Table of Contents

Measures Collected, but not Reported in this Manuscript

The following measures were collected in the studies reported in this manuscript. However, they were measured to speak to a different research question and are analyzed and reported in another manuscript that is more broadly focused on behaviors in a 911 call that might trigger suspicion (Bean et al., in preparation). It is beyond the scope of this article, which focuses specifically on emotionality.

Studies 1-3, S1:
The following scales included five-point response scales ranging from *Not at all* to *Extremely*

To what degree was the caller feeling emotion? (Studies 1-2, S1)

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 2-3, S1)

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 2-3, S1)

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, S1)

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, S1)

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, S1)

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, S1)

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.

Study 3: Measures unique to police sample

These measures were taken to characterize the police sample and to provide qualitative data for a separate more applied project on how police become suspicious of witnesses. They will be reported in a future article (Lawrence et al., in preparation) and to inspire follow-up studies.

Have you currently OR in the past worked in any of the following professions (please check all that apply)?

__ Law enforcement
__ An intelligence officer of any kind

__ Lawyer, Judge, Mediator

__ Forensic Interviewer of any kind not listed above

__ 911 dispatcher

__ None of the above

If select [law enforcement/intelligence/mediator/interviewer]:

Can you please briefly describe what your job in *[law enforcement/intelligence/mediator/interviewer]* is/was? If you've had several jobs in this category, please list each.

In your job (past or present) did you ever have to conduct any kind of interviews with people who were potentially involved with a crime in any way (potential suspects, witnesses, victims, etc.)?

- ○ yes
- ○ no

If Yes:
Did you primarily interview potential suspects, witnesses, or victims?

- ○ Potential suspects
- ○ Potential witnesses
- ○ Potential victims
- ○ A roughly equal mix of at least two of these categories

We are interested in how experts in forensic interviewing come to identify guilty parties during their interviews. In other words, the skills that develop in assessing people and searching out suspicious behavior.

In your job(s), did you ever have an experience when you were interviewing someone who, at the time, was not suspected of a crime, but who you became suspicious of during the interaction, and came to believe was guilty of a crime (or at least was involved in a crime)?

- ○ yes
- ○ no

If yes: Can you please explain what led you to start suspecting the person? Was it something they said? Something they did? Something about their behavior, demeanor, or appearance? Something else you remember? As much detail you can provide would be much appreciated.

When you are talking with someone potentially involved in a crime, are there things that you keep any eye out for that lead you to become suspicious of someone, in general?

Study 4:

These measures are parallel versions of the behavioral scales reported above for Studies 1-3, S1. Because this was a within-subjects design they had to be converted into single-items, but they are also reported in the other manuscript that is more broadly focused on behaviors in a 911 call that might trigger suspicion (Bean et al, in preparation).

The following scales included items with six-point scale from *Not at all* to *Extremely*

- **Urgency for Victim:**
  Caller seemed concerned about getting help for the victim.
- **Urgency for Self*:***
  Caller seemed concerned for their own safety.
- **Cognitive Load:**
  Caller seemed to have trouble thinking clearly.
- **Impression Management:**
  Caller was trying to act normal.
- **Information Management:**
  Caller was trying to control the information they were giving the dispatcher.
- **Fluency Question:**
  Caller's voice went up and down in volume.

Open-Ended Short Answers, Studies 1-3, S1:

1. 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.

2. We are interested in how police might come to suspect that an individual might have played a role in a crime. You might have found the caller to be suspicious or to be credible. What about the caller's behavior made you think they were either suspicious or believable? Please write a few sentences.

3. Is there anything else that you thought was noteworthy about the 911 call that we did not ask about? Or any other comments?

Relationship Check (Studies 1-2, S1)
This measure was collected as an additional manipulation check for studies with manipulated relationship of the caller. After participants indicating the relationship between the caller and the victim (constituting the primary relationship check, see Appendix A), participants answered additional questions about this relationship.

1. To what degree did you believe this relationship? (*Not at all believable, Slightly believable, Believable, Very believable, Extremely, believable*);

> If participants answered the two lowest items on the scale (*Not at all believable, Slightly believable*), they answered an additional open-ended question: Why did you not find this relationship believable?

Supplemental Analyses for Study 1

Treat Caller as a Suspect Scale ANOVA Results Study 1

The following are analyses that are parallel to the ANOVA analyses reported in the main text with the suspicion scale as the dependent variable, but with treat caller as a suspect scale as the dependent variable. The two outcomes were strongly related and produced largely the same pattern of results, reported here.

As hypothesized, there was a significant main effect of emotion expression (Table S1). Bonferroni post-hoc tests confirmed our hypothesis that participants would support treating low-emotion callers as a suspect more (M = 4.87, SD = 1.37) than moderate-emotion callers (M = 3.66, SD = 1.69), M_{diff} = 1.20, $p < .001$, d = .78, and high-emotion callers (M = 3.85, SD = 1.67), M_{diff} = 1.02, $p < .001$, d = .68. However, they thought callers who expressed moderate emotion and high emotion should be treated similarly, M_{diff} = -.18, p = .44, d = .11. There was also a significant main effect of relationship, $F(2, 930) = 22.80$, $p < .0001$, d = .28, such that participants supported treating callers as a suspect more when they were reporting that their parents had been shot (M = 4.40, SD = 1.65) relative to when they reported that a stranger had been shot (M = 3.93, SD = 1.65). The main effect of caller gender was not significant.

The main effects were qualified, however, by one of the hypothesized two-way interactions. There was a significant caller emotion × gender interaction (Table S1, Figure S1). Specifically, the simple effect of emotion was significant for both male and female callers (Table S1), but the patterns were different. As hypothesized, participants thought that women who expressed low emotion (and therefore might have violated gender stereotypes) should be treated as a suspect significantly more than women who expressed moderate M_{diff} = 1.21, d = .80, $p <$ .001, 95% CI [.89, 1.54], and high levels of emotion, M_{diff} = 1.56, d = 1.01, $p < .001$, 95% CI

[1.23, 1.89]. Participants also thought that women who expressed moderate emotion should be treated as a suspect significantly more than women who expressed high levels of emotion, M_{diff} = .35, d = .16, p = .04, 95% *CI* [.02, .67]. Similarly, participants thought that men who expressed low emotion should be treated as a suspect significantly more than men who expressed moderate emotion M_{diff} = 1.24, d = .73, p < .001, 95% *CI* [.91, 1.56]. However, participants thought men who expressed a high level of emotion (and therefore might have violated gender stereotypes) should be treated as a suspect significantly *more* than those who expressed moderate emotion, M_{diff} = -1.00, d = .38, p < .001, 95% *CI* [-1.39, -.61].

In summary, and in line with gender stereotypes, participants thought women should be treated as a suspect more when they violated gender stereotypes: for each increase in emotion level, participants supported treating female callers as a suspect less. In contrast, for men, expressing too little *and* too much emotion justified treating them as a suspect more than a moderate level. No other interactions were significant.

Table S1

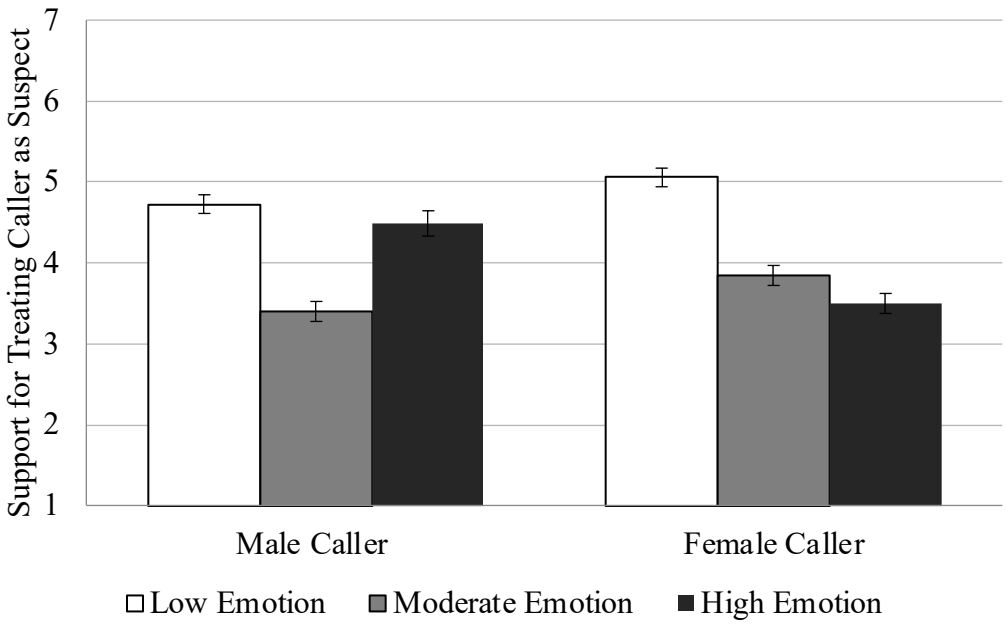
ANOVA Table for the Effects of Caller Gender, Caller Emotion Expression, and Victim Relationship to the Caller on the Treat as a Suspect Scale, Study 1

	Independent Variable	*df*	*F*	*p*	Effect Size
Treat as Suspect	Caller Gender	1	.92	.34	$d = .03$
	Relationship to the Caller	1	22.80	< .001	$d = .28$
	Caller Emotion	2	57.17	< .001	$\eta_p^2 = .11$
	Caller Gender × Relationship	1	.77	.38	$\eta_p^2 = .001$
	Caller Gender × Caller Emotion	1	16.68	< .001	$\eta_p^2 = .04$
	Male caller: Emotion simple effect	1	29.28	<.001	$\eta_p^2 = .06$
	Female caller: Emotion simple effect	1	47.95	<.001	$\eta_p^2 = .09$
	Caller Emotion × Relationship	1	1.16	.32	$\eta_p^2 = .002$
	Gender × Emotion × Relationship	1	.07	.94	$\eta_p^2 < .001$
	Error	930			

Note. Following Lakens' (2013) recommendations, we calculated *Cohen's d* values for all two-group comparisons and *partial eta²* (η_p^2) values for all effects that included more than two groups.

Supplemental Figure S1

Mean (SE) Agreement that the Caller should be Treated as a Suspect as a Function of Caller Emotion Expression and Caller Gender, Study 1.



