Supplementary Materials

This document contains supplementary materials for: The effect of trauma education judicial instructions on decisions about complainant credibility in rape trials

In these supplementary materials, we report the results of our pre-registered analyses, both Bayesian and frequentist, in our pre-registered design. We also report the complete results of our frequentist analyses in our unregistered study design.

Pre-registered Bayesian Analyses on Pre-registered Design

Unless otherwise specified, all analyses are either a Bayesian or frequentist between-subjects ANOVA with complainant emotion and instruction type entered as predictors in the analysis.

Complainant Credibility and Defendant Guilt (H1)

To test H1, we conducted a Bayesian ANOVA on perceptions of the complainant's credibility. There was moderate to strong evidence against including main effects of complainant emotion or instruction or the interaction as predictors (see Table S1 for Bayes factors).

As a corroboratory test of H1, we conducted a frequentist ANOVA on perceptions of the complainant's credibility. There was no significant main effect of either instruction type, complainant emotion or interaction of instruction type or complainant emotion on perceptions of complainant credibility (see Table S2 for ANOVA summary statistics and Table S3 for condition descriptive statistics).

Also to test H1, we also conducted a Bayesian ANOVA on perceptions of the likelihood of the defendant’s guilt. There was anecdotal evidence against including either a main effect of complainant emotion or instruction type as predictors in the model and strong evidence against there being an interaction of complainant emotion (see Table S1 for Bayes factors).

As a corroboratory test of H1, we conducted a frequentist ANOVA on perceptions of the likelihood the defendant. There was a significant main effect of complainant emotion. The defendant was perceived to be more guilty when the complainant was distressed (M = 4.65, SD = 1.09) compared to when the complainant was unemotional (M = 4.46, SD = 1.13). There was no significant main effect of instruction type or interaction of instruction type and complainant emotion (see Table S2 for ANOVA summary statistics and Table S3 for condition descriptive statistics). Given that the Bayes factors and p-value for the main effect of complainant emotion on perceptions of defendant guilt provide contrary interpretations, the effect may not be robust and should be treated with caution (see Table S1). Collectively, these results suggested there was no support for H1.

Table S1

Bayes Factors, p-values and Effect Sizes for Bayesian and Frequentist ANOVAs on Complainant Credibility and Defendant Guilt Outcomes

Outcome Measure	Effect	Bayesian	Frequentist	Inference
Victim Credibility	Effect of complainant emotion	BF_{incl} = 0.32	p = .077, d = 0.14 95% CI [-0.01, 0.28]	- -
	Effect of instruction type	BF_{incl} = 0.32	p = .383, d = -0.14 95% CI [-0.28, 0.01]	- -
	Interaction	BF_{incl} = 0.06	p = .526	- -
Likelihood Defendant Guilty	Effect of complainant emotion	BF_{incl} = 0.82	p = .014, d = 0.17 95% CI [0.03, 0.32]	~ +
	Effect of instruction type	BF_{incl} = 0.37	p = .602, d = 0.14 95% CI [-0.29, 0.002]	- -
	Interaction	BF_{incl} = 0.15	p = .239	- -

Note. Effects marked with + + or - - indicate Bayes and Frequentist analysis both suggest effect present or not present respectively. Effects marked with ~ indicate either Bayesian or Frequentist analysis did not provide support for the effect and these effects should not be considered robust.

Table S2

ANOVA statistics for Registered Design on Key Outcomes Variables

	Outcomes				
	Complainant Credibility	Likelihood of Defendant Guilt	Preferences for Additional Evidence	Cognitive Response Task	Complainant Distress
Main effect of instruction type	$F(1, 730) = 0.76$, $p = .383$	$F(1, 730) = 0.27$, $p = .602$	$F(1, 730) = 1.26$, $p = .262$	$F(1, 730) = 0.06$, $p = .813$	$F(1, 730) = 3.36$, $p = .067$
Main effect of complainant emotion	$F(1, 730) = 3.14$, $p = .077$	$F(1, 730) = 6.13$, $p = .014$	$F(1, 730) = 1.26$, $p = .262$	$F(1, 730) = 3.09$, $p = .079$	$F(1, 730) = 543.72$, $p < .001$
Complainant emotion x Instruction type	$F(1, 730) = 0.40$, $p = .526$	$F(1, 730) = 1.39$, $p = .239$	$F(1, 730) = 0.85$, $p = .358$	$F(1, 730) = 1.21$, $p = .272$	$F(1, 730) = 23.89$, $p < .001$

Table S3

Condition Descriptive Statistics for Registered Design on Key Outcome Variables

			Outcomes				
			Complainant Credibility	Likelihood of Defendant Guilt	Preferences for Additional Evidence	Cognitive Response Task	Complainant Distress
Instruction type	Complainant emotion	*n*	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
Standard	Distress	182	4.67 (0.08)	4.62 (0.08)	1.97 (0.24)	0.34 (0.11)	6.21 (0.10)
	Unemotional	185	4.47 (0.08)	4.34 (0.08)	2.35 (0.24)	0.05 (0.11)	2.79 (0.10)
Trauma Education	Distress	185	4.77 (0.08)	4.68 (0.08)	2.35 (0.24)	0.37 (0.11)	5.94 (0.10)
	Unemotional	182	4.67 (0.08)	4.59 (0.08)	2.29 (0.24)	0.34 (0.11)	3.54 (0.10)

Information Processing (H2)

To test H2, we conducted two Bayesian ANOVAs on indices for the preference for additional evidence measure and cognitive response task. For the preference for additional evidence measure, a Bayesian ANOVA indicated that there was very strong evidence against a main effect of either complainant emotion or instruction type and extreme evidence against an interaction (see Table S4 for Bayes factors). For the cognitive response task, a Bayesian ANOVA indicated there was moderate evidence against a main effect of either complainant emotion or instruction type and very strong evidence against an interaction of complainant emotion and instruction type.

As a corroboratory test of H2, we conducted two frequentist ANOVAs on the indices for the preferences for additional evidence and cognitive response task. For the preferences for additional evidence measure, there were no significant main effects of complainant emotion or instruction type and no significant interaction of complainant emotion and instruction type. For the cognitive response task, there were also no significant main effects of complainant emotion or instruction type and no significant interaction of complainant emotion and instruction type (see Table S2 for ANOVA summary statistics and S3 for condition descriptive statistics). Collectively, this suggested there was no support for H2.

