

A longitudinal investigation of social norms and social networks in science

Holly J. Kalyn Bogard

holly.kalynbogard@uleth.ca

University of Lethbridge

Justin Petluk

University of Lethbridge

Calvin Lloyd

University of Lethbridge

Louise Barrett

University of Lethbridge

Research Article

Keywords: academic social norms, social networks, trainee satisfaction, Bayesian longitudinal analysis, supervisor support

Posted Date: August 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10541446/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Research on scientists' social norms and networks has relied largely on cross-sectional designs, which cannot address whether observed patterns reflect stable features or artifacts of a single measurement occasion. We addressed this gap using four waves of survey data (six-month intervals) collected from 93 Canadian academic trainees (graduate students and postdoctoral fellows), evaluating the robustness of findings from two previous cross-sectional studies on social norms and networks in academia. We examined repeated-measures stability, time-lagged associations, and variance decomposition across five research questions using Bayesian mixed-effects models. Mean responses across 17 social norm items were robust across waves, with substantial between-person variance (25–64% of total variance) but also some detectable within-person fluctuation. Academic support emerged as a central correlate of both current experience and research productivity, with financial support and research productivity contributing additional, smaller effects to current experience. Having more “very” close and fewer difficult relationships predicted academic support more than the number of members providing “a lot” of technical support. Comfort discussing academic failure with the supervisor was the strongest single predictor of trainee satisfaction, with stable differences both between- and within-persons. Group differences emerged across outcomes for gender, international status, underrepresentation, and career stage, though most estimates were modest and imprecise. Network characteristics and satisfaction were poor indicators of perceived adherence to Mertonian norms in our sample. These findings indicate that trainees' social experiences and satisfaction reflect a combination of enduring individual differences and responsiveness to changing conditions, with implications for promoting positive research experiences for trainees.

1.0 Introduction

Empirical research on scientists' beliefs, norms, and social relationships has primarily used cross-sectional survey designs. These are cost-effective, time-efficient and suitable for estimating the ‘current state’ of a population (e.g., the prevalence of a behaviour, attitude, or norm), as well as comparing subgroups, and mapping preliminary associations— especially when a snapshot is more informative than patterns of change or causality within a population. While cross-sectional designs have produced a substantial body of work on scientists' experiences within their academic environments (e.g., Anderson et al., 2007; Brooker & Allum, 2024), how they perceive normative expectations in social behaviour (e.g., Anderson, 2000; Bieliński & Tomczyńska, 2019; Bray & Von Storch, 2017; Kim & Kim, 2018), and the ways that they draw on social resources within their networks (e.g., Deanna et al., 2022; Heffernan, 2021; Li et al., 2022; Liu et al., 2026), there remain well-documented limitations. Without longitudinal evidence, it is difficult to evaluate whether findings represent enduring and stable patterns, and developmentally appropriate career-stage transitions, or whether they represent transient snapshots of a system in dynamic flux.

Here, we treat response stability as an empirical question, rather than an assumption, with direct implications for future investigations of social behaviour in science more broadly. Using four waves of

data collected between 2024 and 2026 from a sample of academic trainees (graduate students and postdoctoral fellows) in Canada, we evaluate the temporal stability of findings from two previous studies: (1) normative perceptions of social norms in science among faculty and trainees (Kalyn Bogard & Barrett, 2026) and; (2) associations between trainee satisfaction and social resource exchange from their academic social networks (Kalyn Bogard et al., 2026). Although our sample size ($n = 93$) precluded the use of highly parameterized dynamic models (e.g., dynamic structural equation models, latent growth models, panel analysis), we leveraged repeated observations to evaluate robustness, stability, and temporal structure of previously observed associations using three complementary longitudinal analytical frameworks:

(1) Repeated measures analyses: to assess whether associations observed in cross-sectional analyses remain stable or vary over time (Gelman & Hill, 2006);

(2) Time-lagged analyses: to examine whether earlier values predict subsequent changes in related outcomes, thereby providing evidence regarding temporal ordering and relative within-person variance (Gelman & Hill, 2006).

3) Variance decomposition: to distinguish stable individual differences between people from within-person change and to avoid conflating these sources of variation (Curran & Bauer, 2011).

Together, these approaches allow us to move beyond static descriptions of association and toward a more nuanced understanding of how social norms, social support, and satisfaction develop during academic training. We evaluated five substantive research questions derived directly from previously reported cross-sectional findings, which we summarize below.

