On the strength of ties that bind: Measuring the strength of norms in romantic relationships

John K. Sakaluk[1] ⓘ, Monica Biernat[2], Bonnie M. Le[3] ⓘ, Sarah Lundy[1], and Emily A. Impett[4] ⓘ

Abstract

In seven studies ($n_{\text{cross-sectional}} = 1,699$, $n_{\text{longitudinal}} = 118$), we developed a measure of relationship norm strength defined as qualities that make the rules and expectations in romantic couples more or less likely to be followed. In our six cross-sectional samples, the resulting Relationship Norm Strength Questionnaire (RNSQ) yielded consistent *norm tractability*, *norm agreement*, *anticipated punishment for deviance*, and *norm explicitness* factors, and estimated factors generally demonstrated evidence of convergent, discriminant, and criterion validity. Meta-analyzed effects across these samples—yielding more reliable and generalizable estimates—indicated that greater norm tractability and norm agreement were strongly linked to higher levels of relationship quality. Further supporting our model of relationship norm functioning, results from our 8-week longitudinal study of community members in relationships indicated that greater levels of norm tractability and agreement resulted in greater subsequent norm conformity. Taken together, our results suggest that relationship norm strength offers a promising new perspective on relational well-being and can add to a more comprehensive account of normative processes in close relationships.

[1] University of Victoria, Canada
[2] University of Kansas, USA
[3] University of Rochester, USA
[4] University of Toronto Mississauga, Canada

Corresponding author:
John K. Sakaluk, University of Victoria, 3800 Finnerty Road (Ring Rd), Victoria, BC British Columbia V8P 5C2, Canada.
Email: jksakaluk@gmail.com

High-quality romantic relationships benefit health and well-being (Diener & Seligman, 2002; Holt-Lundstad, Smith, & Layton, 2010), yet these relationships can be fraught with areas of potential conflict. Couples can find themselves in arguments over topics like money, sex, division of labor, and communication styles, which can lead to relationship dissolution (Kurdek, 1994). One way in which couples may be able to navigate the challenging domains of their relationship, and thereby maintain satisfaction, is by developing rules about acceptable conduct in these contexts. Researchers have termed these "rules" and expectations *relationship norms*, which proscribe and prescribe certain behaviors, roles, beliefs, and attitudes among relationship members (see Clark & Mills, 1979; Sakaluk, Todd, Milhausen, Lachowsky, & the Undergraduate Research Group in Sexuality, 2014). The establishment of these relationship norms—a common area of clinical focus (e.g., Barker, 2012; Hoffman, 2011)—may be one way in which couples are able to maintain satisfying relationships.

There exists an extensive body of literature on the influence of relationship norms on well-being (Clark & Mills, 2011; Muise & Impett, 2016). Researchers within this literature have typically focused on studying the *content* of relationship norms (e.g., the ways in which people provide benefits to one another, Clark & Mills, 1993), as well as *motivation* to follow relationship norms (e.g., to what extent individuals feel compelled to follow particular norms in their relationship for benefit giving, Mills, Clark, Ford, & Johnson, 2004), as important correlates of relationship well-being. In the present article, we introduce a new feature of relationship norms—norm *strength*—that can explain variation in the motivation to follow, and ultimately apply, norms within relationships.

What makes a relationship norm strong?

Even when different romantic relationships adopt identical norm content (e.g., two couples following a norm of monogamy), couples may not be equally successful at following it (e.g., infidelity; see Mark, Janssen, & Milhausen, 2011). Even partners within the same relationship may possess discrepant levels of motivation to follow the same norm in their relationship. What can explain these differences in how norms are applied? Although relationship researchers have thoroughly studied features of relationship norms, including their content (e.g., Clark & Mills, 1979, 2011), motivation (e.g., Mills et al., 2004), and behavior (Argyle & Henderson, 1984), and their links to relationship outcomes (Le, Impett, Lemay, Muise, & Tskhay, 2018), the literature remains unclear on how these features ought to be organized into a coherent theory of norm functioning in close relationships. Within this theoretical gap, we propose that certain features of norms—norm strength—make their content within a particular domain more (or less) compelling, thereby increasing (or decreasing) partner motivation and ultimately behavioral conformity with a given norm. In turn, it is by caring about—and following—relationship norms that couples' norms can influence their relationship well-being (see Figure 1). Further, within a relationship, norms may vary in terms of their strength from one domain to the next (e.g., stronger norms for parenting, but weaker
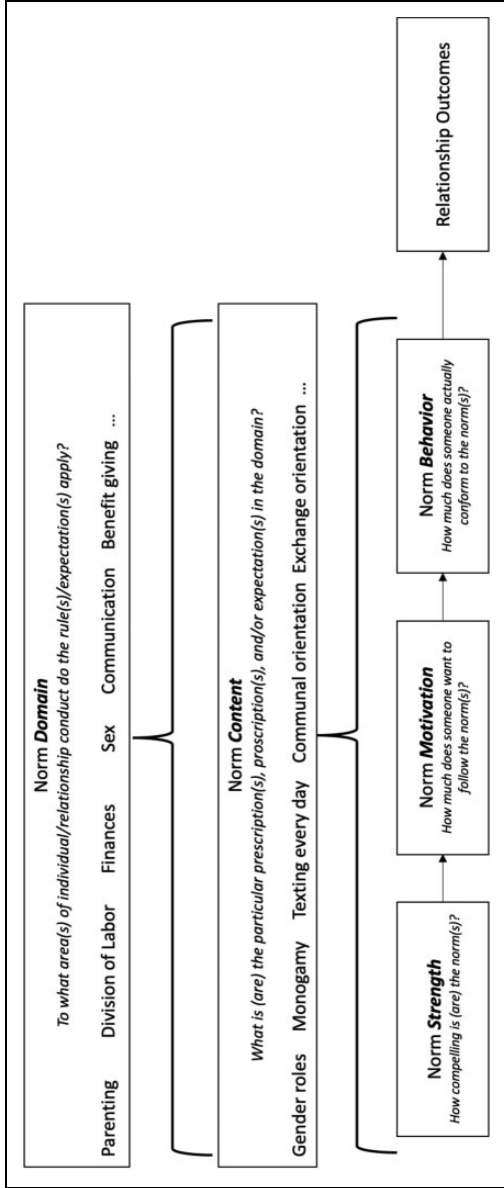
Figure 1. A model of normative processes in close relationships. Norm content defines rules/expectation (e.g., monogamy) in particular domain of relational conduct (e.g., sex). Norm content that manifests more strongly (e.g., RNSQ factors) should result in greater conformity, through which norms transmit their impact on relationship outcomes (e.g., relationship quality).

pertaining to finances), but relationships may also be characterized by more consistent patterns of norm strength across the norms they employ.

Early studies of relationship norms offered hints of the possibility that they might also vary in their strength. In their expose of norms in friendship relationships, for example, Argyle and Henderson (1984) mused that some relationship norms would be "obviously weak" (p. 212), whereas others would be deemed more important. Results from their first study largely supported their predictions, as some rules were evaluated by their participants as highly unimportant for friendships, whereas others were evaluated as very important. The work of Argyle and Henderson (1984) and others (e.g., Acitelli, 1988) also suggests possible avenues through which norms may be made stronger (or weaker), such as by explicit negotiation, and/or by linking normative conformity (or deviance) to positive (or negative) relationship outcomes.

The idea that relationship norms vary in their strength parallels similar discussions in both cross-cultural and ecological psychology. For example, *tightness–looseness* reflects the strength of norms within a particular culture or society (Pelto, 1968), whereas *situational strength* reflects the extent to which individual behavior is constrained based on a specific context (Mischel, 1973). Although there are likely appreciable differences in normative processes occurring at the level of societies, organizations, or relationships, these literature suggest to us that norms that vary in their persuasive potential may be a generalizable phenomenon across levels of social hierarchy.

Are strong norms good?

Research from the cross-cultural and ecological psychology literature suggests that stronger norms may be beneficial for their group members' well-being. Both tightness–looseness (Gelfand et al., 2011; Harrington & Gelfand, 2015) and situational strength (Meyer, Dalal, & Hermida, 2010) are linked to the well-being of members of a given social group. For example, Maass, Cadinu, and Galdi (2013) reviewed evidence suggesting that strong norms promote workplace well-being by helping to limit sexual harassment, due to extensive organizational communication and consistent sanctioning of norm violators.

Relationship-related research also provides a rationale to anticipate that stronger relationship norms may be linked to positive couple well-being. For example, clinicians working with romantic couples increasingly recognize the value in attending to and discussing relationship "rules" as a means of mitigating or preventing conflict and promoting relationship quality (Barker, 2012; Hoffman, 2011). Further, Argyle and Henderson (1984) said of relationship norms:

> we propose that they [the rules] will be functional, i.e. will help people to attain goals that are commonly sought in a situation or relationship . . . co-ordinate behaviour, and so facilitate goal-attainment through a process of group problem-solving . . . [and] will develop where there is a problem. (p. 212)

It is difficult to imagine relationship norms being so positively functional unless they are expressed in a way that can make them more binding, allowing their ostensible positive effects to be realized.

