

Deception Cues During High-Risk Situations: 911 Homicide Calls



Patrick M. Markey, Erika Feeney, Brooke Berry,
Lauren Hopkins, and Isabel Creedon

Department of Psychological and Brain Sciences, Villanova University

Psychological Science

1–8

© The Author(s) 2022

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/09567976221077216

www.psychologicalscience.org/PS



Abstract

During everyday interactions, cues tend to be weakly related to deception. However, there are theoretical reasons to suspect that such cues will be more prominent during high-risk interactions. The current study explored deception cues during one particular high-risk interaction—911 homicide calls placed by adults. In Sample 1, judges coded 911 homicide calls ($n = 82$) by Q-sorting 86 cues. Results indicated that **deceptive callers tended to display emotional cues (e.g., self-dramatizing, moody, worried, emotional, nervous), appeared overwhelmed, and related narratives that lacked structure, clarity, and focus.** Judges coded a separate sample of 911 calls ($n = 64$), and *deception scores* were computed using a template-matching approach based on the findings from Sample 1. Results indicated that deceptive 911 callers had higher deception scores than honest callers. The effect sizes yielded in this study highlight the relevance of deception cues during high-risk interactions and the usefulness of the person-centered Q-sort method.

Keywords

violent crime, judgment, interpersonal interaction, social behavior, open data, preregistered

Received 9/24/21; Revision accepted 12/9/21

It is difficult to accurately detect deception in other people (Bond & DePaulo, 2006). Even specially trained law-enforcement officers have trouble recognizing when deception occurs (DePaulo & Pfeifer, 1986; Köhnken, 1987; Vrij, 1993). Although it may be challenging to correctly judge when someone is lying, the act of deception often leaves a trail of cues. One large-scale meta-analysis examining deception in everyday interactions and university laboratories found that deceivers tended to appear nervous, tense, and uncooperative and told unconvincing narratives that lacked structure and logic (DePaulo et al., 2003). This trail of cues is consistent with some models of deception (cf. DePaulo, 1992; Ekman, 1985/1992; Zuckerman et al., 1981), but the effect sizes linking any single cue to deception during everyday interactions were comparatively low (range $r_{pb} = .00-.31$; median $r_{pb} = .05$).

There are theoretical and empirical reasons to suspect that the trail of cues related to deception will be more pronounced during high-risk interpersonal interactions, such as 911 homicide calls, than in everyday situations. First, cues related to deception may become

more evident when people are lying about serious transgressions (e.g., committing a homicide) than mundane issues (G. R. Miller & Stiff, 1993). Second, high-risk interactions are typically accompanied by strong emotions that are difficult to fake (Porter & ten Brinke, 2010). Third, 911 calls are usually placed soon after a crime, giving deceptive individuals little time to create and rehearse a false narrative (Harpster et al., 2009). Finally, a person who has committed homicide is unlikely to know how an innocent person would typically behave when calling 911.

Harpster and colleagues (Harpster, 2006; Harpster et al., 2009) carried out one of the earliest systematic attempts to examine cues related to deception during 911 homicide calls. Harpster et al. (2009) reported the same data and results as those reported by Harpster (2006) and are therefore discussed here as a single

Corresponding Author:

Patrick M. Markey, Villanova University, Department of Psychological and Brain Sciences

Email: patrick.markey@villanova.edu

study. In the Harpster study, 20 dichotomous cues were related to the deception or honesty of 911 homicide callers. Deceptive callers tended to provide extraneous information, gave conflicting facts, and were resistant to answering questions from the 911 operator (see Table 1). On the other hand, honest callers were more likely to make demanding and urgent pleas while displaying high levels of voice modulation (i.e., emotionally charged speech).