Supplemental Analyses for Study 2

Treat Caller as a Suspect Scale ANOVA Results Study 2

We replicated the main effect of emotion expression on whether the caller should be treated as a suspect (Table S2). Bonferroni post-hoc tests confirmed our hypothesis that participants thought callers who expressed low emotion (M = 5.27, SD = 1.33) should be treated as a suspect significantly more than callers who expressed moderate emotion (M = 4.64, SD = 1.63), M_{diff} = .63, p < .001, d = .42, and high emotion (M = 4.61, SD = 1.56), M_{diff} = .66, p < .001, d = .46. However, moderate- and high-emotion callers did not differ, M_{diff} = .03, p = 1.00, d = .02. Participants thought male callers (M = 4.97, SD = 1.48) should be treated as a suspect more than female callers (M = 4.72, SD = 1.59). The interaction was not significant (Figure S2).

Table S2

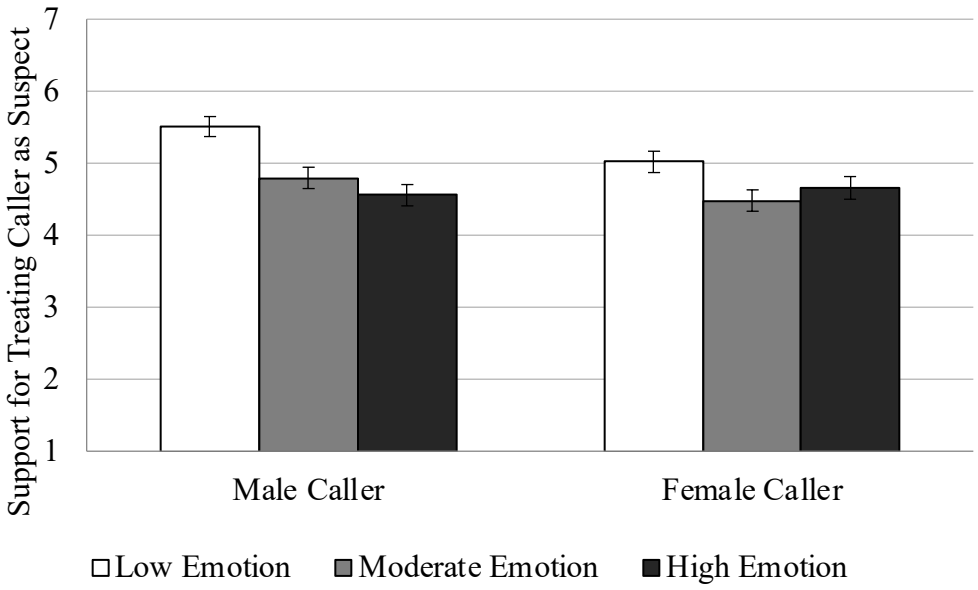
ANOVA Table for the Effects of Caller Gender, Caller Emotion Expression, and Victim Relationship to the Caller on Dependent Measures, Study 2

	Independent Variable	*df*	*F*	*p*	Effect Size
Treat as Suspect Scale	Caller Gender	1	3.87	.050	d = .16
	Caller Emotion	2	12.65	< .001	η_p^2 = .04
	Caller Gender × Caller Emotion	1	2.06	.13	η_p^2 = .01
	Error	618			
Violate Expectations Scale	Caller Gender	1	1.10	.30	d = .08
	Caller Emotion	2	74.07	< .001	η_p^2 = .19
	Caller Gender × Caller Emotion	1	1.12	.33	η_p^2 = .004
	Error	618			
Moral Agency	Caller Gender	1	14.38	<.001	d = .30
	Caller Emotion	2	21.96	< .001	η_p^2 = .07
	Caller Gender × Caller Emotion	1	.11	.90	η_p^2 < .001
	Error	618			
Moral Patiency	Caller Gender	1	.09	.77	d = .03
	Caller Emotion	2	40.96	< .001	η_p^2 = .12
	Caller Gender × Caller Emotion	1	.41	.67	η_p^2 = .001
	Error	618			

Note. Following Lakens' (2013) recommendations, we calculated *Cohen's d* values for all two-group comparisons and *partial eta*^{2} (η_p^2) values for all effects that included more than two groups.

Supplemental Figure S2

Mean (SE) Agreement that the Caller Should be Treated as a Suspect as a Function of Caller Emotion Expression and Caller Gender, Study 2.



Analyses (ANOVAs) of Mediators Study 2

We replicated the main effect of emotion expression on each of our mediators: violation of expectations, perceived moral agency, and perceived moral patiency (Table S2).

Violation of Expectation. We replicated the main effect of emotion expression on suspicion on the violation of expectations scale (Table S2). Bonferroni post-hoc tests confirmed our hypothesis that callers who expressed low emotion violated participants' expectations (M = 4.00, SD = 1.29) significantly more than callers who expressed moderate emotion (M = 3.02, SD = 1.21), M_{diff} = .98, $p < .001$, d = .78, who in turn violated expectations significantly more than

callers who expressed high emotion (M = 2.57, SD = 1.16), M_{diff} = .45 p < .001, d = .38. The gender main effect and the interaction were not significant.

Moral Agency. We also replicated the main effect of emotion expression on perceived moral agency (Table S2). Bonferroni post-hoc tests confirmed our hypothesis that participants thought callers who expressed low emotion (M = 3.72, SD = 1.04) had significantly more moral agency than callers who expressed moderate emotion (M = 3.17, SD = 1.20), M_{diff} = .55, p < .0001, d = .49, and high emotion (M = 3.01, SD = 1.16), M_{diff} = .70, p < .0001, d = .64. However, moderate and high emotion did not differ, M_{diff} = .16, p = .47, d = .13. Further, participants perceived male callers as having more moral agency (M = 3.47, SD = 1.18) than female callers (M = 3.13, SD = 1.15). The interaction was not significant.

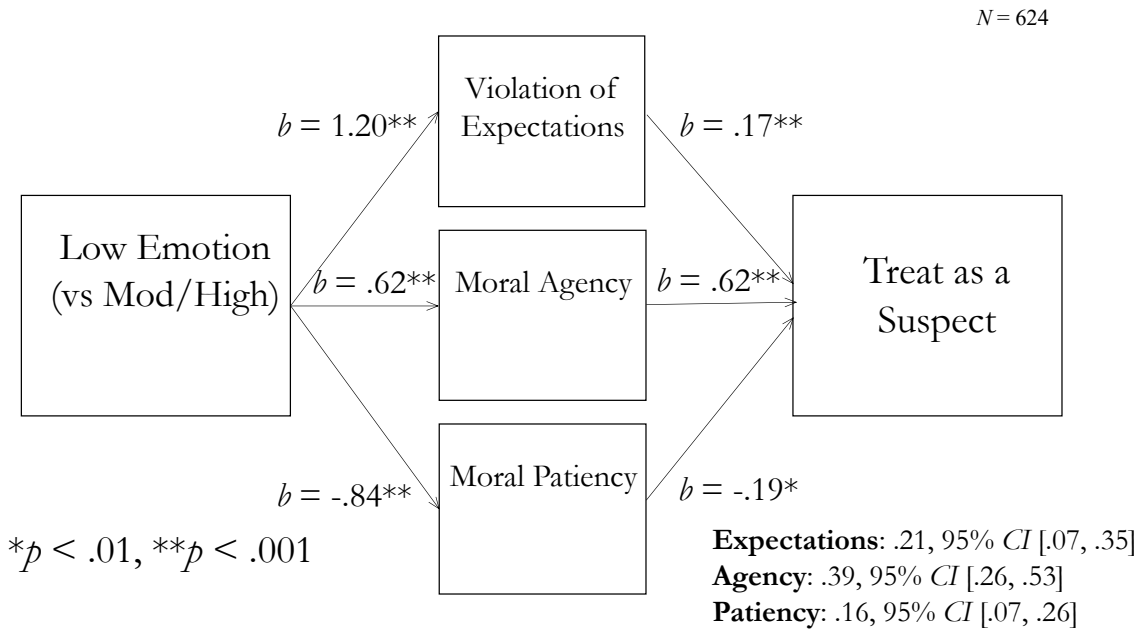
Moral Patiency. We also replicated the main effect of emotion expression on perceived moral patiency (Table S2). Bonferroni post-hoc tests confirmed our hypothesis that participants thought callers who expressed low emotion (M = 3.20, SD = 1.22) had significantly less moral patiency than callers who expressed moderate emotion (M = 3.92, SD = 1.10), M_{diff} = .-.72, p < .001, d = .62, and high emotion (M = 4.17, SD = 1.07), M_{diff} = .-.97, p < .001, d = .64. However, moderate and high emotion did not differ, M_{diff} = -.25, p = .07, d = .23. The gender main effect and the interaction were not significant.

Treat as a Suspect Mediation Results Study 2

Note that caller gender did not moderate the indirect effect through violation of expectations, Index of moderated mediation (*IMM*) = .02, SE = .04, 95% *CI* [-.05, .10], moral agency *IMM* = -.05, *SE* = .12, 95% *CI* [-.29, .17], or moral patiency, *IMM* = -.02, SE = .04, 95% *CI* [-.10, .06]. We dropped gender as a moderator and re-ran the model (Figure S3).

Supplemental Figure S3

Indirect Effects of Emotion Expression on Support for Treating the Caller as a Suspect through Violation of Expectations, Moral Agency, and Moral Patiency, Study 2



We also replicated the three indirect effects of callers expressing low (versus moderate or high) emotion on support for treating the caller as a suspect through violation of expectations, $M_{indirecteffect}$ = .21, *SE* = .07, 95% *CI* [.07, .35], moral agency, $M_{indirecteffect}$ = .39, SE = .07, 95% CI [.26, .53], and moral patiency, $M_{indirecteffect}$ = .17, *SE* = .05, 95% *CI* [.07, .26].

Study 2 Supplemental Analyses Summary

We replicated our findings regarding suspicion reported in the main text with support for treating the caller as a suspect. More specifically, we found that expressing low levels (versus moderate or high levels) of emotion on a 911 call led to increased support for treating the caller as a suspect. Further, expressing low (versus moderate or high) levels of emotion also increased violation of expectations, increased perceptions of moral agency and decreased perceptions of

moral patiency. Similar to suspicion, all three mediators were significantly associated with support for treating the 911 caller as a suspect. In summary, low (versus moderate or high) emotion increased support for investigating the caller as a suspect indirectly through violation of expectations and morally typecasting the caller as more of a moral agent and less of a moral patient.