Still that stronger relationship norms will promote higher quality romantic relationships is not a foregone conclusion. On the one hand, strong relationship norms might help relationship partners better align their goals and coordinate their relationship conduct, leading to increased relationship quality (Gere & Schimmack, 2013), as suggested in previous research (e.g., Argyle & Henderson, 1984). On the other hand, strong relationship norms might lead romantic partners to feel controlled by one another (La Guardia & Patrick, 2008), reducing their sense of autonomy and competency, and thereby threatening their relationship. These two potential outcomes of norm strength can have important, yet very different, downstream consequences for relationship well-being. Thus, it is important to not only establish the features of norms that make them stronger or weaker (and verify that they do, in fact, translate to differences in norm conformity) but also establish how relationship norm strength may uniquely predict relationship well-being in order to understand the costs or benefits of establishing norm strength in romantic relationships.

Overview of the present studies

In seven studies, we developed a measure of relationship norm strength—the Relationship Norm Strength Questionnaire (RNSQ)—and collected evidence regarding aspects of validity for the resulting factors. We first assessed *measurement* validity—whether there was a replicable latent structure underlying relationship norm strength items. We also assessed the *convergent* and *divergent* validity of relationship norm strength factors, in terms of whether they correlated positively (or negatively) with scores from conceptually related (or conceptually opposite) measures or correlated weakly (if it all) with scores from conceptually unrelated measures. And finally, we assessed the *criterion* validity of relationship norm strength factors, in terms of predicting variables that are crucial for establishing the theoretical and clinical importance of relationship norm strength, and *incremental* validity, predicting scores for criterion variables above and beyond other conceptually similar constructs. We have organized the presentation of our results into these areas of validity, as within each study we collected data that were pertinent to each area of validity.

Methodological details and demographic descriptive statistics for each of the seven studies are summarized in Table 1. Our samples were typically 29–36 years old, relatively gender-balanced, mostly Caucasian, and most participants identified as heterosexual. Our cross-sectional sample sizes ranged from roughly 250 to 350 participants and were adequately powered ($1 - \beta = .80$, $\alpha_{\text{two-tailed}} = .05$) to reliably detect relatively small zero-order correlations ($r = .15$–.18) according to a sensitivity power analysis using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009).

We leveraged a number of statistical modeling strategies. These included various forms of latent variable analysis (in order to perform stronger tests of validity claims by removing the influence of error variance), meta-analysis (in order to obtain high-powered, reliable, and more generalizable estimates of effects), and the use of a consistent set of theoretically relevant control variables (i.e., age, gender, and relationship length, in order to consistently establish key associations while accounting for third variable explanations). We analyzed all of our data using *R* (R Core Team, 2019, with the

Table 1. Sample characteristics (all studies).

| | | | Methodology | | | | | Demographics | | | | |
Sample	N	Source	Design	RNSQ version	RQ measure	Analysis	M (SD) Age	M (SD) Rel. length	% Female	% Caucasian	% Heterosexual
Study 1	312/154	Mturk	Corr.	Full	PRQC (Full)	EFA/regression	29.52 (9.12)	—	44.87	73.20	81.90
Study 2	239	Mturk	Exp.	Full	PRQC (Full)	ESEM	34.07 (11.05)	7.99 (9.53)	50.63	75.10	90.70
Study 3	286	Mturk	Exp.	Full	PRQC (Full)	ESEM	34.59 (10.63)	7.34 (7.24)	55.94	74.00	91.30
Study 4	246	Mturk	Corr.	Full	PRQC (Short)	CFA/SEM	32.94 (10.38)	4.92 (5.45)	39.02	78.86	85.77
Study 5	284	Mturk	Corr.	Full	PRQC (Full)	CFA/SEM	35.22 (10.53)	8.60 (8.63)	52.82	76.06	88.38
Study 6	332	Mturk	Corr.	Short	PRQC (Full)	CFA/SEM	35.55 10.61	9.81 (9.40)	56.33	82.83	90.06
Study 7	115	Community	Long.	Short	PRQC (short)	SEM	24.02 (6.94)	3.32 (3.53)	77.97	71.67	74.17

Note. RNSQ = Relationship Norm Strength Questionnaire, RQ = Relationship Quality. ESEM = exploratory structural equation modeling (see Asparouhov & Muthén, 2009). CFA/SEM = traditional confirmatory factor analysis and structural equation modeling. Exp. = experimental, Corr. = correlational, Long. = longitudinal. The subsample of individuals in romantic relationships (vs. other sexual relationship types) from Study 1 was used when testing the models reported in Table 5. Studies 2 and 3 were experimental studies that were omitted due to space constraints; their results will be described in a forthcoming paper, but measurement-related data contributed to the present paper.

exceptions of analyses from Studies 2 and 3, which we conducted using Mplus (Muthén & Muthén, 2012). Finally, we used open science practices throughout our studies to facilitate transparency and replicability in our findings; when doing so, we note throughout the article when we preregistered hypotheses, methods, and data analysis plans, and all of our data and analytic scripts are available in our supplemental materials (https://osf.io/979fu/).

Evaluating the measurement validity of the relationship norm strength questionnaire

Our process of developing the RNSQ first entailed piloting items and determining an adequate number of factors to represent them using exploratory factor analysis (EFA). We then determined a model for the RNSQ items to later test using confirmatory factor analysis (CFA) with the exploratory measurement model results from our first three samples. We then tested this candidate model using CFA in two later samples, while assessing measurement invariance between men and women to ensure that the factors were measured similarly across gender.

Item development

Based on the literature of tightness–looseness, situational strength, and social influence (Gelfand et al., 2011; Maass, Cadinu, & Galdi, 2013; Meyer et al., 2010), the first author drafted 24 items for the RNSQ that assessed different qualities of relationship norms, including volume (e.g., number of norms), saliency (e.g., to what extent the norms are on someone's mind), clarity (e.g., how easily understandable norms are), agreement (e.g., being on "the same page"), and punishment for deviance (e.g., what actions are taken when norms are broken), among others. Participants were asked to think about their relationship as they rated the extent to which each statement described them accurately on a 7-point scale ($1 = $ *not at all accurate*; $7 = $ *completely accurate*).

Factors of the RNSQ

We used EFA in Study 1, via the *psych* package (Revelle, 2016), to determine a plausible number of factors to represent the RNSQ items. We estimated factor solutions using maximum-likelihood extraction and oblimin rotation and used a combination of parallel analysis (Horn, 1965), nested model comparisons, indexes of model fit (root mean square error of approximation [RMSEA] and Tucker–Lewis index; Hu & Bentler, 1999), and solution interpretability to determine the number of factors to retain for our exploratory measurement model (Sakaluk & Short, 2017). We then extracted factor scores using the method described in ten Berge, Krijnen, Wansbeek, and Shapiro (1999) for later tests of our hypotheses regarding relationship norm strength and relationship quality.

The sample in Study 1 consisted of individuals reporting on different relationship experiences (e.g., one-night stands, committed romantic relationships) in order to increase the breadth of the measure's eventual applicability. Parallel analysis suggested that a maximum of six factors was sufficient (see Figure 2); we therefore evaluated
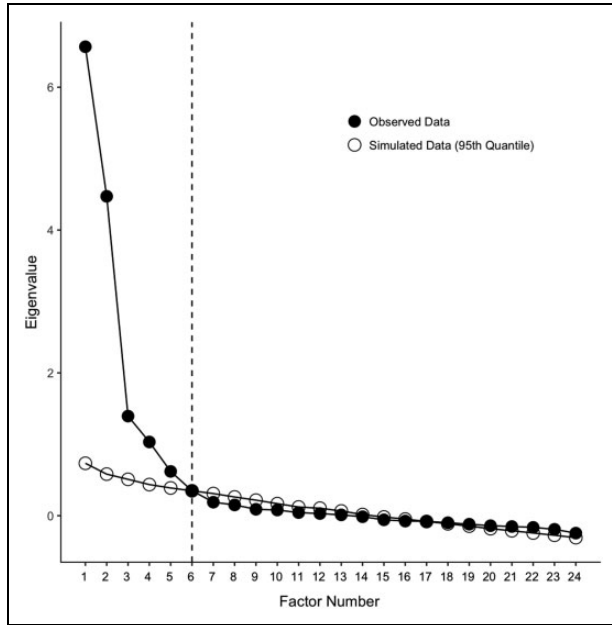
Figure 2. Parallel analysis of RNSQ data suggests that up to six factors outperform simulated "random noise" factors, in terms of explaining variance in RNSQ items.