The results from the Harpster study need to be considered in light of their limitations. Only a single judge, the primary author, seems to have coded the 911 calls. Because this judge also “personally contacted” each lead detective to obtain the 911 calls, it is unclear whether the judge was unaware of the guilt or innocence of the callers (Harpster, 2006). The possibility of nonmasked coding might explain why some of the effect sizes in this study seemed unreasonably large. For example, the single dichotomous cue “extraneous information” was correlated with deception ($r = .81$) and accurately predicted deception 91% of the time. If this effect size is accurate, it would be the largest effect size found examining deceptive behaviors and among the largest effect sizes ever discovered in the social sciences (DePaulo et al., 2003; Richard et al., 2003).

Recently, Cromer et al. (2019) and M. L. Miller et al. (2021) attempted to replicate a set of the findings from the Harpster study using multiple judges unaware of the deception of the 911 callers. Unfortunately, only two of the 22 (9%) replication analyses conducted by Cromer et al. and M. L. Miller et al. produced results similar to those in the Harpster study (see Table 1). However, it is essential to note that the studies by Cromer et al. and M. L. Miller et al. suffered from low power. Assuming a moderate effect size ($r = .30$) with a nondirectional test, the power estimates for these studies, given their unequal sample sizes, ranged between .48 and .62. Therefore, it is unclear whether these null results were due to a lack of a relationship between cues and deception or to inadequate power.

Past research examining 911 homicide calls has suffered from nonmasked-coding issues, low power, and inconsistent results. None of these studies investigated whether the results found in their exploratory analyses could predict deception using an independent sample. Additionally, because most studies on deception have tended to produce low to moderate effect sizes (DePaulo et al., 2003), it seems unlikely that any single cue would be practical for accurately detecting deception. Instead, exploring how a trail of cues combine within deceptive individuals will likely provide a more holistic perspective than examining cues in isolation (Ozer, 1993).

Statement of Relevance

Every day, emergency communication centers across the United States receive numerous 911 calls related to homicides. These calls might be from victims before death, innocent witnesses, or the perpetrator of the crime. Can the cues that callers display during these 911 calls be used to determine which callers are deceptive and guilty of homicide and which are innocent? We found that deceptive callers displayed a pattern of overly emotional cues, acted overwhelmed, and told narratives that lacked clarity. We also found that this unique pattern of deceptive cues can be used to help establish the guilt or innocence of 911 homicide callers. These findings suggest that law-enforcement officers and other people can use the pattern of cues displayed during 911 homicide calls to help identify people and areas of interest.

The Current Study

The current research used masked judges, examined a large sample of 911 homicide calls, employed a person-centered Q-sort methodology, and subsequently examined whether findings from the initial analysis could predict deception in an independent sample. Using the Q-sort methodology, a group of judges Q-sorted 86 cues expressed during 911 calls. These ratings were then used to generate a template of cues that distinguished deceptive 911 callers from honest 911 callers. In a second sample, this template was compared with the cue ratings of 911 callers to assess each caller’s similarity to the prototypical cues of a deceptive 911 caller. In this manner, the probability of a caller being deceptive is viewed as a monotonically increasing function of how well this caller’s cues matched the template of a prototypically deceptive individual (Bem & Funder, 1978; Reise & Oliver, 1994).

This study’s research hypotheses, methodology, and analytic plans were preregistered on OSF before the data were coded (<https://osf.io/pvsm3/>). The data and coding instructions for the study have been made publicly available on OSF as well (<https://osf.io/v4dx7/>). The preregistration for this study discussed four planned analyses, both exploratory and confirmatory.