Table S4

Bayes Factors, p-values and Effect Sizes for Bayesian and Frequentist ANOVAs on Information Processing Outcomes

Outcome Measure	Effect	Bayesian	Frequentist	Inference
Preferences for Additional Evidence	Effect of complainant emotion	$BF_{incl} = 0.07$	$p = .262$, $d = -0.04$ 95% CI [-0.19, 0.09]	- -
	Effect of instruction type	$BF_{incl} = 0.07$	$p = .262$, $d = -0.04$ 95% CI [-0.19, 0.10]	- -

	Interaction	BF_{incl} = 0.006	p = .358	- -
Cognitive Response Task	Effect of complainant emotion	BF_{incl} = 0.15	p = .813, d = 0.10 95% CI [-0.04, 0.25]	- -
	Effect of instruction type	BF_{incl} = 0.16	p = .079, d = -0.11 95% CI, [-0.25, 0.04]	- -
	Interaction	BF_{incl} = 0.03	p = .272	- -

Note. Effects marked with + + or - - indicate Bayes and Frequentist analysis both suggest effect present or not present respectively. Effects marked with ~ indicate either Bayesian or Frequentist analysis did not provide support for the effect and these effects should not be considered robust.

Trauma Knowledge and Misperceptions (H3)

To test H3, we conducted five Bayesian ANOVAs on responses on each trauma knowledge question. There was strong evidence in favour of a main effect of instruction type on participants' perceptions of the proportion of rape complainants who would experience distress. There was also anecdotal evidence against a main effect of complainant emotion and moderate evidence against an interaction of complainant emotion and instruction type (see Table S5 for Bayes factors).

A Bayesian ANOVA indicated that there was extreme evidence in favour of a main effect of instruction type on participants' perceptions of the proportion of complainants who would experience negative feelings after an assault. There was also strong evidence against a main effect of emotion and an interaction of complainant emotion and instruction type (see Table S5 for Bayes factors). These analyses provided partial support for H3.

A Bayesian ANOVA indicated there was anecdotal evidence in favour of a main effect of instruction type on participants perceptions of the proportion of rape complainants who would avoid thoughts and feelings of the assault. There was moderate evidence against a

main effect of complainant emotion and strong evidence against an interaction of complainant emotion and instruction type (see Table S5 for Bayes factors).

Table S5

Bayes Factors, p-values and Effect Sizes for Bayesian and Frequentist ANOVAs on Trauma Knowledge Outcomes

Trauma Symptom	Effect	Bayesian	Frequentist	Inference
Emotionally upset	Effect of complainant emotion	$BF_{incl} = 0.35$	$p = .045$, $d = 0.13$ 95% CI [-0.02, 0.27]	~ +
	Effect of instruction type	$BF_{incl} = 9.42 \times 10^8$	$p < .001$, $d = 0.52$ 95% CI [0.37, 0.67]	+ +
	Interaction	$BF_{incl} = 0.23$	$p = .333$	- -
Negative feelings	Effect of complainant emotion	$BF_{incl} = 0.08$	$p = .340$, $d = 0.04$ 95% CI [-0.10, 0.19]	- -
	Effect of instruction type	$BF_{incl} = 458.33$	$p < .001$, $d = 0.32$ 95% CI [0.17, 0.46]	+ +
	Interaction	$BF_{incl} = 0.05$	$p = .501$	- -
Avoid thoughts/feelings	Effect of complainant emotion	$BF_{incl} = 0.12$	$p = .083$, $d = 0.08$ 95% CI [-0.06, 0.23]	- -
	Effect of instruction type	$BF_{incl} = 1.15$	$p = .006$, $d = 0.18$ 95% CI [0.04, 0.33]	~ +
	Interaction	$BF_{incl} = 0.09$	$p = .182$	- -

Note. Effects marked with + + or - - indicate Bayes and Frequentist analysis both suggest effect present or not present respectively. Effects marked with ~ indicate either Bayesian or Frequentist analysis did not provide support for the effect and these effects should not be considered robust.

Table S6

Bayes Factors, p-values and Effect Sizes for Bayesian and Frequentist ANOVAs on Trauma Knowledge and Misperceptions

Trauma Symptom	Effect	Bayesian	Frequentist	Inference
Feeling Isolated	Effect of complainant emotion	$BF_{incl} = 0.06$	$p = .794$, $d = -0.03$ 95% CI [-0.17, 0.12]	- -
	Effect of instruction type	$BF_{incl} = 0.07$	$p = .307$, $d = 0.05$ 95% CI [-0.09, 0.19]	- -
	Interaction	$BF_{incl} = 0.005$	$p = .467$	- -
Difficulty Feeling anything	Effect of complainant emotion	$BF_{incl} = 0.06$	$p = .721$, $d = 0.01$ 95% CI [-0.13, 0.15]	- -
	Effect of instruction type	$BF_{incl} = 0.29$	$p = .085$, $d = -0.14$ 95% CI [-0.28, 0.01]	- -
	Interaction	$BF_{incl} = 0.01$	$p = .546$	- -
Trauma Misperceptions	Effect of complainant emotion	$BF_{incl} = 0.14$	$p = .498$, $d = 0.10$ 95% CI [-0.05, 0.24]	- -
	Effect of instruction type	$BF_{incl} = 16.33$	$p = .031$, $d = 0.25$ 95% CI [0.11, 0.40]	+ +
	Interaction	$BF_{incl} = 0.08$	$p = .707$	- -

Note. Effects marked with + + or - - indicate Bayes and Frequentist analysis both suggest effect present or not present respectively. Effects marked with ~ indicate either Bayesian or Frequentist analysis did not provide support for the effect and these effects should not be considered robust.

For perceptions of the proportion of complainants who would feel isolated, a Bayesian ANOVA indicated that there was strong evidence against both a main effect of complainant emotion and instruction type and extreme evidence against an interaction of complainant emotion and instruction type (see Table S6 for Bayes factors). For perception of the proportion of complainants who would have difficulty feeling emotions, a Bayesian ANOVA indicated strong evidence against a main effect of complainant emotion, moderate evidence against a main effect of instruction type and very strong evidence against an interaction (see Table S6 for Bayes factors). These analyses did not provide support for H3.

As corroboratory tests of H3, we conducted five frequentist between-subjects' ANOVAs responses on each trauma knowledge question. There was a significant main effect of instruction type on participants perceptions of the proportion of rape complainants would appear distressed. Participants who received the trauma education instruction thought fewer rape complainants would show distress (M = 76.70, SD = 21.53) than participants who received a standard instruction (M = 86.15, SD = 14.23). There was also a significant main effect of complainant emotion such that participants who saw a distressed complainant thought more rape complainants would show distress (M = 82.64, SD = 18.34) compared to those who saw an unemotional complainant at trial (M = 80.22, SD = 19.28). There was no significant interaction of instruction type or complainant emotion.

There was also significant main effect of instruction type on participants' perceptions of the proportion of complainants who would experience negative feelings. Participants who received the trauma education instruction thought fewer complainants would experience strong negative feelings after the assault (M = 79.84, SD = 17.98) compared to those who received the standard instruction (M = 85.15, SD = 15.31). There was no significant main effect of complainant emotion or interaction of instruction type and complainant emotion.

