

SUPPLEMENTAL ONLINE MATERIALS

THE VIGILANTE IDENTITY IN ORGANIZATIONS

Supplemental Analysis 1: Initial version of the VIS

Below we report an earlier version of the VIS we submitted as part of the review process. In this version, a total of 10 items were developed to measure the VIS, where 8 items tapped onto the punishment dimension and 2 items onto the monitoring dimension.

However, as pointed out by the helpful reviewers (along with other psychometric issues), the content validity of this measure was questionable, which prompted us to revisit our scale development effort (as reported in the current manuscript). For the purpose of transparency, we present our initial scale development and validation effort below.

We use this initial scale only for Study 1 in our main manuscript. From Study 2 onwards, we revised the VIS and used the new VIS scale in all subsequent studies (Studies 3 – 4).

Below we report (in verbatim) *the scale development effort of the initial VIS reported in the original submission OBHDP.*

Vigilante Identity Scale Development and Convergent Validity

The goal of this study was to develop and validate a psychometrically valid scale to assess vigilante identity. We constructed a short, easily administered 10-item vigilante identity scale (VIS) and examined its psychometric properties through an exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA). We also sought to establish convergent validity, where we tested the correlations between the VIS and published scales that may influence people's willingness to administer unauthorized punishment to norm violators, but that are not based on the concept of an identity (i.e., social vigilantism and punishment orientation).

Convergent Validity Tests

We selected *social vigilantism* (Saucier & Webster, 2010) and *punishment orientation* (Yamamoto & Maeder, 2019) for tests of convergent validity for the following reasons. First, social vigilantism shares some features with our conceptualization of vigilante identity. While they both reflect one's desire to enact corrective action, the two differ in their manifested behaviors and scope. Social vigilantes are people with tendencies to "assert their 'superior' beliefs onto others to correct others' more 'ignorant' opinions for the 'greater good'" (p. 19; Saucier & Webster, 2010), without referencing behavior or placing importance on identity-driven *punishment* tendencies. In contrast, these punishment tendencies are central to those who adopt a vigilante identity (DeCelles & Aquino, 2020). Thus, we expected the social vigilantism scale (SVS) to be a suitable candidate for convergent validity and we expected that the two scales will be correlated positively.

Second, we selected *punishment orientation* (Yamamoto & Maeder, 2019) as another candidate for examining convergent validity. Punishment orientation contains four dimensions: 1) prohibitive utilitarianism (punishment focuses on prospective benefits), 2) prohibitive retributivism (avoiding punishing innocent people), 3) permissive utilitarianism (punish to ensure public safety and deter crime), and 4) permissive retributive (punish to get back at the offender). Like SVS, however, the punishment orientation scale does not address behavioral demands inherent in vigilantism (i.e., monitoring and punishing others). In contrast, the punishment orientation scale seeks to assess how lay people think about punishment, with the goal of capturing individual-level differences in perceptions of punishment ethics (Yamamoto &

Maeder, 2019). We expected vigilante identity to correlate positively with all four dimensions of punishment orientation, as both scales contain punitive elements and perceptions of those who violate norms.

Nomological Network

We also explored the nomological network (Cronbach & Meehl, 1955) of the vigilante identity by examining whether it is related to certain personality traits and cognitive tendencies that are consistent with the definition. Specifically, individuals who adopt a vigilante identity may tend to be attentive to signs of deviance in one's environment and be chronically predisposed to being *hypervigilant*, thus making them more capable at detecting transgressions. Therefore, dispositions that are associated with chronic hypervigilance (Chan & McAllister, 2013), such as justice sensitivity (Schmitt, Gollwitzer, Maes, & Arbach, 2005), moral attentiveness (Reynolds, 2008), and perception of threats to social order, should all covary positively with vigilante identity (DeCelles & Aquino, 2020).