Exploratory analyses

In Sample 1, 86 cues were correlated to the callers’ deception or honesty to create a deception template of

Table 1. Findings From Three Studies Examining the Link Between Various Cues and Deception During 911 Homicide Calls

Cue	Harpster et al. (2009)	Cromer et al. (2019)	M. L. Miller et al. (2021)
Extraneous information	Deception	Deception	Unrelated
Conflicting facts	Deception	Deception	Unrelated
Resistance to answer	Deception	Unrelated	Unrelated
Acceptance of death with relation	Deception		Unrelated
Inappropriate politeness	Deception	Unrelated	Unrelated
Repetition	Deception		Unrelated
Acceptance of death	Deception	Unrelated	Unrelated
Possession of problem	Deception	Unrelated	
Thinking pause	Deception	Unrelated	Unrelated
Plea for caller only	Deception		
Insulting the victim	Deception		Unrelated
Minimizing just	Unrelated	Unrelated	Unrelated
Minimizing just early	Unrelated		
Verbal reaction	Honesty		
Plea for help	Honesty		
Self-correction	Honesty		
Plea for victim only	Honesty		Unrelated
Location of plea	Honesty		
Voice modulation	Honesty		Deception
Urgency of plea	Honesty		Deception
Demanding plea	Honesty		Unrelated

Note: Cues were coded as being positively related to deception (“deception”), negatively related to deception (“honesty”), or unrelated to deception (“unrelated”).

a prototypical deceptive caller. Our preregistration stated that a randomization test would be utilized to assess whether the number of significant correlations yielded when creating the deception template was greater than expected by chance (Sherman & Funder, 2009). Hypothesis 1 was that the number of significant correlations found in the initial analysis would be significantly larger than the number of significant correlations obtained in the randomization test.

We further preregistered our intention to assess the overall effect size of the deception-template cues in Sample 1, using a randomization test to evaluate the difference in the mean absolute effect size found when creating the deception template and the mean absolute effect size expected compared with chance (Sherman & Funder, 2009). Hypothesis 2 was that the mean absolute effect size yielded in the initial analysis would be significantly larger than the mean absolute effect size obtained in the randomization test.

Confirmatory analysis

In Sample 2, our preregistered intention was to examine a separate set of 911 homicide calls using the deception template and compute deception scores for each caller

by examining the similarity of each caller’s 86 cue ratings to the deception template derived in Sample 1 of the prototypical deceptive caller. Hypothesis 3 was that deceptive callers would yield greater deception scores than honest callers.

Method

Data and sources

Calls placed to 911 were deemed eligible for the study on the basis of criteria similar to those used by Cromer et al. (2019): (a) The call involved the killing of another person; (b) emergency services were notified; (c) the caller was aware of and able to communicate the general nature of the emergency; (d) at least two news sources could verify prosecution, admission of guilt, or another outcome resulting from the call; and (e) the caller did not confess to wrongdoing. Callers claiming extenuating circumstances (e.g., self-defense, accident) that led to the death of another person were also included.

Following prior 911 research (Cromer et al., 2019; M. L. Miller et al., 2021), we obtained audio calls from publicly available open-source data, such as news sources, police department releases, and various archives. An

a priori power analysis determined that a sample size of 82 (41 deceptive callers and 41 honest callers) for Sample 1's exploratory analyses and a sample size of 64 (32 deceptive callers and 32 honest callers) for Sample 2's confirmatory analysis would be necessary to achieve 80% power to detect a moderate effect ($r_{pb} = .30$).

Determination of deception

At least two external sources (usually media reports of the crime) were used to determine callers' deception or honesty. To be coded as "deceptive," the caller was required to have been found guilty in a court of law. When an indictment was not possible, such as the death of the caller, expert opinions were employed (e.g., medical examiner, police investigators, grand jury finding). Because none of the callers used in the study confessed to wrongdoing during the call, callers coded as "deceptive" were both those who lied by commission (i.e., the active use of false statements) and those who lied by omission (i.e., the passive omission of relevant information) during the 911 call. This coding is consistent with past research, which has defined *deception* as the use of statements or acts of omission that intentionally mislead (cf. Gaspar & Schweitzer, 2013; Tenbrunsel & Messick, 2004). The remaining callers not deemed guilty were coded as "honest." This coding method is consistent with past 911 research (Cromer et al., 2019; M. L. Miller et al., 2021).