There was a significant main effect of instruction type on participants perceptions of how many complainants would avoid thoughts or feelings about the assault. Participants who received the trauma education instruction thought fewer rape complainants would avoid thoughts and feelings about the assault (M = 82.02, SD = 16.39) compared to those who received the standard instruction (M = 84.89, SD = 14.93). There was no significant main effect of complainant emotion or interaction of instruction type and complainant emotion. These analyses provided partial support for H3.

There was no significant main effect of instruction type for either participants perceptions of how many complainants would feel isolated after the assault or would have

difficulty feeling any emotion after the assault. There was also no significant main effect of complainant emotion or interaction of complainant emotion and instruction type. Collectively, these results did not provide support for H3. The educational instruction does improve jurors' knowledge of some trauma symptoms.

Manipulation Checks

To check that our manipulation of complainant emotion was effective, we conducted both Bayesian and frequentist ANOVAs on participants' perceptions of how distressed the complainant appeared. There was extreme evidence in favour of including main effects of complainant emotion and instruction type in the model as well as in favour of an interaction of complainant emotion and instruction type (see Table S7 for Bayes factors).

The frequentist ANOVA indicated there was a significant main effect of complainant emotion. Participants perceived the complainant to be more distressed in the distressed complainant condition (M = 6.07, SD = 0.98) compared to the unemotional complainant condition (M = 3.16, SD = 1.77, see Table S7 for effect sizes). This suggests the manipulation of complainant emotion was effective. There was also an unexpected significant interaction of complainant emotion and instruction type (see Table S2 for ANOVA summary statistics, S3 for condition descriptive statistics and Figure S1 for interaction plot).

To check that our manipulation of instruction type was effective, we conducted Bayesian and frequentist ANOVAs on participants perceptions of how objective the judge had told them to be (a key message from the standard witness instruction). The Bayesian ANOVA indicated there was only anecdotal evidence in favour of a main effect of instruction type, such that participants who received the standard witness instruction thought they had to be more objective (M = 5.63, SD = 1.26) compared to participants who read the trauma education instruction (M = 5.35, SD = 1.42). There was strong evidence against a main effect

of complainant emotion or interaction in this analysis. This suggests that the manipulation of instruction type was not very effective.

These results were not corroborated by the frequentist ANOVA on participants perceptions of how objective the judge had told them to be. This analysis indicated there were no significant effects of either instruction type or complainant emotion or an interaction (see Table S7).

Table S7

Bayes Factors, p-values and Effect Sizes for Bayesian and Frequentist ANOVAs on Manipulation Check Items

Manipulation Checks	Effect	Bayesian	Frequentist	Inference
Complainant Distress	Effect of complainant emotion	BF_{incl} = infinite	p <.001, d = 2.04 95% CI [1.85, 2.22]	+ +
	Effect of instruction type	BF_{incl} = 7067.04	p = .067, d = -0.13 95% CI [-0.27, 0.02]	+ ~
	Interaction	BF_{incl} = 21257.00	p <.001	+ +
Extent of Objectivity	Effect of complainant emotion	BF_{incl} = 0.08	p = .259, d = -0.06 95% CI [-0.21, 0.08]	- -
	Effect of instruction type	BF_{incl} = 2.76	p = .147, d = 0.21 95% CI [0.06, 0.35]	- -
	Interaction	BF_{incl} = 0.05	p = .441	- -

Note. Effects marked with + + or - - indicate Bayes and Frequentist analysis both suggest effect present or not present respectively. Effects marked with ~ indicate either Bayesian or Frequentist analysis did not provide support for the effect and these effects should not be considered robust.

Figure S1

Condition Estimates and 95% Confidence Intervals for Complainant Distress

Frequentist Analyses of Key Outcome Variables in Unregistered Design

Analysis Strategy

We also conducted the equivalent frequentist analyses on our unregistered design data. For all continuous outcome variables, we conducted 2 (complainant emotion) x 3 (instruction type) between-subjects ANOVAs.

Follow up contrasts. In our registered design, we had several pre-registered hypotheses that instruction type should moderate the effect of complainant emotion on core outcome variables. When analysing the data for our unregistered design, in which we added a no instruction condition to the design, we extended the logic of our hypotheses to guide how we decomposed main effects of instruction type for the core outcome variables. We used two sets of contrasts in these analyses. Our first contrast compared the no instruction condition to both instruction conditions. Our second contrast represented our pre-registered hypotheses – comparing the standard and educational instruction conditions.

Interactions. We decomposed interactions via estimation methods proposed by Cummings (e.g., Cummings & Finch, 2005; Cummings, 2007; Cummings, 2009). In this

method, means for condition and confidence intervals are plotted to determine which conditions are statistically different from each other. Estimation based decomposition of interactions is recommended when exploratory analyses are undertaken (e.g., Cummings, 2012).

Complainant Credibility and Defendant Guilt

We conducted a frequentist ANOVA on complainant credibility scores. There was a significant main effect of complainant emotion, a significant main effect of instruction type and a significant interaction of complainant emotion and instruction type (see Table S8 for ANOVA summary and contrast statistics). Decomposing the interaction by eye, the distressed complainant was only seen as more credible than the unemotional complainant in the no instruction condition (see Table S9 for condition descriptive statistics). In the instruction conditions, the distressed complainant was not seen as more credible than the unemotional complainant (see Figure S2 for interaction plot).

Table S8

ANOVA Statistics for Unregistered Design on Key Outcomes

	Outcomes			
	Complainant Credibility	Likelihood of Defendant Guilt	Preferences for Additional Evidence	Cognitive Response Task
Main effect of instruction type	$F(2, 1079) = 5.17, p = .005$	$F(2, 1079) = 6.72, p = .001$	$F(2, 1079) = 6.49, p = .002$	$F(2, 1079) = 3.06, p = .047$
No instruction vs instructions	$t = 3.17, p = .001.$	$t = -3.12, p = .002$	$t = -3.54, p < .001$	$t = -2.05, p = .041$
Standard vs education	$t = 1.87, p = .062$	$t = 1.92, p = .053$	$t = 0.68, p = .494$	$t = 1.39, p = .165$
Main effect of complainant emotion	$F(1, 1079 = 20.56, p < .001$	$F(1, 1079) = 13.77, p < .001$	$F(1, 1079) = 0.27, p = .606$	$F(1, 1079) = 12.91, p < .001$
Complainant emotion x Instruction type	$F(2, 1079) = 5.23, p = .005$	$F(2, 1079) = 1.45, p = .235$	$F(2, 1079) = 2.23, p = .108$	$F(2, 1079) = 4.35, p = .013$