Another reason theorized for why an individual who adopts a vigilante identity also tends to adjudicate the transgression and concludes that a wrongdoing has occurred, is because the person has a higher *moral certainty* (i.e., belief about one's moral superiority and capability of determining right and wrong; DeCelles & Aquino, 2020). We posit communal narcissism (Gebauer, Sedikides, Verplanken, & Maio, 2012) as an indicator of moral certainty. Beyond the heightened self-oriented motives of grandiosity, esteem, and entitlement typical of traditional narcissism (Campbell & Foster, 2007; Emmons, 1987; Morf et al., 2011; Raskin & Terry, 1988), communal narcissists also display an exaggerated belief in one's superior prosociality and morality. Those beliefs can facilitate the execution of various moral rationalization or disengagement strategies (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996). This includes imposing unauthorized punishment on others, which might be justified through a belief that one is contributing to the public good. Thus, we predicted a positive relationship between vigilante identity and communal narcissism.

Initial Item Generation

Following the best practice recommendation by Hinkin (1995), we deductively generated a 10-item VIS that reflects the key behavioral demands associated with being a vigilante, following the conceptualization of DeCelles & Aquino (2020). These behavioral demands include the tendencies to monitor the environment for signs of deviance and the willingness to rectify a social system's perceived failure to control it by punishing norm violators, even without having any formal authority to do so. The items we generated are presented in Table 2. Participants were guided by the stem question: "*I am the kind of person who...*" (1 = *completely unlike me*, 7 = *completely like me*). Based on the abovementioned theoretical assumptions that guided item construction, we expected the items to load on a single factor.

Methods

Participants

We used Samples: A) for EFA, B) for CFA and testing the relationships between vigilante identity, social vigilantism, and various dimensions of punishment orientation, C) for nomological network, and D) for test-retest reliability. Sample characteristics (*N*, age, and gender) for all our studies are detailed in Appendix B (Samples Table).

Procedure and Measure

Sample A (CloudResearch). Participants completed the VIS along with other measures unrelated to the present study (included for exploratory purposes). We also assessed a set of demographic characteristics to explore their potential relationships with VIS. These included

gender, age, number of years of work experience, political ideology (1 = *extremely liberal*; 9 = *extremely conservative*), and religiosity (1 = *not at all religious*; 9 = *deeply religious*).

Sample B (Prolific Academic). Sample B was part of a larger longitudinal project, and these measures were included as part of the multi-wave survey. Social vigilantism, punishment orientation, and demographic characteristics were measured at Time 1, whereas VIS was measured 3 weeks later at Time 2.

Sample C (Student Participants). Participants completed a one-time survey that contained VIS and perceived threat to social order, justice sensitivity (Schmitt et al., 2005), moral attentiveness (Reynolds, 2008), and communal narcissism (Gebauer et al., 2012).

Sample D (Student Participants). Participants were recruited across two time points. At Time 1, the VIS were embedded in a large, department-wide collaborative survey. At the follow-up session (Time 2, approximately 2 months later), participants were recruited for an unrelated experiment, and we embedded the measure of VIS at the start of the experiment. We used data from participants who completed both sessions to assess test-retest reliability.

Measures

Unless otherwise noted, all assessments used a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*).

VIS (Samples A – D). Full items are shown in Table 2.

Social Vigilantism (Sample B). We used the 15-item measure of social vigilantism from Saucier and Webster (2010). A sample item was: “*I like to imagine myself in a position of authority at work so that I could make the important decisions around my workplace*”.

Punishment Orientation (Sample B). We used the 17-item Punishment Orientation Questionnaire (POQ) to assess the four dimensions of punishment orientation (Yamamoto & Maeder, 2019, p. 1288). Sample items included: “*Punishment should be about looking forward to improve society, not backward to address the criminal’s misdeeds*” (prohibitive utilitarianism) and “*It is better to let 10 guilty criminals go free than to punish one innocent person*” (prohibitive retribution).

Justice Sensitivity (Sample C). We used the Schmitt et al. (2005) 10-item unidimensional subscale¹ to assess sensitivity to injustice from a perspective of an observer witnessing an injustice. A sample item was: “*I am upset when someone is treated worse than others*” (1 = *not at all*, 5 = *a great deal*).