1.1 Previous findings

Our previous and current research adopts a social epistemological perspective, which views scientific knowledge as a product of organized communities rather than isolated individuals (Longino, 1990). Scientists rely on their peers for collaboration, critique, and validation of their research findings, making social behaviour paramount to maintaining process-level integrity and objectivity in science (McIntyre, 2019; Oreskes, 2021). For instance, there are cases where the erosion of conventional social norms—namely scientists not being permitted to transparently view and scrutinize another’s work to protect profitable research—leads to less trust between individuals (Bright & Heesen, 2023), and more suspicion of scientific outputs. Therefore, in our previous investigations, we explicitly documented social behaviour in science, and did so through two complementary themes: social norms (Kalyn Bogard & Barrett, 2026) and social networks (Kalyn Bogard et al., 2026).

The social norms study asked 615 academic faculty and trainees to rate how well various social norms in science captured how scientists *should* behave versus how scientists *actually* behave. The respondents’ perceptions revealed a clear tension: researchers strongly endorsed traditional normative frameworks in science as ideal (i.e., Mertonian and other academic norms), yet acknowledged that in

practice, these norms are not fully realized. Instead, industrial and counternormative behaviours are more prevalent than researchers would prefer, especially with respect to factors like careerism, professional status, and competition for funding. Exploratory analyses detected subtle group-level response differences by gender, career stage, international status, research discipline and experience with industrial forms of science. The present longitudinal study provides an opportunity to assess whether these response patterns and group differences are stable features of how trainees perceive scientific norms or whether they are artifacts of a single measurement occasion.

The social network study examined whether social resource exchanges and affective relational ties within academic social networks were associated with various aspects of trainee satisfaction among graduate students and postdoctoral fellows. Overall satisfaction with the current experience was associated with greater satisfaction with academic support, research productivity, and financial support, as well as with having more network members with whom the respondent felt comfortable discussing academic struggles or setbacks (i.e., emotional support) and fewer members with whom the respondent found it 'difficult' to work alongside. Interestingly, these affective relational ties were better predictors of academic support satisfaction than the number of network members providing 'a lot' of technical support. Emotional support from the supervisor was especially instrumental in predicting satisfaction with both academic support and the current experience. Career-stage effects were apparent: more senior trainees had higher odds of providing technical support, and postdocs were especially notable for providing high levels of technical and emotional support relative to other professional academic roles. The present study therefore aims to (i) evaluate the robustness of these associations under repeated measurement, (ii) examine whether sustained or improving supervisor emotional support predicts greater satisfaction over time, (iii) test the persistence of group-level differences, and (iv) assess whether trainees with more supportive networks or greater satisfaction hold more positive evaluations of enacted Mertonian norms.

2.0 Methods

2.1 Survey distribution and respondent demographics

Detailed descriptions of the social norms and social network studies are given in Kalyn Bogard & Barrett (2026) and Kalyn Bogard et al. (2026), respectively. Briefly, 93 trainees (45 MSc students, 36 PhD students, 11 postdoctoral fellows, 1 graduated student) participated in both surveys during the initial cross-sectional phase of data collection. Respondents from the Universities of Calgary and Lethbridge completed wave 1 (April 2024) six months before respondents from the University of Alberta (October 2024). Both cohorts were invited by email to complete follow-up versions of both surveys at six-month intervals over 18 months, capturing four waves of data in total. To minimize anchoring or carryover effects, previous responses were cleared before each wave of both surveys. Respondents named 20 alters and provided demographic details for their personal social network. This information was retained between waves to avoid respondent fatigue and potential participation drop-out due to recreating their network at each measurement occasion. Importantly, however, respondents who wished to make

changes to their 20-alter network at each wave could do so if desired. Respondents received a small honorarium upon completion of each survey to encourage retention. Attrition was monitored and documented to account for those who graduated or left graduate school, as well as those lost to follow-up (Kristman et al., 2004).

Overall, 80.7% (n = 75) of respondents completed wave 2 of the survey, 89.3% (n = 67) completed wave 3, and 79.1% (n = 53) of respondents completed wave 4 (Table 1). Attrition across all four waves was modest and partially attributable to transitions out of graduate training (n = 10, 10.8%) and those lost to follow-up (n = 30, 32.3%). Four respondents transferred career stage between waves. No evidence of systematic demographic imbalance was detected across waves for those who dropped out.