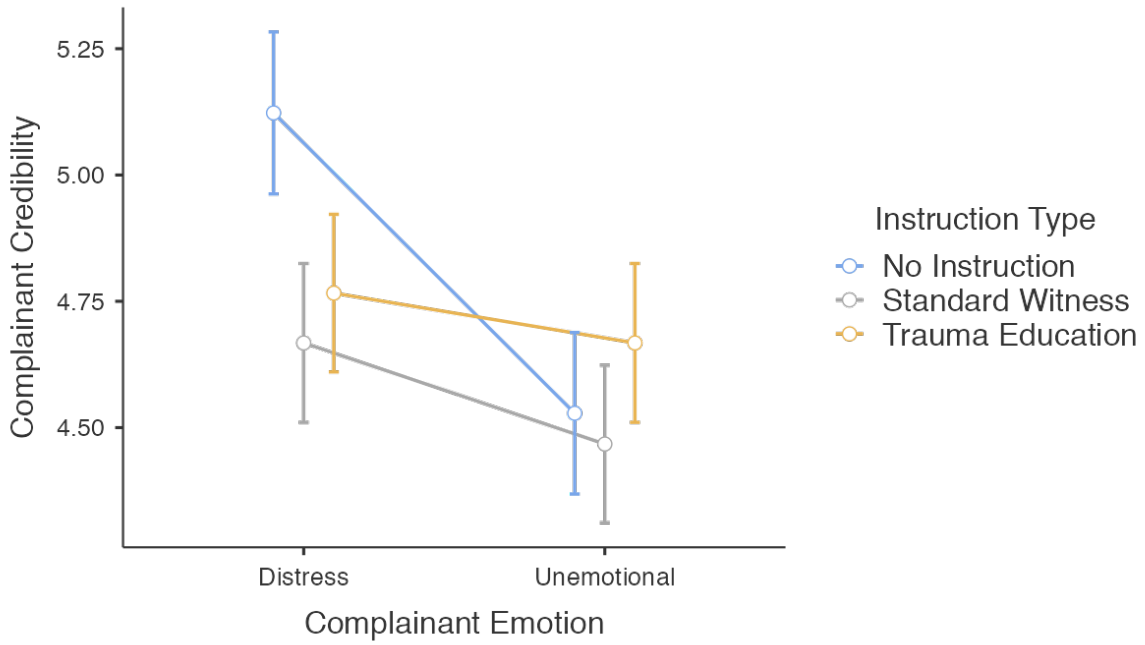
Table S9

Condition Descriptive Statistics for Unregistered Design on Key Outcome Variables

			Outcomes			
			Complainant Credibility	Likelihood of Defendant Guilt	Preferences for Additional Evidence	Cognitive Response Task
Instruction type	Complainant emotion	*n*	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
No instruction	Distress	175	5.12 (0.08)	4.96 (0.08)	3.26 (0.24)	0.85 (0.12)
	Unemotional	176	4.53 (0.08)	4.59 (0.08)	2.65 (0.24)	0.13 (0.12)
Standard	Distress	182	4.67 (0.08)	4.62 (0.08)	1.97 (0.24)	0.34 (0.12)
	Unemotional	185	4.47 (0.08)	4.34 (0.08)	2.35 (0.24)	0.05 (0.12)
Trauma Education	Distress	185	4.77 (0.08)	4.68 (0.08)	2.35 (0.24)	0.37 (0.12)
	Unemotional	182	4.67 (0.08)	4.59 (0.08)	2.29 (0.24)	0.34 (0.12)

Figure S2

Condition Estimates and 95% Confidence Intervals for Complainant Credibility



We conducted a frequentist ANOVA on perceptions of the defendant's guilt. There was a significant main effect of complainant emotion and instruction type but no significant interaction of complainant emotion and instruction type (see Table S8 for ANOVA summary statistics). When the complainant was distressed the defendant was perceived as more likely to be guilty (M = 4.75, SD = 1.07) compared to when the complainant was unemotional (M = 4.50, SD = 1.14). To follow up the main effect of instruction type, we examined our planned contrasts. Only the first contrast was significant. When participants received no instruction, they perceived the defendant to be more likely to be guilty (M = 4.77, SD = 1.10) compared to when they received either judicial instruction (M = 4.56, SD = 1.11). These analyses provided no support for H1.

Information Processing

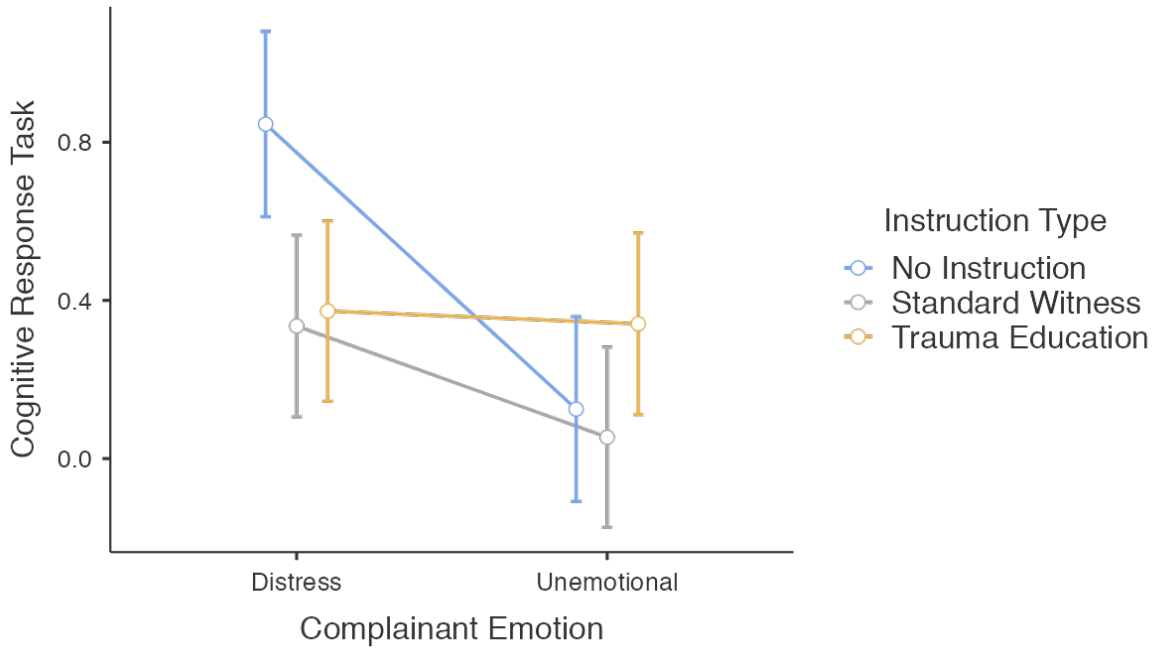
A frequentist ANOVA indicated that there was only a significant main effect of instruction type on participants preferences for additional evidence. Only the first contrast, comparing the effects of no instruction to both instructions was statistically significant. Participants who received no instruction indicated an interest in seeing more items of evidence supporting the complainant's account (M = 2.96, SD = 3.01) compared to those in either of the instruction conditions (M = 2.24, SD = 3.23). Neither the main effect of complainant emotion or interaction of complainant emotion and instruction type were statistically significant (see Table S8 for ANOVA summary statistics).

