Additional Methodological Information

Sample Selection

In each country, we surveyed university students to keep cross-national samples as similar as possible in terms of age and education. We gathered the data in large cities to control for urbanization. The USA sample was drawn from various cities in states on the east and west coast: California (25 respondents), Connecticut (1), Delaware (1), Florida (13), Georgia (10), Illinois (10), Maryland (6), Massachusetts (16), New Jersey (7), New York (22), North Carolina (9), Oregon (3), Rhode Island (1), Virginia (14), and Washington (3).

Translation of the Questionnaire

The original English version of the questionnaire was translated into each country's official language by each country's coauthor or another PhD-level researcher and then back-translated into English by a second translator to ensure compatibility and equivalence in meaning (Brislin, 1986). The back-translated English questionnaire was then compared to the original English questionnaire by the first author. Next the translators discussed the differences between the two translations until they reached consensus. When no consensus could be reached, a third translator was involved. Lastly, the final translation was checked for preservation of meaning and cultural appropriateness by each country's coauthor. In countries with multiple official languages we used the medium of instruction in public schools, for instance, English in Singapore and Urdu in Pakistan.

Power Distance

Apart from collectivism and tightness, we measured an additional cultural dimension, power distance, to explore whether it would influence people's reactions to norm violators. The power distance scale consisted of three 7-point bipolar items that were derived from the GLOBE study (House et al., 2004). The items read as follows: "In this country followers: *1 = obey their*

leader without question to *7 = question their leader when in disagreement*" (reverse-coded);

"How often are followers afraid to express disagreement with their leaders? *1 = very seldom* to *7 = very frequently*"; and "In this country power is: *1 = concentrated at the top* to *7 = shared throughout the society*" (reverse-coded). Similar to the procedure we followed to test the psychometric qualities of collectivism and tightness, we assessed the internal consistency of the power distance scale within each country, the equivalence of factor structures across countries by means of Tucker's phi congruence coefficient, the intraclass correlation coefficient, and the within-country agreement. The results of these analyses showed a mean internal consistency of $\alpha=0.46$ (*SD*=0.17), a congruence coefficient of Tucker's $\varphi=.93$ (*SD*=.12), an intraclass correlation coefficient of ICC(1)=.18, and a mean within-country agreement of $r_{wg(J)}=.65$ (*SD*=0.16). Given the insufficient internal consistency of the scale (i.e., $\alpha<.60$) and the low within-country agreement (i.e., $r_{wg(J)}<.70$), we concluded that the psychometric qualities of power distance were poor and we therefore decided to not explore its effects on people's reactions to norm violators.

Additional Results

Metric Equivalence

Given the experimental design of the study and our interest in the *relative* magnitude of differences between conditions across countries (rather than differences in *absolute* scores across countries), we provided evidence for configural equivalence in the main text to support the assumption that the psychological construct underlying each scale is the same across countries. However, to provide a more complete view of the psychometric qualities of our scales, we also tested for metric equivalence to examine whether respondents across groups attributed the same meaning to the latent constructs under study. We tested metric equivalence by running a

confirmatory factor analysis (CFA) in Mplus for each of the scales we included in the study. For each scale we specified a CFA model that constrained the factor loadings to be equal across countries (van de Schoot et al., 2012). The model fit was evaluated by means of two indices, the comparative fit index (CFI) and the standardized root-mean-square residual (SRMR). The CFI compares the fit of the model under consideration with the fit of a baseline model (i.e., a model with no constrained parameters) and the SRMR evaluates the fit of a model in reference to perfect fit (Cheung & Rensvold, 2002). The values of the CFI and the SRMR indices range from 0 to 1, and they are independent from sample size and model complexity (Meade, Johnson, & Braddy, 2008). Fit is considered adequate if the CFI value is above .90 and the SRMR value is below .08 (van de Schoot et al., 2012). However, more liberal evaluation criteria should be considered when a study includes a large number of comparison groups with unequal sample sizes within each group (Meade, Johnson, & Braddy, 2008; Rutkowski & Svetina, 2014). Given that our study included a large number of comparison countries with unequal sample sizes within each country, the results of the CFAs (see Table S1) indicated sufficient to good fit for all scales providing evidence for metric equivalence.

