassin (2013) content-analyzed twenty false confessions and found that many of them contained not only vivid sensory details about the crime but statements about the confessor’s motivation, assertions that the confession is voluntary, apologies, and expressions of remorse. In short, many false confessions contain cues that inflate perceptions of their credibility.

**Corruptive Effects of Confessions on Other Evidence**

Just as confessions are trusted by judges and juries, often providing a sufficient basis for conviction, basic social cognition research suggests that confessions may also influence the way in which other evidence is interpreted—for example, by tainting the perceptions of eyewitnesses, forensic scientists, and others. Over the years a good deal of research has revealed that top-down influences inform human judgment. Classic studies showed that prior exposure to images of a face or a body, an animal or a human, or letters or numbers, can bias what people see in an ambiguous figure. Indeed, the presence of ambiguous objective evidence, by providing the perception of support, may actually exacerbate the effects of preexisting stereotypes (Darley & Gross, 1983).

In a forensic demonstration of this point, Hasel and Kassin (2009) had participants witness a staged theft and then make an identification decision from a lineup in which the actual perpetrator was not present (akin to real life instances in which the suspect is innocent). Two days later, they were given additional information and an opportunity to change their decision. When told that another suspect had confessed, 61% of participants changed their initial decision and identified the suspect who had allegedly confessed. Those who were told that the individual they had identified confessed became more confident in their decision. Among participants who at first correctly did not make an identification, indicating that the culprit was not present, nearly half went on to identify an innocent person after being told that someone had confessed.

Other research has shown that a strong belief in a suspect’s guilt can also bias lay people’s judgments of handwriting samples (Kukucka & Kassin, 2014),
their willingness to vouch as an alibi for a confederate (Marion et al., 2014), their perceptions of whether degraded speech recordings contain incriminating remarks (Lange, Thomas, Dana, & Dawes, 2011), and the judgments of experts who are presented with inconclusive polygraph charts (Elaad, Ginton, & Ben-Shakhar, 1994) and latent fingerprint samples (Dror & Charlton, 2006).

In addition to the results of laboratory and field experiments, archival data also support the notion that confessions can taint other evidence. Using the Innocence Project’s DNA exoneration files, Kassin, Bogart, and Kerner (2012) tested the hypothesis that confessions yield additional evidentiary errors by examining whether other types of evidence errors were present in DNA exoneration cases containing a false confession. As predicted, additional evidence errors were present in 78% of these cases. Specifically, false confessions were accompanied by invalid or improper forensic science (63%), mistaken eyewitness identifications (29%) and snitches or informants (19%). Consistent with the causal hypothesis that the false confessions had influenced the subsequent errors, the confession was obtained first rather than later in the investigation in approximately two thirds these cases.

To sum up, an emerging body of research has suggested that “forensic confirmation biases” are pervasive and has inspired the recommendation that all lay witnesses and forensic examiners, as a matter of practice, be blinded to case information concerning confessions and other contextual cues (Kassin, Dror, & Kukucka, 2013; Saks, Risinger, Rosenthal, & Thompson, 2003; with similar regard to the importance of having eyewitness lineup identifications conducted by a blind administrator, see Canter, Hammond, & Youngs, 2013; Wells, Small, Penrod, Malpass, Fulero, & Brimacombe, 1998).

**Possible Effects on the Truth-Seeking Process**

There may be an additional pernicious effect of confessions, not only on the substance of a crime investigation but on the truth-seeking process. Currently, an estimated 97% of convicted defendants in the federal criminal justice system plead guilty (Rakoff, 2014). With numbers of this magnitude, the courts have expressed a concern over the possibility of an “innocence problem” in guilty pleas. However, the prevalence of this alleged problem is not known.

On the one hand, role playing and behavioral laboratory experiments have shown that many innocent people will accept a false guilty plea—at rates as high as 33% (Gregory et al., 1978), 43% (Russano et al., 2005), and 56% (Dervan & Edkins, 2013). On the other hand, consistent with the notion that innocent people have faith in their innocence will prevail during an investigation or at trial (Kassin, 2005), it seems that the guilty plea rate is very low among wrongfully convicted innocent defendants. In one analysis, Gross et al. (2005) studied 340 wrongful convictions and found that only 6% had pled guilty; Redlich (2010) and Kassin
(2012) examined the more specific sample of post-conviction DNA exonerations and found that only 8% had pled guilty.