Coding of 911 calls

The 911 Q sort was designed to code audio recordings of 911 homicide calls. Cues in the 911 Q sort were created to capture audio cues at a psychologically meaningful level, requiring as little subjective interpretation from the coders as possible. First, a set of relevant cues was generated from past 911 studies (e.g., Harpster et al., 2009), from research examining deception cues in everyday interactions (e.g., DePaulo et al., 2003), and from the Riverside Behavioral Q-Sort (an assessment designed to examine various cues during dyadic interactions; Funder et al., 2000). Next, a group of researchers eliminated overly redundant cues and reworded cues to be relevant in the context of 911 calls. The final set of 86 cues ranged from items directly relevant to 911 calls (e.g., "Caller makes the dispatcher confused," "Caller quickly asks for help for the victim") to cues that assessed a caller's general behavior (e.g., "Caller acts in a reckless manner," "Caller is talkative"). See <https://osf.io/v4dx7/> for a complete list of the 911 Q-sort cues.

The 911 Q sort was conducted using a Q-sort methodology (Ozer, 1993). Three judges, unaware of the deception of the caller, listed to each 911 call and then independently coded the audio of all 146 calls (82 calls for Sample 1 and 64 calls for Sample 2). See <https://osf.io/9hkaf/> for a copy of the online instructions associated with coding procedures. Judges were graduate research assistants whose training consisted of reviewing the eighty-six 911 Q-sort cues and then receiving directions on the Q-sorting procedures and instructions concerning the practical issues involving accessing the 911 audio clips and how to access the Q-sort program. Before coding the data set, all judges practiced Q-sorting five 911 calls (not included in the final analysis), and discrepancies and concerns among the coders were discussed until resolved. Judges used a modified version of the online software *HTMLQ* (Version 2.0; Killing, 2019) to sort the eighty-six 911 Q-sort cues into nine categories (1 = *extremely uncharacteristic*, 9 = *extremely characteristic*); cues were distributed as follows: 3, 6, 10, 15, 18, 15, 10, 6, and 3, respectively. By forcing judges to compare each cue with other cues, the Q-sort methodology produces a person-focused description (Ozer, 1993). Although time consuming, Q sorts are less susceptible to biases such as extremity bias, midpoint responding, acquaintance bias, and halo effects than other more traditional assessments (Ozer, 1993; Serfass & Sherman, 2013). The median cue reliability was .63, which is similar to the reliability of single items in other behavioral Q sorts (cf. Dunkel et al., 2015; Sherman et al., 2013). More importantly, as discussed later, judges' agreement was .86 when the entire 911 Q-sort deception template was used to compute 911 callers' overall deception scores.

Results

Exploratory analyses

Analyses were first conducted to create a template of 911 Q-sort cues that distinguished deceptive 911 callers from honest 911 callers. Point-biserial correlations were computed between callers' deception (coded 1 = *deceptive caller* and 0 = *honest caller*) and each of the eighty-six 911 Q-sort cues. Table 2 displays the cues that were significantly related to deception. The resulting pattern of correlations between the eighty-six 911 Q-sort cues and deception served as the template for the pattern of cues that differentiated a prototypical deceptive caller from an honest caller (results for all eighty-six 911 Q-sort cues are available at <https://osf.io/v4dx7/>). The replicability of this pattern of correlations was examined using Sherman and Wood's (2014) split-sample (SS)

Table 2. Significant Correlations in Sample 1 Between the 911 Q-Sort Items and Callers' Deception During 911 Homicide Calls ($n = 82$)