Table 2. Model fit indexes and model comparisons for 1–6 exploratory factor solutions.

# of Factors	RMSEA 95% CI	TLI	BIC	χ^2 (df)	$\Delta\chi^2$ (Δdf)	Improves model?
1	.16, .18	.40	1059.37	2506.61*** (252)	—	—
2	.11, .13	.71	−98.79	1216.36*** (229)	1290.25*** (23)	Yes
3	.08, .10	.83	−461.52	727.29*** (207)	489.07*** (22)	Yes
4[a]	.07, .08	.88	−562.56	505.64*** (186)	221.65*** (21)	Yes
5	.04, .06	.94	−641.05	312.28*** (166)	193.36*** (20)	Yes
6	.03, .06	.96	−600.51	243.71*** (147)	68.57*** (19)	Yes

Note. RMSEA = root mean square error of approximation; TLI = Tucker–Lewis index; BIC = Bayesian Information Criterion. [a]Selected model.
***p < .001.

solutions consisting of one-to-six factors. Nested model comparisons indicated that extracting each factor beyond the first significantly improved the fit of the exploratory model. However, descriptive fit indexes (see Table 2) indicated that only four-to-six factor solutions were tenable. After close examination, we selected the four-factor solution, as the five- and six-factor solutions further fragmented the first extracted factor in a way that did not add conceptual or theoretical utility. The four-factor solution (see Table 3) fit the data reasonably well, was conceptually interpretable, accounted for a

Table 3. Oblimin rotated factor loadings for the selected four-factor EFA solution.

Item	F1*	F2	F3	F4*	h^2
1. I feel like our relationship has a lot of "rules."	−.69[a]	.12[b]	.04	.18[a]	.51
2. I don't feel like I have to worry about following too many "rules" in our relationship.	.27	.22[b]	.05	−.35[a]	.29
3. I feel overwhelmed by the number of "rules" in our relationship.[s67]	−.81[a]	.03	−.01	.04[b]	.65
4. The "rules" of our relationship are frequently on my mind.[s7]	−.67[a]	.09	.00	.01[b]	.44
5. I rarely think about the "rules" of our relationship.[s6]	.21[b]	.22[b]	.08	−.29[a]	.20
6. I feel like our relationship has more "rules" than many other relationships I have been in or known of.	−.74[a]	.05	.02	.02[b]	.53
7. I feel like I clearly understand many of the "rules" of our relationship.[s6]	.08	.62[a]	.02	.05	.45
8. I am confused by many of the "rules" of our relationship.[s7]	−.80[a]	−.09[b]	.01	−.04	.71
9. My partner and I have talked openly about the "rules" of our relationship.[s67]	−.15	.43[a]	.09[b]	.56[a]	.61
10. My partner and I don't discuss the "rules" of our relationship; they are just understood.[s67]	−.10[b]	.14	−.08	−.73[a]	.54
11. My partner and I agree on the "rules" of our relationship.[s67]	−.03	.80[a]	.03	.17	.70
12. Disagreements between my partner and I about the "rules" of our relationship occur frequently.	−.66[a]	−.12[b]	.07[b]	−.05	.51
13. Some of the "rules" of our relationship seem to contradict one another.	−.78[a]	−.04[b]	−.01	−.01	.63
14. The "rules" of our relationship are consistently applied to both my partner and me.[s67]	.03	.64[a]	.07	.03	.53
15. It feels like there are different "rules" for me and different "rules" for my partner.	−.78[a]	−.01[b]	.02	−.13	.61
16. The "rules" of our relationship have not changed over time.	.01	.50[a]	.04	−.25[b]	.29
17. The "rules" of our relationship are legitimate.[s67]	.08	.77[a]	.06	−.08[b]	.68
18. I agree with most of the "rules" of our relationship.	.09[b]	.86[a]	−.04	−.13[b]	.77
19. The "rules" of our relationship are ridiculous.[s6]	−.76[a]	−.15[b]	.01	−.07	.68
20. If my partner or I broke one of the "rules" of our relationship, the other would be very upset.	−.04	.15	.78[a]	.08	.77
21. The consequences for breaking one of the "rules" of our relationship would be severe.[s67]	−.00	−.03	.96[a]	−.06	.87
22. My partner and I don't really care if the other follows the "rules" of our relationship.	−.24[a]	.01	−.42[a]	−.28[a]	.33
23. If my partner or I broke one of the "rules" of our relationship, it is likely our relationship would end.[s67]	.05	−.03	.78[a]	.03	.60
24. My partner or I would be punished by the other for breaking one of the "rules" of our relationship.[s67]	−.09	−.04	.79[a]	−.04	.63

Note. F1 = tractability factor; F2 = agreement factor; F3 = punishment factor; F4 = explicitness factor; h^2 = item communality. *Loadings reversed for this factor, to score in direction of norm strength. [a]Reliably strong-loading ($\lambda > .30$, $p < .05$) item from measurement meta-analysis of loadings in Studies 1–3. [b]Reliably weak-loading ($\lambda < .30$, $p < .05$) item from measurement meta-analysis of loadings in Studies 1–3 (see Asparouhov & Muthén, 2009, on importance of modeling weak, yet reliable, loadings). [s6]Short-form adapted for Study 6. [s7]Short-form adapted for Study 7.

Table 4. Model fit for confirmatory and invariance models of the RNSQ.

Model	χ^2 (df)	CFI	RMSEA 90% CI	Scaled $\Delta\chi^2$ (Δdf)
RNSQ (Study 4)	496.06*** (222)	.91	.07, .09	—
Configural Invariance	890.24*** (444)	.87	.09, .11	—
Weak Invariance	861.37*** (488)	.88	.08, .10	18.92 (44)
Strong Invariance	754.12*** (508)	.91	.06, .09	4.35 (20)
RNSQ (Study 5)	504.23*** (222)	.91	.07, .08	—

***p < .001.

large amount of item variation (median $h^2 = .61$), and produced a discriminating pattern of factor loadings.

The first factor, which we labeled *norm tractability*, reflects the manageability and consistency of norms (e.g., "I feel overwhelmed by the number of 'rules' in our relationship" [reversed]; 8 items). The second factor, which we labeled *norm agreement*, reflects the individual (and their partner) being "on the same page" with relationship norms, and evaluating their norms as legitimate (e.g., "My partner and I agree on the 'rules' of our relationship"; 7 items). The third factor, *anticipated punishment*, reflects the likelihood and severity of consequences for deviance from relationship norms (e.g., "The consequences for breaking one of the 'rules' of our relationship would be severe"; 5 items). The fourth factor, which we labeled *norm explicitness*, reflects whether partners have talked openly about relationship norms or, rather, have simply assumed they existed (e.g., "My partner and I have talked openly about the 'rules' of our relationship"; 2 items). The four exploratory RNSQ factors were generally weakly correlated with one another ($-.10 \leq r$s $\leq .37$).

RNSQ factor structure confirmation

Our candidate model for confirmatory testing was based on results meta-analyzing each of the factor loadings of the 24 RNSQ items onto each of the four exploratory factors fit in Studies 1–3 (see supplementary analytic materials on OSF). Doing so allowed us to increase the reliability of the estimates through increased power, in addition to increasing generalizability across participants and relationships. We tested our candidate measurement model, in full, using CFA (Beaujean, 2014) in two subsequent samples (Studies 4 [preregistered] and 5) via the *lavaan* package (Rosseel, 2012), using a robust maximum-likelihood estimator (MLR) and full-information maximum-likelihood (FIML) to manage missing data, evaluating our models with both an absolute index (RMSEA) and a relative index (confirmatory fit index [CFI]) (Hu & Bentler, 1999). All models were identified—and the scale set for each latent variable—by fixing the latent variance to 1 and estimating unique factor loadings for each item.

We also tested whether our candidate measurement model was invariant between men and women (Vandenberg & Lance, 2000; Study 4) using the *semTools* package (sem-Tools Contributors, 2016), by looking at changes in our absolute and relative indices of model fit between each level of invariance constraints that we imposed (Cheung & Rensvold, 2002). Fit indexes for each model are presented in Table 4. In summary, the

measurement model of the RNSQ items demonstrated acceptable fit in two separate samples, and further, invariance analyses indicated that assessment of RNSQ items was comparable between men and women.

Convergent and discriminant validity of the RNSQ

To evaluate the convergent and discriminant validity of RNSQ factors, we focused primarily on evaluating the convergent and discriminant validity of norm *tractability* and *agreement*, as these two factors are the most similar to other constructs in the literature and were found to be the most reliable predictors of relationship quality (described later). Structural equation models (SEMs) were fit using the *lavaan()* package (Rosseel, 2012), using the same selection of MLR estimation and FIML for missing data, and a fixed-factor method of scale-setting and model identification.