In light of the historical, commonsense, and empirically demonstrated power of confessions, one would expect that innocent people who were induced into confession would feel more pressure than normal to plead guilty because they and their attorneys believe that the confession increases the risk of conviction at trial. Archival data are highly consistent with this prediction. Drizin and Leo (2004) assembled for analysis 125 false confession cases and found that of those defendants who were prosecuted, 27% had pled guilty. Examining the DNA exonerations from the Innocence Project, Redlich (2010) found that exonerees who falsely confessed were four times more likely to plead guilty than those who had not confessed. Although the difference was based on a small number of guilty pleas, the pattern has persisted. Through the first 289 DNA exonerations, Kassin (2012) confirmed that false confession cases were far more likely to be resolved by a guilty plea than were nonconfession cases—26% versus 4%. Relative to other innocents, it appears that defendants who confess are later more likely to relinquish their constitutional right to defend themselves at trial—cloaked in a presumption of innocence, with the state burdened to prove guilt beyond a reasonable doubt, and with an opportunity to confront their accusers. At this point, further research is needed, in the laboratory, to test the causal hypothesis that false confessions trap innocent suspects into pleading guilty.

**Social and Policy Implications**

In the landmark case of *Miranda v. Arizona* (1966), the U.S. Supreme Court described custodial police interrogation as “inherently coercive” and ruled that police must inform suspects in custody of their constitutional rights to silence and to counsel. Only if suspects waive these rights “voluntarily, knowingly, and intelligently,” said the Court, can the statements they produce be admitted into evidence.

Although *Miranda* is presumed to have provided a profound safeguard for people who stand accused, its benefits are unclear. For starters, many suspects lack the capacity to understand and apply these rights. Particularly problematic is comprehension among young adolescents (e.g., Goldstein, Condie, Kalbeitzer, Osman, & Geier, 2003) and adults who are mentally retarded (e.g., Clare & Gudjonsson, 1995; Everington & Fulero, 1999). Among normal adults, research further shows that Miranda warnings vary enormously in comprehensibility from one jurisdiction to another (Rogers, Harrison, Shuman, Sewell, & Hazelwood, 2007) and that a suspect’s comprehension may also be compromised by interrogation stress and other situational factors (Rogers, Gillard, Wooley, & Fiduccia, 2011; Scherr & Madon, 2012).
Even among suspects who comprehend their *Miranda* rights, a second reason that the warnings may not adequately protect the accused is that most people tend to waive their rights (Baldwin, 1993; Leo, 1996b). Different explanations have been put forth to explain this phenomenon. Leo (1996b) noted that police detectives are highly effective at persuading suspects to waive their rights. Rogers et al. (2010) added that many suspects waive their rights because they harbor the misconception that invoking *Miranda* will prove ineffective and lead police and others to infer guilt (Rogers et al., 2010). Offering a third reason, Kassin (2005) proposed that innocence itself is a state of mind that would lead innocent people to waive their rights because they believe that they have nothing to fear or to hide. Kassin and Norwick (2004) tested this hypothesis in a mock crime study and found that 81% of innocent participants signed a waiver compared to only 36% of those who were guilty (also see Moore & Gagnier, 2008). To sum up, it appears that *Miranda* warnings do not adequately protect the citizens who need it most—those accused of crimes they did not commit. Therefore, other safeguards are needed.

Research on police interrogations, confessions, and their consequences for people who are wrongfully convicted has inspired calls for reform. In particular, this research has compelled a number of proposals for reform designed to protect highly vulnerable suspect populations (e.g., juveniles, people with cognitive impairments or mental health problems that increase compliance tendencies and suggestibility) and to ban the use of coercive police interrogation practices (e.g., the false evidence ploy, minimization tactics that imply a promise of leniency). I believe that the most important possible safeguard is to require the video recording of interrogations—the entire process, not just the confession. Indeed, this was the primary recommendation in the recent AP-LS White Paper: “Without equivocation, our most essential recommendation is to lift the veil of secrecy from the interrogation process in favor of the principle of transparency” (Kassin et al., 2010).

In 1985 and 1994, respectively, the Supreme Courts of Alaska and Minnesota ruled that police must electronically record all suspect interviews and interrogations in felony cases. Since that time, both as a result of state Supreme Court opinions and legislative statutes, a growing number of states—now up to 17 plus the District of Columbia—have started to require the recording of interrogations in major felony investigations (Sullivan, 2012). In a particularly new and notable milestone, the U.S. Department of Justice also recently reversed its long-standing opposition and refusal and established the presumptive requirement that the FBI and other federal law enforcement agencies record the custodial interrogations of felony suspects (Schmidt, 2014).

In recent years, interviews with police investigators who have started to record full interrogations have shown that their reaction has been uniformly favorable (Sullivan, 2004; Sullivan, Vail, & Anderson, 2008). But what are the actual effects?
There are two sets of reasons for the recommendation that interrogations be electronically recorded. The first is the expectations that the practice of recording will induce an attentional state of self-awareness, increase accountability, and deter the most aggressive police tactics, thereby reducing the risk of false confessions. To test this hypothesis, Kassin, Kukucka, Lawson, and DeCarlo (2014) conducted a field experiment in a mid-sized city police department. Sixty-one investigators inspected a staged crime scene and interrogated a male suspect who was guilty or innocent of a mock crime in sessions that were surreptitiously recorded. By random assignment, half the police participants were informed that the sessions were being recorded; half were not. Coding of the interrogations revealed the use of several common tactics designed to get suspects to confess. As one might predict, police in the camera-informed condition were less likely than those in the camera-uninformed condition to use both maximization and minimization tactics; they were also perceived by suspects—who were uninformed about the camera manipulation—as trying less hard to elicit a confession. The results thus suggested that video recording can affect the process of interrogation—notably, by inhibiting the use of certain sometimes egregious tactics.