A frequentist ANOVA indicated that there were significant main effects of both instruction type and complainant emotion on participants scores in the cognitive response task (see Table S9 for ANOVA summary and contrast statistics). There was also a significant interaction of complainant emotion and instruction type (see Figure S3 for plot). Participants in the no instruction condition, provide more thoughts about the evidence that support the complainant's account when the complainant is distressed compared to when she is

unemotional. There are no differences in the type of thoughts listed about the evidence based on complainant emotion in the standard instruction and educational instruction condition (see Table S9 for condition descriptive statistics).

Figure S3

Condition Estimates and 95% Confidence Intervals for the Cognitive Response Task



Trauma Knowledge and Misperceptions

We conducted a series of frequentist between-subjects ANOVAs with instruction type and complainant emotion entered as predictors and each trauma knowledge item and the trauma misperceptions composite entered separately as outcome variables. As the focal effect of interest was the main effect of instruction type, and these analyses were exploratory in the unregistered design, we used post-hoc comparisons (with Tukey's correction) to follow up significant main effects of instruction type. We decomposed interactions via estimation methods proposed by Cummings (e.g., Cummings & Finch, 2005; Cummings, 2007; Cummings, 2009).

A frequenstist ANOVA indicated there were significant main effects of complainant emotion and instruction type as well as a significant interaction of complainant emotion and

instruction type on perceptions of proportion of complainants who would feel emotionally upset after the assault (see Table S10 for ANOVA and post-hoc comparison statistics). Participants who received the education instruction thought fewer complainants would be emotionally distressed (M = 76.70, SD = 21.53) compared participants in the standard (M = 86.16, SD = 14.23) and no instruction conditions (M = 86.18, SD = 14.86). This suggests that the educational instruction is helping to improve trauma knowledge. Decomposing the interaction via estimation methods by eye, only in the no instruction condition did participants who read about a distressed complainant think a greater proportion of complainants would be emotionally distressed compared to those who read about an unemotional complainant (see Figure S4 and Table S11 for condition descriptive statistics).

Figure S4

Condition Estimates and 95% Confidence Intervals Perceptions of Proportion of Complainants Emotionally Upset

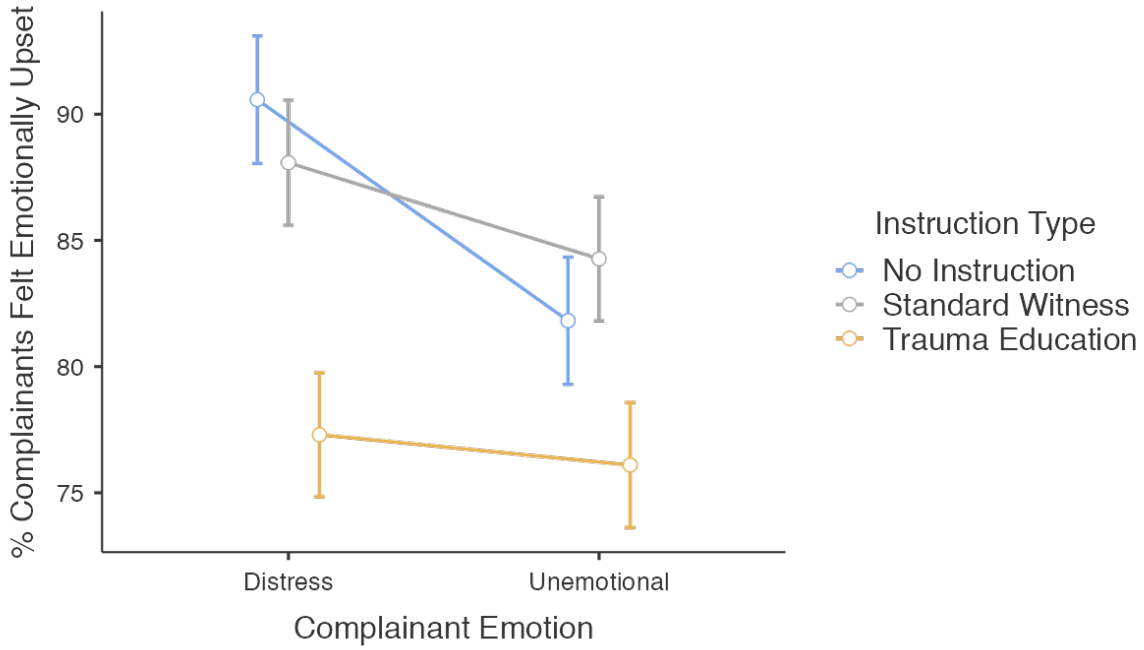


Table S10

ANOVA and Post-hoc Comparison Statistics for Trauma Knowledge and Misperception Outcomes for Unregistered Design

	Trauma Knowledge					Trauma Misperceptions
	Being upset	Avoiding thoughts of assault	Having negative feelings	Feeling isolated	Difficulty feeling emotion	
Main effect of complainant emotion	$F(1, 1079) = 4.58$, $p = .032$ $d = 0.25$, 95% CI [0.13, 0.37]	$F(1, 1079) = 2.95$, $p = .086$ $d = 0.17$, 95% CI [0.05, 0.29]	$F(1, 1079) = 3.92$, $p = .048$ $d = 0.11$, 95% CI [0.00, 0.23]	$F(1, 1079) = 2.33$, $p = .126$ $d = 0.08$, 95% CI [-0.03, 0.21]	$F(1, 1079) = 0.13$, $p = .718$ $d = 0.07$, 95% CI [-0.05, 0.19]	$F(1, 1079) = 0.48$, $p = .488$ $d = 0.13$, 95% CI [0.01, 0.25]
Main effect of instruction type	$F(2, 1079) = 31.22$, $p < .001$	$F(2, 1079) = 5.12$, $p = .003$	$F(2, 1079) = 14.03$, $p < .001$	$F(2, 1079) = 0.28$, $p = .760$	$F(2, 1079) = 3.20$, $p = .041$	$F(2, 1079) = 4.21$, $p = .015$
Education vs No instruction	$t(1079) = -7.47$, $p < .001$ $d = -0.56$, 95% CI [-0.71, -0.41]	$t(1079) = -1.97$, $p = .122$ $d = -0.15$, 95% CI [-0.29, -0.00]	$t(1079) = -4.74$, $p < .001$ $d = -0.35$, 95% CI [-0.50, -0.21]	$t(1079) = -0.54$, $p = .853$ $d = -0.04$, 95% CI [-0.19, 0.11]	$t(1079) = -0.09$, $p = .99$ $d = -0.01$, 95% CI [-0.15, 0.14]	$t(1079) = -3.40$, $p = .002$ $d = -0.25$, 95% CI [-0.40, -0.11]
Education vs Standard	$t(1079) = -7.54$, $p < .001$ $d = -0.56$, 95% CI [-0.70, -0.41]	$t(1079) = -2.47$, $p = .037$ $d = -0.18$, 95% CI [-0.33, -0.04]	$t(1079) = -4\ 39$, $p < .001$ $d = -0.32$, 95% CI [-0.47, -0.18]	$t(1079) = -0.71$, $p = .757$ $d = -0.05$, 95% CI [-0.20, 0.09]	$t(1079) = 1.86$, $p = .152$ $d = 0.14$, 95% CI [-0.01, 0.28]	$t(1079) = -3.50$, $p = .001$ $d = -0.26$, 95% CI [-0.40, -0.11]
Complainant emotion x Instruction type	$F(1, 722) = 4.54$, $p = .011$	$F(1, 1079) = 2.97$, $p = .052$	$F(1, 1079) = 1.52$, $p = .219$	$F(1, 1079) = 4.05$, $p = .018$	$F(1, 1079) = 1.23$, $p = .292$	$F(1, 1079) = 0.35$, $p = .705$

