

From Caller to Suspect: Identifying Behaviors That Trigger Suspicion in 911 Calls

Samantha R. Bean¹, Alia N. Wulff¹, Megan L. Lawrence², Isabelle Reeder³, Nicholas D. Duran⁴,
Saul M. Kassir⁵, and Jessica M. Salerno²

¹ School of Interdisciplinary Forensics, Arizona State University West Valley Campus

² Department of Psychology, Cornell University

³ Department of Psychology, Fordham University

⁴ Department of Psychology, University of Texas at Austin

⁵ John Jay College of Criminal Justice, City University of New York



Objective: Despite research on case factors that can trigger confirmation bias in investigations leading to wrongful convictions, we know little about what sparks this chain reaction and why an innocent person *initially* falls under suspicion. Across four studies, we investigated what perceived behaviors exhibited by 911 callers (urgency, emotionality, cognitive load, impression management, and information management) are related to laypeople's (Studies 1, 2, and 4) and police officers' (Study 3) suspicion toward the caller. **Hypotheses:** We predicted, for lay and police samples, callers perceived as more urgent or emotional would be perceived as less suspicious, whereas callers perceived as more under cognitive load, managing information, or managing impressions of themselves would be perceived as more suspicious. We tested whether these relationships depended on caller gender—predicting gender stereotypes might play a role, particularly with emotionality. **Method:** Participants (Studies 1, 2, and 4: online laypeople; Study 3: police officers) listened to a real 911 call (Studies 1 and 4) or a more controlled, simulated 911 call (Studies 2 and 3) by a male or female caller. We assessed behavioral predictors associated with suspicion via participants' spontaneous impressions and scales assessing these behaviors. Participants indicated how suspicious they found the caller and how much the caller violated their expectations. **Results:** Participants spontaneously mentioned the callers' emotionality most frequently and consistently across studies (77%–85%). Perceptions of callers being more urgent and emotional were associated with less suspicion and violation of expectations, while perceptions of callers engaging in more information management and impression management were associated with more suspicion and violation of expectations. Police, but not laypeople, perceived male callers as more suspicious than female callers—despite holding 911 call scripts constant. **Conclusions:** Citizens engaging in the well-intentioned act of calling 911 risk observers—including police—holding expectations for their behavior and targeting them as a suspect if they violate those expectations.

This article was published Online First November 10, 2025.

Diane Sivasubramanian served as Action Editor.

Samantha R. Bean  <https://orcid.org/0000-0001-9838-218X>

Alia N. Wulff  <https://orcid.org/0000-0001-7271-1539>

Megan L. Lawrence  <https://orcid.org/0000-0002-2323-7719>

Isabelle Reeder  <https://orcid.org/0009-0009-4754-3540>

Nicholas D. Duran  <https://orcid.org/0000-0002-8872-5617>

Saul M. Kassir  <https://orcid.org/0000-0002-0495-8377>

Jessica M. Salerno  <https://orcid.org/0000-0001-7394-6444>

Study materials, data, and analysis code (https://osf.io/yat52/?view_only=8154999c3be5483ba061a9018b39756f) for all studies, preregistration for Study 2 (https://osf.io/pqzqm?view_only=5c71a52caf7e405fa07b9f6116848981) and preregistration for Study 4 (https://osf.io/w38tm?view_only=66b677a1736146258c6aed431fcf0a8a) are available on the Open Science Framework. Data and theory from these studies were presented by the Samantha R. Bean at the Annual Meeting of the American Psychology-Law Society (Bean et al., 2022). Part of this research was conducted while the first five and last author were at Arizona State University.

The authors have no known conflicts of interest to disclose. Funding was

provided by the National Science Foundation (Grant 2146834) awarded to Nicholas D. Duran, Saul M. Kassir, and Jessica M. Salerno and by Arizona State University's Institute for Social Science Research Seed Grant, and New College Scholarship, Research, and a Creative Activities Seed Grant awarded to Nicholas D. Duran and Jessica M. Salerno. The authors acknowledge and thank Breanna Olson, Logan Bussey, and Hannah Phalen for their contribution in coding data for this research.

Samantha R. Bean played a lead role in formal analysis, visualization, writing—original draft, and writing—review and editing and a supporting role in conceptualization, data curation, and methodology. Alia N. Wulff played a supporting role in data curation, formal analysis, and writing—review and editing. Megan L. Lawrence played a supporting role in writing—original draft and writing—review and editing. Isabelle Reeder played a supporting role in data curation, investigation, and writing—review and editing. Nicholas D. Duran played a supporting role in conceptualization, data curation, methodology, project administration, supervision, and writing—review and editing and an equal role in funding acquisition. Saul M. Kassir played a supporting role in methodology and writing—review and editing and an equal role in funding acquisition. Jessica M. Salerno played a lead role in project

continued

Public Significance Statement

Recent police trainings instruct law enforcement to judge whether someone who calls 911 to report a crime could be the perpetrator based on their behavior on the call. In this research, we investigated how perceptions of a 911 caller's behavior can trigger suspicion toward the caller for both laypeople and police officers. Given there is no "one size fits all" reaction to a traumatic event, it is important to understand how others expect (perhaps unreasonably) a 911 caller to behave and how violating those expectations can trigger suspicion toward the caller. Our data suggest that individuals who call 911 to help others might make themselves vulnerable to becoming targeted as a suspect if they violate potentially misguided expectations for what is "normal" behavior during a traumatic event. It is critical to examine why someone is initially suspected because research and anecdotal case evidence suggest that, once someone is a suspect, a chain reaction of confirmation bias can lead to a wrongful conviction.

Keywords: suspicion, confirmation bias, 911 calls, expectancy violation, judgment

Supplemental materials: <https://doi.org/10.1037/lhb0000640.supp>

The Innocence Project, founded in 1992, and the National Registry of Exonerations, founded in 2012, were established to investigate claims of innocence after conviction. These organizations identified types of evidence that have led to wrongful convictions, such as mistaken eyewitness identifications (Wells et al., 1998, 2020), forensic science errors (Kassin et al., 2013; National Research Council, 2009), police-induced false confessions (Kassin et al., 2010, 2025), and jailhouse informants who lied (Natapoff, 2022; Neuschatz & Golding, 2022). These are critical, but prompt a deeper inquiry: Why did these innocent people become suspects in the first place, thereby unleashing a guilt-presumptive series of confirmation biases from which a cumulative disadvantage accrues (Scherr et al., 2020)?

Case files and media coverage of 460 factually innocent defendants revealed eight "origins of implication" leading innocent individuals to be misclassified as suspects (e.g., suspect's race and age, being physically proximate to the crime, prior relationship to the victim, past or present criminal activity, Lowrey-Kinberg et al., 2019). However, these documents did not thoroughly address behavioral cues that may have led police to become suspicious based on a witness's demeanor. Case files and media rarely include officers' explicit admissions of these motivations, potentially because they are unaware or unwilling to admit such motivations. We considered this deeper inquiry, focusing on the first moment a person could become a suspect: when they call 911 to initially report that a crime has occurred. We report four studies investigating what behaviors in a 911 call violate laypeople's and police officers' expectations and trigger suspicion about the caller.

Witness Demeanor and Suspicion

Suspicion is characterized as the state of mind when an individual has multiple, potentially contradictory hypotheses that cause them to


question a person's authenticity or behavioral motives (e.g., Barone et al., 2004; Vonk, 1998). As a result, the observer hesitates to take that person's behavior at face value (Fein, 1996; Toris & DePaulo, 1984). Because feelings of suspicion predict deception judgments (Toris & DePaulo, 1984) and related negative impressions of a target (e.g., Barone et al., 2004; DeCarlo, 2005; Ferguson et al., 2011), it is critical to understand what behaviors spark suspicion toward innocent witnesses at the beginning of an investigation.


Cases in which police targeted innocent individuals reveal that an "inappropriate" level of emotion is often cited as an origin of suspicion. Michael Crowe was coerced into falsely confessing to his sister's murder, in part, because detectives thought he reacted with too little emotion (Johnson, 2003; Kassin & Gudjonsson, 2004). Marty Tankleff also falsely confessed to killing of his parents after a detective thought he was "too calm" (Heath, 2009). In contrast, Jeffrey Deskovic drew suspicion when police saw him as "overly distraught" after a high-school classmate was raped and murdered (Heath, 2009; Santos, 2006). Amanda Knox was described as overly emotional at one moment but not emotional enough at another, following her roommate's murder (Burleigh, 2011). These cases illustrate how perceptions of a suspect's demeanor have triggered suspicion, even among factually innocent, exonerated individuals.

Police are human—and humans routinely form first impressions quickly based on a mere snapshot of information (Todorov & Oh, 2021; Willis & Todorov, 2006) or a "thin slice" of expressive behavior (Ambady & Rosenthal, 1993; N. A. Murphy et al., 2019). Further, police officers in general—including those trained in the so-called Behavioral Analysis Interview of the Reid technique (Inbau et al., 2013)—are both overly confident in their deception detection abilities based on judging witnesses' verbal and nonverbal behavior and biased toward seeing deception (Masip et al., 2005, 2016; Meissner & Kassin, 2002; Vrij, 2008).

administration and supervision, a supporting role in data curation, investigation, writing—original draft, and writing—review and editing, and an equal role in conceptualization, funding acquisition, and methodology.

 The data are available at https://osf.io/yat52/?view_only=8154999c3be5483ba061a9018b39756f.

 The experimental materials are available at https://osf.io/yat52/?view_only=8154999c3be5483ba061a9018b39756f.

 The preregistered design is available in Study 2 at https://osf.io/pqzqm?view_only=5c71a52caf7e405fa07b9f6116848981; Study 4 at https://osf.io/w38tm?view_only=66b677a1736146258c6aed431fcf0a8a.

Correspondence concerning this article should be addressed to Samantha R. Bean, School of Interdisciplinary Forensics, Arizona State University West Valley Campus, 4701 West Thunderbird Road, Glendale, AZ 85036, United States. Email: srbean1@asu.edu

The 911 Paradigm

Today's centralized 911 system was created in 1968 following public concern about bystander nonintervention in the 1964 Kitty Genovese case (Rosenthal, 1964). Yet, an emerging issue is whether individuals who call 911 to report crimes, emergencies, or injuries are socially responsible citizens and victims in need or perpetrators worthy of investigation. Over the years, numerous individuals have found themselves under suspicion after calling 911. For example, Gary Gauger was labeled as affectively "flat" in his 911 call about his parents' deaths, after which he was coerced into a false confession before being exonerated (Gauger & Von Bergen, 2008; Heath, 2009; Shapiro, 1998). Sandra Ortiz's "calm tone" on a 911 call was used against her in trial in which she was wrongfully convicted of murdering her boyfriend (Possley, 2013). More recently, Thomas Perez falsely confessed to his father's murder after a 911 dispatcher felt "something was off," as she perceived him as lacking concern and "rambling" (Clarke & Prokupecz, 2024). Perez's father was found alive days later. Even then, officers continued to search for a possible victim or crime involving Perez because of his behavior on the 911 call.

These cases underscore the importance of understanding the assumptions and judgments made about 911 callers. Since we collected our data, a controversial new "911 linguistic analysis" training ("Is the Caller the Killer?") for police has emerged—which claims to teach attendees behavioral indicators of guilt and innocence on 911 calls (Harpster et al., 2009; for a critique, see B. Murphy, 2022; for a description of this pseudoscience in court, see B. Murphy, 2023). Indeed, the training claims to identify dubious behavioral cues on a 911 call that indicate the caller is concealing their involvement in the crime they are reporting (e.g., opening the call by saying "hello"; see Harpster & Adams, 2017). Importantly, the only study supporting the training contains methodological concerns and unusually large effect sizes (Markey et al., 2025) and has repeatedly failed to replicate (e.g., Cromer et al., 2019; Markey et al., 2022, 2025; Miller et al., 2021; O'Donnell et al., 2022). Given recent attention to this training, the present research—examining what behaviors the public (laypeople and police) expect or find suspicious in a 911 call—is particularly timely and critical.

Behaviors That Violate Expectations

People often have expectations for how an individual should behave based on factors such as who the person is, the social desirability of their behavior, and the norms inherent in the situation they are in (Heider, 1958; Jones & Davis, 1965; Kelley, 1967). When an individual's verbal or nonverbal behaviors deviate from these normative expectations, we may perceive them as suspicious (e.g., Bond et al., 1992) if not outright dishonest (Levine et al., 2000).

Behavioral expectations for the specific context can also shape diverging inferences drawn from the same behavior. For example, when calling 911 immediately after witnessing a horrific crime, a presumed-innocent person is expected to express "appropriately" high emotion—perhaps sounding nervous, fearful, or panic stricken (Salerno et al., 2025). Yet, appearing nervous and fearful in an interview *after* becoming a suspect might *increase* suspicion. For example, the Reid technique advises investigators that behaviors demonstrating anxiousness or nervousness are indicators of guilt

during a suspect's interview (Inbau et al., 2013). Thus, the same emotions (i.e., nervousness, fear) might trigger suspicion of a suspect if they infer the fear is due to getting caught in a lie, but *decrease* suspicion toward a 911 caller if they infer the fear is about witnessing a traumatic event. Therefore, it is important to investigate what sparks initial suspicion of a presumably innocent witness *before* they become a suspect—a critical distinction, as suspicion often precedes a deception judgment (Toris & DePaulo, 1984) and could critically set off a chain reaction of confirmation bias, increasing chances of wrongful conviction.

Individuals experiencing acute stress or a traumatic event can exhibit a range of behavioral responses, from heightened emotional expression and urgency to dissociation, cognitive overload, or controlled speech (Isserlin et al., 2008; Sharpe et al., 2010). These different reactions can shape how observers—including police officers—interpret a caller's demeanor. While some trauma responses might align with societal expectations of distress (e.g., urgency, panic), other reactions may violate these expectations (e.g., excessive detail, lack of affect) and could be misinterpreted as suspicious. It is critical to understand that witnesses to a violent crime may exhibit a wide range of behavioral reactions; there is no "one size fits all" for how to respond to a traumatic event (for an overview, see Salerno, 2021).