Three-way Interaction Among Collectivism, Tightness, and Norm Violation

To explore whether the pattern of results remains the same when the interaction among collectivism, tightness, and norm violation is included in the model, we carried out multilevel SEM testing the effects of collectivism, tightness, and norm violation (main effects, two-way interaction effects, and three-way interaction effect) on power perception, moral outrage, and leader support. The analyses largely overlapped with the results reported in the main text where we tested the model without the three-way interaction effect, except for a now non-significant interaction between actor's behavior and collectivism on power perception (see Table S2 in

comparison to Table 4).

Interestingly, there was a significant three-way interaction among collectivism, tightness, and norm violation on power perception. Probing the interaction revealed that in tighter and more individualistic cultures, norm violators were seen as more powerful than norm abiders (-$1SD_{collectivism}$: b=0.37 (SE=0.16), p=0.012, 95% CI [0.06, 0.69]; $-2SDs_{collectivism}$: b=1.03 (SE=0.36), p<0.001, 95% CI [0.31, 1.74]), whereas in tighter and more collectivistic cultures norm violators were seen as less powerful than norm abiders ($1SD_{collectivism}$: b=-0.21 (SE=0.11), p=0.032, 95% CI [-0.43, 0.01]; $2SDs_{collectivism}$: b=-0.81 (SE=0.32), p<0.001, 95% CI [-1.44, -0.18]). In looser and more individualistic cultures, and in looser and more collectivistic cultures norm violators and norm abiders were not perceived differentially.

This interaction pattern suggests that norm violators are perceived as more powerful in individualistic cultures and as less powerful in collectivistic cultures (as predicted) only or especially when the culture endorses tightness too. It is conceivable that norm-violating behavior is more salient in tight cultures where it occurs less frequently, while it is less salient in loose cultures where it happens more frequently. This would imply that norm violating behavior leads to higher or lower power perceptions especially when it becomes salient against a background of cultural tightness.

Additional References (not included in the main text)

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255. doi:10.1207/S15328007SEM0902_5

Table S1

Fit Indices for the Models Testing Metric Invariance

Scale	CFI	SRMR
Manipulation Check		
Norm Violation Perception	.94	.06
Outcome Variables		
Moral Outrage	.97	.08
Power Perception	.94	.11
Leader Support	.96	.07
Cultural Moderators		
Collectivism	.92	.11
Tightness	.87	.12

Note. CFI = comparative fit index. SRMR = standardized root-mean-square residual.

Table S2

Parameter Estimates for Multilevel Structural Equation Model Testing the Hypotheses that the Positive and Negative Pathways are Moderated by Collectivism, Tightness, and their Interaction