Table S11

Condition Descriptive Statistics for Unregistered Design on Trauma Knowledge and Misperception Outcomes

			Trauma Knowledge					Trauma Misperceptions
			Being upset	Avoiding thoughts of assault	Having negative feelings	Feeling isolated	Difficulty feeling emotion	
Instruction type	Complainant emotion	*n*	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* *(SE)*	*M* *(SE)*
No instruction	Distress	175	90.57 (1.28)	87.09 (1.20)	87.77 (1.24)	80.97 (1.30)	70.23 (1.70)	2.95 (0.10)
	Unemotional	176	81.82 (1.28)	81.59 (1.19)	83.52 (1.24)	75.28 (1.29)	65.85 (1.69)	2.67 (0.10)
Standard	Distress	182	88.08 (1.26)	86.32 (1.17)	85.99 (1.21)	78.57 (1.27)	64.40 (1.66)	2.88 (0.10)
	Unemotional	185	84.27 (1.25)	83.48 (1.17)	84.32 (1.21)	78.11 (1.26)	65.24 (1.65)	2.78 (0.10)
Education	Distress	185	77.30 (1.25)	81.89 (1.16)	79.84 (1.21)	76.76 (1.26)	68.49 (1.65)	2.56 (0.10)
	Unemotional	182	76.10 (1.26)	82.14 (1.17)	79.83 (1.21)	78.12 (1.27)	67.31 (1.66)	2.39 (0.10)

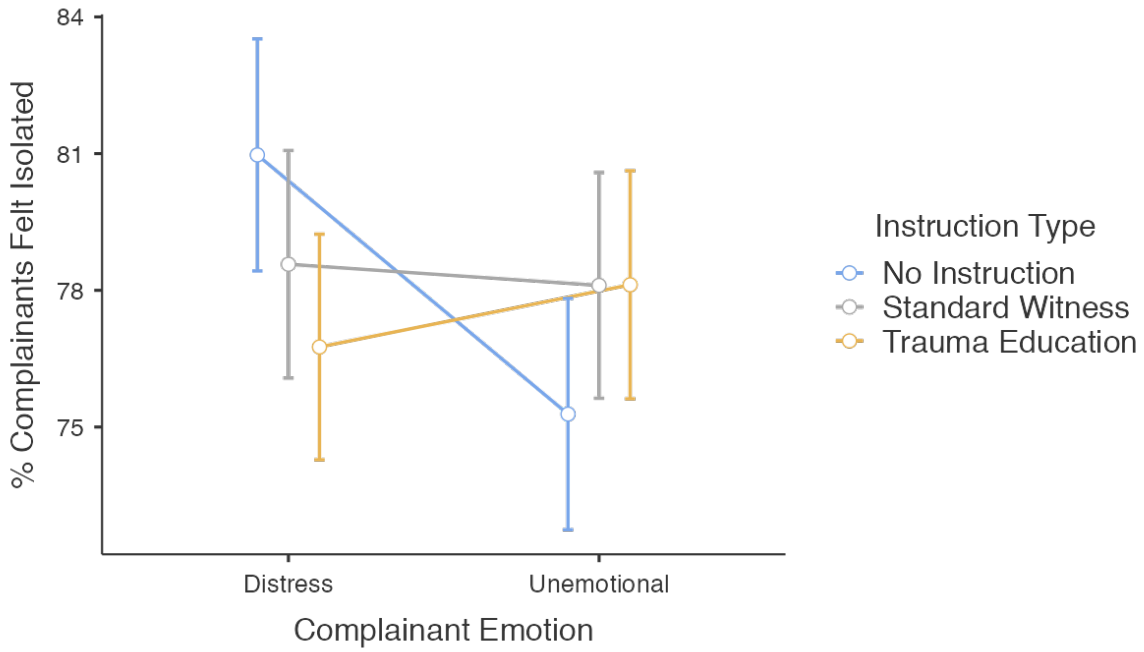
There was a significant main effect of complainant emotion and a significant main effect of instruction type on perceptions of the proportion of complainants who would avoid thoughts and feelings of the assault (see Table S10 for ANOVA and post-hoc comparison statistics). Participants who read about a distressed complainant thought a greater proportion of complainants would avoid thoughts and feelings about the assault (M = 85.06, SD = 15.01) than participants who read about an unemotional complainant (M = 82.42, SD = 16.71). Participants who received the educational instruction thoughts fewer complainants would avoid thoughts and feelings about the assault (M = 82.02, SD = 16.39) compared to those who received the standard instruction (M = 84.89, SD = 14.93).

For perceptions of the proportion of complainants who would experience negative feelings, there was a significant main effect of complainant emotion and a significant main effect of instruction type. Participants who read about a distressed complainant thought more complainants would have negative feelings after an assault (M = 84.46, SD = 16.16) than participants who saw an unemotional complainant (M = 82.56, SD = 17.04). Participants who received the educational instruction thought fewer complainants would have negative feelings after the assault (M = 79.84, SD = 17.99) compared to those who received either the standard instruction (M = 85.15, SD = 15.31) or no instruction (M = 85.64, SD = 15.86).

For perceptions of the proportion of complainants who would feel isolated from others, there was a significant interaction of complainant emotion and instruction type only (see Figure S5 and Table S11 for condition descriptive statistics), participants who received no instruction believe a greater proportion of complainants would feel isolated when the complainant is distressed compared to when the complainant was unemotional. There are no differences based on complainant distress in the instruction conditions.