Yet, people often have expectations for what is considered "appropriate" emotion depending on situational cues. For example, defendants in a murder trial who are accused of killing a spouse are expected to exhibit greater emotionality on the witness stand than those accused of killing a stranger (Heath et al., 2007; Heath & Grannemann, 2015). What is seen as appropriate may also vary with gender, as women are expected to behave more emotionally than men (e.g., Fabes & Martin, 1991; Salerno et al., 2019).

In this study, we identify behavioral categories that might trigger suspicion in the context of reporting a violent crime in a 911 call: perceived urgency, emotionality, cognitive load, impression management, and information management. Although these categories are informed by several reviews and meta-analyses of deception cues and linguistic perception (Bond & DePaulo, 2006; Buller & Burgoon, 1996; DePaulo et al., 2003; Duran et al., 2010; Fuller et al., 2013; Hartwig & Bond, 2014; Vrij et al., 2000)—including research involving 911 calls specifically (Burns & Moffitt, 2014; Cromer et al., 2019; Markey et al., 2022)—they are not meant to be direct diagnostic markers of deception. Instead, they serve as critical elements in exploring how the violation of listeners' expectations may spark suspicion. Ample research in other contexts has shown many of the demeanor cues police are trained to use (e.g., eye contact, changes in posture, facial expressions) are not empirically diagnostic of deception, yielding low accuracy rates. Moreover, training and experience are not associated with significant improvements and often an increased bias toward seeing deception (Bond & DePaulo, 2006; Denault et al., 2022; Hartwig & Bond, 2011; Honts et al., 2021; Luke, 2019; Meissner & Kassin, 2002; Vrij et al., 2019).

Rather than focusing on deception detection accuracy, we investigated what behaviors spark suspicion before individuals become suspects. Given the underlying premise of our approach that suspicion often arises from discrepancies between expected and observed behaviors, we identified behaviors we predicted would violate expectations and be associated with suspicion, described

next. Additionally, we sought to investigate how perceptions of these behaviors might differ based on the caller's gender.

Perceived Emotionality

Although we are not aware of any previous studies specifically addressing expectations of a crime witness's emotionality, researchers have documented an emotional *victim* effect, which demonstrates that individuals who expect victims to be highly emotional perceive less emotional victims to be less credible (e.g., Ask & Landström, 2010; Hackett et al., 2008; McAuliff et al., 2015). Indeed, police and laypeople's beliefs about what is appropriate emotion expression influence their perceptions of a victim's credibility (Bollingmo et al., 2008; Kaufmann et al., 2003). This could be problematic given that humans naturally exhibit a wide range of emotional responses to traumatic events, ranging from numbness to hysteria (Substance Abuse and Mental Health Services Administration, 2014). Yet, anecdotal evidence suggests people harbor an expectation for a "typical" emotional response to traumatic events (Heath, 2009), when those expectations are violated, suspicion is triggered.

In another article based on the current series of studies, we manipulated the level of emotion that 911 callers expressed while holding the script constant. Police officers found callers who exhibited a low level of emotion, compared with moderate and high emotion, as significantly more suspicious (Salerno et al., 2025). In the present research, we analyzed measures of perceived emotionality (rather than manipulated emotion) along with four other behaviors potentially characteristic of 911 callers that were collected but not reported in the prior article. In the context of a 911 call to report a shooting, we predicted that the caller's perceived emotionality would be negatively associated with expectation violations and suspicion.

Perceived Urgency

Individuals who call 911 generally do so because they have observed an emergency that requires police, firefighter, or medical intervention. Thus, one might expect a witness to a violent crime to express urgency in their 911 call, with callers that sound less urgent triggering suspicion. Prior research on 911 linguistic behavior reveals conflicting results for the role of urgency as it relates to deception. Harpster et al. (2009) reported perceived urgency was associated with being a truthful 911 caller. In contrast, other studies have failed to observe an association between urgency and actual deception (Miller et al., 2021; O'Donnell et al., 2022). Given its potential (but mixed) relationship with deception, we tested whether participants expected a caller to be urgent when reporting a violent crime to 911. We predicted the caller's perceived sense of urgency would be negatively associated with expectation violations and suspicion.

Perceived Cognitive Load

Telling a lie is generally more cognitively demanding than telling the truth, so it takes more time and is more easily disrupted (Vrij et al., 2010; for a meta-analysis, see Suchotzki et al., 2017). In the context of an in-person interview or interrogation, people expect suspects to present their narrative easily if they are merely reporting the truth. Yet, there is a critical difference in the context of 911 calls:

the fast pace and traumatic nature of the call might prove cognitively demanding even to innocent bystanders *trying* to provide a coherent description to the dispatcher—because they are upset, distracted, or still witnessing a violent event. One might expect that a caller engaging in deception is in an especially cognitively demanding situation, as they must comprehend what the operator is saying while planning their next response, monitor the operator's reaction, and respond in an appropriate amount of time. In a manifestation of the fundamental attribution error (Gilbert & Malone, 1995; Ross, 1977, 2018), we predicted that perceiving a caller as struggling to produce a coherent narrative might spark suspicion in listeners who attribute the cognitive load to internal efforts of deception rather than to the external and traumatic situation—thereby violating their behavioral expectations for an innocent caller. Thus, we predicted the caller's perceived cognitive load would be positively associated with expectation violations and suspicion.

Perceived Impression Management

Impression management—the attempt to self-present oneself to others in a favorable light—is common among both liars and truth-tellers in social situations (Colwell et al., 2006; DePaulo, 1992; Hartwig et al., 2010). Yet people are inclined to believe that individuals who appear to engage in impression management are more immoral than those who do not (Lafrenière et al., 2016), in part because it suggests an individual has ulterior motives (Ham & Vonk, 2011). In criminal investigations, 911 callers who appear to engage in impression management might be perceived as disingenuous and as having ulterior motives. For example, observers might believe a 911 caller who exhibits high levels of politeness to the dispatcher is trying to make a good impression—perhaps to distract from and conceal their potential involvement in the reported crime. Impression management behaviors have been inconsistently linked to deception in 911 calls. Whereas Harpster et al. (2009) reported politeness to the dispatcher is associated with guilty callers, more recent studies have not replicated this finding (Cromer et al., 2019; Miller et al., 2021). As it relates to *suspicion*, we hypothesize that in reporting a violent crime, perceptions of the caller engaging in impression management might be associated with a violation of expectations and suspicion, as prioritizing actions such as politeness might not be expected in such an urgent situation. However, given past inconsistencies for findings between impression management and actual deception, we sought to test this factor in this new context, but suspected perceived impression management would be positively associated with expectation violations and suspicion.

Perceived Information Management

Individuals sometimes manage how they are perceived by carefully selecting, or gatekeeping, which information they present to others. In the context of in-person interviews and interrogations, research shows that, compared with innocent suspects, perpetrators are more concerned with strategically managing the information they present to others (Hartwig et al., 2010; Strömwall et al., 2006). Intuitively, one might expect that an innocent bystander calling 911 is motivated to share any and all information that might prove beneficial to investigators; thus, any appearance of inhibition would trigger suspicion. Yet, it is important to note that a 911 caller who might *want* to freely provide information to the operator might be

unable to do so, given the trauma of the situation. Similar to perceived cognitive load, it is possible listeners might underestimate this external factor and assume an internal motivation to withhold information. Thus, we predicted the caller's perceived information management would be positively associated with expectation violations and suspicion.

Research Overview

To investigate which behaviors on 911 calls trigger suspicion, we conducted (a) three experiments with an associated mini meta-analysis and then (b) a conceptual replication with a larger, more diverse set of real calls. Across studies, we investigated the relationship between perceptions of behavioral predictors on a 911 call and the degree to which expectations were violated and suspicion was triggered in lay observers (Studies 1, 2, and 4) and police officers (Study 3). Participants listened to real (Studies 1 and 4) and simulated, tightly controlled (Studies 2 and 3) 911 calls, half made by women and half by men. We measured perceptions of five theoretical predictors of suspicion—first, by coding participants' spontaneous mentions of each behavior in an open-ended response and then directly through participants' rating of the caller's urgency, emotionality, cognitive load, impression management, and information management. Participants indicated how much the caller violated their expectations, how suspicious they found the caller, and ultimately whether police should expend resources to investigate the caller's role in the crime.

Hypotheses

We coded open-ended first impressions of 911 callers to determine how frequently listeners spontaneously mentioned the five behavioral predictors before they were primed with these concepts in our survey scales. We hypothesized that increased perceptions of (a) urgency and emotionality would be associated with *decreased* suspicion, as listeners likely expect these behaviors in 911 callers reporting emergency situations; and (b) cognitive load, information management, and impression management would be associated with *increased* suspicion, as these behaviors might be perceived as the caller strategically modifying their behavior. Additionally, given stereotypes of women as more emotional than men, we hypothesized there may be other gender-based expectations for 911 callers. Accordingly, we explored whether the five behaviors predicted suspicion differently based on the caller's gender.

Given our similar methodological approaches and hypotheses across Studies 1–3, we present these studies collectively and note key differences where applicable. We conducted a mini meta-analysis to provide overall conclusions about the hypothesized pattern of results and which behaviors are reliable predictors across studies. We report full descriptions of Studies 1–3 individually in [Supplemental Materials](#). Then, we report Study 4 (a conceptual replication) that tested whether our results generalize to a larger, diverse set of real 911 calls.

Transparency and Openness

We collected the data for the following studies within a larger research project. In another article ([Salerno et al., 2025](#)), we focused only on emotion expression and investigated the causal impact of

experimental manipulations of 911 callers' emotion expression and relationship to the victim on suspicion and tested psychological explanations (mediators) of this effect, including moral typecasting and expectancy violation. We powered the present studies to detect effects for that manipulation with three experimental variables (caller emotion, caller relationship to the victim, and caller gender). In this article, we collapsed across caller emotion and relationship to the victim to more broadly investigate how a variety of perceived behaviors trigger suspicion, beyond just manipulated emotion levels. In the present article, we report data not previously analyzed, including the degree to which each of the behavioral factors (urgency, emotionality, cognitive load, information management, and impression managements) predict suspicion and violation of expectations and whether these predictions depend on the caller's gender. Further, we include quantitative coding of the open-ended first impressions of the caller for spontaneous mentions of these behaviors, not reported in the other article. Materials and data files are available on the Open Science Framework (https://osf.io/yat52/?view_only=8154999c3be5483ba061a9018b39756f; Salerno et al., 2024), along with preregistrations for Study 2 (https://osf.io/pquzm?view_only=5c71a52caf7e405fa07b9f6116848981) and Study 4 (https://osf.io/w38tm?view_only=66b677a1736146258c6aed431fcf0a8a).

Studies 1–3

Across Studies 1–3, we tested our hypotheses that perceived urgency and emotionality would be *negatively* associated with suspicion; whereas perceived cognitive load, impression management, and information management would be *positively* associated with suspicion. Additionally, we conducted exploratory tests of whether any of these behaviors differentially predicted suspicion for male versus female callers. In Studies 2 and 3, we directly measured expectation violations, extending Study 1 to investigate which behaviors are associated with violating listeners' expectations for how a 911 caller should behave.

In Study 3, we extended our findings to a law enforcement sample. Although police are more confident in their ability to judge the veracity of suspects' statements than are laypeople, they demonstrate a predisposition to believe other people are lying ([Kassin et al., 2005](#); [Meissner & Kassin, 2002](#)). Anecdotal reports reveal instances of police officers solving crimes by making quick decisions from a gut feeling ([Pinizzotto et al., 2004](#); [Worrall, 2013](#)) and often citing an "inappropriate" level of emotion as an origin of suspicion in many wrongful conviction cases ([Heath, 2009](#); [Leo & Drizin, 2010](#)). In fact, [Leo and Drizin \(2010\)](#) argued that police confidence in the ability to serve as "human lie detectors" is one of the main sources of original misclassification errors of a witness as a suspect. It is possible that, because of their expertise and experience, police might know that people express different levels and types of behaviors (e.g., emotion or varying degrees of cognitive load), so these behaviors might not violate their expectations or trigger suspicion to the same degree as for naïve laypeople. It is also possible, however, that, like laypeople, law enforcement officers rely on stereotypes for their expectations and interpret behaviors that violate their expectations as suspicious, especially when they are under intense cognitive load to quickly solve serious crimes ([National Institute of Justice, 2019](#); [Van Knippenberg et al., 1999](#)).

Method

Participants

Our final samples were as follows: 943 participants in Study 1 (47% female; 77% White, 10% Black, 6.5% Hispanic/Latinx, 5% Asian, 1% other; Amazon's Mechanical Turk), 624 participants in Study 2 (45% female; 71.5% White, 15.4% Black, 6% Hispanic/Latinx, 5% Asian, and 2% other; CloudResearch), and 298 law enforcement officers in Study 3 (21% women; $M_{\text{age}} = 40$, $SD_{\text{age}} = 10$; 75% White, 2.4% Black, 15% Hispanic/Latinx, 1% Asian, and 5% other; recruited through emails distributed via police department newsletters in a large metropolitan area). We report full participant details and information on exclusionary criteria, see [Supplemental Materials](#).

Materials

The institutional review board at Arizona State University approved this research. Across all studies, participants listened to a 911 call, but the nature of the call varied between studies.

Study 1. In Study 1, participants listened to one of six real 911 audio calls we found on the internet in which the caller reported a shooting to emergency services (13–21 s). We employed stimulus sampling (Wells & Windschitl, 1999), such that three of the calls were from female callers and three were from male callers.