In addition, we found that male callers were perceived to have significantly more moral agency compared to female callers—participants also supported treating male callers as a suspect more than female callers.

Supplemental Analyses for Study 3

Treat Caller as a Suspect Scale ANOVA Results Study 3

We failed to replicate the main effect of emotion expression on whether the caller should be treated as a suspect (Table S3, Figure S4). Although officers thought male callers (M = 5.59, SD = 1.11) should be treated as a suspect more than female callers (M = 5.28, SD = 1.14) overall, caller gender and emotion did not interact. This lack of emotion effect replication in this sample appears to be due to police officers thinking that all of the callers should be investigated as a suspect regardless of emotion level—all of the Suspicion Scale means were high (Figure S4). In hindsight, we realized that police routinely investigate the spouse in a crime like this.

Table S3

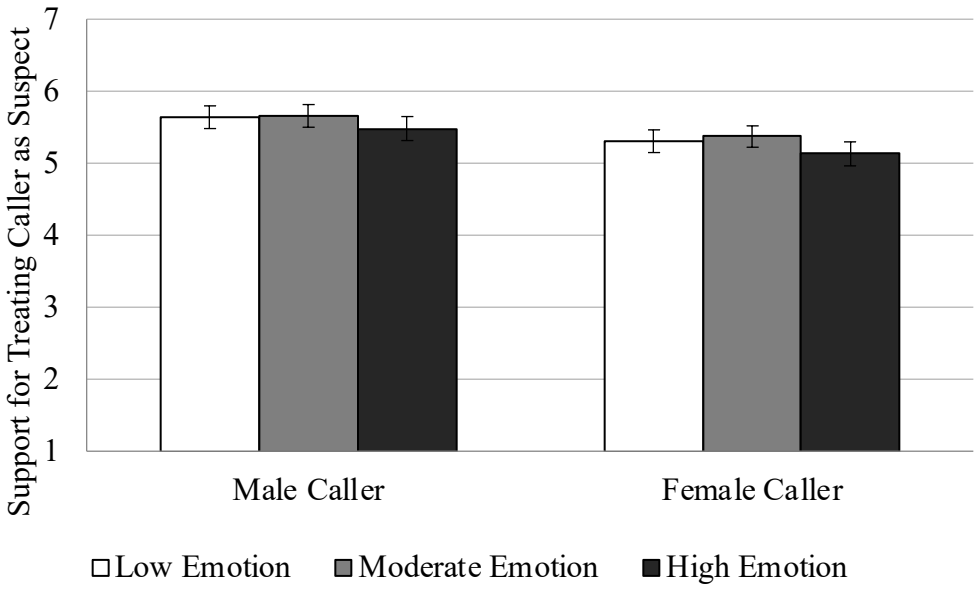
ANOVA Table for the Effects of Caller Gender, Caller Emotion Expression, and Victim Relationship to the Caller on the Treat as a Suspect Scale, and Mediators (Violation of Expectations, Moral Agency, and Moral Patiency), Study 3 (Police Officers)

	Independent Variable	*df*	*F*	*p*	Effect Size
Treat as Suspect Scale	Caller Gender	1	5.98	.01	d = .28
	Caller Emotion	2	.89	.41	η_p^2 = .01
	Caller Gender × Caller Emotion	1	.02	.98	η_p^2 < .001
	Error	292			
Violate Expectations Scale	Caller Gender	1	1.38	.24	d = .11
	Caller Emotion	2	31.70	< .001	η_p^2 = .18
	Caller Gender × Caller Emotion	1	.58	.56	η_p^2 = .004
	Error	292			
Moral Agency	Caller Gender	1	2.52	.11	d = .17
	Caller Emotion	2	8.51	< .001	η_p^2 = .06
	Caller Gender × Caller Emotion	1	.05	.95	η_p^2 < .001
	Error	292			
Moral Patiency	Caller Gender	1	3.68	.06	d = .20
	Caller Emotion	2	30.33	< .001	η_p^2 = .17
	Caller Gender × Caller Emotion	1	1.76	.17	η_p^2 = .01
	Error	292			

Note. Following Lakens' (2013) recommendations, we calculated *Cohen's d* values for all two-group comparisons and *partial eta*^{2} (η_p^2) values for all effects that included more than two groups.

Supplemental Figure S4

Mean (SE) Agreement that the Caller should be Treated as a Suspect as a Function of Caller Emotion Expression and Caller Gender, Study 3.



Analyses (ANOVAs) of Mediators Study 3

We replicated the main effect of emotion expression on each of our mediators: violation of expectations, perceived moral agency, and perceived moral patiency Table S3.

Violate Expectations. There was again a significant main effect of emotion expression (Table S3). Bonferroni post-hoc tests revealed that callers who expressed low emotion violated officers' expectations significantly more (M = 4.42, SD = 1.02) than callers who expressed moderate emotion (M = 3.81, SD = 1.13), M_{diff} = .62, p < .001, d = .57, and high emotion (M = 3.21, SD = 1.03), M_{diff} = 1.21, p < .001, d = 1.18. Callers who expressed moderate emotion also

violated officers' expectations more than callers who expressed high emotion, M_{diff}= .60, p < .0001, d = .55. Officers also reported that male callers (M = 3.91, SD = 1.20) violated their expectations for 911 callers significantly more than female callers (M = 3.77, SD = 1.13) overall. Caller gender and emotion did not interact (Table S3).

Moral Agency. There was again a significant main effect of emotion expression (Table S3). Bonferroni post-hoc tests revealed that officers perceived callers who expressed low emotion (M = 3.96, SD = 0.74) to have similar moral agency as callers who expressed moderate emotion (M = 3.81, SD = 0.82), M_{diff}= .15, p = .56, d = .19. They perceived callers who expressed high emotion, however, to have significantly less moral agency (M = 3.48, SD = 0.86) than callers who expressed moderate emotion, M_{diff}= -.33, p = .02, d = .39, and low emotion, M_{diff}= -.47, p < .001, d = .59.

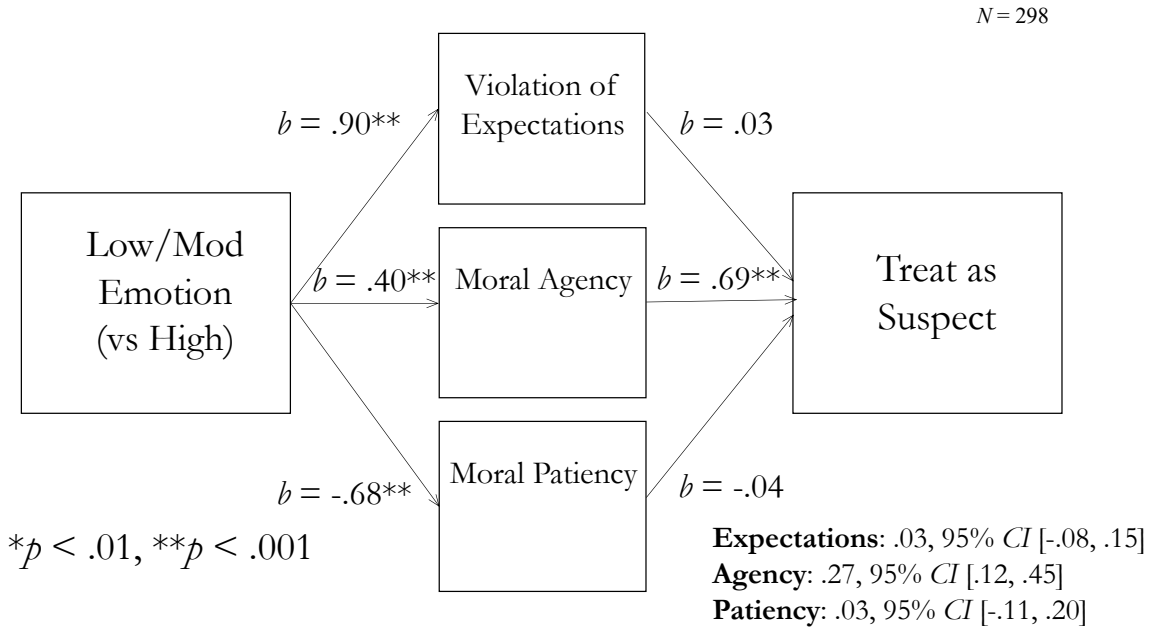
Moral Patiency. There was again a significant main effect of emotion expression (Table S3). Bonferroni post-hoc tests revealed that officers perceived callers who expressed low emotion (M = 2.94, SD = 0.70) to have significantly less moral patiency than callers who expressed moderate emotion (M = 3.28, SD = .81), M_{diff}= -.34, p = .004, d = .45, and compared to callers who expressed high emotion (M = 3.80, SD = 0.82), M_{diff}= -.86, p < .001, d = 1.13. Further, officers perceived callers who expressed high emotion to also have significantly more moral patiency than callers who expressed moderate emotion, M_{diff}= .51, p < .001, d = .63.

Treat as a Suspect Mediation Results Study 3

Note that caller gender did not moderate the indirect effect through violation of expectations, Index of moderated mediation (*IMM*) = -.01, *SE* = .02, 95% *CI* [-.06, .04], moral agency *IMM* = .03, *SE* = .15, 95% *CI* [-.26, .33], or moral patiency, *IMM* = -.01, *SE* = .05, 95% *CI* [-.12, .07]. We dropped gender as a moderator and re-ran the model (Figure S5).

Supplemental Figure S5

Indirect Effects of Emotion Expression on Support for Treating the Caller as a Suspect through Violation of Expectations, Moral Agency, and Moral Patiency, Study 3 (Police Officers)



We also replicated one of the three indirect effects of callers expressing low/moderate (versus high) emotion on support for treating the caller as a suspect through moral agency, $M_{indirecteffect}$ = .27, *SE* = .08, 95% *CI* [.12, .45]—but did not replicate the indirect effects through violation of expectations, $M_{indirecteffect}$ = .03, *SE* = .06, 95% *CI* [-.08, .15], or moral patiency, $M_{indirecteffect}$ = .03, *SE* = .08, 95% *CI* [-.11, .20].