Table 1

Descriptive summary of respondents across four survey waves. Values represent counts and percentages (%) within each wave.

Variable	Category	Wave 1	Wave 2	Wave 3	Wave 4
		N (%)	N (%)	N (%)	N (%)
Total		93	75	67	53
Gender	Woman	52 (55.9)	47 (62.7)	39 (58.2)	32 (60.4)
	Man	38 (40.9)	26 (35.7)	26 (38.8)	19 (35.9)
	Other	1 (1.1)	-	-	-
	Prefer not to answer	2 (2.2)	2 (2.7)	2 (3.0)	2 (3.8)
Academic position	Master's student	45 (48.4)	39 (52.0)	31 (46.3)	25 (47.2)
	PhD student	36 (38.7)	28 (37.3)	28 (41.8)	21 (39.6)
	Post-doctorate fellow	11 (11.8)	8 (10.7)	8 (11.9)	7 (13.2)
	Other	1 (1.1)	-	-	-
International students	No	59 (63.4)	49 (65.3)	43 (64.2)	36 (67.9)
	Yes	33 (35.5)	25 (33.3)	23 (34.3)	16 (30.2)
	Prefer not to answer	1 (1.1)	1 (1.3)	1 (1.5)	1 (1.9)
Underrepresented students	Yes	47 (50.5)	41 (54.7)	36 (53.7)	28 (52.8)
	No	30 (32.3)	23 (30.7)	20 (29.9)	15 (28.3)
	I'm not sure	11 (11.8)	8 (10.7)	7 (10.5)	7 (13.2)
	Prefer not to answer	5 (5.4)	3 (4.0)	4 (6.0)	3 (5.6)

2.2 Research questions for longitudinal investigation

We examined five central research questions (RQ) using one or more longitudinal analytic frameworks described below:

RQ1) To what extent are the responses to the 17 'ideal' and 'actual' social norm survey items stable over time, and how much of the variance in these responses is attributable to between-person versus within-

person differences?

RQ2) Do associations between social network characteristics and trainee satisfaction persist over time in predicting trainees' satisfaction with their current experience in their research program?

RQ3) To what extent do between-person and within-person differences in trainees' comfort levels discussing academic failure with their supervisor(s) predict satisfaction with academic support and their current experience?

RQ4) Do gender, career stage, international student status, and perceived underrepresentation predict response variation in social norm survey responses, social network characteristics, and satisfaction outcomes?

RQ5) Which aspects of research culture are associated with more positive evaluations of enacted Mertonian norms, including network characteristics, satisfaction, perceived value congruency among colleagues, frequency of epistemic-value discussions, and concern about academic misbehaviour?

2.3 Modelling framework

All analyses were conducted using a Bayesian framework, in which parameter estimates are drawn from posterior distributions that combine the observed data with prior information. This framework allows uncertainty around each estimate to be explicitly captured and is well-suited for use with small or unbalanced samples. All models were multivariate, multilevel, and at times, jointly estimated to account for relatedness among outcome variables using the *brms* package in R (Bürkner, 2018) via Markov Chain Monte Carlo (MCMC) sampling. Outcome variables were modelled using a cumulative distribution for ordinal outcomes with a logit link (e.g., social norm items, measures of satisfaction), a binomial distribution (e.g., proportion-based counts of ties out of 20 alters), or a skew (normal) distribution when modelling network characteristic variables as outcome variables, handling the bounded and skewed nature of count data. All models used four chains using a minimum of 6,000 iterations per chain (3,000 warmup). Convergence was assessed using standard diagnostics, including inspection of model summary statistics, $\hat{R}$ (< 1.01), effective sample sizes, trace plots, and posterior predictive checks. Posterior summaries are reported as means and 95% credible intervals (CrI). For interpretability, log-odds estimates were reported as odds ratios in the text where appropriate.