Communal Narcissism (Sample C). Participants responded to the 16-item scale (Gebauer et al., 2012).

Moral Attentiveness (Sample C). We assessed the extent to which participants showed attentiveness to ethical dilemmas in their daily lives using a 7-item perceptive dimension subscale (Reynolds, 2008, p. 1031).²

Perception of Threats to Social Order (Sample C). We assessed participants’ perceptions of threat to social order with four face-valid items based on previous literature (Creppell, 2011). We tailored our items to our student-based context. Items are: “*It seems to me that the way my fellow peers are behaving threatens the moral order of society*”, “*More of my*

¹ The second subscale assesses justice sensitivity from a perspective of an individual as a victim of injustice (“*It bothers me when others receive something that ought to be mine*”). Because we made no speculations that sensitivity to own injustice would relate to VIS, we left it out to reduce participant fatigue.

² We selected the 7-item perceptive scale due to its relevance with VIS. *Reflective* moral attentiveness, which was not measured in our study, indicates the extent to which an individual ruminates on matters of morality. A sample item is: “*I like to think about ethics*”.

fellow peers seem to be breaking the social rules that used to govern our relationships with one another in this society”, “I think that our society’s way of life is being threatened by different ideas and values”, and “My fellow peers are becoming incapable of determining right from wrong”.

Results and Discussion

Tables 3 – 5 show the descriptive statistics, reliabilities coefficients, and correlations among study variables for Samples A – C, respectively.

Exploratory Factor Structure (Sample A)

EFA results based on a maximum likelihood extraction method showed that a one-factor solution best fit the data, with the first latent factor accounting for 64% of the total variance. Standardized factor loadings ranged from .72 to .85. The items also had high internal consistency (Cronbach’s $\alpha = .95$). These results provided preliminary evidence for scale unidimensionality.

Confirmatory Factor Structure and Convergent Validity (Sample B)

We conducted a CFA on the 10 items of VIS. No correlated residuals were specified. Corroborating the unidimensional solution found in Sample A, results showed that a one-factor solution fit the data well, $\chi^2(35) = 131.93$, $p < .0001$, CFI = .94, RMSEA = .10, SMSR = .04, and did not fit worse than any 2- or 3-factor models. Favoring parsimony, this result provided further evidence for treating VIS as a unidimensional construct. Moreover, as shown in Table 4, VIS was positively and moderately correlated with SVS and three of the punishment orientation subscales (except for the prohibitive utilitarianism dimension), supporting our assertion that VIS and these measures share common variances due to overlapping punitive sentiments.

Nomological Network (Sample C)

Consistent with our expectation, VIS correlated positively and significantly with justice sensitivity ($r = .38$), communal narcissism ($r = .51$), moral attentiveness, ($r = .24$), and perception of threat to social order ($r = .17$), all $ps < .05$.

Test-retest Reliabilities (Sample D)

The correlations between VIS collected from Sample D across two time points was $r = .65$. This correlation is smaller than the median test-retest correlations of trait-like constructs, such as big-five personality traits (range $r = .83 - .93$), trait hostility ($r = .80 - .83$), trait assertiveness ($r = .82 - .91$), need for order ($r = .80 - .90$); (for a comprehensive review of reliability coefficients in personality assessments, see McCrae, Kurtz, Yamagata, & Terracciano, 2010). Our result provided evidence for the notion that VIS should be treated as an identity rather than a highly stable individual-level trait. That is, as a dispositional construct that has some trait-like stability, but that is also dynamic and susceptible to being activated or deactivated by situational factors (Aquino, Freeman, Reed, Felps, & Lim, 2009).

Tables for the initial scale development study**Table 2***Items of the Vigilante Identity Scale*

Item	
1	Pays attention to make sure that others do the right thing.
2	Stays alert for any signs of corruption and wrongdoing in society.
3	Makes sure other people act morally.
4	Rights wrongs whenever I see them.
5	Makes sure people are held accountable when they do something wrong.
6	Ensures that people who do something wrong get punished for it.
7	Protects people who are mistreated by others by punishing wrongdoers.
8	Enforces society's rules.
9	Makes sure that people who do wrong don't get away with it.
10	Enforces justice in society.