Convergent validity of norm tractability

We used the behavioral norm uncertainty subscale from the Relationship Uncertainty Scale (Knobloch & Solomon, 1999) in Study 5. The behavioral norm uncertainty contains 4 items (e.g., "How certain are you about how you can or cannot behave around your partner) that are rated on a 6-point scale (1 = *completely or almost completely uncertain*; 6 = *completely or almost completely certain*). We expected more tractable relationship norms to be associated with behavioral norms that were less uncertain, as (un)certainty is reflected in the item content of the RNSQ tractability factor, though the tractability factor assesses a broader array of cognitive elements (e.g., (in)consistency, overwhelming frequency) beyond certainty. Supporting these preregistered predictions, highly tractable norms were associated with less behavioral norm uncertainty ($r = -.52$, 95% confidence interval [95% CI] $-.63$, $-.41$), but not to an extent where the two constructs were singular ($R^2 = .27$).

We also assessed several aspects of relational manipulation (Buss, 1992), selected a priori, in Study 4. Specifically, we assessed the manipulation tactics (i.e., "To get me to do what they want, my partner will . . .") of *coercion* (3 items; e.g., "criticize me for not doing it"), *regression* (3 items; e.g., "pout until I do it"), *hardball* (5 items; e.g., "tell me that they will leave me if I don't do it"), *silent treatment* (3 items; e.g., "ignore me until I do it"), and *social comparison* (3 items; e.g., "tell me that other partners would do it"). We expected more tractable norms to be found in relationships characterized by lower levels of relational manipulation, as manipulation tactics might be found among relationship members feeling like their relationship norms are confusing, inconsistent, and overwhelming. Supporting these predictions, tractable norms were associated with lower levels of all measured relationship manipulation factors ($rs = -.58$ to $-.31$, 95% CI $-.70$, $-.16$) though norm intractability remained empirically distinguishable from each ($R^2 = .10$–$.34$).

Convergent validity of norm agreement

We expected perceived similarity between partners and norm agreement scores to be positively associated, but distinct (i.e., sharing less than 50% variance or $r \leq .70$), in

Studies 1 and 5 (pre-registered) because of prior theorizing, suggesting that stronger norms ought to produce more homogenous (i.e., similar) groups (Triandis, 1989). We assessed perceived similarity using a singleton ad hoc, face-valid item created for this study: "In general, how similar are you and your partner?" ($1 = $ *not at all similar*, $7 = $ *extremely similar*). Our meta-analytic estimate across the two samples suggested that strong norm agreement is associated with greater perceived similarity to one's partner ($r = .17$, 95% CI .08, .26), while the amount of overlap between these measures in our two samples ($R^2 = .01–.10$) indicates that our measure of norm agreement is distinguishable from a general sense of similarity to one's romantic partner.

Dyadic consensus (Spanier, 1976), meanwhile, reflects the extent to which relationship members agree with their partner about general areas of their relationship (e.g., handling of finances, amount of time spent together, household tasks); the measure consists of 13 areas which are rated on a 6-point scale ($1 = $ *always disagree*, $6 = $ *always agree*). We expected that norm agreement would positively correlate with dyadic consensus, as agreement specific to relationship norms should bear some correspondence to general tendencies to agree within a relationship. Supporting our preregistered predictions, strong norm agreement was also positively associated with dyadic consensus in Study 5 ($r = .60$, 95% CI .51, .70), but not so strongly to suggest the two constructs are singular ($R^2 = .36$).

Finally, we adapted a measure of procedural justice (Bauer et al., 2001) for use in Study 4, in order to assess the extent to which participants felt like they were treated fairly and with respect within their relationship. We selected, a priori, factors pertaining to *openness* (4 items, e.g. "I am treated honestly and openly by my partner in our relationship"), *treatment*, (4 items, e.g. "My partner is considerate of me"), and *two-way communication* (5 items, e.g. "I feel comfortable with the idea of expressing my concerns about our relationship to my partner"), all of which were rated on a 5-point scale ($1 = $ *strongly disagree*; $5 = $ *strongly agree*). We expected norm agreement to be positively associated with procedural justice, as relationships that are procedurally fair and respectful likely also have more highly agreed upon norms, but that sense of fairness within a relationship should encompass a variety of influences beyond those that are norm-specific. As predicted, norm agreement was positively correlated with procedural justice factors, with stronger norm agreement associated with greater *openness* ($r = .56$, 95% CI .43, .69), *treatment* ($r = .56$, 95% CI .42, .69), and *two-way communication* ($r = .57$, 95% CI .46, .69). However, as with the other correlates of norm agreement, the empirical overlap between norm agreement and the procedural justice factors ($R^2 = .31–.32$) was not sufficient to indicate a singular factor.

Discriminant validity of RNSQ factors

As a test of discriminant validity for the entire RNSQ, we assessed the extent to which RNSQ factors correlated with dimensions of socially desirable responding in Studies 5 (preregistered) and 6. We used Paulhus's (1998) Balanced Inventory of Desirable Responding, which assesses two distinct aspects of socially desirable responding. *Self-deception* (20 items; e.g., "My first impressions of people usually turn out to be right") captures an individual's propensity to deceive themselves, whereas *impression*

management (20 items; e.g., "I don't gossip about other people's business") captures an individual's propensity to present themselves to others in a disingenuously favorable manner; all items were rated on a 7-point scale (1 = *not true*; 7 = *very true*). We predicted that correlations between latent RNSQ factors and latent self-deception and impression management factors would be small and negligible.

Meta-analytic estimates across Study 5 and Study 6 generally supported our predictions. Norm tractability ($r = -.31$, 95% CI $-.38$, $-.24$) and norm agreement ($r = -.31$, 95% CI $-.38$, $-.24$) correlated to only small degrees with self-deception. We observed a similar pattern of correlations for the associations with impression management: for norm tractability ($r = -.21$, 95% CI $-.29$, $-.14$) and norm agreement ($r = -.18$, 95% CI $-.25$, $-.10$). Norm explicitness was not reliably associated with either self-deception ($r_{Study 5} = .08$, 95% $CI_{Study 5} -.11$, .26) or impression management ($r_{Study 5} = .14$, 95% $CI_{Study 5} -.04$, .32). Anticipated punishment for deviance, meanwhile, was weakly associated with self-deception ($r = .13$, 95% CI .05, .21) and not significantly associated with impression management ($r = .06$, 95% CI $-.02$, .14). Significant or not, all levels of overlap between relationship norm strength and social desirability factors were trivial ($R^2 = .0001–.17$).

Summary of convergent and discriminant validity

Results from our analyses generally provided evidence for the convergent and discriminant validity of RNSQ factors. As expected, norm tractability was associated with less behavioral norm uncertainty and relationship manipulation, whereas norm agreement was associated with greater perceived similarity, dyadic consensus, and relational procedural justice—all of these correlations were never large enough to suggest RNSQ factors were tapping the same constructs as these other measures. Also supporting our predictions, RNSQ factors were generally weakly associated (or not reliably associated at all) with social desirability factors.

Criterion validity of RNSQ factors

We next tested the criterion validity of the RNSQ factors. Our analyses included tests of two relatively proximal criterion variables in our model depicted in Figure 1—*motivation* to follow relationship norms and *behavioral conformity* to relationship norms—as well as repeated testing of our downstream criterion of relationship quality.

For each criterion, we first report its measurement and analyses in the relevant study(ies) with cross-sectional data. We then report on our longitudinal analyses, which we conducted using generalized cross-lagged panel models (see Zyphur et al., 2019), a newly developed method that provides better means of evaluating causality in longitudinal designs, including both short-term and long-term changes. Again, all results come from SEMs that we fit using the *lavaan()* package (Rosseel, 2012).

Cross-sectional analyses

Norm motivation. In Studies 5 and 6, we assessed norm-relevant motivation of a particular kind: communal strength (Mills et al., 2004)—an individual's motivation to follow his or

her relationship's norm pertaining to providing benefits on the basis of need. We selected this measure, given that the communal strength is one of the most well-researched forms of norm motivation (see Le et al., 2018) and because we expected higher levels of general relationship norm strength to be associated with greater levels of norm-specific motivation. Norms that are generally more tractable, agreed upon, explicit, and enforced should lead to greater levels of motivation to abide by these norms, including for communalism, specifically. We used Mills, Clark, Ford, and Johnson' (2004) measure (10 items; e.g., "I would incur a large cost in order to help my partner") rated on a 7-point scale (1 = *strongly disagree*; 7 = *strongly agree*) to model a latent factor of communal strength in Studies 5 (preregistered) and 6.