2.4 Random Effects

We explored the appropriate level of model complexity by comparing alternative models that differed in their random-effects structure: a simpler random intercept by respondent model and a more complex random intercept plus random slope model. Random slope variances in the models were small with wide credible intervals, and intercept-slope correlations were poorly identified. Fixed-effect estimates were nearly identical across models, indicating that the slope terms were not contributing meaningfully to explain longitudinal patterns. Model fit was also evaluated using approximate leave-one-out cross-validation (LOO) with Pareto-smoothed importance sampling, as implemented in the *loo* package

(Vehtari et al., 2017). In some cases, the more complex specification yielded a higher estimated expected log predictive density (ELPD), but its LOO estimates exhibited extensive Pareto-k failures ($k > 0.7$ for a large proportion of observations), indicating unreliability. These cases also produced inflated estimates for the random effects, likely due to the relatively small sample size. In contrast, the simpler model showed no diagnostic issues and produced stable estimates. Thus, we retained the simpler intercept-only random-effects structure for all analyses, which included wave as a categorical fixed effect to allow for non-linear changes across measurement occasions.

2.5 Sensitivity analyses

Three main sensitivity analyses were conducted. First, to detect potential response differences between respondents who remained in the study versus those who dropped out after the initial measurement, we compared mean responses for all survey items (i.e., social norms, social networks, and satisfaction). Next, all responses were checked in each wave for an abundance of non-responses and unusual patterns across responses. We detected 19 of 3,162 instances where the respondent registered 2–3 extreme responses in one direction (e.g., 10) on an ‘ideal’ item, and then flipped to the opposite pole (e.g., 0) in only one instance. Although we cannot be certain, we considered these to be accidental responses due to inattention rather than such an extreme change in perspective on only one occasion. Consequently, comparisons were conducted on both the raw and ‘cleaned’ datasets for the repeated measures and variance composition analyses for all social norm items. Lastly, sensitivity analyses across a reasonable range of priors were conducted for both fixed (i.e., normal (0, 0.5), normal (0,1), normal (0, 2.5)) and random effects (i.e., exponential (1), student-t (3, 0, 2.5), student-t (3, 0, 10)), to ensure that model estimates were not overly influenced by fluctuations in model specification.

2.6 Repeated measures analyses

Repeated measures analyses maintained a core structure across all five research questions. The measurement occasion (wave) was included as a fixed effect (factor) to account for wave-level changes in the outcomes without imposing a linear assumption. Models included multilevel random effects to account for the non-independence of repeated observations and to estimate between-person differences.

We assessed response stability of 17 social norm items (RQ1) by fitting a wave-only model with random effects, separately for ‘ideal’ and ‘actual’ ratings per item. Item responses were measured on an 11-point ordinal scale (0–10) and exhibited moderate skewness for several ‘ideal’ items.

Sample size precluded the use of a dynamic structural equation model (SEM) to replicate the SEM from previous work (RQ2). Instead, we fit jointly estimated multilevel models capturing the strongest associations from that model: current experience was predicted by academic support, financial support, research productivity, technical support, emotional support, difficulty, and closeness; academic support was predicted by technical support, emotional support, difficulty, and closeness; and research productivity was predicted by academic support, financial support, difficulty, and closeness. Wave-only models were also fit for each outcome and predictor separately.

The association between supervisor emotional support and satisfaction outcomes (RQ3) was tested with comfort discussing academic failure with one's supervisor (not comfortable, somewhat comfortable, very comfortable) as a fixed effect alongside wave and random effects.

We examined whether responses on social norms items, network characteristics and satisfaction measures differed by gender, career stage, international student status, and perceived underrepresentation in the research field (RQ4). Given the exploratory nature of these analyses and the increased risk of detecting statistically certain but practically negligible effects, attention was restricted to the Mertonian norm items, for which theoretical and empirical expectations regarding group differences were strongest. Where mean 'actual' ratings exceeded corresponding 'ideal' ratings within a group, this was noted as indicating the behaviour was not viewed as normatively desirable, relative to non-group counterparts.

Finally, we tested whether trainees with more supportive social networks, more positive evaluations of satisfaction variables, or who perceive more value congruency with faculty or peers, more frequent discussions of epistemic values, or lower concern with academic misbehaviour reported higher rates of enacted Mertonian norms (RQ5). Given the large number of associations considered simultaneously, interpretation was limited to patterns that were either consistent across multiple items or involved single-item differences that were both substantively large and estimated with relatively high posterior certainty.