Table 3*Means, Standard Deviations, Reliabilities, and Correlations among Variables for Sample A*

	1	2	3	4	5	6
1 VIS	1	-.01	-.07	-.27***	.34***	.45***
2 Gender (1= Male)		1	-.15*	-.15*	.17*	.03
3 Age			1	.55***	-.10	-.06
4 Work experience (years)				1	-.24***	-.28***
5 Conservatism					1	.55***
6 Religiosity						1
<i>M</i>	4.81	-	33.50	8.79	5.33	4.71
<i>SD</i>	1.31	-	10.17	8.75	2.56	2.89
<i>α</i>	0.95	-	-	-	-	-
<i>N</i>	229	226	229	225	228	229

Table 4*Descriptive Statistics, Reliabilities, and Correlations Among Variables for Sample B*

	1	2	3	4	5	6	7	8	9	10
1 VIS	1	.48***	.05	-.20***	.37***	.41***	-.07	-.11	.27***	-.14*
2 Social Vigilantism		1	.18**	-.07	.53***	.47***	.10*	-.16**	.30***	-.26***
3 Prohibitive Utilitarianism			1	.37**	-.15**	-.27***	.11*	-.10*	-.21***	-.11*
4 Prohibitive Retribution				1	-.45***	-.42***	.02	-.13**	-.26***	.10*
5 Permissive Utilitarianism					1	.80***	.04	-.19**	.49***	-.19**
6 Permissive Retribution						1	.03	-.09	.44***	-.10
7 Gender (1 = Male)							1	-.17**	.05	-.11*
8 Age								1	-.10	.65**
9 Conservatism									1	-.11*
10 Work experience (years)										1
<i>M</i>	4.44	4.48	5.10	4.87	4.17	4.55	-	34.70	4.78	11.19
<i>SD</i>	1.26	0.98	1.20	1.19	1.47	1.41	-	9.85	2.54	10.91
α	0.93	0.89	0.86	0.59	0.85	0.87	-	-	-	-
<i>N</i>	255	396	397	397	397	397	396	397	396	392

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. Different *N*s due to missing data. Only 255 participants completed the T2 follow-up, when VIS was measured.

Table 5*Descriptive Statistics, Reliabilities, and Correlations among Variables for Sample C*

	1	2	3	4	5	6	7
1 VIS	1	.38***	.51***	.24***	.17**	.10*	.03
2 Justice Sensitivity		1	.35***	.30***	.16**	-.07	-.13*
3 Communal Narcissism			1	.32***	.33***	.10*	.06
4 Moral Attentiveness				1	.31***	.01	-.01
5 Perceived Threats to Social Order					1	.02	.01
6 Age						1	.54***
7 Gender (1 = Male)							1
<i>M</i>	4.91	5.22	4.24	3.87	3.75	22.45	-
<i>SD</i>	0.86	0.84	1.02	1.08	1.20	1.72	-
α	0.90	0.88	0.94	0.90	0.82	-	-
<i>N</i>	365	365	365	365	365	364	364

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

Sample table for the initial scale development effort

Sample	Two-wave Study	Participants	<i>N</i> recruited	<i>N</i> retained	Exclusion Criteria	% Men	<i>M</i> Age	<i>SD</i> Age
A	No	Cloud Research	358	229	Removed if failed attention check question.	65.9	33.50	10.17
B	Yes	Prolific Academic	T1: 483	T1: 397	Did not complete both time points or failed attention check.	T1: 65.0	T1: 34.70	T1: 9.85
			T2: 368	T2: 255		T2: 62.7	T2: 36.38	T2: 10.31
C	No	Singapore: Students	377	365	Removed if failed attention check question.	37.8	22.45	1.71
D	Yes	Canada: Students	T1: 447	-	-	T1: 45.3	-	-
			T2: 412	T2: 237	Did not complete both time points or failed attention check.	T2: 41.6	T2: 20.40	T2: 1.81

Supplemental Analysis 2: Extreme punishment condition in Study 3.