Consistent with our predictions, meta-analytic estimates across Study 5 and Study 6 suggested that more tractable ($r = .55$, 95% CI .49, .60) and agreed upon ($r = .57$, 95% CI .51, .62) norms were positively associated with communal strength. However, contrary to our expectations, norm explicitness ($r_{Study 5} = -.18$, 95% $CI_{Study 5}$ −.37, .01) was unassociated with communal strength, and anticipated punishment for deviance was negatively associated with communal strength ($r = -.24$, 95% CI −.31, −.17).

Relationship norm conformity. We used two ad hoc measures of relationship norm conformity in Study 5. The first measure asked participants to use two sliding scales (0–100%) to capture a holistic sense of how consistently they and their romantic partner follow the "rules" and expectations of their relationship, without referring to any particular domains of relationship norms. These two items were used to form a latent perception of domain-ambiguous relationship norm conformity. We also asked participants to indicate how consistently they conformed to the "rules" and expectations of their relationship in nine particular relationship domains in which couples might establish norms (e.g., sex, finances, division of labor), using a 5-point rating scale (1 = *0%–19% of the time* to 5 = *80%–00% of the time*). Participants could also indicate that their relationship did not have "rules" and expectations in each domain, which was recoded as missing data. These 9 items were used to form a latent measure of self-perceived domain-specific relationship norm conformity.

Supporting our preregistered predictions, participants in relationships with greater levels of domain-ambiguous conformity were also characterized by having greater norm tractability ($r = .57$, 95% CI .41, .73) and norm agreement ($r = .60$, 95% CI 47, .73). After controlling for our standard set of covariates, higher levels of domain-ambiguous conformity were uniquely associated with greater levels of norm tractability ($b = 0.67$, 95% CI 0.27, 1.06), norm agreement ($b = 0.46$, 95% CI 0.16, 0.76), and anticipated punishment for deviance ($b = 0.24$, 95% CI 0.09, 0.38). Norm explicitness was not significantly associated with domain-ambiguous relationship norm conformity levels on its own or alongside the standard covariates and other RNSQ factors. Together, RNSQ factors explained nearly half the latent variance in domain-ambiguous relationship norm conformity ($R^2 = .47$).

Participants in relationships with greater levels of domain-specific conformity were also characterized by having much greater norm tractability ($r = .43$, 95% CI .30, .57) and norm agreement ($r = .47$, 95% CI .36, .59), although levels of explicitness ($r = -.12$, 95% CI −.32, .07) and anticipated punishment ($r = -.12$, 95% CI −.05, .17) for

Table 5. Meta-analyzed zero-order correlations between norm strength factors and relationship quality/satisfaction (Studies 1–6).

RNSQ Factor	k	r (95% CI)	Mdn Power	R-Index	TIVA	p-curve	b (95% CI)
Tractability	6	.48*** (.38, .56)	1.00	1.12	12.22	−8.61***	0.19* (0.04, 0.34)
Agreement	6	.59*** (.54, .64)	1.00	1.00	5.84	−10.45***	0.65*** (0.53, 0.77)
Punishment	6	−.14 (−.31, .03)	.85	.88	1.21	−4.91***	−0.12 (−0.25, .02)
Explicitness[a]	5	−.09 (−.21, .04)	.23	0.06	9.38	−4.16***	−0.01 (−0.11, 0.09)

Note. Correlation estimates have been transformed and back-transformed using Fisher's Z during meta-analysis. Replicability indexes refer to zero-order correlations from Studies 1–6. [a]One non-converging sample excluded.
**p< .01; ***p<.001.

deviance were no different. After controlling for our standard covariates, higher levels of domain-specific conformity were uniquely associated with greater levels of norm tractability ($b = 0.33$, 95% CI 0.04, 0.62) and norm agreement ($b = 0.37$, 95% CI 0.12, 0.62), with the RNSQ factors explaining about a third of the latent variance in domain-specific relationship norm conformity ($R^2 = .28$).

Relationship quality. The final area of criterion validity that we evaluated for the RNSQ factors was the extent to which they were associated with relationship quality. We predicted that stronger relationship norms would be associated with greater relationship quality. As a follow-up test of the robustness of these predictions, we evaluated whether RNSQ factors could demonstrate incremental validity by uniquely predicting this relationship quality even after controlling for conceptually similar constructs identified during the analyses related to convergent validity.

Relationship quality was assessed in all studies using the Perceived Relationships Quality Components (PRQC) measure (Fletcher, Simpson, & Thomas, 2000). The PRQC captures six dimensions of relationship quality using 18 items, rated on a 7-point scale (1 = *not at all*, 7 = *extremely*): *satisfaction, commitment, intimacy, trust, passion,* and *love.* We operationalized relationship quality as a higher order factor of the six dimensions. In one case (Study 4), we used the short form of the PRQC whereby relationship quality was operationalized as an amalgam of the first item from each factor.[1]

Cross-sectional evidence: Zero-order correlations. We meta-analyzed all latent zero-order correlations between individual RNSQ factors and relationship quality across Studies 1–6. These estimates provided mixed support for our predictions (see Table 5). As expected, norm tractability and norm agreement were strongly associated with greater relationship quality. Contrary to our predictions, however, associations with relationship quality were not significant for anticipated punishment for deviance or norm explicitness. Replicability indexes produced by Schönbrodt's (2015) *p-checker* application suggested that these analyses were high-powered, replicable, reported without bias, and contained evidential value, with the exception of the norm explicitness factor, which was hampered by convergence problems in two of our samples.

Cross-sectional evidence: Unique associations. We then examined the unique associations of RNSQ factors, as they predicted relationship quality in unison across Studies 1–2 and 4–6.[2] In Studies 4–6, we also provide evidence of the incremental validity of particular RNSQ factors as predictors of relationship quality, above and beyond competitor constructs identified during measure development. We used a consistent set of control variables throughout these analyses: gender, age, and relationship length, except in Study 1, where relationship length was not consistently measured.

Basic model. Meta-analyzed estimates of unstandardized regression slopes for our basic model are presented in Table 5. Similar to our analyses of zero-order correlations, these results provide partial support for our predictions, as they suggest that the association between RNSQ factors and relationship quality is specifically driven by levels of norm tractability and norm agreement, as both uniquely predict increased relationship quality. Across these five studies, RNSQ factors combined to consistently explain a considerable amount of the variance in relationship quality (R^2 = .34–.46), above and beyond our selected control variables.

Incremental validity. In order to maximize our confidence in our effects for relationship quality, we tested the incremental validity of norm tractability and norm agreement, after controlling for various competitor constructs. Norm tractability significantly predicted greater relationship quality, even after first simultaneously controlling for all five relationship manipulation tactics (b = 0.43, 95% CI 0.18, 0.67), as well as when separately controlling for behavioral norm uncertainty (preregistered, b = 0.63, 95% CI 0.35, 0.91). In addition, higher levels of norm agreement significantly predicted greater relationship quality even after first controlling for perceived similarity ($b_{Study 1}$ = .22, 95% $CI_{Study 1}$ 0.11, 0.33; $b_{Study 5}$ = .89, 95% $CI_{Study 5}$ 0.63, 1.15), as well as separately controlling for all three procedural justice factors (b = 0.33, 95% CI 0.07, 0.58), and separately, dyadic consensus levels (preregistered, b = 0.63, 95% CI 0.34, 0.91). We interpret these patterns across multiple samples to suggest that the unique predictive effects of norm tractability and norm agreement for relationship quality to be robust.

Longitudinal analyses

Measures. We adopted an entirely norm-specific approach to assessment in Study 7, which would allow us to test predictions of our Model (Figure 1) at the finest level of concept resolution (i.e., the strength of a particular norm predicting the level of motivation for the same norm). Participants in Study 7 first read descriptions of communal and exchange norms for benefit-giving and indicated whether their relationship followed a communal (91%) or exchange norm (9%) for benefit-giving. Then, each week, participants reported on their perceived strength, motivation, and behavioral conformity for their relationship's (communal or exchange) norm for benefit-giving. We assessed RNSQ factors using an 11-item short-form of the measure with items for each factor that could be easily adapted for a specific norm of benefit-giving (e.g., "I feel overwhelmed by the [communal or exchange] approach in our relationship"; "The consequences for failing to live up to the [communal or exchange] approach in our relationship would be severe."). We then assessed participants' motivation to follow their communal or

exchange norm using 3 items (e.g., "Following [the communal or exchange] approach in my relationship is a high priority for me.") rated on a 7-point scale (1 = *not at all accurate*; 7 = *completely accurate*), while we assessed participants' behavioral conformity using 2 items (e.g., "Approximately what percent of the time . . . are you successful at following [the communal or exchange] approach in your relationship?") using sliders from 0% to 100%.