911 Q-sort item	r_{pb}
Positive correlations (indicating deception)	
10. Caller acts in a reckless manner.	.40**
83. Caller is self-dramatizing; histrionic. Theatrical; exaggerates emotions.	.38**
75. Caller's mood fluctuates frequently.	.38**
24. Caller uses many non sequiturs.	.36**
54. Caller seems to be purposely trying to sabotage or obstruct the victim getting help.	.35**
72. Caller's behavior is unpredictable.	.33**
65. Caller is self-defensive.	.33**
27. Caller is evasive.	.32**
20. Caller expresses worry that they might be blamed.	.32**
78. Caller interprets simple situations in unnecessarily complicated ways.	.32**
58. Caller expresses self-pity or feelings of victimization.	.29**
21. Caller expresses ingratiating remarks to gain favor.	.28*
22. Caller blames the victim for the situation.	.28*
29. Caller self-interrupts.	.26*
77. Caller sounds depressed.	.26*
85. Caller is overwhelmed by the situation.	.25*
48. Caller is awkward.	.25*
38. Caller sounds emotional.	.24*
33. Caller makes attempts to convince dispatcher of innocence.	.24*
25. Caller makes the dispatcher confused.	.24*
9. Caller is nervous.	.22*
49. Caller talks at rather than with the operator.	.22*
Negative correlations (indicating honesty)	
74. Caller is forthright and candid.	-.48**
76. Caller focuses on the important issues and provides relevant details.	-.46**
63. Caller is working hard and is focused on the event.	-.41**
7. Caller is cooperative and helpful.	-.39**
86. Caller expresses distrust (at anyone).	-.38**
4. Caller responds in ways that are direct, relevant, and clear.	-.38**
2. Caller's message is plausible and believable.	-.36**
11. Caller corrects any errors they made in the call.	-.36**
67. Caller behaves and acts quickly.	-.36**
3. Caller's message is consistent and provides a coherent sequence of events.	-.30**
12. Caller admits to not knowing information.	-.28*
57. Caller blames others (for anything).	-.27*
64. Caller appears to have a high degree of intellectual capacity.	-.26*
71. Caller is moralistic.	-.26*
15. Caller accepts the death of the victim.	-.24*
53. Caller acts irritated (toward anyone).	-.23*
40. Caller apologizes to the victim.	-.22*

* $p < .05$. ** $p < .01$.

methodology, which utilized 1,000 random samples to estimate the template's reliability ($\hat{\rho}_{ss} = .80$, $SE = .06$). This finding indicates that if the current study were replicated, the pattern of correlations in the replication study would be expected to have a correlation of .80 with the current study's deception template.

The previous set of 86 analyses found 39 significant effects linking 911 Q-sort cues to deception. However, given the number of nonindependent analyses conducted, some of these effects might be significant simply because of chance. Sherman and Funder's (2009) randomization method was therefore employed to

Table 3. Results From Randomization Tests of 911 Q-Sort Correlates of Deception in Sample 1

Statistic	Observed value	Value expected by chance	SE	<i>p</i>	95th percentile
Number of significant ($p < .05$) results	39	4.247	3.802	< .0001	12
Mean absolute r_{pb}	.193	.089	.016	< .0001	.119

determine the probability of obtaining 39 significant results under a random model of no association between the deception and the 911 Q-sort cues. This was done by (a) randomly redistributing deceptive and honest codes and (b) computing the number of significant ($p < .05$) correlations between deceptiveness and the 911 Q-sort cues in this new sample. This procedure was repeated 10,000 times. The resulting values were used to form a sampling distribution indicating the number of expected significant effects under the null hypothesis of no relation between deception and the 911 Q-sort cues. Consistent with Hypothesis 2, this analysis found that the probability of the current study finding 39 significant ($p < .05$) correlations linking deception to the 911 Q-sort cues by simple chance was $p < .0001$ (see Table 3).

Randomization tests were again used to estimate the significance of the mean effect size found linking deception to the entire set of 911 Q-sort cues (Sherman & Funder, 2009). The absolute correlation value between caller deceptiveness and the entire set of eighty-six 911 Q-sort cues was first computed (average $r_{pb} = .19$). As in the previous analysis, 10,000 random samples were used to estimate the probability of obtaining this observed effect size under a random model of no association between deception and the 911 Q-sort cues. Consistent with Hypothesis 3, this analysis indicated that the chance of the current study finding this observed effect size linking deception to the 911 Q-sort cues was $p < .0001$ (see Table 3). Taken together, the results from Hypotheses 2 and 3 suggest that the 911 Q-sort deception template displayed in Table 2 is real (i.e., beyond chance) and can likely be used to predict deception during 911 calls.