Studies 2 and 3. Next, we aimed to replicate Study 1 but in a more controlled design using simulated 911 calls. We created a scripted dialog between a 911 operator and caller based on a set of real 911 calls and replicated specific phrases (see [Appendix A](#) for call script). We hired voice actors to perform a 911 call reporting a shooting, allowing us to control the dialog, duration of the call, details of the crime, operator's responses, and make the caller's guilt relatively ambiguous—thereby better isolating the moderating role of caller gender.

Fourteen professional voice actors (seven female, seven male) performed the same 911 call script with the first actor's recordings as examples and we provided additional feedback as needed. We created a single female operator recording to combine with the voice actors' recordings to hold the operator audio constant. We edited the caller and operator recordings together and enhanced the authenticity of the calls by adding a dial tone at the beginning and a telephone noise filter. All audio recordings were between 73 and 101 s in length due to natural variation in the different actors' pacing. We pilot tested these calls by asking participants from Amazon's Mechanical Turk ($N = 332$) to report open-ended initial impressions of the calls and eliminated any calls recorded by actors that 25% or more of participants spontaneously mentioned sounded fake. This resulted in a final set of nine voice actors (five female, four male).

Given emotionality was a key focus of the original research, each voice actor performed the same script three times, varying their emotional intensity for each one. The effect of this manipulation is reported in the other article (Salerno et al., 2025) and we collapse across it here. This manipulation enabled the present analysis of how important listeners' perceptions of the five behavioral predictors—including but not limited to emotionality—are for suspicion across varying levels of emotionality.

Measures. The following measures were used in every study, with exceptions noted. Full measures for each study are available in [Appendix B](#).

Attention and Manipulation Checks. After reading the instructions, but before listening to the 911 call, participants had to correctly answer three multiple choice attention checks before listening to the call. At the end of each study, participants also indicated the caller's gender as a manipulation check.

Open-Ended First Impressions. Immediately after listening to the 911 call, participants answered an open-ended question about their impression of the caller. Participants were asked, "Please tell us about your first impressions of the caller. Please write a few sentences describing your impression of the person and behaviors that led you to that impression."

Behavioral Predictors. Each of the five behavioral predictor scales comprised five or six items with 5-point response scales ranging from *not at all* to *extremely*. Note, we attempted an additional scale for verbal fluency, but it suffered lower reliability and participants noted interpretability issues in open-ended comments, so we do not report it further here (see [Supplemental Materials](#)). The urgency scale (Cronbach's α : Study 1 = .88) assessed the degree to which participants thought the caller made the situation sound pressing or urgent on behalf of the victim (e.g., *caller made the situation sound urgent, caller seemed concerned about getting help for the victim*). After reading open-ended comments in Study 1, we realized it was important to split the urgency scale into two new behavioral predictors: (a) Urgency for the victim, assessing the degree to which participants felt urgent for the injured person (e.g., *caller seemed concerned about getting help for the victim*) with four items (Cronbach's α : Study 2 = .82, Study 3 = .87); and (b) urgency for the self, assessing the degree to which participants felt urgent for themselves (e.g., *caller seemed concerned for their own safety*) with four items in Study 2 and three items in Study 3 (Cronbach's α : Study 2 = .83, Study 3 = .89).

The emotionality scale (Cronbach's α : Study 1 = .92, Study 2 = .89, Study 3 = .89) assessed the degree to which participants thought the caller was emotional (e.g., *caller was hysterical, caller kept their composure* [reverse scored], *caller seemed very upset*). The cognitive load scale (Cronbach's α : Study 1 = .87, Study 2 = .83, Study 3 = .86) assessed the degree to which participants thought the caller seemed to be under mental load (e.g., *caller seemed to have trouble thinking clearly, caller had difficulty getting the right words out*). The impression management scale (Cronbach's α : Study 1 = .74, Study 2 = .74, Study 3 = .74) assessed the degree to which participants thought the caller was trying to be polite or make a positive impression to the dispatcher (e.g., *caller was actively trying to make a good impression on the dispatcher, caller was polite to the dispatcher*). The information management scale (Cronbach's α : Study 1 = .58, Study 2 = .71, Study 3 = .80) assessed the degree to which participants thought the caller was hiding information or managing what they shared with the dispatcher (e.g., *caller freely provided details about the incident* [reverse scored], *caller gave vague and unclear responses*). Although information management had lower reliability compared with others in Study 1, we elected to keep it based on participants' clear interpretations of the concept in the open-ended impression comments and the literature identifying it as a key variable; fortunately, its reliability improved in subsequent studies.

Outcome Measures. Our primary dependent variable was a five-item suspicion scale (Cronbach's α : Study 1 = .90, Study 2 = .91, Study 3 = .89). Participants rated how suspicious they thought the caller was behaving on the following items: "How likely do you

think it is that the caller is hiding something?" (7-point scale from *extremely unlikely* to *extremely likely* with *neither unlikely nor likely* at the midpoint), "How suspicious was the caller acting?" (6-point scale from *not at all suspicious* to *extremely suspicious*), "How likely do you think it is that the caller played a role in the crime?" (7-point scale from *extremely unlikely* to *extremely likely* with *neither unlikely nor likely* at the midpoint), "How sincere do you think the caller is being?" (reverse scored; 6-point scale from *extremely insincere* to *extremely sincere*), and "How normal do you think the caller is acting?" (reverse scored; 6-point scale from *extremely strange* to *extremely normal*). Higher scores indicate higher levels of suspicion.

Our second dependent variable was a treat as suspect scale (Spearman–Brown coefficient: Study 1 = .87, Study 2 = .88, Study 3 = .62). Participants reported their level of support for police to investigate the caller on two items, on a 7-point scale from *strongly disagree* to *strongly agree*. Items included, *the police should use their resources to investigate the caller's involvement in the crime further* and *the police should treat the caller as a suspect*. Given our results predicting the treat as suspect scale largely replicated our results predicting suspicion, we do not report analysis of this outcome in the main text but include full details and any exceptions in [Supplemental Materials](#).

Additionally, for Studies 2 and 3 we developed another outcome scale directly assessing the degree to which participants thought that the caller's behavior violated their expectations. The violation of expectations scale (Cronbach's α : Study 2 = .83, Study 3 = .80) had four items: *I was genuinely surprised about how the caller acted*, *the caller behaved about how I expected* (reverse scored), *the caller's behavior violated my expectations for what a 911 call would sound like*, and *if I was in the same situation I think I would behave similarly to the caller* (reverse scored).

Procedure

Participants were randomly assigned to listen to one of six real 911 calls (Study 1) or one of the 27 versions of the voice actor calls (Studies 2 and 3). We informed participants that the calls were short, and they might not feel they had enough information, but their gut impressions were informative. Participants were instructed to listen to the call three times before moving on to the next page of the study and were able to listen to it again if they chose to throughout the study. Participants first answered an open-ended question about their initial impressions of the caller and then completed scales for the behavioral predictors assessing perceptions of the caller, the suspicion scale, the treat as suspect scale, violations scale (Studies 2 and 3), the manipulation check, and demographic measures. Lay participants were compensated \$1.50 (Studies 1 and 2) and police officers were compensated with a \$25 Amazon gift card (Study 3) for their participation.

Results

Open-Ended Coding of First Impressions

Independent raters (blinded to which call the participant heard) coded the participants' first impressions of the caller for spontaneous mentions of each behavioral predictor (i.e., urgency, emotionality, cognitive load, impression management, information management;

see codebook in [Supplemental Materials](#)). Following established best practices for quantitative coding (Syed & Nelson, 2015; see also Lombard et al., 2002), the four raters were first trained on a set of 158 comments from pilot data to reach at least 80% agreement. Next, raters worked together on the first 50 comments from the Study 1 actual data to ensure they maintained agreement, after which they coded the remaining impressions on their own. We repeated this process for Studies 2 and 3: three of the four raters from Study 1 began on an additional "refresher" training set with 26 comments (Study 2) and 45 comments (Study 3), to ensure they had retained at least 80% reliability before individually coding the remaining impressions in each study. We confirmed adequate reliability through analyses that took into account chance agreement: across studies, Fleiss' κ ranged from moderate (.55) to almost perfect (.95) agreement (see [Supplemental Materials](#) for full analyses).

Participants spontaneously cited emotionality the most frequently, which included specifically noting the caller was expressing emotion or a lack of emotion. In fact, the majority of participants mentioned emotionality (Study 1: 85%, Study 2: 79%, Study 3 [police]: 77%). After emotionality, participants' mentions of other behavioral predictors dropped substantially, with the next most common being urgency (Study 1: 31%, Study 2: 33%, Study 3 [police]: 20%), information management (Study 1: 18%, Study 2: 22%, Study 3 [police]: 32%), cognitive load (Study 1: 16%, Study 2: 22%, Study 3 [police]: 14%), and impression management (Study 1: 3%, Study 2: 9%, Study 3 [police]: 14%; see [Figure 1](#)).

Of note, participants' frequency of mentioning emotionality did not depend on which 911 call they heard, which varied naturally in emotion level (Study 1) or through stimulus sampling (Studies 2 and 3). Participants frequently cited emotionality, regardless of whether they heard a call with lower or higher emotion, Study 1: $\chi^2(2, N = 880) = 1.78, p = .41$; Study 2: $\chi^2(2, N = 578) = 1.53, p = .47$; Study 3 [police]: $\chi^2(2, N = 293) = 3.51, p = .17$.

Across studies, we found some differences in the frequency of participants' spontaneous mentions of each behavior for male versus female callers. However, we did not find consistent results across our studies as a whole, so we recommend caution in interpreting these differences (see [Supplemental Materials](#) for study-specific findings).

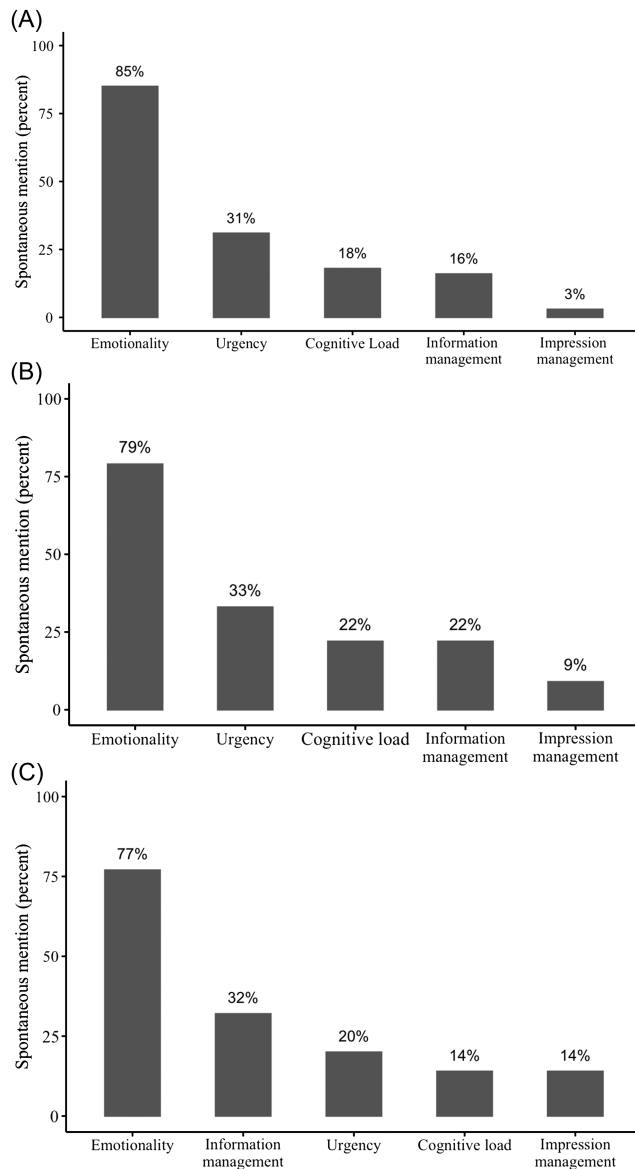
Relationships Among Behavioral Predictors

All descriptive information on the behavioral predictors and correlations among measures are reported in [Supplemental Tables S1–S3](#). Most behavioral predictors were significantly correlated with each other across studies. Most correlations between behavioral predictors were small-to-moderate in size indicating related, but not redundant, constructs. Of note, the strongest relationship was the positive relationship between urgency and emotionality (Study 1: $r = .73, p < .001$; Study 2: $r = .64, p < .001$; Study 3: $r = .70, p < .001$; see [Supplemental Materials](#) for detailed results).

Behavioral Predictors of Suspicion and Violations of Expectations

We began by analyzing each study independently to identify which behavioral predictors were associated with suspicion and found highly consistent patterns, with minor deviations. Given the

Figure 1
Percentage of First Impressions of 911 Callers That Spontaneously Mentioned Each Behavior, Study 1 (A), Study 2 (B), and Study 3 (C)



Note. Spontaneous mentions of each behavioral predictor in participants' first impressions of the 911 caller.

overall similarity of results across studies, we opted to present the main findings via a mini meta-analysis that combines data across all three studies, which provides a more concise and integrative summary of effect sizes and conclusions. Next, we then provide a brief summary of the individual study findings, noting the deviations (see Table 1 and Figures 2–3 for main results and Supplemental Materials for more detailed reporting of individual study results).

Mini Meta-Analysis

Although we found largely consistent relationships across studies, there were minor exceptions. To take advantage of the

similar methodology and measures across Studies 1–3, and to optimize statistical power, we conducted mini meta-analyses. This allowed us to synthesize across studies and get the best estimation of effect size for each behavioral predictor (Goh et al., 2016). Using a fixed-effects approach, we estimated weighted effect sizes based on each study's sample size. Below we share the meta-analytic results for our suspicion outcome variable (see Supplemental Materials for meta-analytic results for violations of expectations).