Generalized cross-lagged panel models. With longitudinal data where the number of waves is more limited (i.e., <20), researchers often employ a cross-lagged panel model (CLPM), where later instances of a given *X* or *Y* variable are regressed onto their respective immediately preceding instances (i.e., "auto-regressive" effects, e.g., *X1* → *X2*), and later instances of *X* or *Y* variables are regressed onto the other variable's immediately preceding instances (i.e., "cross-lagged" effects, e.g., *Y2* → *X1*). The generalized cross-lagged panel model (GCLPM) improves on the CLPM by introducing unit effects, moving average effects, and cross-lagged moving average effects, and in doing so, allows the estimation of a model which affords stronger causal claims, and a more complicated constellation of pathways through which earlier processes may affect later outcomes (Zyphur et al., 2019). In particular, the GCLPM affords partitioning the testing of *short-run effects* (direct effects of earlier instances of *X* on the immediately subsequent instances of *Y*, through the combined influence of cross-lagged effects and cross-lagged moving averages) and *long-run effects* (indirect effects from earlier instances of *X* on subsequent instances of *Y*, through all possible cross-lagged effects and cross-lagged moving averages, across longer lag periods).

We evaluated different patterns of causality for short-run effects, with models evaluating whether (1) earlier levels of relationship norm strength impacted subsequent relationship norm motivation (or conformity)[3], (2) earlier relationship norm motivation (or conformity) impacted subsequent relationship norm strength, or (3) a reciprocal pattern of causality between relationship norm strength (or conformity) was best supported. We also tested long-run effects for the downstream impact of norm strength factors on relationship norm motivation and relationship norm conformity at 4 weeks (i.e., halfway through our study) and again at 8 weeks (see Figure 3).

Longitudinal evidence. Our results for norm motivation largely challenged our conceptualization link between norm strength and norm motivation in Figure 1. Specifically, earlier levels of norm strength factors for benefit-giving norms predicted very little in terms of subsequent levels of norm motivation for benefit-giving, either in the immediate short-term or at 4- or 8 weeks. Indeed, only greater anticipated punishment for deviance from benefit-giving norms yielded a short-term reciprocally causal effect of increased motivation to follow relationship benefit-giving norms, $b = 0.19$, $p = .001$, while increased motivation to follow relationship norms for benefit-giving also resulted in greater anticipated punishment for deviance from benefit-giving norms, $b = 0.19$, $p < .001$. However, over the long-term, anticipated punishment for deviance from benefit-giving norms actually ended up reducing norm motivation for benefit-giving norms over 8 weeks ($b = -0.63$, $p = .04$).
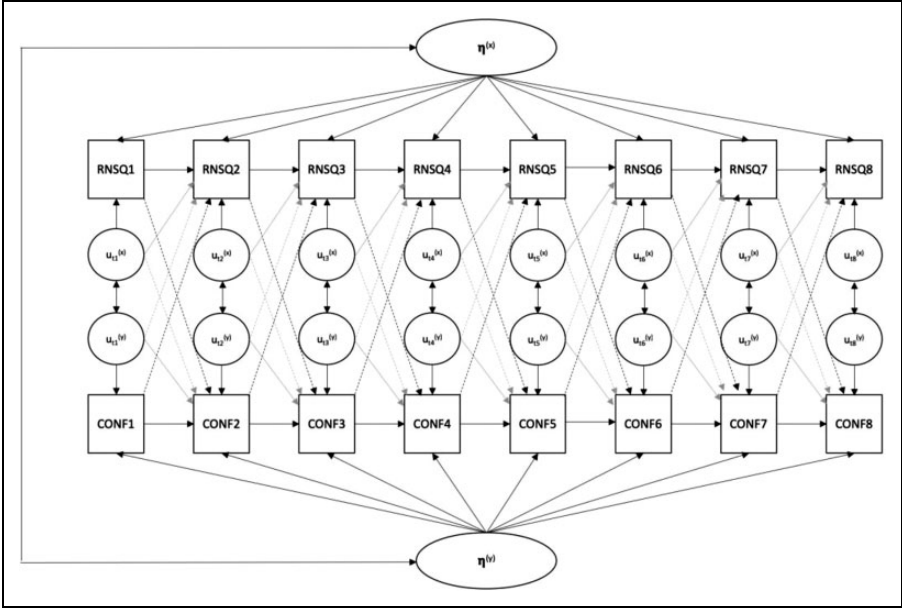
Figure 3. Example generalized cross-lagged panel model depicting longitudinal effects of a given norm strength factor (RNSQ) on one of the validity-related variables (e.g., conformity, CONF). Solid black paths between observations of the same variable (e.g., CONF1 → CONF2) represent auto-regressive effects. Dashed black paths between observations of different variables (e.g., RNSQ1 → CONF2) represent cross-lagged effects. Solid gray paths between impulses and observations of the same variable (e.g., $u_{t1}^{(y)}$ → CONF2) represent moving averages. Dashed gray paths between impulses and observations of different variables (e.g., $u_{t1}^{(x)}$ → CONF2) represent cross-lagged moving averages. Factor loadings capture unit effects.

In contrast, the effects of norm strength factors on subsequent conformity to benefit-giving norms were much more aligned with our theorizing. For norm tractability of benefit-giving norms, short-term effects suggested that greater conformity to benefit-giving norms increased the tractability of benefit-giving norms, $b = 0.01$, $p = .01$. However, greater earlier tractability of benefit-giving norms promoted considerably increased conformity to benefit-giving norms in the intermediate-term (i.e., 4 weeks), $b = 7.47$, $p = .001$, and in the long-term (8 weeks), $b = 13.65$, $p = .03$. Likewise, for norm agreement for benefit-giving norms, short-term effects suggested that greater conformity to benefit-giving norms increased the agreement of benefit-giving norms, $b = 0.02$, $p < .001$. However, greater earlier norm agreement for benefit-giving norms promoted considerably increased conformity to benefit-giving norms in the intermediate-term (i.e., 4 weeks), $b = 10.92$, $p = .001$, and in the long-term (8 weeks), $b = 21.28$, $p = .04$. Conversely, anticipated punishment for deviance from benefit-giving norms produced unreliable increases in levels of conformity to benefit giving-norms when assessed in the short-term, $b = 1.08$, $p = .42$, but over time its influence on levels of conformity to benefit-giving norms trended to become increasingly negative in the intermediate-term, $b = -1.26$, $p = .28$, and the long-term, $b = -5.74$, $p = .07$.

Summary of criterion validity

Evidence for the criterion validity of RNSQ factors was mixed, though mostly corroborative. In our cross-sectional studies, all of the norm strength factors (with the exception of explicitness) were reliably associated with greater self-reported conformity to relationship norms. Norm tractability and agreement were also positively associated with motivation to follow communal norms.

Meta-analyzed multi-sample evidence, meanwhile, strongly supported our predictions of the positive and strong associations between norm tractability and agreement, and relationship quality—which in two key samples remained significant and substantial even after controlling for intuitive competitor constructs already available in the close relationship literature. Multiple replicability-related indexes further suggest that these patterns of association are likely to replicate in future investigations. Associations between norm explicitness and anticipated punishment for deviance and our criterion variables, meanwhile, did not support our predictions to various extents—a pattern of results that we unpack in greater detail in the general discussion.

Our longitudinal study, meanwhile, suggests a more complicated picture, in which norm tractability and agreement for benefit-giving norms, specifically, may operate to increase conformity levels for benefit-giving norms more directly—irrespective of motivation levels—while anticipated punishment for violation of benefit-giving norms might increase motivation in the short-term, but be ultimately deleterious to conformity levels in the long run.

General discussion

In seven studies we developed the RNSQ, confirmed its measurement structure, and provided preliminary evidence of its validity. Regarding construct measurement, our studies suggest that relationship norm strength manifests in four different, weakly correlated ways, including the extent to which norms are tractable, agreed upon, subject to punishment, and explicitly discussed. Further, the RNSQ was invariant between men and women, and the effects observed for its factors appear to be distinguishable from those of key competitor constructs.

Our predictions regarding the roles of norm strength in promoting norm-related motivation and conformity received mixed support. Few RNSQ factors predicted subsequent changes in norm-relevant motivation, but as we expected, stronger norms generally resulted in long-term boosts in subsequent levels of conformity. Stronger norms may therefore help to shape relationship behavior through environmental (e.g., Meagher, 2019) and/or automatic cognitive channels (e.g., Bargh, Schwader, Hailey, Dyer, & Boothby, 2012), as opposed to deliberated cognitive processes.

Our prediction about a positive association between relationship norm strength and relationship quality also received some support. Positive associations between norm tractability and relationship quality, and norm agreement and relationship quality, were consistent across our studies and not attributable to a myriad of competitor constructs. For the other aspects of relationship norm strength, however, our data were not so clear.

The association between anticipated punishment for deviance and relationship quality, though consistently significant, was not consistent in its direction.