In Study 3, we also included an extreme punishment condition for exploratory reason, where the university fired the instructor, and the instructor was unable to find a new job because he was banned by the whole academia.

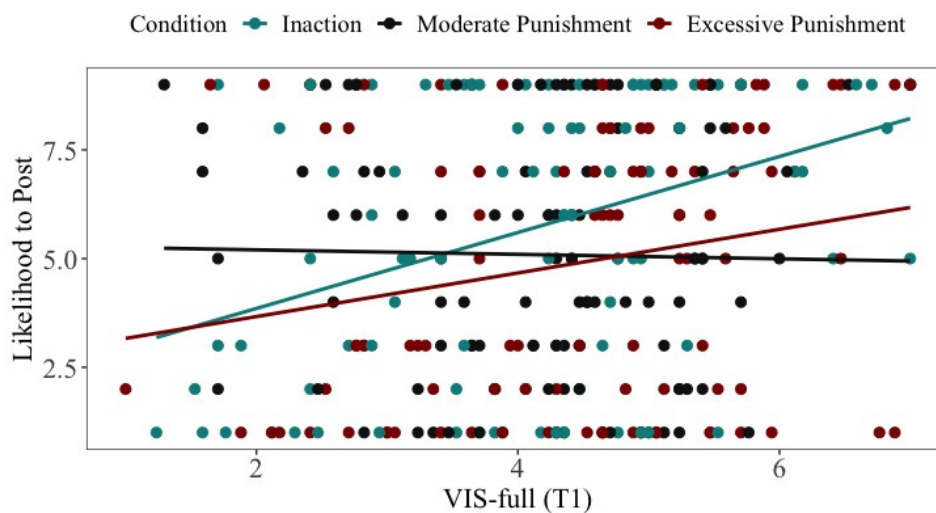
In terms of manipulation check, subjects in this extreme punishment condition ($M = 3.03$, $SD = 1.74$) perceived a lower level of justice failure by the university as compared to the justice-upheld condition ($M = 3.83$, $SD = 1.69$) and to those in the justice failure condition ($M = 6.12$, $SD = 1.18$), $t(208) = -11.34$, both $ps < .001$. All conditions differed significantly from each other, all $p < .001$.

When we used an Analysis of Covariance (ANCOVA) to examine the condition $\times$ VIS interaction on likelihood to post (continuous measure), we observed a significant interaction effect, $F(2, 312) = 3.65$, $p = .027$. Simple slope analysis (see Figure S1) revealed that the VIS had a positive effect in both the justice failure condition, $b = .92$, $SE = .22$, $t(102) = 4.18$, $p < .001$, and extreme punishment condition, $b = .51$, $SE = .22$, $t(106) = 2.27$, $p = .0252$, but had no effect in the justice-upheld condition, $b = -.01$, $SE = .25$, $t(104) = .00$, $p = .99$. Results remained unchanged when controlling for SVS, punishment orientation, and moral identity.

One potential explanation for the surprising findings in the excessive punishment condition is that the nature of this context might inadvertently signal a highly punitive norm in the environment. Such an external punitive norm then aligned nicely with the internal behavioral demands imposed by a vigilante identity. This fit between one's normative environment and internal behavioral demands may facilitate the manifestation of relevant behaviors (i.e., punishment), which then partly explain why VIS continued to have a positive effect on punishment even in the excessive punishment condition.

Figure S1

Effect of the VIS on Likelihood to Post for all Three Conditions.



Supplemental Analysis 3: Additional information in Study 4.