Behavioral Predictors' Bivariate Relationships With Suspicion

To estimate effect size, we used the bivariate correlation between each behavioral predictor and suspicion. We transformed each correlation to Fisher's z for analysis but transformed back to Pearson r values for interpretations (see Table 2). Using the Stouffer's Z test, we determined that five of the six behavioral predictors were significantly correlated with suspicion, as indicated by meta-analytic values. Across studies, emotionality ($M_r = -.48, p < .001$) had a medium, negative correlation with suspicion. Measured in two studies, urgency for others ($M_r = -.61, p < .001$) had a large, negative correlation with suspicion, and urgency for the self ($M_r = -.07, p = .01$) had a small, negative correlation. Across studies, information management ($M_r = .56, p < .001$) and impression management ($M_r = .43, p < .001$) exhibited large and medium positive correlations with suspicion, respectively. Given the variability in the relation between cognitive load and suspicion across studies (i.e., significant positive correlation in Study 1, no significant relation in Study 2, and significant negative correlation in Study 3, see Table 2), cognitive load did not have a significant meta-analytic correlation with suspicion ($M_r = .002, p = .19$). Of note, cognitive load was not meta-analytically correlated with suspicion but was a unique meta-analytic predictor of suspicion when statistically adjusting for other behavioral predictors (described next).

Behavioral Predictors' Unique Relations to Suspicion

Additionally, we conducted a mini meta-analysis of each behavioral predictor's unique relation to suspicion, controlling for the other behavioral predictors. To estimate effect sizes, we used the semipartial, or "part" correlation, from each study's model predicting suspicion, in which each behavioral predictor's part correlation represents its unique relation to suspicion, statistically adjusting for the other behavioral predictors in the model (Nakagawa & Cuthill, 2007; van den Berg, n.d.). We transformed each part correlation to Fisher's z for analysis but transformed back to Pearson r for interpretations (see Table 3).

Using the Stouffer's Z test, we determined all six behavioral predictors were significantly and uniquely related to suspicion, as indicated by the meta-analytic values. Across studies, emotionality had a small negative effect ($M_r = -.06, p = .01$). Measured in two studies, urgency for others had a medium negative effect ($M_r = -.22, p < .001$) and urgency for the self had a small negative effect ($M_r = -.12, p < .001$). Across studies, information management demonstrated the strongest unique effect with a medium positive value ($M_r = .24, p < .001$), followed by a smaller unique positive for cognitive load ($M_r = .12, p < .001$) and impression management ($M_r = .04, p = .04$).

Table 1

Regression Model Results (Step 1) for Behavioral Predictors Uniquely Predicting Suspicion and Violations of Expectations, Studies 1–3

Behavioral predictor	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Predicting suspicion					
Study 1 model: $R^2 = .55$, $F(6, 934) = 187.01$, $p < .001$					
Urgency	−0.62	0.04	934	−15.93	<.001
Emotionality	−0.12	0.05	934	−2.68	.01
Cognitive load	0.39	0.04	934	10.44	<.001
Information management	0.41	0.06	934	7.55	<.001
Impression management	0.07	0.05	934	1.34	.18
Caller gender	−0.10	0.06	934	−1.59	.11
Study 2 model: $R^2 = .54$, $F(7, 616) = 102.63$, $p < .001$					
Urgency for the victim	−0.55	0.07	616	−8.39	<.001
Urgency for the self	−0.21	0.05	616	−4.36	<.001
Emotionality	−0.18	0.06	616	−3.13	.002
Cognitive load	0.03	0.05	616	0.60	.55
Information management	0.75	0.07	616	11.05	<.001
Impression management	0.10	0.08	616	1.27	.20
Caller gender	−0.04	0.09	616	−0.47	.64
Study 3 model: $R^2 = .55$, $F(7, 290) = 50.53$, $p < .001$					
Urgency for the victim	−0.32	0.06	290	−5.27	<.001
Urgency for the self	−0.18	0.06	290	−2.85	.01
Emotionality	−0.03	0.07	290	−0.48	.63
Cognitive load	−0.06	0.06	290	−0.99	.33
Information management	0.51	0.06	290	8.55	<.001
Impression management	0.19	0.07	290	2.66	.01
Caller gender	−0.33	0.09	290	−3.60	<.001
Predicting violations of expectations					
Study 2 model: $R^2 = .60$, $F(7, 616) = 135.77$, $p < .001$					
Urgency for the victim	−0.40	0.05	616	−7.61	<.001
Urgency for the self	−0.13	0.04	616	−3.46	<.001
Emotionality	−0.48	0.05	616	−10.44	<.001
Cognitive load	0.08	0.04	616	1.76	.08
Information management	0.47	0.05	616	8.63	<.001
Impression management	−0.02	0.06	616	−0.28	.78
Caller gender	−0.07	0.07	616	−1.07	.29
Study 3 model: $R^2 = .52$, $F(7, 290) = 46.55$, $p < .001$					
Urgency for the victim	−0.32	0.06	290	−5.03	<.001
Urgency for the self	−0.15	0.06	290	−2.37	.02
Emotionality	−0.39	0.07	290	−5.61	<.001
Cognitive load	0.11	0.06	290	1.83	.07
Information management	0.33	0.06	290	5.30	<.001
Impression management	−0.01	0.07	290	−0.10	.92
Caller gender	−0.003	0.09	290	−0.03	.98

Note. *SE* = standard error.

Summary of Study-Specific Suspicion Findings

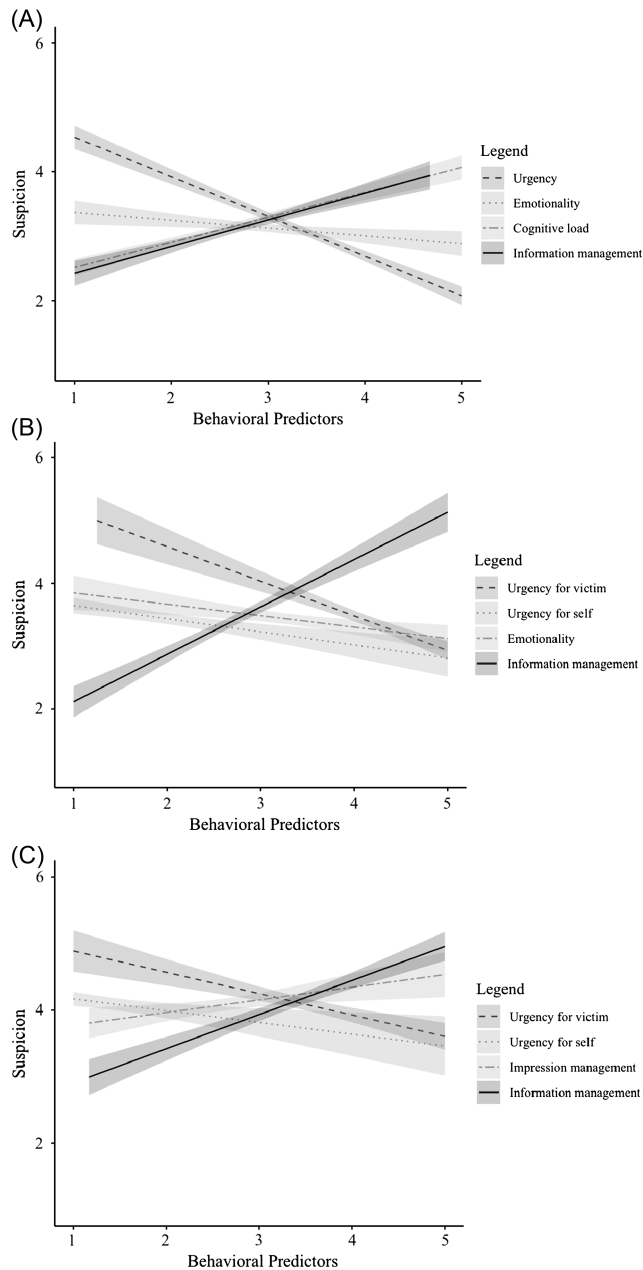
Although we encourage readers to rely on the meta-analytic results, we summarize the study-specific findings here for transparency. Across three studies, and consistent with the meta-analytic findings, bivariate correlations (Supplemental Tables S1–S3) revealed support for our hypotheses that emotionality and urgency for the victim were negatively related to suspicion, while perceived information management and impression management were positively related to suspicion. Urgency for the self and cognitive load were more mixed: They were related to suspicion in some studies, but not others. Hierarchical regressions (Table 1) that included all behavioral predictors and caller gender revealed that urgency, emotionality, and information management largely remained unique predictors while controlling for other behaviors. One of the rare differences across studies was that lay samples (Studies 1 and 2) had

similar suspicion toward male versus female callers, but police officers (Study 3) were significantly less suspicious of female versus male callers. Across studies, we did not find evidence that the relationship between the behavioral predictors and suspicion depended on caller gender (see Supplemental Materials for models with two-way interaction terms for each behavioral predictor and caller gender).

Summary of Study-Specific Violation of Expectations Findings

In Studies 2 and 3, we largely replicated our results predicting suspicion summarized above. Measured in two studies, bivariate correlations (Supplemental Tables S1–S3) revealed support for our hypotheses that emotionality and urgency for the victim were negatively related to violation of expectations, while perceived

Figure 2
Significant and Unique Behavioral Predictors of Suspicion Toward 911 Callers, Study 1 (A), Study 2 (B), and Study 3 (C)



Note. Significant behavioral predictors uniquely associated with suspicion, over and above other predictors and gender. The shading around each line represents the 95% confidence interval.

cognitive load, information management, and impression management were positively related to violation of expectations. Urgency for the self was more mixed: It was related to violation of expectations in Study 3 but not in Study 2. Hierarchical regressions (Table 1) that included all behavioral predictors and caller gender revealed that urgency (for the victim and the self), emotionality, and information management remained unique

predictors while controlling for the other behaviors. Caller gender did not play a significant role in violation of expectations in any of the studies (see Supplemental Materials for models with two-way interaction terms for each behavioral predictor and caller gender).

Studies 1–3 Discussion

Our combined results enable an overall conclusion spanning the first three studies: across real and simulated 911 calls, urgency for the victim, urgency for the self, and emotionality were significantly and uniquely associated with less suspicion; while cognitive load, information management, and impression management were significantly and uniquely associated with greater suspicion. Participants' spontaneous first impressions of the caller revealed the caller's emotionality was *very* salient—77%–85% of participants spontaneously noted the callers' emotion or lack of emotion. We found similar patterns for violation of expectations: Behaviors that increased suspicion were those that violated expectations more (i.e., less emotion and urgency, and more impression management and information management). Finally, laypeople and police officers shared similar judgments for behaviors associated with suspicion and violations of expectations, but officers appeared to find female callers generally less suspicious than male callers—despite male and female callers describing the same incident with the same script.

Study 4

Given that the majority of our results thus far were based on reactions to different versions of one scripted, simulated call, we conducted a conceptual replication to test whether our findings generalized to a larger, more diverse set of real 911 calls. In a within-subjects design, participants listened to eight real 911 calls and rated each caller on the six behavioral predictors and reported how suspicious they found the callers to be. We predicted replication: urgency for the victim, urgency for the self, and emotionality would be uniquely associated with *less* suspicion; cognitive load, information management, and impression management would be uniquely associated with *more* suspicion. We also tested whether these relationships depended on gender.

Method

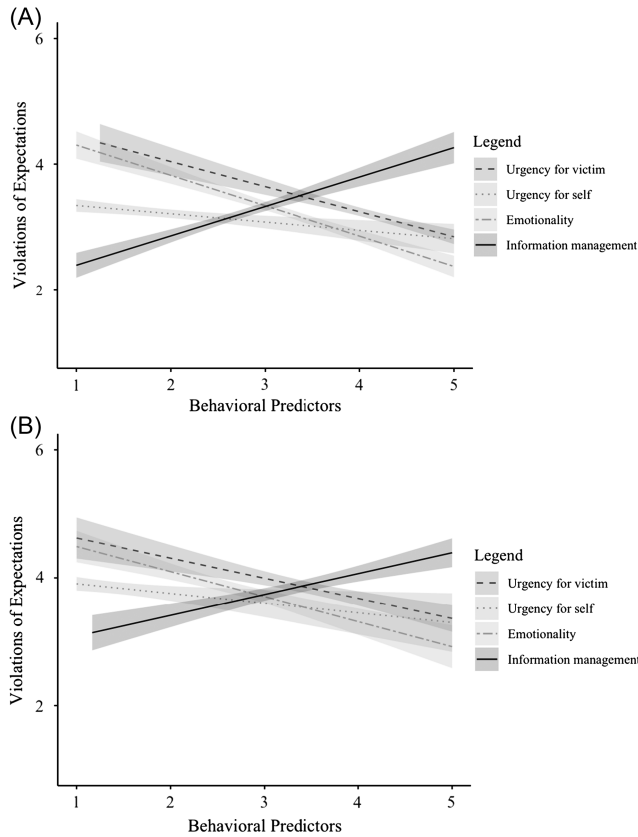
Participants

We recruited 250 participants from Connect CloudResearch. Based on prior research recommending 20 ratings per stimulus (Lloyd et al., 2018), we determined that, for our sample of 88 calls (described below), if each participant rated eight calls, we would need a minimum of 220 participants (i.e., 220 participants times eight calls is 1,760 ratings; and 1,760 ratings divided by 88 calls is 20 ratings per call). All participants passed attention checks, leading to a final sample of 250 participants (43.2% women, 53.2% men, 3.6% nonbinary, or other; $M_{\text{age}} = 38.71$, $SD_{\text{age}} = 10.94$; 69.2% White, 10.4% Black, 9.6% Hispanic/Latinx, 8% Asian, 2.8% other).