We suspect that the negative content of the anticipated punishment items may elicit thoughts of low relationship quality, at least under some circumstances. It will therefore be important to evaluate whether there are moderators that determine when anticipating punishment for deviance benefits, or hurts, the quality of relationships. The perceived severity of the deviance may be one such moderator, as romantic partners might feel that punishment restores a feeling of fairness in the relationship when normative trespasses are severe, whereas punishment in the face of trivial deviances could lead to partners feeling unduly controlled. It might also be the case that norm tractability and agreement predict positive aspects of relationship quality, such as those included in the PRQC, whereas anticipated punishment for deviance may be more successful in predicting negative indicators of relationship quality (e.g., conflict, tension, anxiety, breakup thoughts).

Norm explicitness was generally unrelated to relationship quality. However, measurement problems likely hindered tests of this association (we later discuss this issue).

Cumulatively, the results of our studies suggest that relationship norm strength factors—particularly tractability and agreement—may be of rather large importance to the process of promoting relationship conformity and thereby maintaining high-quality romantic relationships. Indeed, relationship norm factors consistently explained a considerable amount of variation—in the realm of 30–50%—in relationship quality. We interpret the explanatory power of relationship norm strength as an encouraging sign that it will make a valuable contribution to normative perspectives on relationships and lead to fruitful investigations of the underpinnings of relationship quality.

Future research on relationship norm strength

From our examination of relationship norm strength and other features of relationship norms and relationship quality emerge a number of avenues for promising future research. In pursuing these and other applications of relationship norm strength, researchers may find it useful to employ the entire RNSQ measure, or conversely, may find it more efficient to employ only the items for the particular RNSQ factor(s) most pertinent to their research question (see https://osf.io/979fu/ reproducible measure and scoring resources).

The epistemic development of relationship norm tractability. Our first factor of norm strength—norm tractability—is unique in that it captures a heavily epistemic flavor of how relationship norms manifest (see also Knobloch & Solomon, 1999). As our results suggest, feeling confused and overwhelmed about one's relationship norms is not only deleterious for relational conformity, but it is also reliably linked to lower relationship quality. More formative epistemic processes regarding relationship norms, however, remain poorly understood. How do relationship members determine the content of their norms in a given domain, and how do they identify that they require norms in this domain in the first place? Do relationship members anticipate areas of friction and preemptively identify rules to help them avoid it, or are norms formed in a more reactionary manner in

response to an occurring relational breach? Theory abounds attempting to explain the developmental precursors to the establishment of norms in groups like relationships (e.g., Argyle & Henderson, 1984; Sherif, Harvey, White, Hood, & Sherif, 1988), each centering a different process as the key to norm formation, but strong tests of these specific claims are hard to come by. How norms in relationships come to be and how relationship members come to understand them, however, are both interesting and important questions needing further attention.

The unique importance of relationship norm agreement. The unique importance of relationship norm agreement for relationship quality was repeatedly supported throughout our studies. In essence, we liken the perception of relationship norm agreement as being akin to feeling "on the same page" with one's partner, with respect to the major proscriptions and prescriptions of their relationship. There are, however, a number of ways in which relationship partners might perceive themselves to be similar (e.g., attractiveness levels, hobbies, political identification). We speculate that relationship norm agreement may be such an important area of perceived relationship agreement because perceived relationship norm agreement reflects not only fairness in the relationship but also a sense of shared relational values. Thus, relationship partners perceiving high norm agreement may feel as though they and their partner are subject to the same relational rules and expectations, and moreover, that these rules are good and reflective of the kind of relationship both partners wanted to be in. These perceptions should translate to a positive perception of the relationship itself (La Guardia & Patrick, 2008). This process may be one way in which relationship norm strength could ultimately facilitate relational goal congruence between partners (Gere & Schimmack, 2013).

Clarifying norm explicitness. In the current work, we consistently used analytic techniques for increasing the precision of our hypothesis tests. These analytic approaches helped us to maximize statistical power and allowed us to identify areas of improvement in the RNSQ. In particular, measurement of one factor—norm explicitness—was a consistent challenge throughout our studies. Although this factor emerged repeatedly across our measurement development studies, it was primarily defined by only 2 items, making it difficult to get reliable estimates for its association with relationship quality. Thus, while we are confident that norm explicitness is a reliably detectable feature of norm strength, its association with relationship quality must be clarified. Moving forward, adding relevant items to improve the measurement resolution of norm explicitness, or using other methods (e.g., experimental, therapeutic observation), will be fruitful in examining its association with relationship quality. However, whereas factors for tractability, agreement, and anticipated punishment were reliably extractable and amenable to modeling across our studies, the problems introduced when attempting to model the explicitness factor lead us encourage caution for researchers using the RNSQ to assess the explicitness factor.

Reinforcement in relationships and anticipated punishment for deviance. Of our four norm strength factors, anticipated punishment appeared to have the most complicated role within our theoretical model of normative relationship processes. In particular, it seemed as though there was a precarious balance in which anticipated punishment could improve

norm-relevant motivation (and, unreliably, conformity) in the short run, but in the long-term might ultimately have a negative impact on norm-relevant motivation and behavior. Inducing the fear of punishment for rule breaking in relationships may therefore be a rather short-cited and risky method of increasing desirable relationship conduct (e.g., Argyle & Henderson, 1984).

Broader directions for future research. Our investigation of norm strength adds to the already large literature on normative processes in close relationships. Researchers have stressed the importance of other qualities of relationship norms, including their content (e.g., Clark & Mills, 1979, 2011), and individuals' motivation to follow these norms (e.g., Mills et al., 2004). The study of these constructs implies that norm-relevant behavior is also likely to be an important feature of relationships. Currently, however, the relationship norms literature does not make plain how these features of norms—content, motivation, and behavior—ought to be organized into a coherent theory. With our introduction of relationship norm strength, we expand the broader research on relationship norm content and motivation, laying the foundation for the creation, testing, and revision of a theoretical model of how relationship norms impact close relationships, wherein relationship norm strength—and other aspects of norms—can be situated, and around which more specific interventions could be designed to help struggling couples. Given, for example, the considerable downstream effects of stronger norms on conformity levels in the absence of any corresponding effects on norm motivation, the question of how stronger norms transmit their effects to behavior is a compelling target of subsequent research.

A second generative route for future research on relationship norm strength, broadly construed, would be to evaluate whether norms—and thereby forms of norm strength—serve particular functions within a given relationship, such as by helping partners to better understand their shared social environment (Sherif, 1936), bringing their attitudes and beliefs into uniformity (Festinger, Schachter, & Back, 1950), and/or coordinating their collective action to better achieve relationship goals (Sherif et al., 1988). We think that longitudinal and experience sampling studies (see Mehl & Connor, 2012) would likely prove illuminating to evaluate possible functions of relationship norms and their strength. If strong relationship norms help to coordinate group action, for example, we might expect the association between norm strength and relationship quality to strengthen during times of collective stress for couples (e.g., after childbirth) and to facilitate congruence of goals between partners, more broadly, across the span of the relationship (Gere & Schimmack, 2013).

Finally, it remains unclear whether and to what extent relationship norm strength causally impacts relationship quality, as opposed to opposite (or even simultaneous/reciprocal) patterns of causation. Although our longitudinal sample and modelling strategy would have ideally enabled us to evaluate different patterns of short-term causation (Zyphur et al., 2019), relationship quality was so stable over our 8-week study that none of its effects—including its autoregressive pathways—were significant. Therefore, future research should consider the use of experimental manipulations of both/either relationship norm strength factors and/or relationship quality, and longitudinal designs of longer duration that might afford the evaluation of a more variable window of relationship quality.

Conclusion

Relationship partners establish norms to guide their conduct in a number of important domains, including parenting, sexuality, benefit-giving, and communication. In the present study, we provided evidence to suggest that *strength* of these norms is both an important determinant of conformity within relationships and a reliable correlate of the quality of romantic relationships. Looking beyond the present investigation, we foresee the emergence of the relationship norm strength construct not only as an important contributor to the well-being of relationships but also to a richer theoretical understanding of the organization of normative processes within close relationships.

ORCID iD

John K. Sakaluk ⬤ https://orcid.org/0000-0002-2515-9822
Bonnie M. Le ⬤ https://orcid.org/0000-0002-1857-6716
Emily A. Impett ⬤ https://orcid.org/0000-0003-3348-7524

Open research statement

As part of IARR's encouragement of open research practices, the author(s) have provided the following information: This research was a mix of not pre-registered (Studies 1–4, 7) and pre-registered (Studies 5 and 6) research. The aspects of the research that were pre-registered include hypotheses, measures, and data analysis plans for Studies 5 and 6. All data, materials, and the relevant preregistrations for these studies can be found at our Open Science Framework Page: https://osf.io/979fu/. The data used in the research are available. The data can be obtained at the respective OSF component for each study. The materials used in Studies 1–4 are not available at our OSF page, but most can be found in the appendixes of the first author's dissertation.

Supplemental material

Supplemental material for this article is available online https://osf.io/979fu/.