2.6 Time-lagged effects analyses

To examine temporal ordering, we fit a series of wave-specific multivariate cumulative ordinal regression models in which each outcome variable (i.e., current experience, academic support, and research productivity) at Wave_(i+1) was regressed on its Wave_i counterpart (RQ2). Trainees' perceived comfort level discussing academic failure with their supervisor(s) was also regressed at Wave_i by the prior level (RQ3). This offered complementary evidence to the repeated-measures analyses regarding the presence of lagged influences across measurement occasions.

2.8 Variance decomposition analyses

To partition response variance between and within respondents across measurement occasions for social norm items (RQ1), we estimated variance partition coefficients (VPC, Goldstein et al., 2002; Grilli & Rampichini, 2007) to represent the proportion of total latent variance attributable to stable between-person differences. This approach aligns with our goal of variance decomposition and is better suited to the non-Gaussian nature of our data. VPC was estimated from cumulative logit multilevel models (Section 2.6) that included a random intercept for respondent, treating within-person variance as fixed (i.e., a standard logistic distribution variance: $\pi^2/3 \approx 3.29$) and estimating between-person variance from the posterior distribution of the random-intercept standard deviation. The VPC for each item was then computed as a ratio of between-person variance to total variance (between + within) across all posterior draws and summarized as the posterior median with 95% credible intervals.

For each fixed effect tested in RQ2-3, we decomposed each predictor into (1) a person-specific mean across waves and (2) a wave-level deviation from that mean (Hoffman & Stawski, 2009). Outcome variables were then modelled using fixed effects for both components with wave as a factor, and random respondent-level intercepts. The person-mean term captures whether respondents who are generally higher or lower on the predictor tend to report systematically different average outcomes. In contrast, the deviation-from-mean term indicates whether an individual's outcome changes on waves when they are above or below their own typical level.

3.0 Results

3.1 Sensitivity analyses

Effect sizes were small, with no statistically significant differences detected across any wave for satisfaction with academic support, financial support, research productivity, overall current experience in the research program, or any social network variables (i.e., emotional and technical support exchanged, or the number of close or difficult members reported) when comparing those who dropped out versus remaining in the study.

For the social norms items, a very small proportion (10.4%) of all possible cross-wave within-item comparisons (i.e., five of 34 cases) revealed small but statistically credible differences. The estimated range of odds ratios indicated these differences in mean response corresponded to a 23–58% increase in odds of a one-unit change on the 11-point numeric scale, which we considered trivial given the large number of differences tested across items and waves and the inconsistency of the effect across waves ($n = 48$). One exception emerged. The item describing Organized Skepticism 2 (i.e., that the claims of well-known scientists *actually* receive less scrutiny than their lesser-known peers) changed consistently across waves, such that respondents who remained in the study reported lower mean values compared to the first measurement occasion. Subsequent waves had between 24–51% higher odds of being in lower categories. Model summaries for all analyses here are reported in Table S1.

Sensitivity analyses comparing our raw and cleaned social norms item datasets (with the single aberrant response in the raw dataset (e.g., 10, 10, 0, 10) adjusted to match the majority of extreme responses in the cleaned version) indicated no meaningful differences in repeated-measures analyses (Table S2) or in VPC estimates across both datasets. We report the VPC results for both the raw and cleaned versions in Tables S3 and S4, respectively, and used the cleaned values for final analyses, as we believe this likely represents more 'true' responses and does not meaningfully influence the estimates or the ordering of items.

Prior-based sensitivity analyses indicated that model estimates for fixed effects were insensitive to variations in prior scale (Figure S1). Similarly, random-effect variance estimates were stable across alternative prior scales with only minor shifts in posterior standard deviations (Figure S2). These

changes did not alter the relative ordering of respondent-level effects or the overall hierarchical structure, suggesting that model conclusions are primarily driven by the data rather than prior assumptions.

3.2 RQ1: Social norms survey item responses

Mean responses to social norm items were very stable across waves (Fig. 1, Table S5), with no consistent patterns of change emerging across 53 repeated-measures analyses (Table S6). Only one 'ideal' (Universalism 1: $\beta = -0.70$, 95% CrI [-1.36, -0.04]) and three 'actual' (Disinterestedness 2: $\beta = -0.62$, 95% CrI [-1.21, -0.03]; Organized Skepticism 1: $\beta = 0.62$, 95% CrI [0.05, 1.20]; Transparency: $\beta = 0.65$, 95% CrI [0.10, 1.19]) items showed marginally different mean values in a single subsequent wave compared to wave 1. Overall, fixed effects were small and imprecisely estimated effects with 95% credible intervals overlapping zero for nearly all items. In contrast, random effects were consistently large across all items (weighted mean = 1.77, 95% CrI [1.29, 2.31]), indicating substantial individual differences in response.