Confirmatory analysis

Deception scores were computed in Sample 2 using a template-matching approach (Bem & Funder, 1978; Reise & Oliver, 1994). This approach entails correlating (i.e., matching) the observed pattern of a caller's 911 Q-sort cues with the 911 Q-sort deception template derived from Sample 1. Specifically, each judge's sort of each 911 call was correlated with the matched 86 effect sizes of the 911 Q-sort deception template (see Table 2). To ease interpretation, we standardized the resulting correlations before conducting any analyses.

Therefore, high deception scores indicate that the caller's 911 Q-sort pattern was similar to the pattern of deceptive individuals, and low deception scores indicate that the caller's 911 Q-sort pattern was similar to the pattern of honest callers. The three judges' deception scores were then aggregated for each 911 call; judges' reliability for deception scores was .86. Consistent with Hypothesis 3, results for Sample 2 showed that deceptive 911 callers were significantly more likely to have higher deception scores than honest callers, $t(62) = 5.29$, $p < .001$, $d = 1.32$, 95% confidence interval (CI) = [0.78, 1.86] (see Fig. 1).

Discussion

Every day, emergency communication centers across the United States receive numerous 911 calls related to homicides. These calls might be from victims before death, innocent witnesses, or the perpetrator of the crime. Such calls could be a crucial investigative tool by providing a unique insight into the guilt or innocence of a suspect. Consistent with this notion, results of the current study showed that cues expressed during 911 homicide calls were related to the deception of the caller.

Although we made no specific predictions concerning individual cues in the initial exploratory analysis, many of the study's findings were consistent with models of deception that suggest that the disconnect between a deceiver's narrative and reality causes the leakage of emotionally related cues (DePaulo, 1992; Ekman, 1985/1992; Zuckerman et al., 1981). For example, as seen in Table 2, deceptive callers were self-dramatizing, moody, reckless, worried, depressed, emotional, and nervous. Likewise, these models predict that the cognitive load and uncertainty inherent in crafting false narratives cause deceivers to tell less-than-compelling narratives. Again, this conjecture is consistent with the current study's findings that deceptive individuals appeared overwhelmed and related narratives that lacked structure, clarity, and focus.

Such findings have implications for understanding how high-risk situations might alter the importance of cues during deception and may serve as a helpful tool for detecting criminal deception. Unfortunately, past research has found that even specially trained law-enforcement officers tend to be poor at detecting

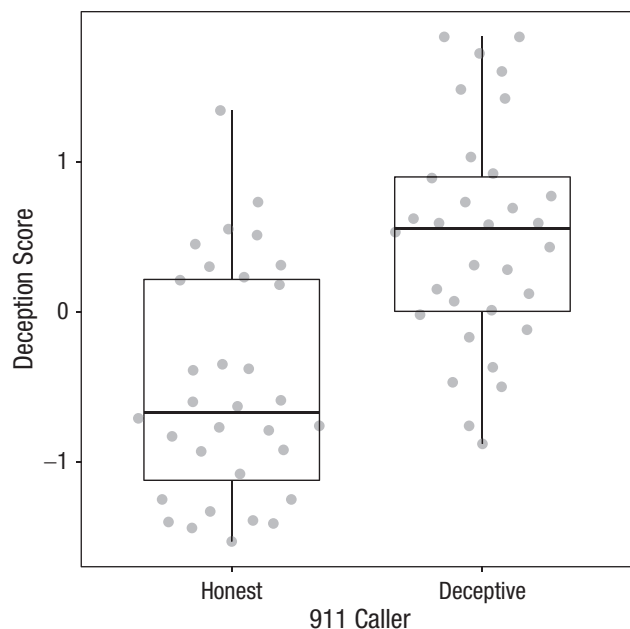


Fig. 1. Deception scores for honest and deceptive 911 callers in Sample 2. Each box represents the interquartile range (IQR), the whiskers represent the range of values within 1.5 times the IQR, and the horizontal line represents the median. Dots represent individual data.