Study 3 Supplemental Analyses Summary

We partially replicated our suspicion findings with the treat as suspect scale dependent variable—although also found some interesting differences among this police officer sample. Consistent with the prior studies and suspicion findings in this study, we again found that hearing low emotion from a 911 call significantly violated police officers' expectations and increased their perceptions of the caller as a moral agent relative to hearing high levels of emotion from

911 callers. However, in contrast to lay samples, police found the moderate-emotion calls to be similar to low-emotion callers. That is, police had a higher threshold for emotion level: police needed to see high emotion before their expectations were violated less and they perceived the caller to have lower moral agency. Similar to laypeople, police also perceived low-emotion callers to have more moral patiency than moderate-emotion and high-emotion callers.

We did not replicate the emotion expression effect on support for treating the caller as a suspect. This was because police officers thought that all of the callers should be investigated as a suspect regardless of emotion level—support for treating the caller as a suspect was high across all versions of the call. In hindsight, we realized that police routinely investigate the spouse in a crime like this.

Finally, we found that police had a gender bias to think that male callers should be investigated as a suspect more than female callers—despite the details of the crime and call script being identical.

Supplemental Analyses for Study 4

Treat Caller as a Suspect Scale Results Study 4

We re-ran the model reported in the main text of Study 4 testing the relationship between perceived emotionality and suspicion, but this time with support for treating the caller as a suspect as the outcome variable. Thus, we again conducted a multilevel model that included perceived emotionality (grand mean centered; level 1 variable), caller gender (male as the reference group; level 2 variable), and the interaction as predictors of support treating the caller as a suspect. We entered perceived emotionality was as a polynomial variable with three potential relationships with treating the caller as a suspect: linear, quadratic, and cubic. The random effects were participant at level 1 and call at level 2. The model was fitted with restricted maximum likelihood using the lme4 package (Bates et al., 2015) in R and the assessment of fixed effects was conducted using *t*-tests in the lmerTest package (Kuznetsova et al., 2017). We ran one model with gender (Table S4) as the moderator and another with relationship to the victim as the moderator (Table S5).

Gender moderation. See Table S4 for the full model results. We found no caller gender effect and no relationship between perceived emotion and support for treating the caller as a suspect.

Table S4

Multilevel Model of Perceived Emotionality of the Caller and Caller Gender Predicting Support for Treating Caller as a Suspect, Study 4 (Real 911 Calls)

	β	*SE*	*df*	*t*	*p*
Intercept					
Level 1 Variable	2.15	0.11	121.82	19.10	< .001
Perceived Emotionality (Linear)	-4.45	2.70	1120.46	-1.65	.099
Perceived Emotionality (Quadratic)	-0.62	1.96	1781.52	-0.32	.750
Perceived Emotionality (Cubic)	-0.71	1.77	1681.40	-0.40	.690
Level 2 Variable					
Caller Gender	-0.27	0.14	79.37	-1.86	.067
Interaction Term					
Perceived Emotionality* Caller Gender (Linear)	-2.12-3	3.57	1209.85	-0.60	.551
Perceived Emotionality* Caller Gender (Quadratic)	3.58	2.69	1742.79	1.33	.184
Perceived Emotionality* Caller Gender (Cubic)	-2.52	2.43	1667.93	-1.04	.300
Random Effects		Variance		*SD*	
Participant (Level 1)		0.68		0.82	
Call (Level 2)		0.36		0.60	
Residual		1.23		1.11	

Relationship to the victim moderation. To investigate the potential moderating role of the caller's relationship to the victim, we ran a similar multilevel model, replacing caller gender with relationship to the victim as the potential moderator. See Table S5 for the full model results. This model included perceived emotionality (grand mean centered; level 1 variable), relationship to the victim (stranger as the reference group; level 2 variable), and the interaction as predictors of support for treating the caller as a suspect. We entered perceived emotionality as a polynomial variable with three potential relationships with treating the caller as a suspect: linear, quadratic, and cubic.

We again found a significant effect of caller relationship, $\beta = 1.11$, $p < .001$. Participants supported treating callers as a suspect more when they reported that someone close to them had been hurt than those who called about a stranger. This model revealed a significant and negative linear relationship between emotionality and treating the caller as a suspect, such that lower emotionality significantly predicted increased support for treating the caller as a suspect, conceptually replicating Studies 1-3 (and S1) and the suspicion findings in the main text. Consistent with suspicion, this effect was qualified by the hypothesized interaction between perceived emotionality and caller relationship on treating the caller as a suspect, $\beta = -12.13$, $p = .01$. There were no significant quadratic or cubic relationships between emotion and no significant quadratic or cubic interactions between caller gender and emotion.

Simple slopes analysis (that dropped the quadratic and cubic effects and associated interaction terms), revealed that when the caller knew the victim (i.e., when more emotion would be expected), increased perceived emotionality predicted a larger decrease in treating the caller as a suspect, $\beta = -0.26$, $p < .01$, than when the caller did not know the victim, $\beta = -0.07$, $p = .02$—although both were significant.

Table S5.

Multilevel Model of Perceived Emotionality of the Caller and Caller Relationship with the Victim Predicting Support for Treating Caller as a Suspect, Study 4 (Real 911 Calls)

	β	*SE*	*df*	*t*	*p*
Intercept	1.85	0.08	176.56	22.32	< .001
Level 1 Variable					
Perceived Emotionality (Linear)	-4.34	1.97	841.24	-2.20	.028
Perceived Emotionality (Quadratic)	1.42	1.47	1781.41	0.96	0.33
Perceived Emotionality (Cubic)	-1.55	1.33	1715.19	-1.17	.244
Level 2 Variable					
Caller Relationship	1.11	0.17	94.00	6.54	< .001
Interaction Term					
Perceived Emotionality* Caller Relationship (Linear)	-12.13	4.84	748.70	-2.50	.012
Perceived Emotionality* Caller Relationship (Quadratic)	-2.30	3.82	1747.18	-0.60	.547
Perceived Emotionality* Caller Relationship (Cubic)	-1.25	3.50	1742.20	-0.36	.721
Random Effects		Variance		*SD*	
Participant (Level 1)		0.68		0.83	
Call (Level 2)		0.23		0.48	
Residual		1.23		1.11	

Study 4 Supplemental Analyses Summary

We conceptually replicated the relationship between caller emotion and suspicion found in Studies 1-3 (and S1) on support for treating the caller as a suspect in the model that included caller relationship. More specifically, perceived emotionality of the caller was significantly and negatively related to support for treating the caller as a suspect—a relationship that was significantly stronger when the caller knew the victim. We also replicated the finding that people overall supported treating callers who were reporting that someone they knew had been hurt as a suspect more than callers who were reporting that a stranger had been hurt.

In the caller gender model, we did not replicate the negative relationship between perceived emotionality and treating the caller as a suspect. This could be due to the fact that this model controlled for caller gender, which was somewhat correlated with perceived emotionality ($r = .24$, $p < .001$). The discrepant perceived emotionality relationship between the gender and relationship model might also be due to the fact that although the gender model was run on all participants, the relationship model did not include participants' observations about the four calls in which the relationship was unknown. Thus, the perceived emotionality result from the relationship model was run on a subset of observations (call $n = 84$, observation $n = 1{,}909$), while the gender model was run on the complete set of observations (call $n = 88$, observation $n = 2{,}000$).

Supplemental Study S1: The Moderating Impact of Caller Gender and Relationship to the Victim

We conducted this additional study to further investigate whether the impact of emotion expression on 911 calls on suspicion differs for men and women. In Study 1, men (but not women) were more suspicious when expressing high emotion compared to moderate emotion about a stranger or parent victim, but in Studies 2-3 this gender interaction vanished when they called about a spouse. Perhaps people expect greater emotion when a spouse is harmed—even for men. We tested whether emotion expression depended on gender in the context of a stranger relationship (like Study 1), but not in the context of a spouse (like Studies 2-3). More specifically, we predicted a three-way interaction, such that we would find a significant caller gender × caller emotion two-way interaction in the stranger condition, but not in the spouse condition.

Supplemental Study S1 Method

Participants and Design

We recruited 976 Amazon Mechanical Turk workers. A total of 156 (16%) participants were excluded for failing the gender manipulation check (n = 64, 7%), failing the relationship manipulation check (n = 38, 4%), and/or indicating in an open-ended question that they thought the call was fake (n = 66, 7%). The remaining 818 (63% women; M_{age} = 42, SD_{age} = 13; 74% White, 10% Black, 4% Hispanic/Latinx, 10% Asian, and 1.5% Other) were randomly assigned to condition in a 2 (caller gender: male, female) × 3 (emotion expression: low, moderate, high) × 2 (relationship: spouse, stranger) between-subjects design. Our sample size was determined by the power analysis we preregistered for Study 2 (i.e., N = 976).

Materials

Stimuli. We used the same 911 call recordings from Studies 2 and 3. However, for this study we had to modify the calls to mask information about the caller's relationship to the victim to enable a successful relationship manipulation. To do so, we used an audio clip from Adobe's Audition that sounded like a "beep" to cover over specific segments (e.g., when the caller referred to the victim as their wife or husband). To provide a cover story for these beep tones, the study instructions informed participants that all of the calls had been edited for timing and that personal information was removed to protect confidentiality.

Participants were again randomly assigned to hear a version of the 911 call delivered (a) by one of the male or female callers, (b) with either low, moderate, or high emotion, and (c) we added a relationship manipulation.

***Relationship Manipulation*.** Participants read a short description of their task in which they were told that they would listen to a short clip from a real 911 call in which the caller reports that their spouse or a stranger had been shot. Participants were then asked to confirm their understanding that they listen to a 911 audio and answer questions about someone who had been shot. They were able to proceed once they correctly answered the attention checks by identifying the relationship of the victim to the caller and these other filler attention checks.

Measures

Participants completed the same measures as Studies 2 and 3. Once again, our scales had acceptable-to-good reliability: Suspicion Scale ($\alpha = .93$), Treat as Suspect Scale (Spearman Brown Coefficient = .91), Violation of Expectations Scale ($\alpha = .90$), Moral Agency Scale ($\alpha = .80$), and Moral Patiency Scale ($\alpha = .69$).

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations and measures. All data, analysis code and research materials are accessible via the OSF link in the author's note in the main text of the article. This study's design and analysis were not preregistered.

Supplemental Study S1 Results

Descriptive information and correlations among measures are reported in Table S6.

Table S6

Descriptive Statistics and Pearson Correlation Matrix for Mediators and Dependent Variables, Study S1.