Notes

1. The short form of the PRQC did not exhibit acceptable local model fit, so we opted to parcel the measure by combining the first and second, third and fourth, and fifth and sixth items, given that the full measure fit well in our other samples, and the validation and refinement of the PRQC was not a focus of our research (see Little, Cunningham, Shahar, & Widaman, 2002, for a discussion on parceling).

2. These samples were selected because the causal direction of the association in each could conceivably be from RNSQ to relationship quality (excluding Study 3, in which relationship quality was manipulated), as well as each using the PRQC as the measure of relationship quality.

3. We were not able to evaluate longitudinal prediction of relationship quality because there was so little variability in week-to-week relationship quality; even the auto-regressive paths for relationship from one week to the next were nonsignificant.

References

Acitelli, L. K. (1988). When spouses talk to each other about their relationship. *Journal of Social and Personal Relationships*, *5*, 185–199. doi:10.1177/026540758800500204

Argyle, M., & Henderson, M. (1984). The rules of friendship. *Journal of Social and Personal Relationships*, *1*, 211–237. doi:10.1177/0265407584012005

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397–438. doi:10.1080/10705510903008204

Bargh, J. A., Schwader, K. L., Hailey, S. E., Dyer, R. L., & Boothby, E. J. (2012). Automaticity in social-cognitive processes. *Trends in Cognitive Sciences*, *16*, 593–605. doi:10.1016/j.tcs.2012.10.002

Barker, M. (2012). *Rewriting the rules: An integrative guide to love, sex, and relationships*. New York, NY: Routledge.

Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001). Applicant reactions to selection: Development of the selection procedural justice scale (SPJS). *Personnel Psychology*, *54*, 387—419.

Beaujean, A. A. (2014). *Latent variable modeling using R: A step-by-step guide*. New York, NY: Routledge.

Buss, D. M. (1992). Manipulation in close relationships: Five personality factors in interactional context. *Journal of Personality*, *60*, 477–499.

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

Clark, M. S., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology*, *37*, 12–24. doi:10.1037/0022-3514.37.1.12

Clark, M. S., & Mills, J. (1993). The difference between communal and exchange relationships: What it is and is not. *Personality and Social Psychology Bulletin*, *19*, 684–691. doi: 10.1177/0146167293196003

Clark, M. S., & Mills, J. (2011). A theory of communal (and exchange) relationships. In P. A. M. van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 232–250). Los Angeles, CA: SAGE.

Diener, E., & Seligman, M. E. P. (2002). Very happy people. *Psychological Science*, *13*, 81–84.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. doi:10.3758/BRM.41.4.1149

Festinger, L., Schachter, S., & Back, K. (1950). *Social pressures in informal groups: A study of human factors in housing*. Stanford, CA: Stanford University Press.

Fletcher, G. J. O., Simpson, J. A., & Thomas, G. (2000). The measurement of perceived relationship quality components: A confirmatory factor analytic approach. *Personality and Social Psychology Bulletin*, *26*, 340–354. doi:10.1177/0146167200265007

Gelfand, M. J., Rana, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., . . . Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*, 1100–1104. doi: 10.1126/science.1197754

Gere, J., & Schimmack, U. (2013). When romantic partners' goals conflict: Effects on relationship quality and subjective well-being. *Journal of Happiness Studies*, *14*, 37–49. doi:10.1007/s10902-011-9314-2.

Harrington, J. R., & Gelfand, M. J. (2015). Tightness and looseness across the 50 United States. *Proceedings of the National Academy of Sciences*, *111*, 7990–7995. doi:10.1073/pnas.1317937111

Hoffman, J. (2011). *Relationship rules: 12 strategies for creating a love that lasts*. Boulder, CO: Venus Publishing.

Holt-Lundstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Medicine*, *7*, e1000316.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185. doi:10.1007/BF02289447

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structureal Equation Modeling*, *6*, 1–55. doi:10.1080/10705519909540118

Knobloch, L. K., & Solomon, D. H. (1999). Measuring the sources and content of relational uncertainty. *Communication Studies*, *50*, 261–278. doi:10.1080/10510979909388499.

Kurdek, L. A. (1994). Areas of conflict for gay, lesbian, and heterosexual couples: What couples argue about influences relationship satisfaction. *Journal of Marriage and Family*, *56*, 923–934. doi:10.2307/353603

La Guardia, J. G., & Patrick, H. (2008). Self-determination theory as a fundamental theory of close relationships. *Canadian Psychology/Psychologie Canadienne*, *49*, 201–209. doi:10.1037/a0012760

Le, B. M., Impett, E. A., Lemay, E. P., Jr., Muise, A., & Tskhay, K. O. (2018). Communal motivation and well-being in interpersonal relationships: An integrative review and meta-analysis. *Psychological Bulletin*, *144*, 1–25. doi:10.1037/bul0000133

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*, 151–173. doi:org/10.1207/S15328007SEM0902_1

Maass, A., Cadinu, M., & Galdi, S. (2013). Sexual harassment: Motivations and consequences. In M. K. Ryan & N. R. Branscombe (Eds.), *Handbook of gender and psychology* (pp. 341–358). London, UK: SAGE.

Mark, K. P., Janssen, E., & Milhausen, R. R. (2011). Infidelity in heterosexual couples: Demographic, interpersonal, and personality-related predictors of extradyadic sex. *Archives of Sexual Behavior*, *40*, 971–982. doi:10.1007/s10508-011-9771-z

Meagher, B. R. (2019). Ecologizing social psychology: The physical environment as a necessary constituent of social processes. *Personality and Social Psychology Review*. doi:10.1177/1088868319845938.

Mehl, M. R., & Connor, T. S. (2012). *Handbook of research methods for studying daily life*. New York, NY: Guilford Press.

Meyer, R. D., Dalal, R. S., & Hermida, R. (2010). A review and synthesis of situational strength in the organizational sciences. *Journal of Management*, *36*, 121–140. doi:10.1177/0149206309349309

Mills, J., Clark, M. S., Ford, T. E., & Johnson, M. (2004). Measurement of communal strength. *Personal Relationships*, *11*, 213–230. doi:10.1111/j.1475-6811.2004.00079.x

Mischel, W. (1973). Towards a cognitive social learning reconceptualization of personality. *Psychological Review*, *80*, 252–283. doi:10.1037/h0035002

Muise, A., & Impett, E. A. (2016). Applying theories of communal motivation to sexuality. *Social and Personality Psychology Compass*, *10*, 455–467. doi:10.1111/spc3.12261

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide: Seventh edition*. Los Angeles, CA: Muthén & Muthén.

Paulhus, D. L. (1998). *Manual for the balanced inventory of desirable responding (BIDR-7)*. Toronto, Canada: Multi-Health Systems.

Pelto, P. J. (1968, April). The difference between "tight" and "loose" societies. *Transaction*, 37–40.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Revelle, W. (2016). *psych: Procedures for personality and psychological research (version 1.6.6)*. Evanston, IL: Northwestern University.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36.

Sakaluk, J. K., & Short, S. D. (2017). A methodological review of exploratory factor analysis in sexuality research: Used practices, best practices, and data analysis resources. *Journal of Sex Research*, *54*, 1–9. doi:10.1080/00224499.2015.1137538

Sakaluk, J. K., Todd, L. M., Milhausen, R., Lachowsky, N. J., & the Undergraduate Research Group in Sexuality. (2014). Dominant heterosexual scripts in emerging adulthood: Conceptualization and measurement. *Journal of Sex Research*, *51*, 516–531. doi:10.1080/00224499.2012.745473

Schönbrodt, F. D. (2015). *p-checker: One-for-all p-value analyzer*. Retrieved from http://shinyapps.org/apps/p-checker/.

semTools Contributors. (2016). *semTools: Useful tools for structural equation modeling*. Retrieved from https://cran.r-project.org/web/packages/semTools/index.html.

Sherif, M. (1936). *The psychology of social norms*. New York, NY: Harper.

Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1988). *The robbers cave experiment: Intergroup conflict and cooperation*. Middletown, CT: Wesleyan University Press.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and Family*, *38*, 15–28. doi:10.2307/350547

ten Berge, J. M. F., Krijnen, W. P., Wansbeek, T., & Shapiro, A. (1999). Some new results on correlation-preserving factor scores prediction methods. *Linear Algebra and its Applications*, *289*, 311–318. doi:10.1016/S0024-3795(97)10007-6

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, *96*, 506–520. doi:10.1037/0033-295X.96.3.506

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*, 4–70.

Zyphur, M. J., Allison, P. D., Tay, L., Voelkle, M. C., Preacher, K. J., Zhang, Z., . . . Diener, E. (2019). From data to causes I: Building a general cross-lagged panel model (GCLM). *Organizational Research Methods*. doi:10.1177/1094428119847278.