Figure S5

Condition Estimates and 95% Confidence Intervals for Proportion of Complainants Feeling Isolated



For perceptions of the proportion of complainants who would have difficulty feeling anything, there was no significant main effects or interaction of instruction type and complainant emotion.

For trauma misperceptions, there was a significant main effect of complainant emotion and instruction type (see Table S10 for ANOVA and post-hoc comparison statistics). There was no significant interaction of complainant emotion and instruction type. Participants who saw a distressed complainant thought complainant emotion was more reliable to judge credibility (M = 2.80, SD = 1.34) than participants who saw an unemotional complainant (M = 2.62, SD = 1.38). Participants who received the standard instruction (M = 2.83, SD = 1.41) and no instruction (M = 2.82, SD = 1.29) incorrectly thought complainant emotion was more reliable to judge credibility than participants in the education instruction condition (M = 2.48, SD = 1.36).

Manipulation of Complainant Emotion

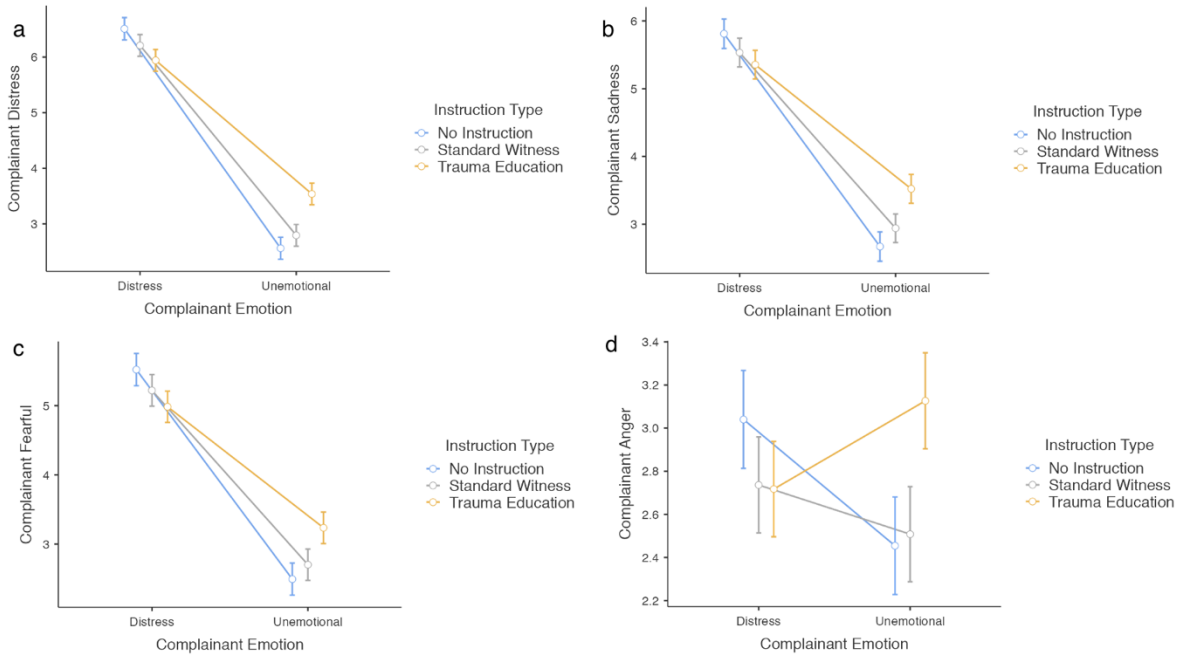
We asked participants to respond to continuous and dichotomous manipulation check measures. First, we asked participants to indicate to what extent the complainant was distressed. Then we asked participants to recall key aspects of the complainant's demeanour on dichotomous response scales.

We conducted a frequentist ANOVA on participants' perceptions of the complainant's distress. There was a significant main effect of complainant emotion. The main effect of complainant emotion indicates that the manipulation of complainant emotion was successful. The distressed complainant was rated as more distressed (M = 6.21, SD = 0.95) than the unemotional complainant (M = 2.97, SD = 1.72). There was also a significant main effect of instruction type as well as a significant interaction of complainant emotion and instruction type. Decomposing the interaction, in no instruction and standard instruction conditions, the distressed complainant was rated as strongly more distressed than the unemotional complainant. In the educational instruction condition, the distressed complainant was rated as more distressed than the unemotional complainant (see Figure S6, panel a and Table S12 for condition descriptive statistics).

For perceptions of the complainant as sad, there was a significant main effect of complainant emotion. The distressed complainant was seen as sadder (M = 5.56, SD = 1.25) than the unemotional complainant (M = 3.05, SD = 1.69). There was also a significant interaction of complainant emotion and instruction type (see Figure S6, panel b). Similar to perceptions of complainant distress, the difference between ratings of the distressed and unemotional complainant was smallest in the educational instruction condition. This suggested that the manipulation was largely effective.

Figure S6

Condition Estimates and 95% Confidence Intervals for Complainant Distress, Sadness, Fear and Anger



For exploratory purposes, we also examined to what extent participants saw the complainant as fearful or angry. For perceptions of the complainant as fearful, there was a significant main effect of complainant emotion and a significant interaction of complainant emotion and instruction type (see Figure S6, panel c). The distressed complainant was consistently rated as more fearful than the unemotional complainant. However, this relationship was strongest in the no instruction condition, followed by the standard instruction condition and smallest in the educational instruction condition.

For perceptions of the complainant as angry, there was a significant main effect of instruction type and interaction of instruction type and complainant emotion. In the no instruction condition, the distressed complainant was seen as angrier than the neutral complainant. In the standard instruction condition, there was no difference in ratings of anger between the distressed and neutral complainant. However, in the educational instruction

condition, the neutral complainant was rated as angrier than the distressed complainant (see Figure S6, panel d)

We conducted several chi-square analyses to assess whether participants correctly remembered the key features of the complainant's emotional demeanour. The first item asked participants to remember the complainant's demeanour. 97.6% of participants who read about a distressed complainant correctly indicated the complainant was distressed and 98.0% of participants who read about an unemotional complainant correctly indicated that the complainant was calm, $\chi^2(1) = 990.14$, $p < .001$. The second item whether the complainant was crying. Similarly, 86.9% of participants who read about a distressed complainant correctly indicated the complainant was crying and 97.8% of participants who read about an unemotional correctly indicated that the complainant was not crying, $\chi^2(1) = 787.69$, $p < .001$. The third item asked about the complainant's voice. 98.0% of participants who read about a distressed complainant correctly indicated that the complainant's voice trembled and 98.0% of participants who read about an unemotional complainant correctly indicated that the complainant had a steady voice $\chi^2(1) = 997.79$, $p < .001$. This suggested that the complainant emotion manipulation was effective.