A second benefit to the recommendation that interrogations be recorded is to provide an accurate factual record for judges and juries needing to assess the voluntariness and credibility of the confessions that are produced. As noted earlier, neither laypeople nor police can accurately discriminate between true and false confessions (Honts et al., 2014; Kassin et al., 2005). Part of the problem is that the commonsense of the fundamental attribution error leads people to infer guilt from confession despite coercion. Another part of the problem is that false confessions often contain accurate crime details and other credibility cues (Appleby et al., 2013; Garrett, 2010). Lacking access indications of coercion and the source of the crime details appearing in the ultimate confession, judges and juries are denied the very information needed for accurate fact finding.

Over the years, a number of pragmatic and logistical concerns have been raised about recording interrogations as a matter of policy (e.g., what conditions should activate a recording requirement; what should happen if the equipment malfunctions or if the suspect refuses to make a recorded statement; what evidentiary consequences would follow from the failure to record). As a matter of practice, however, research suggests that it is important not only that entire sessions be recorded but that the camera adopt a neutral “equal focus” perspective that shows both the accused and his or her interrogators. In a number of studies on illusory causation effects in attribution, Lassiter and his colleagues have taped mock interrogations from three different camera angles so that the suspect, the interrogator, or both were visible. Consistently, participants who see only the suspect judge the situation as less coercive than those also focused on the interrogator. By directing visual attention toward the accused, the camera can thus lead jurors
to underestimate the amount of pressure actually exerted by the “hidden” detective (Lassiter & Irvine, 1986; for a review, see Lassiter, Geers, Munhall, Handley, & Beers, 2001). Under these more balanced circumstances, juries—and judges—make more informed attributions of voluntariness and guilt when they see not only the final confession but the conditions under which it was elicited (Lassiter, Diamond, Schmidt, & Elek, 2007; Lassiter, Geers, Handley, Weiland, & Munhall, 2002).

Finally, it is important to note that although the video recording of interrogations is a reform designed to help prevent the occurrence of false confessions, two additional measures should be taken to minimize the rippling effects of these confessions once taken. The first problem concerns the way in which confessions can corrupt other evidence from lay witnesses and experts alike. The simplest way to protect against the biasing effects of confessions—and other contextual variables as well—is to ensure that eyewitnesses and crime lab examiners are not informed of the presence or absence of a confession. To ensure that an eyewitness’s memory-based identification and a forensic examiner’s perceptual judgments are based solely on the stimuli presented, they should be blind as to whether a lineup member, handwriting sample, polygraph chart, or fingerprint belonged to a suspect who had confessed (see Kassin, Dror, & Kukucka, 2013; Saks, Risinger, Rosenthal, & Thompson, 2003).

The second added safeguard concerns the use of expert testimony at trial. There is now an ample body of research—derived from basic principles of social psychology, recent research specifically focused on confessions, and case studies of wrongful convictions—to inform the courts on the dispositional and suspect factors that put innocent people at risk to confess as a function of interrogation (for a three-tiered framework for expert testimony, see Kassin, 2007a). A good deal of research has also shown that the laypeople do not intuitively understand false confessions and their risk factors as a matter of common knowledge. In 1988, social psychologist Elliott Aronson testified as an expert in a murder trial on how someone could be induced to confess to a crime he did not commit (Davis, 2010; Tavris & Aronson, 2007). Over the years, however, the U.S. courts have varied a great deal in their willingness to admit such testimony (Fulero, 2004). Hence, the American Psychological Association (APA) has opined in three recently submitted amicus briefs that judges and juries have difficulty assessing confession evidence, that the phenomenon of false confession is counterintuitive, and that psychological experts should be permitted to testify at trial because their testimony would draw from generally accepted research and that it would assist the trier of fact (Michigan v. Kowalski, 2012; People v. Thomas, 2013; Rivera v. Illinois, 2011). As always, more research will prove useful to address specific questions. Already, however, the extant literature is sufficient in this author’s opinion, as well as APA’s, for expert testimony to be admitted into court.
References


Kassin


SAUL M. KASSIN (Ph.D., University of Connecticut) is Distinguished Professor of Psychology at the John Jay College of Criminal Justice, New York, NY. As cited by the U.S. Supreme Court, his research is focused on social-psychological processes in the law and the prevention of wrongful convictions.