Materials

We used 88 calls reporting violent encounters that resulted in bodily harm or death (41 female callers, 47 male callers) obtained from two police departments in a large city (see Appendix C for

Figure 3
Significant and Unique Behavioral Predictors of Violation of Expectations Toward 911 Callers, Study 2 (A) and Study 3 (B)



Note. Significant behavioral predictors uniquely associated with violation of expectations, over and above other predictors and gender. The shading around each line represents the 95% confidence interval.

details). We created eight bins of calls based on call duration, with 11 calls per bin, and randomly assigned participants to listen to one call from each of the eight bins to balance the length of the study. Call lengths ranged from 34 s to 9 min and 21 s ($M = 3:31$, $SD = 1:54$). Participants took 54 min on average to complete the study.

Measures

Attention and Honesty Checks. Participants listened to an audio recording specifying a number and then had to accurately

report that number to continue with the study. At the end of the study, participants answered an honesty check about whether they responded to the best of their ability (without risk of losing compensation).

Behavioral Predictors. Participants rated each behavioral predictor for each call they heard. For time sake, participants responded to a single item for each behavioral predictor rather than the full scales used in Studies 1–3 (see Appendix B for full list). The single item was chosen by determining the item closest to the mean for each scale from Study 1, which also used real 911 calls. Participants rated each item on 5-point Likert scales from *not at all* to *extremely*.

Outcome Measures. Participants indicated how suspicious they perceived the caller to be and how much they thought police should investigate the caller as a suspect on 6-point Likert scales from *not at all* to *extremely*. Results for the “treat as a suspect” item largely replicated suspicion results, so we report those results in Supplemental Materials.

Procedure

After listening to each call, participants provided an open-ended, one-sentence impression of the call they heard and then reported their perceptions of the six behavioral predictors and outcome measures. Participants were compensated \$12.

Results

Relationships Among Behavioral Predictors

Descriptives and correlations are reported in Supplemental Table S4. Like Studies 1–3, the strongest relationship was between emotionality and urgency for the victim.

Behavioral Predictors’ Unique Relations to Suspicion

Because each of the 88 calls were rated by a subset of participants, we analyzed the data via a multilevel model with behavioral predictor ratings nested within call. This model included the six behavioral predictors (grand mean centered) as Level 1 variables and the caller gender at Level 2. Random effects included participant at Level 1 and call at Level 2. The model included the cross-level interaction between each behavioral predictor and caller gender. We fit this model with restricted maximum likelihood using the lme4 package in R (Bates et al., 2015). The *t* tests to assess fixed effects were conducted using the lmerTest package (Kuznetsova et al., 2017). The full summary of the final model can be found in Table 4.

Table 2
Correlations of Each Behavioral Predictors Relation to Suspicion in Studies 1–3

Behavioral predictor	Study 1		Study 2		Study 3		Combined	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Urgency for the victim			-.62	<.001	-.56	<.001	-.61	<.001
Urgency for the self			.03	.42	-.27	<.001	-.07	.01
Emotionality	-.46	<.001	-.51	<.001	-.49	<.001	-.48	<.001
Cognitive load	.10	.001	-.02	.61	-.27	<.001	.002	.19
Information management	.51	<.001	.60	<.001	.60	<.001	.56	<.001
Impression management	.38	<.001	.49	<.001	.46	<.001	.43	<.001

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

Table 3
Semipartial (“Part”) Correlations for Each Behavioral Predictor’s Unique Relation to Suspicion in Studies 1–3

Behavioral predictor	Study 1		Study 2		Study 3		Combined	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Urgency for the victim			-.23	<.001	-.21	<.001	-.22	<.001
Urgency for the self			-.12	<.001	-.11	.005	-.12	<.001
Emotionality	-.06	.008	-.09	.002	-.02	.63	-.06	.01
Cognitive load	.23	<.001	.02	.55	-.04	.33	.12	<.001
Information management	.17	<.001	.30	<.001	.34	<.001	.24	<.001
Impression management	.03	.18	.04	.20	.11	.008	.04	.04

Consistent with the meta-analytic results, two behavioral predictors uniquely predicted less suspicion: urgency for the victim, $b = -0.24$, $SE = 0.03$, $p < .001$, and emotionality, $b = -0.07$, $SE = 0.03$, $p = .02$, and three behavioral predictors uniquely predicted more suspicion: cognitive load, $b = 0.02$, $SE = 0.03$, $p < .001$; information management, $b = 0.43$, $SE = 0.03$, $p < .001$; and impression management, $b = 0.05$, $SE = 0.02$, $p = .04$. Contrary to Studies 1–3, urgency for the self ($p = .70$) was not a significant unique predictor of suspicion. Consistent with the lay sample studies, gender ($p = .18$) was also not a significant unique predictor.

Although there was only one significant cross-level interaction, the addition of the Level 2 predictor and cross-level interactions significantly improved model fit, $\chi^2(1) = 14.38$, $p = .05$, so they were included in the final model. There was a significant interaction between cognitive load and caller gender, $b = -0.09$, $SE = 0.04$, $p = .03$. Simple slopes analysis revealed that cognitive load was a significant predictor of suspicion for both genders, but this relationship was stronger for male callers, $b = 0.21$, $SE = 0.03$, $p < .01$, than female callers, $b = 0.12$, $SE = 0.03$, $p < .01$. All other interactions were not significant.

Table 4
Multilevel Model Summary Predicting Suspicion, Study 4

Behavioral predictor	<i>b</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>p</i>
Intercept	1.83	0.08	105	24.45	<.001
Level 1 variables					
Urgency for the victim	-0.24	0.03	1,953	-9.52	<.001
Urgency for the self	0.01	0.03	1,852	0.39	.70
Emotionality	-0.07	0.03	1,708	-2.33	.02
Cognitive load	0.21	0.03	1,724	7.28	<.001
Information management	0.43	0.03	1,951	16.12	<.001
Impression management	0.05	0.02	1,948	2.12	.04
Level 2 variable					
Caller gender	-0.14	0.10	81	-1.36	.18
Interaction terms					
Urgency for Victim × Gender	0.05	0.04	1,870	1.30	.19
Urgency for Self × Gender	-0.04	0.04	1,770	-0.94	.35
Emotionality × Gender	0.00	0.04	1,725	0.09	.93
Cognitive Load × Gender	-0.09	0.04	1,825	-2.18	.03
Information × Gender	-0.01	0.04	1,851	-0.31	.76
Impression × Gender	-0.06	0.03	1,858	-1.75	.08
Random effect					
	Variance		<i>SD</i>		
Participant (Level 1)	0.20		0.44		
Call (Level 2)	0.19		0.44		
Residual	0.73		0.85		

Note. *SE* = standard error.

Discussion

We found the mini meta-analytic results from Studies 1–3 generalized to a larger, more diverse set of real 911 calls. The only substantial deviation from previous studies was that in response to the diverse set of calls, when participants perceived the caller to be more urgent on behalf of themselves they were *more* suspicious; while in previous studies urgency for the self was related to *less* suspicion.

General Discussion

Police often form hunches that a witness is concealing something based on unexpected behavior—hunches that can trigger cumulative confirmation biases throughout the investigation and ultimately wrongful convictions (Scherr et al., 2020). Across four studies, we identified several theoretically relevant behaviors that might violate expectations for what is considered “normal” and trigger suspicion toward 911 callers among laypeople and law enforcement officers. In three studies, we found that the caller’s level of emotion was spontaneously mentioned the most (77%–85%). This effect was consistent across stimuli (real and simulated 911 calls), samples (laypeople, law enforcement officers) and regardless of whether the caller was expressing low, moderate, or high emotion. This consistent and robust finding suggests the importance and salience of emotionality in 911 calls for both lay and law enforcement samples—even without being primed to think about these concepts. We also found several other behaviors that triggered suspicion, described next.

Behavioral Predictors of Suspicion

When laypeople and police were directly asked to consider the behaviors we expected might trigger suspicion, we found several that were consistently related to suspicion. Across all studies, samples, and types of calls, a mini meta-analysis ($N = 1,856$) revealed that the degree to which participants perceived a 911 caller to be (a) urgent (for the victim or themselves) and emotional were correlated with *less* suspicion and (b) engaging in information management and impression management were correlated with *more* suspicion. We replicated these relationships when participants judged a larger set of diverse 911 calls—the only exception being that urgency for the self was no longer related to suspicion. The mini meta-analytic results also identified these relationships in a model that included all predictors simultaneously—meaning the

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

behavioral predictors are not redundant and each behavior is uniquely related to suspicion.

Findings about how cognitive load was associated with suspicion were less consistent. The mini meta-analysis revealed greater perceived cognitive load was uniquely related to more suspicion, but the bivariate correlation was not significant. Looking at the individual studies, cognitive load was not significantly related to suspicion in two of four studies, however, both studies that used real 911 calls (Studies 1 and 4), found callers perceived to be under more cognitive load were perceived as more suspicious.

Behaviors that were associated with suspicion for laypeople were generally the same behaviors that triggered suspicion among police officers, with one exception. In lay samples, callers perceived as more emotional were less suspicious—a relationship that remained even when controlling for other behavioral predictors. This was also the case for law enforcement (i.e., emotion was correlated with suspicion), but emotionality was not a *unique* predictor in our police sample when controlling for other behavioral predictors. Officers might be confounding emotionality and urgency for the victim—this is consistent with the high correlations among urgency for the victim and emotionality. Although future research should more systematically disentangle these two concepts, it is likely these two behaviors co-occur in real-world 911 calls.

Law enforcement officers were more suspicious of male, compared with female, callers—a pattern not found in the samples with lay individuals. This gender bias manifested in the simulated calls, meaning police officers thought male callers were more suspicious than female callers despite all callers describing the same situation with the same exact words. This finding is somewhat consistent with findings in the lie detection literature, suggesting women may be seen as “easier to read” (e.g., Lloyd et al., 2018). However, we did not find any consistent effects that caller gender affected which behaviors triggered suspicion in either sample. Thus, the same behaviors appear to be similarly suspicious for male and female callers.

Violation of Expectations

In two studies, we directly measured how 911 callers' behaviors violated listeners' expectations and found the pattern of results largely mirrored what behaviors predicted suspicion. When laypeople (Study 2) and law enforcement officers (Study 3) perceived callers to be more urgent for the victim or more emotional, they reported their expectations were less violated—suggesting they expected callers to be emotional and urgent. When laypeople and officers perceived callers to be engaging in more information management, they reported their expectations were more violated—suggesting they expected callers would not manage the flow of information. These relationships again replicated with all behaviors included in the model, suggesting these behaviors are not redundant, but uniquely explain violation of expectations while controlling for other behaviors. Although our correlational design does not allow for causal conclusions, the data suggest behaviors that violate expectations might also trigger suspicion.

Are Expectations for Behaviors on a 911 Call Warranted?

An important consideration is whether the behavioral predictors that we found were associated with suspicion *should* be associated

with suspicion. Our findings suggest people have expectations for how 911 callers should behave, yet it may be unrealistic to have these behavioral expectations, as there is no “one size fits all” for responses to traumatic events (see Salerno, 2021). Individuals who witness acts of violence against another person often experience just as much trauma as victims themselves (Substance Abuse and Mental Health Services Administration, 2014). Moreover, witnesses who call 911 are expected to effectively communicate details while behaving in ways consistent with listeners' expectations. Yet, like victims, witnesses experience a wide range of emotional reactions, both negative and positive (Matsumoto & Hwang, 2014). Some individuals' initial reactions to trauma are characterized by intense emotions (e.g., anxiety, agitation, anger) that manifest in various physical (e.g., shaking, rapid breathing) and behavioral (e.g., startled reaction, restlessness) ways, while other individuals initially react to trauma with numbness, dissociation, and withdrawal (Substance Abuse and Mental Health Services Administration, 2014; U.S. Department of Veterans Affairs, 2022).

Thus, a witness's immediate reaction is complicated and impacted by an array of sociocultural factors: including personal history with trauma, experience with emotional expression in their family and community, relationship to the victim, physical proximity to the event, and national culture (Substance Abuse and Mental Health Services Administration, 2014; Matsumoto & Hwang, 2014). These expectations become even more problematic when applied to those who are neurodivergent. For example, many are fighting to save Robert Roberson from execution based on now-debunked “shaken baby evidence.” He was originally targeted because nurses and detectives thought he was not emotional enough—failing to recognize that he is autistic and does not express emotion in expected ways. We believe that no responses to traumatic events should be considered more natural or genuine than others.

Just as there is no “one size fits all” for “appropriate” emotional reactions to trauma, expectations for other behaviors in 911 calls (i.e., urgency, cognitive load, information management, and impression management) might also be based on unrealistic expectations that not everyone is able to follow. Finding these behaviors suspicious is likely based on assuming intentionality (i.e., the caller is trying to conceal their involvement) when, in reality, variation in these behaviors might indicate *inability* (i.e., the caller is not able to provide information due to trauma putting them under cognitive load). This is consistent with findings from trauma-informed investigative interviewing (e.g., Fisher & Geiselman, 1992), which demonstrate that trauma-exposed individuals often struggle to provide coherent, chronological accounts. In other words, someone who does not provide a clear narrative of what happened in a 911 call might do so because they *cannot* due to trauma, or because they *choose not to* in order to hide or obscure information. This is a common problem noted in lie detection literature: Although humans can detect these behaviors (e.g., emotion, information management), they cannot reliably determine the cause of the behaviors (e.g., whether intentional or resulting from trauma). Despite the many reasons a person might exhibit these behaviors, police still erroneously rely on these behaviors as indicators of guilt or innocence.

Real-World Implications

Calling 911 is typically a well-intentioned act of citizenry, yet our findings show that these calls can trigger suspicion. Violations of

socially constructed—but not scientifically supported—expectations can lead listeners to infer that the caller may be involved in the crime they report, opening them up to scrutiny and investigation. As noted earlier, several tragic cases have come to light in which innocent 911 callers were treated with suspicion based on their behavior during the call, resulting in arrest, prosecution, wrongful conviction, and incarceration.