	M (*SD*)	1	2	3	4
1. Suspicion	2.99 (1.61)				
2. Treat as suspect	4.27 (1.75)	.72*			
3. Violation of expectations	2.99 (1.40)	.74*	.51*		
4. Moral patient	3.80 (0.97)	-.53*	-.37*	-.54*	
5. Moral agent	2.81 (1.08)	.80*	.65*	.67*	-.64*

*$p < .001$

For each dependent measure, we conducted a 2 (Caller gender: male, female) × 3 (Caller emotion: low, moderate, high) × 2 (Relationship to caller: stranger, spouse) between-subjects ANOVA (see Table 8 for ANOVA results).

Emotion manipulation check. Our emotion manipulation was again successful in that there was a significant main effect of the emotion condition on perceived emotion (Table S7). Bonferroni post-hoc tests revealed that participants again rated low-emotion callers as having expressed significantly less emotion ($M = 2.56$, $SD = 0.98$) than moderate-emotion callers ($M = 3.81$, $SD = 1.05$), $M_{diff} = -1.26$, $p < .001$, $d = 1.23$; and in turn, they rated moderate-emotion

callers as significantly less emotional than high-emotion callers (M = 4.31, SD = .82), M_{diff} = -.50, $p < .001$, d = .53. They also perceived women, overall, to be significantly more emotional (M = 3.67, SD = 1.18) than men (M = 3.45, SD = 1.23)—something we did not find in prior studies. Interestingly, they also perceived strangers to be significantly more emotional (M = 3.86, SD = 1.08) than spouses (M = 3.29, SD = 1.25)—despite the fact that these two groups heard the exact same recordings (i.e., the emotion levels were identical). However, similar to Study 1 (but unlike Studies 2-3), the strength of the manipulation depended on the caller's gender (Table S7). The emotion manipulation was again significant for both male and female callers—but the effect was stronger for male callers. There were no other significant interactions with the emotion manipulation.

Suspicion Scale. Support for our hypothesis that the caller gender × caller emotion two-way interaction would re-emerge in the stranger condition, but not in the spouse condition was not supported—the three-way interaction was not significant (Table S7).

We did, however, replicate other aspects of our findings. There was again a significant main effect of emotion expression (Table S7). Bonferroni post-hoc tests confirmed our focal hypothesis that participants again found low emotion (M = 3.56, SD = 1.61) to be significantly more suspicious than moderate emotion (M = 2.77, SD = 1.53), M_{diff} = .79, $p < .001$, d = .51, and high emotion (M = 2.67, SD = 1.54, M_{diff} = .89, $p < .001$, d = .56. Moderate and high emotion again did not differ, M_{diff} = .09, p = 1.00, d = .06. Participants found male callers (M = 3.17, SD = 1.64) more suspicious than female callers (M = 2.85, SD = 1.57) overall, and caller gender and emotion did not interact (Table S7, Figure S6a).

Participants also again found callers reporting that a spouse had been shot (M = 3.41, SD = 1.71) as more suspicious than those reporting about a stranger (M = 2.58, SD = 1.38). This

main effect was again qualified by a caller emotion × relationship interaction (Table S7, Figure S6b).

Table S7

ANOVA Table for the Effects of Caller Gender, Caller Emotion Expression, and Victim Relationship to the Caller on Dependent Variables, Study S1

	Independent Variable	*df*	*F*	*p*	Effect Size
Perceived Emotion (MC)	Caller Gender	1	7.45	.01	$d = .18$
	Relationship to the Caller	1	70.00	< .001	$d = .49$
	Caller Emotion	2	268.13	< .001	$\eta_p^2 = .40$
	Caller Gender × Relationship	1	.34	.56	$\eta_p^2 < .001$
	Caller Gender × Caller Emotion	1	5.78	.003	$\eta_p^2 = .01$
	Male caller: Emotion simple effect	1	148.96	< .001	$\eta_p^2 = .27$
	Female caller: Emotion simple effect	1	119.77	< .001	$\eta_p^2 = .23$
	Caller Emotion × Relationship	1	1.48	.23	$\eta_p^2 = .004$
	Gender × Emotion × Relationship	1	1.26	.28	$\eta_p^2 = .003$
	Error	806			
Suspicion Scale	Caller Gender	1	5.59	.02	$d = .20$
	Relationship to the Caller	1	58.29	< .001	$d = .54$
	Caller Emotion	2	30.80	< .001	$\eta_p^2 = .07$
	Caller Gender × Relationship	1	<.001	.98	$\eta_p^2 < .001$
	Caller Gender × Caller Emotion	1	2.20	.11	$\eta_p^2 = .01$
	Caller Emotion × Relationship	1	4.79	.01	$\eta_p^2 < .01$
	Stranger: Emotion simple effect	1	7.52	< .001	$\eta_p^2 = .02$
	Spouse: Emotion simple effect	1	27.89	< .001	$\eta_p^2 = .06$
	Gender × Emotion × Relationship	1	1.57	.21	$\eta_p^2 = .004$
	Error	806			
Treat as Suspect	Caller Gender	1	13.45	< .001	$d = .29$
	Relationship to the Caller	1	65.19	< .001	$d = .62$
	Caller Emotion	2	8.96	< .001	$\eta_p^2 = .02$
	Caller Gender × Relationship	1	.55	.46	$\eta_p^2 = .001$
	Caller Gender × Caller Emotion	1	.98	.38	$\eta_p^2 = .002$
	Caller Emotion × Relationship	1	.31	.74	$\eta_p^2 < .001$
	Gender × Emotion × Relationship	1	.18	.84	$\eta_p^2 < .001$
	Error	806			

Note. Following Lakens' (2013) recommendations, we calculated *Cohen's d* values for all two-group comparisons and *partial eta*^{2} (η_p^2) values for all effects that included more than two groups.

As hypothesized, when calling about a spouse being shot (i.e., when high emotion would be expected), 911 callers who might have violated expectations by expressing low emotion were significantly more suspicious than 911 callers expressing moderate emotion, $M_{diff} = 1.06$, $p < .001$, 95% *CI* [.71, 1.42], or high emotion, $M_{diff} = 1.31$, $p < .001$, 95% *CI* [.94, 1.68]. Callers expressing high (versus moderate) emotion were less suspicious when reporting their spouse had been shot, but this effect was not significant, $M_{diff} = .24$, $p = .18$, 95% *CI* [-.11, .60]. When the 911 caller reported that a *stranger* had been shot (i.e., when relatively lower emotion might be expected), those who expressed low emotion were again significantly more suspicious than those who expressed moderate emotion, $M_{diff} = .72$, $p < .001$, 95% *CI* [.34, 1.09] and high emotion, $M_{diff} = .50$, $p = .01$, 95% *CI* [.14, .86]. In contrast to the spouse condition, however, callers expressing high (versus moderate) emotion were *more* suspicions when a stranger was shot—but again this effect was not significant, $M_{diff} = -.22$, $p = .23$, 95% *CI* [-.58, .14]. Thus, the interaction is driven by high (versus moderate) emotion being less suspicious when reporting a spouse has been shot, but high (versus moderate) being more suspicious when reporting a stranger was shot. Although these simple effects were in opposing directions, neither of them was significant.

In summary, participants were more suspicious of 911 callers who violated expectations by expressing low emotion relative to moderate or high emotion. We found very tentative evidence that expressing high levels of emotion, relative to moderate levels of emotion, in a 911 call had opposing effects on suspicion in line with the hypothesis that participants would be more

suspicion when expectations were violated. More specifically, they were less suspicious of high (versus moderate) emotion when it is more expected (i.e., when a spouse had been shot) and more suspicious of high (versus moderate) emotion when it is less expected (i.e., when a stranger was shot). However, we hesitate to draw conclusions from this interaction because neither of these simple effects were significant and do not discuss them further.

Figure S6

Mean (SE) Suspicion as a Function of Caller Emotion Expression and (a) Caller Gender and (b) Relationship to the Victim, Study S1

(a) (b)

Mediators

We replicated the main effect of emotion expression on each of our mediators: violation of expectations, perceived moral agency, and perceived moral patiency. However, there were some unpredicted higher-order interactions between emotion expression and the manipulations that were not found in previous studies and were not relevant to our hypotheses. For interested readers, these more detailed analyses can be found later in the Supplemental Analysis section for this study (pp. 38-44).

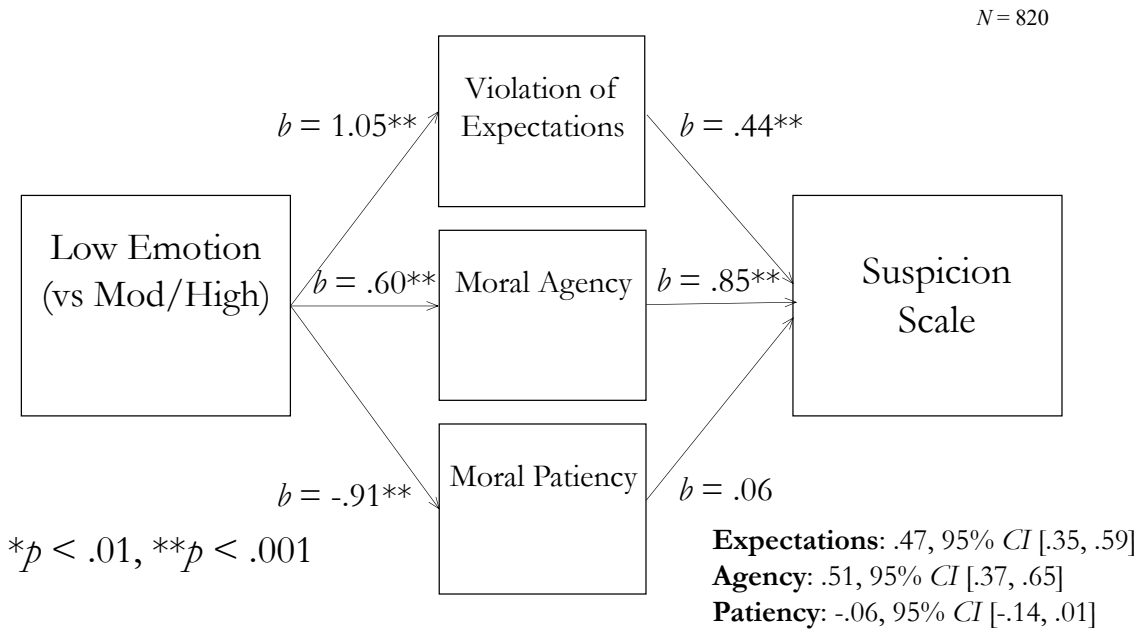
Mediation

Next, we again tested whether we replicated the significant indirect effect of low emotion expression on suspicion through (a) violating the listeners' expectations for 911 callers reporting a violent crime, and (b) triggering moral typecasting of the caller as less capable of being a moral patient and more capable of being a moral agent, relative to those who express at least moderate emotion that we found in Studies 2-3. More specifically, we tested a simultaneous mediator model with low (versus moderate or high) emotion predicting 3 simultaneous mediators (expectation violation, moral agency, and moral patiency), and suspicion as our dependent measure.