For exploratory purposes, we presented subjects in Study 4 with additional information after they completed the VIS. Specifically, subjects learned that the business owners were actually Jewish and that their decision to expel the ‘vigilante-like’ customer was motivated by prosocial motives to protect the confronted customers. We then assessed subjects’ moral self-appraisal (e.g., “*My ratings of the restaurant provided earlier would not change; 5 items in total*”) after reading this new information. No significant relationship between the VIS and moral self-appraisal was found, $r = -.063$, $p = .469$. The only significant predictor of moral self-appraisal was subject’s ratings of the restaurant

An ancillary question we explore in this study is whether the VIS is correlated with others' perceptions that the employee displays other kinds of behaviors that are generally viewed positively (i.e., organizational citizenship, helpfulness, in-role performance) or negatively (i.e., counterproductive work behavior) by management researchers. To do so, we asked the supervisors to rate the focal employee on several measures (in addition to the VIS as mentioned in the main manuscript) that we reported below.

OCB (Supervisor-rated). We used an 8-item OCB scale by Lee and Allen (2002). Supervisors were asked to report how often, in the past year, they had observed the focal employee do each of the noted behaviors in their workplace. Sample behaviors included: “*helped others who have been absent*” and “*willingly given your time to help others who have had work-related problems*” (1 = *never*, 5 = *regularly*).

In-Role Performance (Supervisor-rated). We used a 7-item sub-scale of in-role performance from Williams and Anderson (1991). Supervisor rated how often, in the past year, they had observed the focal employee do each of the noted behaviors in their workplace. Sample behaviors included: “adequately completes assigned duties” and “meets formal performance requirements of the job” (1 = never, 5 = regularly).

Results

Table 6
Descriptive Statistics and Correlations among Variables in Study 2e

[illegible]

We examined how employee VIS relates to other prosocial and antisocial behaviors in the organization, as reported by the supervisor. Replicating Study 1, employee VIS significantly and positively predicted supervisor-rated employee OCB, $b = .27$, $SE = .05$, $t(211) = 5.30$, $p < .001$, in-role performance, $b = .11$, $SE = .05$, $t(211) = 2.20$, $p = .029$, and perceived helpfulness in upholding organization's ethical structures, $b = .40$, $SE = .10$, $t(211) = 3.98$, $p = .029$. However, employee VIS was unrelated to supervisor-rated CWB, $b = .01$, $SE = .07$, $t(211) = 0.05$, $p = .957$. We also included demographic variables as control variables to test the robustness of these findings. We controlled for employee's age, gender, organization tenure, and the duration (in years) they had worked with the supervisor. All results remained significant and directionally consistent with our results reported above. We present these results in Table S2 below.

Together, these findings provide consistent evidence that high VIS scorers were perceived as being helpful, good organizational citizens, and better performers. These latter results suggest that people who internalize the vigilante identity are not perceived by their supervisors as antisocial or uncooperative, though we cannot say whether these perceptions would be corroborated by peers.

Table S2
Additional analysis and model summary for Study 2e

	Supervisor-rated OCB				Supervisor-rated CWB				Supervisor-rated In-role Performance				Supervisor-rated Helpfulness			
	Model 1		Model 2		Model 1		Model 2		Model 1		Model 2		Model 1		Model 2	
	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>	<i>b</i>	<i>SE</i>
Intercept	2.47**	.26	4.34**	.41	1.92**	.34	1.41*	.57	3.1	.1	3.39**	.43	3.13	.1	3.22***	.85
Employee self-rated VIS	.27**	.05	.26**	.05	.01	.07	.00	.07	.11*	.05	.13*	.05	.40**	.1	.39***	.10
Employee Gender (1= Male)			.01	.07			.02	.10			-.13	.08			-.30	.15
Employee Age			-.07**	.01			.02	.02			.00	.01			-.01	.02
Employee Organization Tenure			.05	.05			.06	.07			-.09	.05			.16	.11
Supervisor Tenure			.06	.06			-.12	.08			.03	.06			.01	.12
<i>R</i>²	.12		.27		0		.02		.02		.07		.07		.12	

Notes. VIS = Vigilante Identity Scale; OCB = Organization citizenship behavior; CWB = Counterproductive work behavior;

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