Training on 911 call analysis, developed by a retired police deputy, has grown in popularity despite a lack of empirical support (B. Murphy, 2022). The course and instructor claim to teach police and prosecutors how to determine whether a caller is the perpetrator based on behavioral cues. Although the cues that are supposed to signal guilt have failed to replicate in controlled experiments (Burns & Moffitt, 2014; Markey et al., 2022; Miller et al., 2021), police officers are still taking this training, and using their interpretations of these behaviors as evidence of a caller's innocence or guilt. A very recent example is the case of Jessica Logan, in which Jessica slept through her alarm meant to prompt her to give her sick child his medicine, only to find her infant son had stopped breathing. She called 911 and the detective used her call as proof that she had something to do with her son's death, citing multiple indicators of supposed guilt on the checklist (B. Murphy, 2022). However, Markey et al. (2025) investigated the validity of Harpster and Adams (2017) Considering Offender Probability Statements checklist and found it failed to differentiate between deceptive "false allegation" and truthful "true report" callers. Despite its growing popularity in real world cases, a growing body of research raises serious questions about the so-called 911 linguistic analysis.

Investigators' hunches about suspicious demeanor risk setting off guilt-presumptive chain reactions. By identifying specific behaviors linked to suspicion (e.g., urgency, emotionality, information management), this article provided a foundation for testing the veracity of these intuitions. It is important to investigate these intuitions, because as other scholars have pointed out,

Police need to learn that the scientific research literature does not support their belief in their superior human lie detection abilities; they cannot reliably intuit whether a suspect is innocent or guilty on the basis of their perceptions of his or her demeanor ... and nonverbal behavior. (Leo & Drizin, 2010, p. 26)

Limitations

This research has several limitations. We relied on both real and simulated 911 calls, each with tradeoffs: real calls provide ecological validity but less control, while simulated calls allow for tighter control but may lack realism. Encouragingly, results converged across both approaches. Additionally, Study 4's statistical models (using 88 real calls) accounted for random effects including participant and call, enabling us to conceptually generalize to the larger universe of unsampled calls (Judd et al., 2012). Still, our findings are correlational, so we cannot draw causal inferences regarding how specific behavioral predictors directly lead to suspicion. We believe these findings are an important step and future research could manipulate behavioral predictors to investigate potential causal effects. Specifically with emotionality, we demonstrated the causal effect of 911 caller emotion in another article related to the present studies, and found that failing to express expected levels of emotion *caused* increased suspicion via moral

inferences about the caller (Salerno et al., 2025). Doing so with the other behaviors is important.

It is also worth noting that not all behavioral predictors were unique once entered into the same model, reflecting potential overlap in how observers perceive emotion, urgency, and cognitive load. While our mini meta-analytic results suggest all behavioral predictors uniquely predicted suspicion, we encourage future research to explore potential interactions among behaviors (e.g., whether urgency exacerbates the effects of cognitive load on suspicion). Finding significant correlates but not unique prediction might suggest some behaviors are indicators of an underlying mental state—for example, increased cognitive load might be perceived as a result of emotionality, or emotionality might be perceived as an expression of urgency, or vice versa. While our focus was on examining the unique contributions of each behavior, controlling for shared variance, we hope future research will explore possible interactions between and among behaviors and use experimental manipulations to tease apart constructs and investigate the psychological processes more clearly.

Conclusion

Across four studies, we identified behaviors on 911 calls that relate to suspicion. For lay and law enforcement samples, with real and simulated 911 calls, callers (regardless of gender) were seen as *less* suspicious when they expressed greater urgency or were more emotional; callers were seen as *more* suspicious when they were perceived to be managing information or impressions of them. Our findings suggest that Good Samaritans who call 911 to help others might be targeted for suspicion if their perceived behavior violates potentially misguided expectations for what constitutes "normal" or expected behavior during a traumatic event.

References

- Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, *64*(3), 431–441. <https://doi.org/10.1037/0022-3514.64.3.431>
- Ask, K., & Landström, S. (2010). Why emotions matter: Expectancy violation and affective response mediate the emotional victim effect. *Law and Human Behavior*, *34*(5), 392–401. <https://doi.org/10.1007/s10979-009-9208-6>
- Barone, M. J., Manning, K. C., & Miniard, P. W. (2004). Consumer response to retailers' use of partially comparative pricing. *Journal of Marketing*, *68*(3), 37–47. <https://doi.org/10.1509/jmkg.68.3.37.34769>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bean, S. R., Salerno, J. M., Duran, N., Kassim, S., Bussey, L., Olson, B., & Phalen, H. (2022). *What makes people suspicious? Identifying behavioral predictors of suspicion toward 911 callers* [Paper presentation]. Annual American Psychology-Law Society Conference, Denver, CO, United States.
- Bollingmo, G. C., Wessel, E. O., Eilertsen, D. E., & Magnussen, S. (2008). Credibility of the emotional witness: A study of ratings by police investigators. *Psychology, Crime & Law*, *14*(1), 29–40. <https://doi.org/10.1080/10683160701368412>
- Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*(3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2

- Bond, C. F., Jr., Omar, A., Pitre, U., Lashley, B. R., Skaggs, L. M., & Kirk, C. T. (1992). Fishy-looking liars: Deception judgment from expectancy violation. *Journal of Personality and Social Psychology*, 63(6), 969–977. <https://doi.org/10.1037/0022-3514.63.6.969>
- Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, 6(3), 203–242. <https://doi.org/10.1111/j.1468-2885.1996.tb00127.x>
- Burleigh, N. (2011). *The fatal gift of beauty: The trials of Amanda Knox*. Broadway Books. <https://www.penguinrandomhouse.com/books/197882/the-fatal-gift-of-beauty-by-nina-burleigh/>
- Burns, M. B., & Moffitt, K. C. (2014). Automated deception detection of 911 call transcripts. *Security Informatics*, 3(1), Article 8. <https://doi.org/10.1186/s13388-014-0008-2>
- Clarke, R., & Prokupecz, S. (2024, September 5). Fontana police “psychologically tortured” man to falsely confess to killing his father. CNN. <https://www.cnn.com/2024/09/05/us/fontana-pressured-murder-confession>
- Colwell, K., Hiscock-Anisman, C., Memon, A., Woods, D., & Michlik, P. M. (2006). Strategies of impression management among deceivers and truth-tellers: How liars attempt to convince. *American Journal of Forensic Psychology*, 24(2), 31–38.
- Cromer, J. D., Brewster, J., Fogler, K., & Stoloff, M. (2019). 911 calls in homicide cases: What does the verbal behavior of the caller reveal? *Journal of Police and Criminal Psychology*, 34(2), 156–164. <https://doi.org/10.1007/s11896-018-9282-0>
- DeCarlo, T. E. (2005). The effects of sales message and suspicion of ulterior motives on salesperson evaluation. *Journal of Consumer Psychology*, 15(3), 238–249. https://doi.org/10.1207/s15327663jcp1503_9
- Denault, V., Talwar, V., Plusquellec, P., & Larivière, V. (2022). On deception and lying: An overview of over 100 years of social science research. *Applied Cognitive Psychology*, 36(4), 805–819. <https://doi.org/10.1002/acp.3971>
- DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, 111(2), 203–243. <https://doi.org/10.1037/0033-2909.111.2.203>
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74–118. <https://doi.org/10.1037/0033-2909.129.1.74>
- Duran, N. D., Hall, C., McCarthy, P. M., & Mcnamara, D. S. (2010). The linguistic correlates of conversational deception: Comparing natural language processing technologies. *Applied Psycholinguistics*, 31(3), 439–462. <https://doi.org/10.1017/S0142716410000068>
- Fabes, R. A., & Martin, C. L. (1991). Gender and age stereotypes of emotionality. *Personality and Social Psychology Bulletin*, 17(5), 532–540. <https://doi.org/10.1177/0146167291175008>
- Fein, S. (1996). Effects of suspicion on attributional thinking and the correspondence bias. *Journal of Personality and Social Psychology*, 70(6), 1164–1184. <https://doi.org/10.1037/0022-3514.70.6.1164>
- Ferguson, J. L., Ellen, P. S., & Piscopo, G. H. (2011). Suspicion and perceptions of price fairness in times of crisis. *Journal of Business Ethics*, 98(2), 331–349. <https://doi.org/10.1007/s10551-010-0552-8>
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Charles C Thomas Publisher.
- Fuller, C. M., Biros, D. P., Burgoon, J. K., & Nunamaker, J. F. (2013). An examination and validation of linguistic constructs for studying high-stakes deception. *Group Decision and Negotiation*, 22(1), 117–134. <https://doi.org/10.1007/s10726-012-9300-z>
- Gauger, G., & Von Bergen, J. (2008). *In spite of the system: A personal story of wrongful conviction & exoneration*. Fourcatfarm Press. <https://books.google.com/books?id=Bf6sPQAACAAJ>
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, 117(1), 21–38. <https://doi.org/10.1037/0033-2909.117.1.21>
- Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, 10(10), 535–549. <https://doi.org/10.1111/spc3.12267>
- Hackett, L., Day, A., & Mohr, P. (2008). Expectancy violation and perceptions of rape victim credibility. *Legal and Criminological Psychology*, 13(2), 323–334. <https://doi.org/10.1348/135532507X228458>
- Ham, J., & Vonk, R. (2011). Impressions of impression management: Evidence of spontaneous suspicion of ulterior motivation. *Journal of Experimental Social Psychology*, 47(2), 466–471. <https://doi.org/10.1016/j.jesp.2010.12.008>
- Harpster, T., & Adams, S. H. (2017). *Analyzing 911 homicide calls: Practical aspects and applications*. CRC Press. <https://doi.org/10.1201/9781315386508>
- Harpster, T., Adams, S. H., & Jarvis, J. P. (2009). Analyzing 911 homicide calls for indicators of guilt or innocence: An exploratory analysis. *Homicide Studies*, 13(1), 69–93. <https://doi.org/10.1177/1088767908328073>
- Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, 137(4), 643–659. <https://doi.org/10.1037/a0023589>
- Hartwig, M., & Bond, C. F., Jr. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology*, 28(5), 661–676. <https://doi.org/10.1002/acp.3052>
- Hartwig, M., Granhag, P., Strömwall, L., & Doering, N. (2010). Impression and information management: On the strategic self-regulation of innocent and guilty suspects. *The Open Criminology Journal*, 3(1), 10–16. <https://doi.org/10.2174/1874917801003010010>
- Heath, W. P. (2009). Arresting and convicting the innocent: The potential role of an “inappropriate” emotional display in the accused. *Behavioral Sciences & the Law*, 27(3), 313–332. <https://doi.org/10.1002/bsl.864>
- Heath, W. P., & Grannemann, B. D. (2015). Expectations for defendant emotion. *Applied Psychology in Criminal Justice*, 11(2), 126–145.
- Heath, W. P., Grannemann, B. D., & Grohosky, A. R. (2007). *Mock jurors’ expectations for defendants’ emotional responses*. Association for Psychological Science.
- Heider, F. (1958). *The psychology of interpersonal relations*. Wiley. <https://doi.org/10.1037/10628-000>
- Honts, C., Thurber, S., & Handler, M. (2021). A comprehensive meta-analysis of the comparison question polygraph test. *Applied Cognitive Psychology*, 35(2), 411–427. <https://doi.org/10.1002/acp.3779>
- Inbau, F. E., Reid, J., Buckley, J., & Jayne, B. (2013). *Criminal interrogation and confessions*. Jones & Bartlett Publishers.
- Isserlin, L., Zerach, G., & Solomon, Z. (2008). Acute stress responses: A review and synthesis of ASD, ASR, and CSR. *American Journal of Orthopsychiatry*, 78(4), 423–429. <https://doi.org/10.1037/a0014304>
- Johnson, M. (2003). “The interrogation of Michael Crowe”: A film review focused on education and training. *American Journal of Forensic Psychology*, 21(3), 71–79.
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions the attribution process in person perception. In I. L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–266). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60107-0](https://doi.org/10.1016/S0065-2601(08)60107-0)
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103(1), 54–69. <https://doi.org/10.1037/a0028347>
- Kassin, S. M., Cleary, H. M. D., Gudjonsson, G. H., Leo, R. A., Meissner, C. A., Redlich, A. D., & Scherr, K. C. (2025). Police-induced confessions, 2.0: Risk factors and recommendations. *Law and Human Behavior*, 49(1), 7–53. <https://doi.org/10.1037/lhb0000593>
- Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and