The model replicated two significant indirect effects of callers expressing low (versus moderate/high) emotion on suspicion through violation of expectations and moral agency (Figure 8). The path coefficients again confirmed our hypothesis: expressing low (versus moderate or high) emotion increases violation of expectations and perceptions of moral agency—both of which were associated with greater suspicion. However, although moral patiency had a strong bivariate correlation with suspicion (r = -.53, p < .001), when we controlled for the other two mediators, we did not replicate the unique indirect effect through moral patiency found in Studies 2-3.

Figure S7

Indirect Effects of Emotion Expression on Suspicion through Violation of Expectations, Moral Agency, and Moral Patiency, Study 4



This lack of an indirect effect through moral patiency is likely due to a strong negative correlation between moral agency and moral patiency (which is consistent with the Dyadic Theory of Morality), resulting in moral patiency being related to suspicion, but not *uniquely* related in this model that statistically controls for the relationship between moral agency and suspicion. In summary, low (versus moderate or high) emotion increased suspicion indirectly through violation of expectations and morally typecasting the caller as more of a moral agent—but in this study, not through moral patiency. Indeed, when we tested a *single* mediator model with emotion expression predicting suspicion through violation of expectations and moral patiency (but not moral agency), the indirect effect through moral patiency was significant, $M_{indirecteffect}$ = .31, *SE* = .05, 95% *CI* [.21, .42]. This suggests that although we did replicate the

indirect effect through patiency, we did not replicate the *unique* effect beyond what moral agency explained that we found in Studies 2-3.

In contrast to Studies 2 and 3, however, we found that the indirect effects were moderated by caller gender and caller relationship. In both models, the two significant indirect effects of low (versus moderate/high) emotion on suspicion through violation of expectations and moral agency were significant in all conditions, but they were stronger in the male (versus female) caller condition and in the spouse (versus stranger) condition. Given these are not directly related to our hypotheses (and gender did not moderate the mediation models in Studies 2-3 and therefore are not reliable) we do not discuss them further here, but these moderated mediation model results are detailed later in the Supplemental Analysis section for this study for interested readers (pp. 45-47).

Treat as suspect

We replicated the main effect of caller emotion expression on suspicion and our mediation patterns when we analyzed the treat as a suspect scale as the outcome. More specifically, low (versus moderate or high) emotion expression increased support for treating the caller as a suspect indirectly through increased violation of expectations and morally typecasting the caller as more of a moral agent and less of a moral patient. There was, however, significant moderation of the strength of these effects that were unexpected because we did not find them in Studies 2-3; these moderated mediation model results are detailed later in the Supplemental Analysis section for this study (pp. 47-49).

Study S1 Discussion

We again replicated our findings that low (versus moderate/high) emotion is more suspicious because it violates expectations and leads to moral typecasting of the caller as having

greater moral agency and reduced moral patiency—although the latter finding regarding moral patiency was somewhat weaker than in Studies 2-3. We also replicated our Study 1 finding that the effect of emotionality on suspicion depended on the relationship with the victim—but the follow-up tests were not significant.

The effect of emotion expression did not, however, depend on gender—regardless of whether the caller reported on a stranger/parents (like Study 1 where gender moderated) or a spouse (like Studies 2 and 3 where gender did not moderate). Thus, our proposed explanation that perhaps the gender moderation exists outside of the spousal context was not supported.

Supplemental Analyses for Study S1

Treat Caller as a Suspect Scale ANOVA Results Study S1

We replicated the main effect of emotion expression on whether the caller should be treated as a suspect (Table S4). Bonferroni post-hoc tests confirmed our hypothesis that participants thought callers who expressed low emotion (M = 4.60, SD = 1.65) should be treated as a suspect significantly more than callers who expressed moderate emotion (M = 4.23, SD = 1.69), M_{diff} = .37, p = .03, d = .22, and high emotion (M = 4.00, SD = 1.87), M_{diff} = .60, p < .001, d = .34. However, moderate and high emotion did not differ, M_{diff} = .23, p = .32, d = .13. Participants thought male callers (M = 4.56, SD = 1.70) should be treated as a suspect more than female callers (M = 4.05, SD = 1.77) overall. Participants also thought callers reporting that a spouse had been shot (M = 4.78, SD = 1.62) should be treated as a suspect more than callers reporting that a stranger had been shot (M = 3.77, SD = 1.74). There were no significant interactions (Table S8, Figure S7).

Table S8

ANOVA Table for the Effects of Caller Gender, Caller Emotion Expression, and Victim Relationship to the Caller on Mediator Variables, Study 4.

	Independent Variable	*df*	*F*	*p*	Effect Size
Treat as Suspect Scale	Caller Gender	1	13.45	< .001	$d = .29$
	Relationship to the Caller	1	65.19	< .001	$d = .62$
	Caller Emotion	2	8.96	< .001	$\eta_p^2 = .02$
	Caller Gender × Relationship	1	.55	.46	$\eta_p^2 = .001$
	Caller Gender × Caller Emotion	1	.98	.38	$\eta_p^2 = .002$
	Caller Emotion × Relationship	1	.31	.74	$\eta_p^2 < .001$
	Gender × Emotion × Relationship	1	.18	.84	$\eta_p^2 < .001$
	Error	806			
Violation of Expectations	Caller Gender	1	5.06	.02	$d = .18$
	Relationship to the Caller	1	30.42	< .001	$d = .37$
	Caller Emotion	2	70.27	< .001	$\eta_p^2 = .15$
	Caller Gender × Relationship	1	.72	.40	$\eta_p^2 = .001$
	Caller Gender × Caller Emotion	1	3.12	.045	$\eta_p^2 = .01$
	Male caller: Emotion simple effect	1	42.29	< .001	$\eta_p^2 = .09$
	Female caller: Emotion simple effect	1	28.85	< .001	$\eta_p^2 = .07$
	Caller Emotion × Relationship	1	14.57	< .001	$\eta_p^2 = .03$
	Stranger: Emotion simple effect	1	12.24	< .001	$\eta_p^2 = .03$
	Spouse: Emotion simple effect	1	71.87	< .001	$\eta_p^2 = .15$
	Gender × Emotion × Relationship	1	3.40	.03	$\eta_p^2 = .01$
	Error	806			
Moral Agency	Caller Gender	1	10.68	.001	$d = .26$
	Relationship to the Caller	1	84.90	< .001	$d = .65$
	Caller Emotion	2	39.03	< .001	$\eta_p^2 = .09$
	Caller Gender × Relationship	1	.02	.88	$\eta_p^2 < .001$
	Caller Gender × Caller Emotion	1	5.07	.01	$\eta_p^2 = .01$
	Male caller: Emotion simple effect	1	28.35	<.001	$\eta_p^2 = .07$
	Female caller: Emotion simple effect	1	13.24	<.001	$\eta_p^2 = .03$
	Caller Emotion × Relationship	1	3.09	.046	$\eta_p^2 = .01$
	Stranger: Emotion simple effect	1	11.05	<.001	$\eta_p^2 = .03$
	Spouse: Emotion simple effect	1	30.79	<.001	$\eta_p^2 = .07$
	Gender × Emotion × Relationship	1	2.98	.051	$\eta_p^2 = .01$
	Error	806			

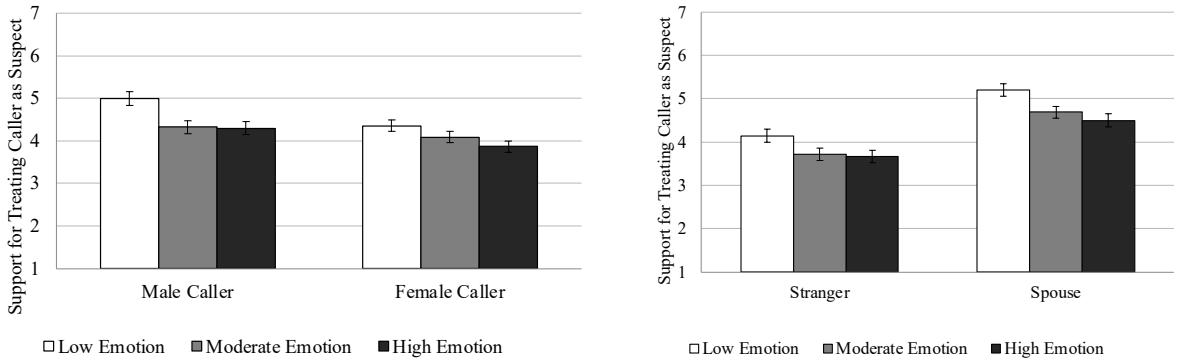
Moral Patiency	Caller Gender	1	7.95	.005	$d = .21$
	Relationship to the Caller	1	19.07	< .001	$d = .28$
	Caller Emotion	2	111.29	< .001	$\eta_p^2 = .22$
	Caller Gender × Relationship	1	3.91	.048	$\eta_p^2 = .01$
	Caller Gender × Caller Emotion	1	1.86	.16	$\eta_p^2 = .005$
	Male caller: Emotion simple effect	1	57.48	< .001	$\eta_p^2 = .12$
	Female caller: Emotion simple effect	1	55.52	< .001	$\eta_p^2 = .12$
	Caller Emotion × Relationship	1	5.47	.01	$\eta_p^2 = .01$
	Stranger: Emotion simple effect	1	34.03	<.001	$\eta_p^2 = .08$
	Spouse: Emotion simple effect	1	82.48	<.001	$\eta_p^2 = .17$
	Gender × Emotion × Relationship	1	1.71	.18	$\eta_p^2 = .004$
	Error	806			

Note. Following Lakens' (2013) recommendations, we calculated *Cohen's d* values for all two-group comparisons and *partial eta*^{2} (η_p^2) values for all effects that included more than two groups.