	Power Perception (positive pathway)			Moral Outrage (negative pathway)			Leader Support		
	b (*SD*)	CI	*p*	*b* (*SD*)	CI	*p*	*b* (*SD*)	CI	*P*
				Fixed Parameters					
Age	-0.01 (0.01)	[-0.02, 0.01]	.047	-0.01 (0.01)	[-0.02, 0.01]	.084	<-0.01 (0.01)	[-0.01, 0.01]	>.250
Gender	-0.01 (0.03)	[-0.06, 0.04]	>.250	0.03 (0.03)	[-0.03, 0.09]	.199	0.08 (0.03)	[0.03, 0.13]	.002
Education	0.04 (0.05)	[-0.07, 0.15]	.220	0.06 (0.07)	[-0.06, 0.19]	.160	0.05 (0.05)	[-0.06, 0.16]	.172
Socioeconomic status	0.03 (0.02)	[-0.01, 0.08]	.071	0.01 (0.03)	[-0.04, 0.07]	>.250	0.01 (0.02)	[-0.03, 0.06]	>.250
Religiosity	0.01 (0.01)	[-0.01, 0.04]	.140	-0.01 (0.01)	[-0.04, 0.02]	>.250	0.04 (0.01)	[0.01, 0.06]	.002
Administration means	-0.15 (0.09)	[-0.32, 0.02]	.039	-0.31 (0.10)	[-0.51, -0.11]	.001	-0.13 (0.08)	[-0.29, 0.04]	.072
Compensation	0.01 (0.03)	[-0.05, 0.06]	>.250	0.03 (0.03)	[-0.04, 0.09]	.196	0.01 (0.03)	[-0.04, 0.06]	>.250
Length of residence	-0.19 (0.12)	[-0.43, 0.05]	.057	-0.10 (0.14)	[-0.38, 0.18]	>.250	0.04 (0.12)	[-0.18, 0.27]	>.250
Intercept	<0.01 (0.03)	[-0.07, 0.07]	>.250	<0.01 (0.04)	[-0.08, 0.07]	>.250	3.32 (0.12)	[3.09, 3.55]	<.001
Actor's Behavior	0.06 (0.08)	[-0.09, 0.21]	.221	0.93 (0.08)	[0.78, 1.09]	<.001	-0.74 (0.07)	[-0.87, -0.61]	<.001
Collectivism	<0.01 (0.06)	[-0.12, 0.12]	>.250	<0.01 (0.07)	[-0.14, 0.14]	>.250	-0.19 (0.21)	[-0.61, 0.23]	.162
Tightness	<0.01 (0.15)	[-0.30, 0.31]	>.250	<-0.01 (0.18)	[-0.36, 0.33]	>.250	0.53 (0.51)	[-0.48, 1.57]	.136
Collectivism x Tightness	-0.01 (0.23)	[-0.47, 0.46]	>.250	0.01 (0.27)	[-0.52, 0.54]	>.250	0.11 (0.80)	[-1.49, 1.70]	>.250
Actor's Behavior x Collectivism	**-0.19** (0.14)	[-0.48, 0.09]	.079	**0.29** (0.15)	[0.01, 0.57]	.025	**0.11** (0.12)	[-0.14, 0.36]	.186
Actor's Behavior x Tightness	**0.08** (0.35)	[-0.62, 0.77]	>.250	**0.15** (0.35)	[-0.54, 0.84]	>.250	**-0.56** (0.30)	[-1.17, 0.04]	.033
Actor's Behavior x Collectivism x Tightness	**-1.11** (0.55)	[-2.20, -0.02]	.024	**0.12** (0.55)	[-1.00, 1.20]	>.250	**0.25** (0.47)	[-0.68, 1.17]	>.250
Power Perception							**0.30** (0.02)	[0.26, 0.34]	<.001
Moral Outrage							**-0.34** (0.02)	[-0.37, -0.31]	<.001
				Random Parameters					
$\sigma_{residual}$	1.28 (0.04)	[1.20, 1.35]	<.001	1.70 (0.05)	[1.61, 1.79]	<.001	1.18 (0.04)	[1.11, 1.25]	<.001
$\sigma_{intercept}$	<0.01 (0.01)	[<0.01, 0.02]	<.001	0.01 (0.01)	[<0.01, 0.03]	<.001	0.15 (0.12)	[0.06, 0.47]	<.001
σ_{slope}	0.07 (0.04)	[0.03, 0.18]	<.001	0.07 (0.04)	[0.03, 0.18]	<.001	0.05 (0.03)	[0.02, 0.13]	<.001
				Explained Variance					
$R^2_{fixed\ slope}$	.07 (0.01)	[.05, .09]	<.001	.34 (0.01)	[.31, .36]	<.001	.54 (0.01)	[.51, .56]	<.001
$R^2_{random\ intercept}$	.22 (0.20)	[.02, .75]	<.001	.24 (0.21)	[.02, .79]	<.001	.14 (0.12)	[.01, .45]	<.001
$R^2_{random\ slope}$	.28 (0.12)	[.07, .52]	<.001	.25 (0.13)	[.05, .54]	<.001	.23 (0.14)	[.04, .54]	<.001

Note. Actor's behavior was coded as -1 for the norm adherence condition and 1 for the norm violation condition. CI stands for the Bayesian credibility interval. Parameters in bold highlight the focal effects that are tested in the model.