- recommendations. *Law and Human Behavior*, 34(1), 3–38. <https://doi.org/10.1007/s10979-009-9188-6>
- Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1), 42–52. <https://doi.org/10.1016/j.jarmac.2013.01.001>
- Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, 5(2), 33–67. <https://doi.org/10.1111/j.1529-1006.2004.00016.x>
- Kassin, S. M., Meissner, C. A., & Norwick, R. J. (2005). “I’d know a false confession if I saw one”: A comparative study of college students and police investigators. *Law and Human Behavior*, 29(2), 211–227. <https://doi.org/10.1007/s10979-005-2416-9>
- Kaufmann, G., Drevland, G., Wessel, E., Overskeid, G., & Magnussen, S. (2003). The importance of being earnest: Displayed emotions and witness credibility. *Applied Cognitive Psychology*, 17(1), 21–34. <https://doi.org/10.1002/acp.842>
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 15, pp. 192–238). University of Nebraska Press.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lafrenière, M.-A. K., Sedikides, C., Van Tongeren, D. R., & Davis, J. (2016). On the perceived intentionality of self-enhancement. *The Journal of Social Psychology*, 156(1), 28–42. <https://doi.org/10.1080/00224545.2015.1041447>
- Leo, R. A., & Drizin, S. A. (2010). The three errors: Pathways to false confession and wrongful conviction. In G. D. Lassiter & C. A. Meissner (Eds.), *Police interrogations and false confessions: Current research, practice, and policy recommendations* (pp. 9–30). American Psychological Association. <https://doi.org/10.1037/12085-001>
- Levine, T., Anders, L., Banas, J., Baum, K., Endo, K., Hu, A., & Wong, N. (2000). Norms, expectations, and deception: A norm violation model of veracity judgments. *Communication Monographs*, 67(2), 123–137. <https://doi.org/10.1080/03637750009376500>
- Lloyd, E. P., Summers, K. M., Hugenberg, K., & McConnell, A. R. (2018). Revisiting perceiver and target gender effects in deception detection. *Journal of Nonverbal Behavior*, 42(4), 427–440. <https://doi.org/10.1007/s10919-018-0283-6>
- Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4), 587–604. <https://doi.org/10.1111/j.1468-2958.2002.tb00826.x>
- Lowrey-Kinberg, B., Senn, S. L., Dunn, K., Gould, J. B., & Hail-Jares, K. (2019). Origin of implication: How do innocent individuals enter the criminal justice system? *Crime & Delinquency*, 65(14), 1949–1975. <https://doi.org/10.1177/0011128718793618>
- Luke, T. J. (2019). Lessons from Pinocchio: Cues to deception may be highly exaggerated. *Perspectives on Psychological Science*, 14(4), 646–671. <https://doi.org/10.1177/1745691619838258>
- Markey, P. M., Dapice, J., Goldman, S., Nicolas, T., Plickys, E., & Saj, S. (2025). Validity concerns of the Considering Offender Probability Statements Scale in 911 homicide call analysis: An empirical study. *Psychology, Public Policy, and Law*, 31(1), 22–30. <https://doi.org/10.1037/law0000432>
- Markey, P. M., Feeney, E., Berry, B., Hopkins, L., & Creedon, I. (2022). Deception cues during high-risk situations: 911 homicide calls. *Psychological Science*, 33(7), 1040–1047. <https://doi.org/10.1177/09567976221077216>
- Masip, J., Blandón-Gitlin, I., Martínez, C., Herrero, C., & Ibabe, I. (2016). Strategic interviewing to detect deception: Cues to deception across repeated interviews. *Frontiers in Psychology*, 7, Article 1702. <https://doi.org/10.3389/fpsyg.2016.01702>
- Masip, J., Sporer, S. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law*, 11(1), 99–122. <https://doi.org/10.1080/10683160410001726356>
- Matsumoto, D., & Hwang, H. C. (2014). Judgments of subtle facial expressions of emotion. *Emotion*, 14(2), 349–357. <https://doi.org/10.1037/a0035237>
- McAuliff, B. D., Lapin, J., & Michel, S. (2015). Support person presence and child victim testimony: Believe it or not. *Behavioral Sciences & the Law*, 33(4), 508–527. <https://doi.org/10.1002/bsl.2190>
- Meissner, C. A., & Kassin, S. M. (2002). “He’s guilty!”: Investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26(5), 469–480. <https://doi.org/10.1023/A:1020278620751>
- Miller, M. L., Merola, M. A., Opanashuk, L., Robins, C. J., Chancellor, A. S., & Craun, S. W. (2021). “911 what’s your emergency?”: Deception in 911 homicide and homicide staged as suicide calls. *Homicide Studies*, 25(3), 256–272. <https://doi.org/10.1177/1088767920948242>
- Murphy, B. (2022, December 28). *They called 911 for help. Police and prosecutors used a new junk science to decide they were liars*. ProPublica. <https://www.propublica.org/article/911-call-analysis-fbi-police-courts>
- Murphy, B. (2023, February 24). *Prosecutors and judges push for conviction reviews, ban on junk science of 911 call analysis*. ProPublica. <https://www.propublica.org/article/911-call-analysis-ban-review-convictions>
- Murphy, N. A., Hall, J. A., Ruben, M. A., Fraundorfer, D., Schmid Mast, M., Johnson, K. E., & Nguyen, L. (2019). Predictive validity of thin-slice nonverbal behavior from social interactions. *Personality and Social Psychology Bulletin*, 45(7), 983–993. <https://doi.org/10.1177/014616167218802834>
- Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biological Reviews*, 82(4), 591–605. <https://doi.org/10.1111/j.1469-185X.2007.00027.x>
- Natapoff, A. (2022). *Snitching: Criminal informants and the erosion of American justice*. New York University Press. <https://doi.org/10.18574/nyu/9781479807741.001.0001>
- National Institute of Justice. (2019). *National Institute of Justice annual report*. <https://www.ojp.gov/pdffiles1/nij/304993.pdf>
- National Research Council. (2009). *Strengthening forensic science in the United States: A path forward*. National Academies Press. <https://doi.org/10.17226/12589>
- Neuschatz, J. S., & Golding, J. M. (2022). *Jailhouse informants: Psychological and legal perspectives*. New York University Press. <https://nyupress.org/9781479803354/jailhouse-informants/>
- O’Donnell, D. E., Shelton, J. L., Shaffer, S. A., Isom, A., Bowlin, J., & Wood, E. (2022). “My child is missing”: 911 calls in mysterious disappearances of children. *Aggression and Violent Behavior*, 67, Article 101795. <https://doi.org/10.1016/j.avb.2022.101795>
- Pinizzotto, A. J., Davis, E. F., & Miller, C. E. I. (2004). Intuitive policing: Emotional/rational decision making in law enforcement. *FBI Law Enforcement Bulletin*, 73(2), 1–6. <https://leb.fbi.gov/file-repository/archives/feb04leb.pdf/view>
- Possley, M. (2013). *Sandra Ortiz*. National Registry of Exonerations. <https://exonerationregistry.org/cases/11355>
- Rosenthal, A. M. (1964, March 27). 37 who saw murder didn’t call the police: Apathy at stabbing of queens woman shocks inspector. *The New York Times*. <https://www.nytimes.com/1964/03/27/archives/37-who-saw-murder-didnt-call-the-police-apathy-at-stabbing-of.html>
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60357-3](https://doi.org/10.1016/S0065-2601(08)60357-3)

- Ross, L. (2018). From the fundamental attribution error to the truly fundamental attribution error and beyond: My research journey. *Perspectives on Psychological Science*, 13(6), 750–769. <https://doi.org/10.1177/1745691618769855>
- Salerno, J. M. (2021). The impact of experienced and expressed emotion on legal factfinding. *Annual Review of Law and Social Science*, 17(1), 181–203. <https://doi.org/10.1146/annurev-lawsocsci-021721-072326>
- Salerno, J. M., Bean, S. R., Duran, N. D., Wulff, A. N., Kassin, S. M., & Lawrence, M. L. (2024). *Falling under the lens of suspicion: Expressing “just enough” emotion in 911 calls*. <https://osf.io/lyat52>
- Salerno, J. M., Bean, S. R., Duran, N. D., Wulff, A. N., Reeder, I., & Kassin, S. M. (2025). Failing to express emotion on 911 calls triggers suspicion through violating expectations and moral typecasting. *Journal of Personality and Social Psychology*, 128(4), 765–789. <https://doi.org/10.1037/pspa0000412>
- Salerno, J. M., Peter-Hagene, L. C., & Jay, A. C. V. (2019). Women and African Americans are less influential when they express anger during group decision making. *Group Processes & Intergroup Relations*, 22(1), 57–79. <https://doi.org/10.1177/1368430217702967>
- Santos, F. (2006, September 21). DNA evidence frees a man imprisoned for half his life. *The New York Times*. <https://www.nytimes.com/2006/09/21/nyregion/21dna.html>
- Scherr, K. C., Redlich, A. D., & Kassin, S. M. (2020). Cumulative disadvantage: A psychological framework for understanding how innocence can lead to confession, wrongful conviction, and beyond. *Perspectives on Psychological Science*, 15(2), 353–383. <https://doi.org/10.1177/1745691619896608>
- Shapiro, J. (1998). The wrong men on death row. *U.S. News & World Report*, 125(18), 22–26.
- Sharpe, M., Baldwin, D., & Walker, J. (2010). Neurotic, stress-related and somatoform disorders. In E. C. Johnstone, D. C. Owens, S. M. Lawrie, A. M. McIntosh, & M. D. Sharpe (Eds.), *Companion to psychiatric studies* (8th ed., pp. 453–491). Elsevier. <https://doi.org/10.1016/B978-0-7020-3137-3.00017-6>
- Strömwall, L. A., Hartwig, M., & Granhag, P. A. (2006). To act truthfully: Nonverbal behaviour and strategies during a police interrogation. *Psychology, Crime & Law*, 12(2), 207–219. <https://doi.org/10.1080/10683160512331331328>
- Substance Abuse and Mental Health Services Administration. (2014). *Trauma-informed care in behavioral health services: Treatment Improvement Protocol (TIP) Series No. 57* (Chapter 3, “Understanding the Impact of Trauma”, pp. 59–85). U.S. Department of Health and Human Services.
- Suchotzki, K., Verschuere, B., Van Bockstaele, B., Ben-Shakhar, G., & Crombez, G. (2017). Lying takes time: A meta-analysis on reaction time measures of deception. *Psychological Bulletin*, 143(4), 428–453. <https://doi.org/10.1037/bul0000087>
- Syed, M., & Nelson, S. C. (2015). Guidelines for establishing reliability when coding narrative data. *Emerging Adulthood*, 3(6), 375–387. <https://doi.org/10.1177/2167696815587648>
- Todorov, A., & Oh, D. (2021). The structure and perceptual basis of social judgments from faces. *Advances in Experimental Social Psychology*, 63, 189–245. <https://doi.org/10.1016/bs.aesp.2020.11.004>
- Toris, C., & DePaulo, B. M. (1984). Effects of actual deception and suspiciousness of deception on interpersonal perceptions. *Journal of Personality and Social Psychology*, 47(5), 1063–1073. <https://doi.org/10.1037/0022-3514.47.5.1063>
- U.S. Department of Veterans Affairs. (2022). *Common reactions after trauma*. National Center for PTSD. https://www.ptsd.va.gov/understand/isitpsd/common_reactions.asp
- van den Berg, R. G. (n.d.). *Effect size—A quick guide*. <https://www.spss-tutorials.com/effect-size/>
- Van Knippenberg, A., Dijksterhuis, A., & Vermeulen, D. (1999). Judgement and memory of a criminal act: The effects of stereotypes and cognitive load. *European Journal of Social Psychology*, 29(2–3), 191–201. [https://doi.org/10.1002/\(SICI\)1099-0992\(199903/05\)29:2<191::AID-EJSP923>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1099-0992(199903/05)29:2<191::AID-EJSP923>3.0.CO;2-O)
- Vonk, R. (1998). The slime effect: Suspicion and dislike of likeable behavior toward superiors. *Journal of Personality and Social Psychology*, 74(4), 849–864. <https://doi.org/10.1037/0022-3514.74.4.849>
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Wiley.
- Vrij, A., Edward, K. E., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior*, 24(4), 239–263. <https://doi.org/10.1023/A:1006610329284>
- Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest*, 11(3), 89–121. <https://doi.org/10.1177/1529100610390861>
- Vrij, A., Hartwig, M., & Granhag, P. A. (2019). Reading lies: Nonverbal communication and deception. *Annual Review of Psychology*, 70(1), 295–317. <https://doi.org/10.1146/annurev-psych-010418-103135>
- Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification evidence. *Law and Human Behavior*, 44(1), 3–36. <https://doi.org/10.1037/lhb0000359>
- Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, 22(6), 603–647. <https://doi.org/10.1023/A:1025750605807>
- Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, 25(9), 1115–1125. <https://doi.org/10.1177/01461672992512005>
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17(7), 592–598. <https://doi.org/10.1111/j.1467-9280.2006.01750.x>
- Worrall, J. L. (2013). The police sixth sense: An observation in search of a theory. *American Journal of Criminal Justice*, 38(2), 306–322. <https://doi.org/10.1007/s12103-012-9176-0>

(Appendices follow)

Appendix A

Controlled Script for 911 Call Used in Studies 2 and 3

911 Call Script

911: 911, what's your emergency?

Caller: I just heard shots outside my house and I think my husband/wife has been shot.

911: You said someone has been shot?

Caller: Yes my husband/wife, he/she needs help please.

911: Okay, what's your address?

Caller: 9718 East Cherry

911: Alright. Is your husband/wife breathing?

Caller: I think so but he's/she's gasping for air. Can you send help?

911: Yes, I'm sending medical units to your location now. Did you see where the shots came from?

Caller: No... no I didn't see anything I was inside the house till I heard them. My husband/wife was getting the mail outside...

911: Do you see anything right now?

Caller: Uhhh ... no I don't see anything.

911: Okay, stay on the phone with me. I need you to stay in a safe spot until help arrives. Where are you right now?

Caller: I'm outside my house, my husband/wife is laying right here.

911: Can you go back inside the house?

Caller: Yeah I'm near the door... I'm not sure if I should leave him/her. Are you almost here?

911: Help is almost to you, just stay on the phone with me. You said you heard a noise, did you see anything or anyone?

Caller: He/She needs help - can you send someone?

911: Did you see anyone?

Caller: He's/She's been shot - she can't breathe...

911: Ma'am/Sir, please answer me: did you see anyone?

Caller: No.... My husband/wife is gasping for air...

911: Help should be arriving right now ma'am/sir.

[End]

Note. Call script between caller and 911 operator. Participants heard either "he" or "she" pronouns representing the female or male caller discussing their husband or wife, respectively. Some participants heard a male caller reporting his wife had been shot, while others heard a female caller reporting her husband had been shot.

(Appendices continue)

Appendix B

Study Measures

Open-Ended Short Answer (Studies 1–3)

Please tell us about your first impressions of the caller. Please write a few sentences describing your impression of the person and behaviors that led you to that impression.