Supplemental Figure S7

Mean (SE) Agreement that the Caller should be Treated as a Suspect as a Function of Caller Emotion Expression, Caller Gender, and Caller Relationship, Study S1.

(a) (b)



Analyses (ANOVAs) of Mediators Study S1

We replicated the main effect of emotion expression on each of our mediators: violation of expectations, perceived moral agency, and perceived moral patiency (Table S8).

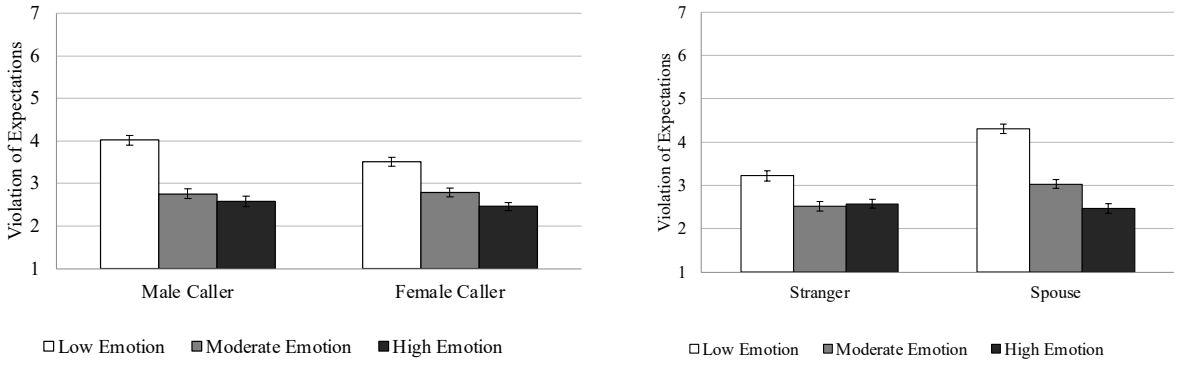
Violate Expectations. There was again a significant main effect of emotion expression on the violation of expectations scale. Bonferroni post-hoc tests confirmed our hypothesis that callers who expressed low emotion violated participants' expectations significantly more (M = 3.70, SD = 1.42) than callers who expressed moderate emotion (M = 2.80, SD = 1.30), M_{diff} = .90 $p < .001$, d = .66, who in turn violated expectations significantly more than callers who expressed high emotion (M = 2.51, SD = 1.20), M_{diff} = .29, p = .02, d = .23. Participants also found male callers (M = 3.14, SD = 1.42) violated their expectations significantly more than female callers (M = 2.89, SD = 1.38) overall. Strangers (M = 2.74, SD = 1.22) violated participants' expectations significantly more than did spouses (M = 3.25, SD = 1.52)—despite actually listening to the exact same set of calls.

There was a significant caller gender × emotion interaction (Table S8), such that the emotion expression effect was significant for both male and female callers, but the emotion expression effect was stronger for male callers than for female callers (See Table S8, Figure S8a).

Supplemental Figure S8

Mean (SE) Violation of Expectations as a Function of Caller Emotion Expression, Caller Gender, and Caller Relationship, Study S1.

(a) (b)



There was also a significant caller relationship × emotion interaction (Table S8, Figure 8b), such that the emotion expression effect was significant for both callers reporting that a stranger or spouse was shot, but the emotion expression effect was stronger for callers reporting that their spouse had been shot than for callers reporting a stranger had been shot (See Figure 8b).

There was also an unexpected significant three-way interaction, but it is difficult to interpret. This interaction did not call our results into question because our hypothesis that low levels of emotion on a 911 call would violate expectations more than higher levels was supported: the simple effect of emotion expression was significant for all versions of the call (See Table S9), however the interaction reflected that the strength of this effect varied across versions in ways that are difficult to interpret.

Table S9

The Simple Effects of Caller Emotion on Violation of Expectations, as a Function of Caller Gender and Caller Relationship to the Victim, Study S1.

Independent Variable	*df*	*F*	*p*	Effect Size
Man calling about a stranger	2	11.50	< .001	$\eta_p^2 = .03$
Man calling about a spouse	2	64.60	< .001	$\eta_p^2 = .09$
Woman calling about a stranger	2	3.74	.02	$\eta_p^2 = .01$
Woman calling about a spouse	2	30.91	< .001	$\eta_p^2 = .07$
Error	806			

Moral Agency. There was again a significant main effect of emotion expression on moral agency (Table S8). Bonferroni post-hoc tests revealed that officers perceived callers who expressed low emotion ($M = 3.22$, $SD = 1.09$) had significantly more moral agency as callers

who expressed moderate emotion (M = 2.73, SD = 1.02), M_{diff} = .49, p < .001, d = .46, and high emotion (M = 2.52, SD = 1.02), M_{diff} = .70, p < .001, d = .66. They perceived callers who expressed moderate emotion were also seen as having significantly more moral agency than high emotion, M_{diff} = .21, p = .03, d = .21.

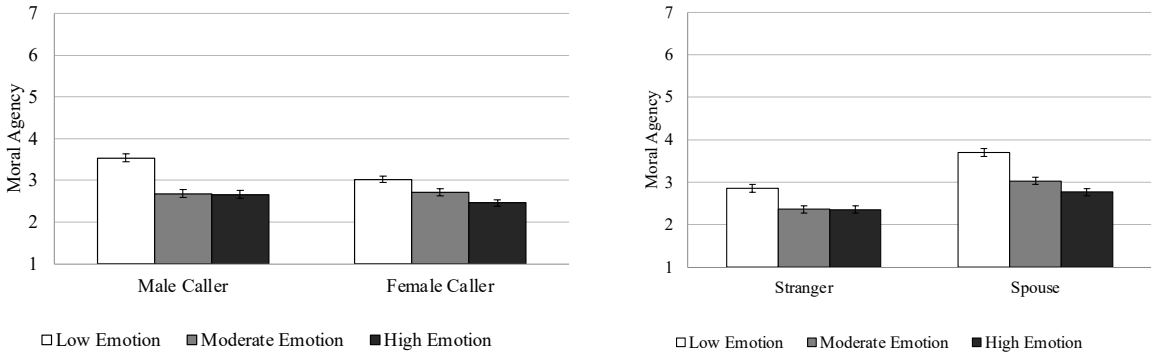
Participants also found male callers (M = 2.98, SD = 1.12) to have more moral agency than female callers (M = 2.70, SD = 1.04) overall. Strangers (M = 2.49, SD = .90) were perceived as having less moral agency than spouses (M = 3.15, SD = 1.15) —despite actually listening to the exact same set of calls.

There was a significant caller gender × emotion interaction (Table S8), such that the emotion expression effect was significant for both male and female callers, but the emotion expression effect was again stronger for male callers than for female callers—similar to the interaction effect on violation of expectations (See Figure 9a).

Supplemental Figure S9

Mean (SE) Perceived Moral Agency as a Function of Caller Emotion Expression, Caller Gender, and Caller Relationship, Study S1.

(a) (b)



There was also a significant caller relationship × emotion interaction (Table S8), such that the emotion expression effect was significant for both callers reporting that a stranger or

spouse was shot, but the emotion expression effect was again stronger for callers reporting that their spouse had been shot than for callers reporting a stranger had been shot—similar to the interaction effect on violation of expectations (See Figure 8b).

The three-way interaction was not significant (Table S8).

Moral Patiency. There was again a significant main effect of emotion expression. Bonferroni post-hoc tests revealed that participants perceived callers who expressed low emotion (M = 3.18, SD = .91) to have significantly less moral patiency than callers who expressed moderate emotion (M = 3.95, SD = .88), M_{diff} = -.07, p < .001, d = .85, and compared to callers who expressed high emotion (M = 4.24, SD = .82), , M_{diff} = -1.05, p < .001, d = 1.22. Further, participants perceived callers who expressed moderate emotion to also have significantly less moral patiency than callers who expressed high emotion, M_{diff} = -.29, p < .001, d = .36.

Participants also found male callers (M = 3.67, SD = .99) to have less moral patiency than female callers (M = 3.88, SD = .96) overall (Table S4). Strangers (M = 3.93, SD = .85) were perceived as having more moral patiency than spouses (M = 3.67, SD = 1.07) —despite actually listening to the exact same set of calls.

There was a significant caller gender × emotion interaction (Table S8, Figure 9a), such that the emotion expression effect was significant for both male and female callers, but the emotion expression effect was again stronger for male callers than for female callers—similar to the interaction effect on violation of expectations and moral agency (See Figure 10a).

Supplemental Figure S10

Mean (SE) Perceived Moral Patiency as a Function of Caller Emotion Expression, Caller Gender, and Caller Relationship, Study S1.

(a) (b)

There was also a significant caller relationship × emotion interaction (Table S8, Figure 10b), such that the emotion expression effect was significant for both callers reporting that a stranger or spouse was shot, but the emotion expression effect was again stronger for callers reporting that their spouse had been shot than for callers reporting a stranger had been shot—similar to the interaction effect on violation of expectations and moral agency (See Figure 9b).

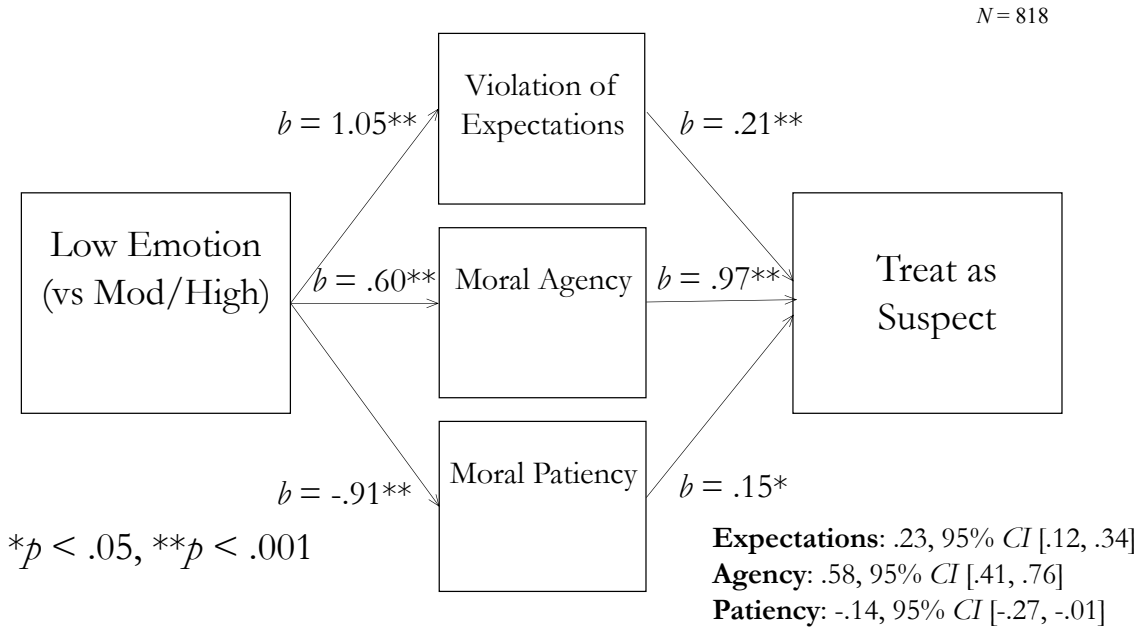
The three-way interaction was not significant (Table S8).