Variance decomposition indicated substantial between-person variance across items, holding within-person variance constant (Fig. 2). VPC median values are interpreted on a scale where 0.50 marks the point at which stable individual differences and within-person fluctuations in response are equal contributors to total variance. Our items ranged from 0.246 (Governance 'ideal') to 0.635 (Disinterestedness 2 'ideal'), meaning that stable between-person differences accounted for approximately 25 to 64% of the total latent variance, with most items clustering near the parity point (VPC = 0.40–0.60). This suggests that scientific norm endorsement reflected a mix of both stable individual orientations and within-person fluctuations across waves. Five 'actual' (Communality 2, Governance, Quality, Transparency, and Expert) and two 'ideal' items (Quality and Transparency) had notably lower median VPC values (VPC ≤ 0.40). Inspecting the response distributions for these items revealed that between-person variation for the 'ideal' responses was genuinely low due to ceiling effects, whereby nearly all respondents strongly agreed with these items. In contrast, the 'actual' items had more normally distributed responses and are more likely to represent higher within-person variation.

3.3 RQ2: Predictors of satisfaction with current experience over time

Descriptively, mean network characteristics and satisfaction variables showed little variation across waves (Fig. 3, Table S7). In simple wave-only models, mean academic support decreased modestly but consistently after each subsequent wave (wave 2: $\beta = -0.70$, 95% CrI [-1.30, -0.12], wave 3: $\beta = -0.64$, 95% CrI [-1.23, -0.06]; wave 4: $\beta = -0.54$, 95% CrI [-1.22, 0.13], Table S8) although the latter effect was uncertain. Mean satisfaction with personal research productivity also decreased slightly across waves (wave 2: $\beta = -0.64$, 95% CrI [-1.22, -0.07], wave 3: $\beta = -0.11$, 95% CrI [-0.70, 0.49]; wave 4: $\beta = -0.70$, 95% CrI [-1.39, -0.03]), although not all effects were certain. Satisfaction with financial support showed a marginal improvement only in wave 4 compared to wave 1 ($\beta = 0.75$, 95% CrI [0.08, 1.43]). Lastly, current experience scores were lower in wave 4 compared to wave 1 ($\beta = -0.67$, 95% CrI [-1.35, 0.00]). In more

complex models, only the pattern observed for academic support and current experience remained certain. In contrast, no meaningful changes over time occurred for network characteristics (Table S9).

Across jointly estimated multivariate models, academic support was the central correlate of both current experience and research productivity (Fig. 4, Table S10). Trainees who reported greater academic support had higher odds of giving more positive ratings of current experience ($\beta = 0.42$, 95% CrI [0.24, 0.60]) and research productivity ($\beta = 0.51$, 95% CrI [0.34, 0.69]). Higher research productivity was also associated with more positive current experience ($\beta = 0.66$, 95% CrI [0.47, 0.85]), while greater satisfaction with financial support also predicted higher current experience ($\beta = 0.27$, 95% CrI [0.15, 0.39]) and research productivity ($\beta = 0.18$, 95% CrI [0.05, 0.30]). Random effects estimates indicated substantial between-person differences for all three outcome variables with consistent autoregressive effects across waves (Table S11).

Indicators of social network quality showed more selective effects, primarily on academic support (Fig. 4, Table S10). Feeling 'very close' to more people was associated with higher perceived academic support ($\beta = 0.22$, 95% CrI [0.09, 0.36]), whereas experiencing more difficult relationships predicted lower perceived academic support ($\beta = -0.14$, 95% CrI [-0.27, -0.01]). Technical support ($\beta = 0.08$, 95% CrI [-0.04, 0.19]) and emotional support ($\beta = 0.08$, 95% CrI [0.03, 0.19]) had smaller, positive, but uncertain effects on academic support. Social network measures showed the same direction of effect for research productivity and current experience, but these estimates were all much smaller and uncertain.