deception (DePaulo & Pfeifer, 1986; Köhnken, 1987; Vrij, 1993). One possible explanation for this finding is that judges sometimes use invalid cues when determining whether a person is honest. For example, speech pauses, “ums,” and “huhs” are often believed to be cues related to deception, but research (and the current study; see <https://osf.io/v4dx7/>) shows that such speech patterns are not valid cues to deception (Davis et al., 2005; DePaulo et al., 2003). However, this does not imply that judges are inaccurate at detecting cues, only that they have trouble understanding how the pattern of cues they observe is related to deception.

As in research examining clinical and personality judgments, the most accurate predictions regarding deception likely result from using judge-rated cues as input into a statistical model that accounts for multiple cues (e.g., template matching; see Wiggins, 1973, for a review). Consistent with this notion, our results showed that when 911 homicide calls were scored by applying the deception template to judge-rated cues, deceptive callers received substantially higher deception scores than honest callers. Furthermore, given past attempts at predicting deception during 911 calls using nonmasked judges (Cromer et al., 2019; M. L. Miller et al., 2021) or during everyday interactions (DePaulo et al., 2003), the effect size yielded from this analysis was larger than expected ($d = 1.32$) and highlights the usefulness of the person-centered Q-sort method.

Although the current study predicted a 911 caller's deception using the 911 Q sort, caution is warranted for

anyone basing a caller's guilt or innocence solely on cues expressed during a 911 call. Such information might help law-enforcement officers identify people and areas of interest, but it would be a mistake to use it to make a definitive conclusion concerning criminal activity. Additionally, results from the current study should be considered within the context of its limitations. Because archival 911 calls were used, there was no transparent chain of custody for these calls. Therefore, it is unknown whether some calls were edited (e.g., names might have been deleted for privacy reasons). The generalizability of these results might be limited because some states do not release 911 calls to the public and all calls were from English speakers. Future researchers might consider obtaining calls directly from law enforcement or examining calls from a more diverse geographic area. Finally, the current study operationalized deception on the basis of whether a 911 homicide caller who did not confess to wrongdoing during the call was later convicted of the homicide. Although this operationalization is consistent with past 911 research (cf. Cromer et al., 2019; M. L. Miller et al., 2021), given the imperfect nature of the criminal justice system, caution is warranted when using criminal conviction as a proxy for deception.

The Q-sort methodology employed in the current study was time consuming, with judges taking approximately 25 min to listen and code each 911 call. It might, therefore, be of practical importance to reduce the number of 911 Q-sort cues to make this coding process more efficient. For example, an auxiliary analysis found that the deception scores computed using all eighty-six 911 Q-sort cues in Sample 2 were highly similar, $r(62) = .98$, $p < .001$, to the deception scores computed when only the 39 significant cues were used. Future researchers might also consider using automatic coding methodologies (e.g., voice-prosody analysis, natural-language processing) to compute deception scores more quickly. Finally, although most of the 911 Q-sort cues apply to a wide variety of crimes, the current study focused only on 911 homicide calls. It is hoped that others employ the person-centered Q-sort method presented here to examine the possibility of detecting deception in other high-risk criminal situations, such as missing-persons cases, aggravated assault, or arson.

