Terms and Conditions Apply:

The Effect of Probation Length and Obligation Disclosure on True and False Guilty Pleas

Supplemental Results

Subsample Analyses

Our first set of sensitivity analyses replicated the main analyses (Tables 4 and 5 in the main manuscript), but used only participant-defendants who passed all six manipulation and attention checks (n = 704). Similar to the main analysis, we controlled for gender in all regression models presented in this Supplemental File due to non-balance among the experiment conditions. We present the results in Tables A1 and A2. All the main effects were similar in magnitude and significance. While the interaction effects on d(WTAP) were similar to those we found in the main analysis, we noted some differences in the interaction effects for the plea acceptance variable. Specifically, the interaction effects between guilty and the other two variables were negative, which contradicted the positive main effects in the main analysis. In this subsample, the effect for both the 5-year sentence and detailed disclosure were *larger* for guilty participant-defendants (since both main effects were negative). However, despite the opposite signs, both interaction effects were small in magnitude and not significant. Overall, our findings were sufficiently robust to replicate across these participant subsets.

Main and Interaction Effects Based on Predicted Plea Acceptance Probabilities

We used logistic regression to estimate all the effects on plea acceptance. Although our presentation and interpretation of the findings have been consistent with most studies so far, some recent methodological research has pointed out potential pitfalls of the approach (Long & Mustillo, 2021; Mustillo et al., 2018). Mize (2019) argued that although it is necessary to include

the product terms in logistic models to estimate the interaction effects (e.g., guilty*5-year or guilty*detailed disclosure), the corresponding coefficients—both the size and significance—should be ignored because they "[do] not necessarily provide accurate information about the significance, magnitude, or even the direction of the underlying interaction effect on the predictions" (p. 112). We followed the recommendation of Mize (2019), and re-estimated the main and interaction effects using predicted probabilities of plea acceptance from the model presented in Table 4. We conducted this analysis using the -margins- command in Stata 17.0.[1]

We present the results in Table A3. Each panel focuses on one pair of variables. We calculated the average marginal effect (AME) sizes as the difference between each pair of experimental conditions (e.g., the AME of the 5-year sentence for guilty was -0.157 = 0.525-0.682, which means that for guilty participant-defendants, facing a 5-year sentence led to a 15.7-percentage-point decrease in probability of pleading guilty). Finally, we calculated the interaction effect as the difference between the two differences (-0.022 = -0.157-(-0.136, last decimal not accurate due to rounding), suggesting that the effect of the 5-year sentence was 2.2 percentage points larger (rather than smaller, because the main effects were negative) for guilty participant-defendants.

The reanalysis largely confirmed our findings on the interaction effects. Neither of the two interaction effects involving detailed disclosure of probation obligations was significant ($p > .12$ for both), and both were very small (1.3 and 0.3 percentage points, respectively). Contrary to our main analysis, this analysis did find a significant interaction between guilt status and sentence length. However, the size of the interaction effect was small as well (2.2 percentage

[1] We did not conduct this analysis for the WTAP outcomes because those models used OLS specification and were therefore not subject to the issues identified by methodological studies.

points, $p = .01$). Overall, we found no major disagreement on the interaction effects among the specifications.

Logistic Regression Models Showing Odds Ratios

The last part of the Supplemental File presents the logistic regression findings, but uses odds ratios instead of coefficients or marginal effects. These analyses do not affect the findings of the study.

Table A1. Logistic Regression Models Explaining Plea Acceptance, Participants Passing All Checks Only

	(1) Main Effects Only	(2) With Interactions
Guilty	1.72***	1.77***
	(0.17)	(0.29)
	[0.42]	[0.43]
Probation Length 5 Years	-0.53**	-0.21
	(0.17)	(0.30)
	[-0.13]	[-0.05]
Detailed Disclosure	-0.42*	-0.15
	(0.17)	(0.31)
	[-0.10]	[-0.04]
Guilty * 5 Years		-0.10
		(0.35)
		[-0.02]
Guilty * Detailed Disclosure		-0.03
		(0.35)
		[-0.01]
5 Years * Detailed Disclosure		-0.51
		(0.34)
		[-0.13]
Observations	704	704

Standard errors in parentheses, marginal effects in brackets,
*** $p < .001$, ** $p < .01$, * $p < .05$

Table A2. OLS Regression Models Explaining WTAP Outcomes, Participants Passing All Checks Only

	(1) d(WTAP) Main Only	(2) d(WTAP) w Interactions	(3) WTAP1 Main Only	(4) WTAP2 Main Only
Guilty	1.51	5.55	27.84***	29.35***
	(1.73)	(2.96)	(2.29)	(2.50)
Probation Length 5 Years	-5.38**	-1.73	-2.06	-7.43**
	(1.72)	(2.98)	(2.28)	(2.48)
Detailed Disclosure	-6.24***	0.27	-1.29	-7.53**
	(1.72)	(3.15)	(2.28)	(2.49)
Guilty * 5 Years		-1.83		
		(3.46)		
Guilty * Detailed Disclosure		-6.76		
		(3.48)		
5 Years * Detailed Disclosure		-5.59		
		(3.44)		
Observations	704	704	704	704