Attention Checks

We asked participants to recall two facts about the case synopsis to check they had read and understood the case. We asked two questions about specific details of the evidence in the trial synopsis. Most participants correctly indicated that the defendant testified that the complainant said yes to having sex (99.17%) and that the defendant testified the complainant did not push him away (96.96%). We also asked participants whether the complainant and defendant's testimonies were mostly consistent. Most participants indicated that the complainant and defendant's testimony was not consistent (72.99%). This indicated that

participants were able to recall specific facts of the case, suggesting that participants attended to the trial synopsis.

Table S12

ANOVA and Contrast Statistics for Unregistered Design on Complainant Emotion and Case Evidence Ratings

	Complainant Emotion				Case Ratings	Interest in Case	Effort to Evaluate Case	Objectivity due to Judge's Direction
	Distress	Sadness	Fear	Anger				
Main effect of instruction type	$F(2, 1079) = 8.00$, $p < .001$ $d = 2.34$, 95% CI [2.19, 2.50]	$F(2, 1079) = 2.26$, $p = .105$ $d = 1.69$, 95% CI [1.55, 1.83]	$F(2, 1079) = 0.87$, $p = .421$ $d = 1.53$, 95% CI [1.39, 1.66]	$F(2, 1078) = 3.55$, $p = .029$ $d = 0.08$, 95% CI [-0.04, 0.20]	$F(1, 1079) = 25.10$, $p < .001$	$F(1, 1079) = 5.70$, $p = .017$	$F(1, 1079) = 1.40$, $p = .236$	$F(1, 1079) = 0.05$, $p = .823$
Main effect of complainant emotion	$F(1, 1079) = 589.67$, $p < .001$	$F(1, 1079) = 805.96$, $p < .001$	$F(1, 1079) = 652.06$, $p < .001$	$F(1, 1078) = 2.11$, $p = .147$	$F(2, 1079) = 5.82$, $p = .003$	$F(2, 1079) = 0.56$, $p = .571$	$F(2, 1079) = 4.98$, $p = .007$	$F(2, 1079) = 13.96$, $p < .001$
No instruction vs Instruction	$t(1079) = 0.97$, $p = .330$	$t(1079) = 1.02$, $p = .308$	$t(1079) = 0.28$, $p = .780$	$t(1078) = 0.25$, $p = .803$	$t(1079) = -2.38$, $p = .017$	$t(1079) = -1.02$, $p = .308$	$t(1079) = 3.13$, $p = .002$	$t(1079) = 4.47$, $p < .001$
Standard vs Educational instruction	$t(1079) = 2.39$, $p = .017$	$t(1079) = 1.87$, $p = .062$	$t(1079) = 1.29$, $p = .199$	$t(1078) = 2.65$, $p = .008$	$t(1079) = 2.45$, $p = .015$	$t(1079) = -0.31$, $p = .760$	$t(1079) = 0.39$, $p = .694$	$t(1079) = -2.81$, $p = .005$
Complainant emotion x Instruction type	$F(2, 1079) = 30.59$, $p < .001$	$F(2, 1079) = 18.10$, $p < .001$	$F(2, 1079) = 15.18$, $p < .001$	$F(2, 1078) = 9.75$, $p < .001$	$F(2, 1079) = 0.97$, $p = .379$	$F(2, 1079) = 0.27$, $p = .723$	$F(2, 1079) = 0.47$, $p = .624$	$F(2, 1079) = 1.86$, $p = .156$

Table S13

Condition Descriptive Statistics for Perceptions of Complainant's Emotions in Unregistered Design

			Outcomes			
			Complainant Distress	Complainant Sadness	Complainant Fearful	Complainant Anger
Instruction type	Complainant emotion	*n*	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
No instruction	Distress	175	6.51 (0.10)	5.81 (0.11)	5.52 (0.12)	3.04 (0.11)
	Unemotional	176	2.56 (0.10)	2.67 (0.11)	2.49 (0.12)	2.45 (0.11)
Standard	Distress	182	6.21 (0.10)	5.53 (0.11)	5.21 (0.12)	2.73 (0.11)
	Unemotional	185	2.79 (0.10)	2.94 (0.11)	2.70 (0.12)	2.50 (0.11)
Trauma Education	Distress	185	6.51 (0.10)	5.36 (0.11)	4.98 (0.12)	2.71 (0.11)
	Unemotional	182	2.56 (0.10)	3.52 (0.11)	3.24 (0.12)	3.12 (0.11)

Manipulation Check for Standard Witness Instruction

To check that participants' correctly understood the standard witness instruction, we asked them to what extent did you feel you had to be objective because of the judge's instructions. A frequentist ANOVA indicated there was only a significant main effect of judicial instruction type on participants' perceptions of how objective the judge told them to be. Participants in both instruction conditions felt they had to be more objective because of the judge's instructions (M = 5.49, SD = 1.35) than participants in the no instruction condition (M = 5.10, SD = 1.37). Participants in the standard instruction condition also felt they had to be more objective (M = 5.63, SD = 1.26) than participants in the educational instruction condition (M = 5.35, SD = 1.42).

Pre-conditions for Systematic Information Processing.

We also checked to make sure that the pre-conditions for bias-correcting and biased systematic processing to take place were present. To do this, we asked participants to evaluate their motivation to engage information processing (via two questions) and to evaluate the strength of evidence in the case.

There was only a significant main effect of complainant emotion for how interested participants were in the case. Participants who viewed a distressed complainant reported being slightly more interested in the case (M = 5.64, SD = 1.18) than participants who read about an unemotional complainant (M = 5.46, SD = 1.34).

There was only a significant main effect of instruction type on how much effort participants reported putting into evaluating the evidence. Analyses of planned contrasts indicated that participants in the no instruction conditions reported using less effort to evaluate the case evidence (M = 5.74, SD = 1.01) compared to participants in either instruction condition (M = 5.94, SD = 1.02).

There was a significant main effect of complainant emotion on participants' evaluation of the case evidence. Participants who read about a distressed complainant rated the case evidence as more compelling (M = 4.06, SD = 1.24) compared to participants who read about an unemotional complainant (M = 3.69, SD = 1.26). There was also a significant main effect of instruction type. Participants who received no instruction from the judge about the complainant rated the case evidence as more compelling (M = 4.00, SD = 1.32) compared to participants who received either instruction about the complainant from the judge (M = 3.81, SD = 1.23). Participants who received the educational instruction rated the case evidence as more compelling (M = 3.93, SD = 1.25) than participants who received the standard instruction (M = 3.70, SD = 1.20). There was no significant interaction of complainant emotion and instruction type.

Demand Probe

Only one third (31.2% or 339) of participants responded to the demand probe. No participants accurately guessed at the experimental hypotheses and there was no indication of demand characteristics in the data.