Transparency

Action Editor: Kate Ratliff

Editor: Patricia J. Bauer

Author Contributions

P. M. Markey developed the study concept. L. Hopkins and I. Creedon contributed to the study design. Data collection was supervised by E. Feeney, B. Berry, and I. Creedon. P. M. Markey analyzed the data and drafted the manuscript. E. Feeney and B. Berry provided critical revisions to the manuscript. All the authors approved the final manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Open Practices

All data and coding instructions have been made publicly available via OSF and can be accessed at <https://osf.io/v4dx7/>. The design and analysis plans for the study were preregistered at <https://osf.io/pvsm3/>. This article has received the badges for Open Data and Preregistration. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>.



References

- Bem, D. J., & Funder, D. C. (1978). Predicting more of the people more of the time: Assessing the personality of situations. *Psychological Review*, *85*, 485–501.
- Bond, C. F. J., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*, 214–234.
- 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.
- Davis, M., Markus, K. A., Walters, S. B., Vorus, N., & Connors, B. (2005). Behavioral cues to deception vs. topic incriminating potential in criminal confessions. *Law and Human Behavior*, *29*, 683–704.
- DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, *111*, 203–243.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*, 74–118.
- DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology*, *16*, 249–267.
- Dunkel, C. S., Summerville, L. A., Mathes, E. W., & Kesslering, S. N. (2015). Using the California Q-sort measure of life history strategy to predict sexual behavioral outcomes. *Archives of Sexual Behavior*, *44*, 1705–1711.
- Ekman, P. (1992). *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. W. W. Norton. (Original work published 1985)
- Funder, D. C., Furr, R. M., & Colvin, C. R. (2000). The Riverside Behavioral Q-Sort: A tool for the description of social behavior. *Journal of Personality*, *68*(3), 451–489.
- Gaspar, J. P., & Schweitzer, M. E. (2013). The emotion deception model: A review of deception in negotiation and the role of emotion in deception. *Negotiation and Conflict Management Research*, *6*, 160–179.
- 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.
- Harpster, T. K. (2006). *The nature of 911 homicide calls: Using 911 homicide calls to identify indicators of innocence and guilt* (Document No. ucin1163534092) [Master's thesis, University of Cincinnati]. http://rave.ohiolink.edu/etdc/view?acc_num=ucin1163534092
- Killing, M. (2019). *HTMLQ* (Version 2.0) [Computer software]. GitHub. <https://github.com/aproxima/htmlq>
- Köhnken, G. (1987). Training police officers to detect deceptive eyewitness statements: Does it work? *Social Behaviour*, *2*(1), 1–17.
- Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication*. SAGE.
- 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.
- Ozer, D. J. (1993). The Q-sort method and the study of personality development. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time: Personality and development* (pp. 147–168). American Psychological Association.
- Porter, S., & ten Brinke, L. (2010). The truth about lies: What works in detecting high-stakes deception? *Legal and Criminological Psychology*, *15*, 57–75.
- Reise, S. P., & Oliver, C. J. (1994). Development of a California Q-set indicator of primary psychopathy. *Journal of Personality Assessment*, *62*, 130–144.
- Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*, 331–363.
- Serfass, D. G., & Sherman, R. A. (2013). A methodological note on ordered Q-Sort ratings. *Journal of Research in Personality*, *47*, 853–858.
- Sherman, R. A., Figueredo, A. J., & Funder, D. C. (2013). The behavioral correlates of overall and distinctive life history strategy. *Journal of Personality and Social Psychology*, *105*, 873–888.
- Sherman, R. A., & Funder, D. C. (2009). Evaluating correlations in studies of personality and behavior: Beyond the number of significant findings to be expected by chance. *Journal of Research in Personality*, *43*, 1053–1063.
- Sherman, R. A., & Wood, D. (2014). Estimating the expected replicability of a pattern of correlations and other measures of association. *Multivariate Behavioral Research*, *49*, 17–40.
- Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research*, *17*, 223–236.
- Vrij, A. (1993). Credibility judgments of detectives: The impact of nonverbal behavior, social skills, and physical characteristics on impression formation. *The Journal of Social Psychology*, *133*, 601–610.
- Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Addison-Wesley.
- Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 14, pp. 1–59). Academic Press.