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table A3. Average Marginal Effects and Interaction Effects Using Predicted Probabilities

	Estimate	*SE*	*p*
Guilt Status and Sentence Length			
Predicted Probabilities			
P(plea), innocent and 1-year	0.368	0.036	<.001
P(plea), innocent and 5-year	0.233	0.029	<.001
P(plea), guilty and 1-year	0.682	0.035	<.001
P(plea), guilty and 5-year	0.525	0.041	<.001
AME of 5-year Sentence			
Innocent	-0.136	0.030	<.001
Guilty	-0.157	0.034	<.001
Second Difference			
Guilty & 5-year interaction	-0.022	0.008	0.010
Guilt Status and Disclosure			
Predicted Probabilities			
P(plea), innocent and general discl.	0.330	0.032	<.001
P(plea), innocent and detailed discl.	0.271	0.032	<.001
P(plea), guilty and general discl.	0.640	0.038	<.001
P(plea), guilty and detailed discl.	0.568	0.038	<.001
AME of Detailed Disclosure			
Innocent	-0.059	0.029	0.046
Guilty	-0.071	0.034	0.033
Second Difference			
Guilty & detailed discl. interaction	-0.013	0.008	0.120
Sentence Length and Disclosure			
Predicted Probabilities			
P(plea), 1-year and general discl.	0.542	0.033	<.001
P(plea), 1-year and detailed discl.	0.499	0.034	<.001
P(plea), 5-year and general discl.	0.402	0.035	<.001
P(plea), 5-year and detailed discl.	0.362	0.034	<.001
AME of Detailed Disclosure			
1-year	-0.043	0.032	0.179
5-year	-0.040	0.030	0.187
Second Difference			
5-year & detailed discl. interaction	0.003	0.007	0.676

Table A4. Logistic Regression Analyses on Plea Acceptance, Odds Ratios (ORs) Presented

	OR	*z*	*p*	95% CI
Full Sample, No Interaction (Table 4, *n* = 906)				
Guilty	4.04	9.62	<.001	[3.04, 5.38]
Probation Length 5 Years	0.51	-4.59	<.001	[0.39, 0.68]
Detailed Disclosure	0.75	-2.01	0.045	[0.56, 0.99]
Full Sample, With Interaction (Table 4, *n* = 906)				
Guilty	3.64	5.28	<.001	[2.25, 5.88]
Probation Length 5 Years	0.53	-2.56	0.011	[0.33, 0.86]
Detailed Disclosure	0.79	-0.95	0.343	[0.49, 1.28]
Guilty * 5 Years	1.13	0.43	0.669	[0.64, 2.01]
Guilty * Detailed Disclosure	1.09	0.31	0.757	[0.62, 1.94]
5 Years * Detailed Disclosure	0.81	-0.74	0.458	[0.46, 1.42]
All-pass Sample, No Interactions (Table A1, *n* = 704)				
Guilty	5.59	9.90	<.001	[3.98, 7.86]
Probation Length 5 Years	0.59	-3.17	0.002	[0.42, 0.82]
Detailed Disclosure	0.66	-2.48	0.013	[0.47, 0.92]
All-pass Sample, With Interactions (Table A1, *n* = 704)				
Guilty	5.90	6.09	<.001	[3.33, 10.44]
Probation Length 5 Years	0.81	-0.72	0.473	[0.45, 1.45]
Detailed Disclosure	0.86	-0.47	0.637	[0.47, 1.60]
Guilty * 5 Years	0.90	-0.29	0.773	[0.46, 1.79]
Guilty * Detailed Disclosure	0.97	-0.10	0.921	[0.49, 1.92]
5 Years * Detailed Disclosure	0.60	-1.51	0.130	[0.31, 1.16]

References

Long, J. S., & Mustillo, S. A. (2021). Using predictions and marginal effects to compare groups in regression models for binary outcomes. *Sociological Methods & Research*, *50*(3), 1284-1320. https://doi.org/10.1177/0049124118799374

Mize, T. (2019). Best practices for estimating, interpreting, and presenting nonlinear interaction effects. *Sociological Science*, *6*, 81-117. https://doi.org/10.15195/v6.a4

Mustillo, S. A., Lizardo, O. A., & McVeigh, R. M. (2018). Editors' comment: A few guidelines for quantitative submissions. *American Sociological Review*, *83*(6), 1281-1283. https://doi.org/10.1177/0003122418806282