Treat as a Suspect Mediation Results Study S1

We next tested whether we could replicate the significant indirect effects of emotion expression on suspicion through the three mediators if we ran the same model on support for treating the caller as a suspect (Figure S11). We replicated all of the three indirect effects of callers expressing low (versus moderate/high) emotion on support for treating the caller as a suspect through violation of expectations, $M_{indirecteffect}$ = .23, *SE* = .06, 95% *CI* [.12, .34], moral agency, $M_{indirecteffect}$ = .58, *SE* = .09, 95% CI [.41, .76], and moral patiency, $M_{indirecteffect}$ = -.14, SE = .07, 95% CI [-.27, -.01].

Supplemental Figure S11

Indirect Effects of Emotion Expression on Support for Treating the Caller as a Suspect through Violation of Expectations, Moral Agency, and Moral Patiency, Study S1



Suspicion Moderated Mediation Results Study S1

In contrast to Studies 2 and 3, we found that the indirect effects of caller emotion expression were moderated by caller gender and caller relationship. Given gender did not moderate the mediation models in Studies 2-3—calling their reliability into question—and the patterns were not related to our hypotheses, these moderated mediation model results are detailed here for transparency, rather than in the main text.

The indirect effects of low (versus moderate/high) emotion on suspicion moderated by caller gender. In contrast to Studies 2-3, caller gender moderated the strength of the indirect effect of low (versus moderate/high) emotion on suspicion through violation of expectations, Index of moderated mediation (*IMM*) = -.24, *SE* = .09, 95% *CI* [-.43, -.07] and moral agency, *IMM* = -.40, *SE* = .14, 95% *CI* [-.66, -.13]. The indirect effects were significant in all conditions,

but the indirect effect was stronger in the male (versus female) caller condition for both violation of expectations (male: $M_{indirecteffect}$ = .61, *SE* = .08, 95% *CI* [.45, .78]; female: $M_{indirecteffect}$ = .37, *SE* = .07, 95% *CI* [.24, .51]) and moral agency (male: $M_{indirecteffect}$ = .75, *SE* = .11, 95% *CI* [.53, .96]; female: $M_{indirecteffect}$ = .34, *SE* = .09, 95% *CI* [.17, .52]).

The strength of the indirect effect of low (versus moderate/high) emotion on suspicion through patiency, however, did not differ based on caller gender, *IMM* = .02, *SE* = .01, 95% *CI* [-.01, .05]. The conditional indirect effects were both not significant, male: $M_{indirecteffect}$ = -.07, *SE* = .04, 95% *CI* [-.16, .02]; female: $M_{indirecteffect}$ = -.05, *SE* = .03, 95% *CI* [-.12, .01].

The indirect effects of low (versus moderate/high) emotion on suspicion moderated by caller relationship. Next, we tested whether caller relationship moderated these patterns, and found that it did. Caller relationship moderated the indirect effect through violation of expectations, Index of moderated mediation (*IMM*) = .41, *SE* = .09, 95% *CI* [.24, .59] and moral agency, *IMM* = .29, *SE* = .13, 95% *CI* [.04, .54]. The indirect effects were significant in all conditions, but the indirect effect was stronger in the spouse (versus stranger) condition for both violation of expectations (stranger: $M_{indirecteffect}$ = .27, *SE* = .06, 95% *CI* [.15, .40]; spouse: $M_{indirecteffect}$ = .68, *SE* = .08, 95% *CI* [.52, .85]) and moral agency (stranger: $M_{indirecteffect}$ = .37, *SE* = .08, 95% *CI* [.22, .54]; spouse: $M_{indirecteffect}$ = .66, *SE* = .10, 95% *CI* [.46, .88]).

The indirect effect through moral patiency was not moderated by caller relationship, *IMM* = -.03, *SE* = .02, 95% *CI* [-.07, .01]. The conditional indirect effects were both not significant, stranger: $M_{indirecteffect}$ = -.05, *SE* = .03, 95% *CI* [-.11, .02]; spouse: $M_{indirecteffect}$ = -.07, *SE* = .05, 95% *CI* [-.17, .02].

In summary, we found support for the impact of low (versus moderate/high) emotion decreasing suspicion through both violation of expectations and greater perceived moral agency

in all conditions—but these two indirect effects were stronger when the caller was male (versus female) or a spouse (versus stranger) to the victim. The unique indirect effect through perceived moral patiency was not significant in any of the conditions—though see the main text (p. 44) for alternative single-mediator model suggesting that the indirect effect of moral patiency is significant when not controlling for moral agency.

Treat as a Suspect Moderated Mediation Results Study S1

The indirect effects of low (versus moderate/high) emotion on support for treating the caller as a suspect moderated by caller gender. In contrast to Studies 2-3, caller gender moderated the strength of the indirect effect of low (versus moderate/high) emotion on support for treating the caller as a suspect through violation of expectations, Index of moderated mediation (*IMM*) = -.12, *SE* = .05, 95% *CI* [-.23, -.03] and moral agency, *IMM* = -.46, *SE* = .16, 95% *CI* [-.78, -.15]. The indirect effects were significant in all conditions, but the indirect effect was stronger in the male (versus female) caller condition for both violation of expectations (male: $M_{indirecteffect}$ = .29, *SE* = .08, 95% *CI* [.15, .46]; female: $M_{indirecteffect}$ = .18, *SE* = .05, 95% *CI* [.08, .29]) and moral agency (male: $M_{indirecteffect}$ = .85, *SE* = .13, 95% *CI* [.59, 1.12]; female: $M_{indirecteffect}$ = .39, *SE* = .10, 95% *CI* [.20, .60].

The strength of the indirect effect of low (versus moderate/high) emotion on support for treating the caller as a suspect through patiency, however, did not differ based on caller gender, *IMM* = .04, *SE* = .03, 95% *CI* [-.003, .10]. The conditional indirect effects were both significant, male: $M_{indirecteffect}$ = -.16, *SE* = .08, 95% *CI* [-.16, -.01]; female: $M_{indirecteffect}$ = -.12, *SE* = .06, 95% *CI* [-.24, -.01].

The indirect effects of low (versus moderate/high) emotion on treating the caller as a suspect moderated by caller relationship. Next, we tested whether caller relationship

moderated these patterns, and found that it did. Relationship with the victim moderated the strength of the indirect effect of low (versus moderate/high) emotion on support for treating the caller as a suspect through violation of expectations, Index of moderated mediation (*IMM*) = .20, *SE* = .06, 95% *CI* [.09, .33]; moral agency, *IMM* = .35, *SE* = .15, 95% *CI* [.06, .63]; and moral patiency, *IMM* = -.06, *SE* = .04, 95% *CI* [-.15, -.004]. The indirect effects were significant in all conditions, but the indirect effect was stronger in the spouse (versus stranger) condition for violation of expectations (stranger: $M_{indirecteffect}$ = .13, *SE* = .04, 95% *CI* [.06, .22]; spouse: $M_{indirecteffect}$ = .33, *SE* = .08, 95% *CI* [.17, .49]), moral agency (stranger: $M_{indirecteffect}$ = .41, *SE* = .10, 95% *CI* [.23, .61]; spouse: $M_{indirecteffect}$ = .76, *SE* = .12, 95% *CI* [.53, 1.02], and moral patiency (stranger: $M_{indirecteffect}$ = -.22, *SE* = .05, 95% *CI* [-.22, -.01]; spouse: $M_{indirecteffect}$ = .76, *SE* = .12, 95% *CI* [.53, 1.02].

Study S1 Supplemental Analyses Summary

We replicated our findings regarding suspicion reported for Study S1 (and earlier studies) with support for treating the caller as a suspect. More specifically, we found that expressing low levels (versus moderate or high levels) of emotion on a 911 call led to increased support for treating the caller as a suspect. Further, with each increase in emotion expression (low versus moderate versus high) there was a significant increase in violation of expectations, an increase in perceptions of moral agency and a decrease in perceptions of moral patiency. Similar to suspicion, all three mediators were significantly associated with support for treating the 911 caller as a suspect. In summary, low (versus moderate or high) emotion increased support for treating the caller as a suspect indirectly through violation of expectations and morally typecasting the caller as more of a moral agent and less of a moral patient.

In addition, we found that male callers were perceived as having more moral agency, less moral patiency and as more worthy of treating as a suspect than female callers. We also replicated the relationship main effect on suspicion, such that people thought callers reporting a spouse was hurt were more worthy of treating as a suspect than callers reporting that a stranger had been hurt.

Finally, we found some caller gender × emotion interactions on mediators that were unexpected because we did not find them on the mediators in Studies 2-3, nor on treating the caller as a suspect in this study. We also found caller gender × relationship interactions on mediators that we did not find on treating the caller as a suspect in this study. They followed a similar pattern: The effect of emotion expression on the three mediators was significant in all conditions, but the simple effects of emotion expression were stronger when the caller was male (versus female) and calling about a spouse (versus a stranger). The moderated mediation analyses were consistent with this: the significant indirect effects of emotion expression on both suspicion and treating the caller as a suspect were significant through violation of expectations and moral agency in all conditions, but they were stronger in the male (versus female) caller condition and in the spouse (versus stranger) condition. The indirect effect of caller emotion expression on suspicion and treating the caller as a suspect through moral patiency did not, however, depend on caller gender or relationship with the victim. Instead, the indirect effect of caller emotion expression on suspicion through moral patiency was not significant (regardless of gender or relationship), but it was significant on support for treating the caller as a suspect.