Behavioral Predictors

Studies 1–3

The following scales included items with 5-point response scales ranging from *not at all* to *extremely*

Perceived Urgency Scale (Study 1)

1. Caller made the situation sound urgent.
2. Caller seemed concerned about getting help for the victim.
3. Caller seemed worried about the victim.
4. Caller pled for help for the victim.
5. Caller did not seem that troubled by what was happening (reverse scored).

Perceived Urgency for the Victim Scale (Studies 2 and 3)

1. Caller seemed concerned about getting help for the victim.
2. Caller did not seem all that worried about the victim (reverse scored).
3. Caller pled for help for the victim.
4. Caller seemed panicked for their spouse.

Perceived Urgency for the Self (Studies 2 and 3)

1. Caller seemed concerned for their own safety.
2. Caller was scared for their own life.
3. Caller seemed afraid of getting hurt.
4. Caller did not seem worried for their own safety (reverse scored).

Perceived Emotionality Scale (Studies 1–3)

1. Caller was hysterical.
2. Caller kept their composure (reverse scored).
3. Caller remained calm (reverse scored).
4. Caller seemed very upset.
5. Caller seemed emotionally rattled.

Perceived Cognitive Load Scale (Studies 1–3)

1. Caller seemed to have trouble thinking clearly.
2. Caller had difficulty getting the right words out.
3. Caller did not seem to understand what the dispatcher was saying.
4. Caller seemed distracted.
5. Caller did not seem to be listening to what the dispatcher was saying.

Information Management Scale (Studies 1–3)

1. Caller was actively trying to make a good impression on the dispatcher.
2. Caller was polite to the dispatcher.
3. Caller did not seem to have much control over their behavior (reverse scored).
4. Caller seemed concerned what the dispatcher might think of him or her.
5. Caller was trying to act normal and caller was trying to be likable.

Information Management Scale (Studies 1–3)

1. Caller freely provided details about the incident (reverse scored).
2. Caller gave vague and unclear responses.
3. Caller gave superficial rather than in-depth responses.
4. Caller was skillful in managing the conversation.
5. Caller was making a lot of effort to control the information they were giving the dispatcher.
6. Caller seemed to be concerned about how much detailed information they should provide.

Suspicion Scale (Studies 1–3)

1. How suspicious is the caller acting? (*Not suspicious at all, a little suspicious, somewhat suspicious, moderately suspicious, really suspicious, extremely suspicious*)
2. How sincere do you think the caller is being? (Reverse scored; *extremely insincere, moderately insincere, somewhat insincere, somewhat sincere, moderately sincere, extremely sincere*)
3. How normal do you think the caller is acting? (Reverse scored; *extremely strange, moderately strange, somewhat*

(Appendices continue)

strange, somewhat normal, moderately normal, extremely normal)

- How likely do you think it is that the caller is hiding something? (*Extremely unlikely, moderately unlikely, somewhat unlikely, neither unlikely nor likely, somewhat likely, moderately likely, extremely likely*)
- How likely do you think it is that the caller played a role in the crime? (*Extremely unlikely, moderately unlikely, somewhat unlikely, neither unlikely nor likely, somewhat likely, moderately likely, extremely likely*)

Manipulation/Attention Checks

Gender (Studies 1–3)

- What was the gender of the caller? (*Male, female, unsure*)

Relationship (Study 1)

- At the beginning of the survey you were told the relationship between the caller and the victim. Which relationship were you told? (*Parents, brother, coworker, neighbor, the victim was someone the caller did not know, I was not given this information*)

Relationship (Studies 2 and 3)

- At the beginning of the survey you were told the relationship between the caller and the victim. Which relationship were you told? (*Parent, sibling, coworker, spouse, the victim was someone the caller did not know, I was not given this information*)

Treat as a Suspect Scale (Studies 1–3)

- For each statement below, participants chose between: *Strongly disagree, disagree, somewhat disagree, neither agree nor disagree, somewhat agree, agree, strongly agree*
Please indicate your level of agreement with the following statements.
- The police should use their resources to investigate the caller's involvement in the crime further.
 - The police should treat the caller as a suspect.

Violation of Expectations Scale (Studies 1–3; Items Presented in a Randomized Order)

- For each statement below, participants chose between: *strongly disagree, disagree, somewhat disagree, neither agree nor disagree, somewhat agree, agree, strongly agree*
- I was genuinely surprised about how the caller acted.
 - The caller behaved about how I expected (reverse scored).
 - The caller's behavior violated my expectations for what a 911 call would sound like.

- If I was in the same situation I think I would behave similarly to the caller (reverse scored).

Study 4 Measures

Suspicion

Do you think the caller was acting suspicious?
Responded on a scale from 1 (*not at all*) to 6 (*extremely*)

Treat as a Suspect

Do you think the police should use their resources to investigate the caller's involvement in the crime further?

Responded on a scale from 1 (*not at all*) to 6 (*extremely*)

Behavioral Predictors

Because this was a within-subjects design they had to be converted into single-items.

Emotionality

Caller seemed very upset.
Responded with *not at all, a little, somewhat, moderately, extremely*

Urgency for the Victim

Caller seemed concerned about getting help for the victim.
Responded with *not at all, a little, somewhat, moderately, extremely*

Urgency for the Self

Caller seemed concerned for their own safety.
Responded with *not at all, a little, somewhat, moderately, extremely*

Cognitive Load

Caller seemed to have trouble thinking clearly.
Responded with *not at all, a little, somewhat, moderately, extremely*

Impression Management

Caller was trying to act normal.
Responded with *not at all, a little, somewhat, moderately, extremely*

Information Management

Caller was trying to control the information they were giving the dispatcher.
Responded with *not at all, a little, somewhat, moderately, extremely*

Appendix C

Study 4 Call Details

Call no.	Reason for call	Length of call	Caller gender
1	Caller reported a man bleeding in the street	6:03	Male
2	Adolescent caller reported a suicide in a neighbor's home	4:26	Male
3	Caller heard gunshots with at least one victim	3:48	Female
4	Caller heard gunshots and sees people on the ground	3:32	Female
5	Security officer reported the alleged assault of an unconscious man in bathroom	1:45	Male
6	Caller reported that his friend has been shot and provided descriptions of the assailants	3:25	Female
7	Store employee called 911 on behalf of a man who claimed his friend was shot outside the store	4:59	Female
8	Caller reported that a man has been stabbed outside of his apartment and someone is banging on his door	5:03	Male
9	Father reported that his 11-year-old child deliberately shot himself in the head	6:09	Male
10	Caller heard a drive-by shooting on a nearby street	2:03	Female
11	Caller reported an alleged shooting of a bleeding man outside of his place of employment	2:42	Male
12	Caller reported an alleged shooting outside her home involving her son, though she is not sure who was shot	3:36	Female
13	Caller reported a shooting of a man who is not breathing and describes the assailant	1:54	Female
14	Caller reported a shooting outside of a store, reporting that both the victim and assailant fled the scene	1:32	Male
15	Caller reported that he heard five gunshots outside of a store	2:08	Male
16	Security officer reported hearing gunshots outside of a store	1:21	Male
17	Caller reported that her mother was dead in her ex-boyfriend's home	5:18	Female
18	Caller reported that her daughter's boyfriend stole her car and threatened to kill the caller	3:37	Female
19	Caller reported that a women who appears to have been assaulted jumped out of a vehicle that then fled	2:04	Male
20	Caller reported a drive-by shooting and described the vehicle	1:24	Male
21	Caller reported that a man was shot outside of her place of business	2:13	Female
22	Caller reported that a young man, and potentially two others, has been shot but abruptly ends the call	0:48	Male
23	Caller reported that a woman has been shot and can see people running from the scene	2:46	Female
24	Caller reported that three victims have been shot and described the injuries	4:42	Female
25	Caller reported that several people have been shot in a hotel stairwell and can hear screaming	1:58	Male
26	Witness called 911 again after a shooting to report that he now sees a vehicle fleeing the scene	1:04	Male
27	Caller reported that a woman is dead and police are on the scene but the victim needs an ambulance	2:01	Female
28	Caller reported that a man was shot at a motel	1:02	Female
29	Caller reported that a man was shot and killed, describing details of a driver and his vehicle who fled the scene	2:18	Male
30	Caller reported that people are dying and urgently requested an ambulance in addition to police officers	0:41	Female
31	Silent witness reported that she knew who has the gun that was used in a shooting the prior week	4:38	Female
32	Two individuals reported a fight that developed into a shooting at a hotel and that people are running	7:06	Male
33	Caller reported that four victims have been shot	3:32	Female
34	Caller reported hearing shots and screaming, adding that another witness saw that someone on the ground	2:31	Male
35	Employee reported a hearing shooting outside of a store, describing the victim and the car that fled the scene	5:24	Female
36	Caller reported a shooting and urges the dispatcher to send help	2:45	Male
37	Caller reported that her parents heard shots fired and found a victim bleeding in an alley	8:48	Female
38	Caller reported a man bleeding in an alley who appears to have been shot in the head	6:47	Female
39	Caller reported a home invasion and that her roommate and an intruder were shot	7:45	Male

(Appendices continue)

Appendix C (continued)

Call no.	Reason for call	Length of call	Caller gender
40	Caller heard a gunshot and watched victim fall to the ground	5:55	Male
41	Caller reported a potential kidnapping after witnessing a struggle between a driver and bleeding passenger	5:11	Male
42	Caller reported finding a deceased victim near her place of employment	3:04	Female
43	Caller reported coming home to discover his family had been shot	1:01	Male
44	Caller reported shots fired and a man laying on floor	2:53	Female
45	Caller reported that a construction worker fell from a large height and is badly injured	1:47	Male
46	Caller reported that a construction worker fell from a large height and is badly injured	0:35	Male
47	Security guard reported that a man was brought into the hospital who was found at a construction site and may have overdosed	2:05	Male
48	Caller heard gunshots in parking lot	3:03	Female
49	Caller reported that, while on the phone with her friend, she heard gunshots and her friend screaming	1:08	Female
50	Caller witnessed a shooting and reported one man down and bleeding	3:48	Male
51	Security officer reported that a customer was shot in a store	3:01	Female
52	Caller heard gunshots and people yelling for help	3:30	Female
53	Caller reported that someone has been shot in a neighborhood	4:20	Female
54	Caller reported that a man was shot in the head outside of his store and appears to be deceased	5:06	Male
55	Caller reported that a man came to her door and informed her that there is a deceased victim in the street	6:28	Female
56	Caller reported that a theft victim hit the assailant with her vehicle in a park near the caller's home, leaving the assailant severely injured	6:47	Female
57	Employee reported gunshots at a gas station and believed that a victim was shot in a car	3:16	Male
58	Caller reported that their friend was shot and the suspects fled the scene	4:51	Male
59	Employee heard gunshots and reported that a victim who was shot entered his store	4:10	Male
60	Callers reported gunshots at their apartment complex with at least one known victim	1:50	Male
61	Caller reported that a victim was shot in the mouth outside of her apartment complex	1:20	Male
62	Caller reported that his cousin was shot	1:19	Male
63	Caller reported that someone was shot but does not provide details regarding the incident	1:07	Male
64	Caller reported that a victim was shot and killed at a house party	2:17	Female
65	Caller reported that a victim was shot in her apartment complex	1:59	Female
66	Caller reported hearing gunshots in her apartment complex, noting that there is a drug dealer living in a building in the direction of the shots	2:13	Female
67	Caller reported she heard shots right outside her building and she is hiding in closet	2:29	Female
68	Caller reported hearing shots and screaming outside her building, and a second witness saw someone took the weapon from the scene	3:30	Female
69	Caller reported several shots outside her apartment with a victim on the ground	2:19	Female
70	Caller reported hearing shots fired outside his apartment	2:23	Male
71	Caller reported an alleged suicide, according to a suicide note at the scene	5:05	Female
72	Caller reported a stabbing, and the assailant fled on a bicycle	3:35	Male
73	Caller reported finding a victim of an alleged stabbing	1:53	Male
74	Caller reported that he found a deceased victim who appears to have been stabbed in a mobile home	3:14	Male
75	Caller found a young male appears to have been shot in the neck and is struggling to breathe	3:33	Male
76	Caller reported that a bleeding victim has been shot and is claiming that someone is trying to kill him	4:40	Male
77	Caller reported an armed robbery and shooting	2:08	Male
78	Caller reported that a man on the street is not moving and who looks beat up	3:02	Male
79	Caller reported that a victim was shot in the head and killed in a yard	7:23	Male
80	Caller reported two people lying in the road, both with apparent gunshot wounds to the head	4:01	Male
81	Caller stated a young boy came to her house, claiming his father shot his mother and brother.	8:14	Female
82	Caller reported that a man in his condominium has assaulted two people, noting that they already have a restraining order against the assailant	4:25	Female
83	Employee reported that two people have been shot at a club	4:02	Male

(Appendices continue)

Appendix C (continued)

Call no.	Reason for call	Length of call	Caller gender
84	Caller reported that an unconscious victim is bleeding profusely	2:22	Male
85	Caller reported following a trail of blood that led to a deceased individual located outside their residence	6:19	Female
86	Caller reported that a victim was shot in the head in her front yard	5:23	Female
87	Caller reported discovering a deceased man in a park, with indications of either a gunshot wound to the head or severe physical assault	5:19	Male
88	Caller reported a shooting at a church involving the injury of a child	5:43	Female

Note. The reason for the call reflects the information based on only the 911 call, rather than information that emerged later in the investigation.

Received October 14, 2024
Revision received September 11, 2025
Accepted September 15, 2025 ■

Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write APA Journals at Reviewers@apa.org. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, “social psychology” is not sufficient—you would need to specify “social cognition” or “attitude change” as well.
- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

APA now has an online video course that provides guidance in reviewing manuscripts. To learn more about the course and to access the video, visit <https://www.apa.org/pubs/journals/resources/review-manuscript-ce-video.aspx>.