For variance decomposition, those reporting higher mean academic support ($\beta = 0.89$, 95% CrI [0.62, 1.17]), research productivity ($\beta = 0.55$, 95% CrI [0.30, 0.81]) and financial support ($\beta = 0.23$, 95% CrI [0.06, 0.41]) had higher odds of being more satisfied with their overall current experience (Table 2). Within-person fluctuations also predicted current experience: increases within a particular wave relative to a person's average in research productivity ($\beta = 0.58$, 95% CrI [0.37, 0.79]) and financial support ($\beta = 0.19$, 95% CrI [0.04, 0.35]) were positively associated with higher current experience, whereas within-person variation in academic support was not credibly associated ($\beta = 0.16$, 95% CrI [-0.04, 0.35]). Random effects continued to add explanatory power to the model, with notable individual differences between respondents' current experience.

Table 2

Mixed-effects model predicting current experience from the person-centred mean and within-person deviations from the mean effects for academic support, research productivity and financial support, controlling for survey wave (reference category: Wave 1) and random effects of the respondents (id). Posterior estimates are reported alongside the 95% credible intervals (CrI). Estimates with an asterisk (*) have a 95% CrI that excludes zero.

Outcome	Predictor	Estimate	95% CrI Lower	95% CrI Upper
<i>Model: Current experience ~ Academic support(mean) + Academic support(within) + Research productivity(mean) + Research productivity(within) + Financial support(mean) + Financial support(within) + factor(wave) + (1 id)</i>				
Current experience	Academic support (mean)	0.887*	0.624	1.171
Current experience	Academic support (within)	0.155	-0.041	0.354
Current experience	Research productivity (mean)	0.550*	0.299	0.808
Current experience	Research productivity (within)	0.578*	0.365	0.791
Current experience	Financial support (mean)	0.228*	0.059	0.410
Current experience	Financial support (within)	0.191*	0.036	0.349
Current experience	Wave 2	0.085	-0.485	0.654
Current experience	Wave 3	-0.079	-0.658	0.490
Current experience	Wave 4	-0.617	-1.269	0.030
	--			
Current experience	SD (Intercept) ID	1.408	0.978	1.861

Variance decomposition of network characteristics was generally a weak predictor of academic support (Table 3). Within-person increases in the number of network members providing emotional support ($\beta = 0.13$, 95% CrI [0.00, 0.26]), closeness ($\beta = 0.18$, 95% CrI [0.03, 0.33]), and technical support ($\beta = 0.08$, 95% CrI [-0.05, 0.22]) were all associated with higher academic support scores, although the latter effect was uncertain. In contrast, having a greater mean difficulty score was associated with lower academic support ($\beta = -0.33$, 95% CrI [-0.57, -0.10]). Other fixed effects all moved in the predicted direction of effect but showed either uncertain or negligible associations.

Table 3

Mixed-effects model predicting academic support from the person-centred mean and within-person deviations from the mean effects for the number of network members whom the respondent feels 'very comfortable' receiving emotional support from, gets 'a lot' of technical support from, finds difficult to interact with, or feels 'very' close to, controlling for survey wave (reference category: Wave 1) and random effects of the respondents (id). Posterior estimates are reported alongside the 95% credible intervals (CrI). Estimates with an asterisk (*) have a 95% CrI that excludes zero.

Outcome	Predictor	Estimate	95% CrI Lower	95% CrI Upper
<i>Model: Academic support ~ Emotional support(mean) + Emotional support(within) + Technical support(mean) + Technical support(within) + Difficulty(mean) + Difficulty(within) + Closeness (mean) + Closeness (within) + factor(wave) + (1 id)</i>				
Academic support	Emotional support (mean)	0.104	-0.093	0.309
Academic support	Emotional support (within)	0.130*	0.001	0.260
Academic support	Technical support (mean)	-0.029	-0.238	0.177
Academic support	Technical support (within)	0.083	-0.051	0.218
Academic support	Difficulty (mean)	-0.328*	-0.568	-0.097
Academic support	Difficulty (within)	-0.053	-0.201	0.096
Academic support	Closeness (mean)	0.128	-0.118	0.373
Academic support	Closeness (within)	0.182*	0.033	0.331
Academic support	Wave 2	-0.707*	-1.286	-0.138
Academic support	Wave 3	-0.819*	-1.421	-0.220
Academic support	Wave 4	-0.703*	-1.361	-0.049
	--			
Academic support	SD (Intercept) ID	2.522	1.948	3.194

3.4 RQ3: Emotional support from supervisors

Comfort level discussing academic failure (i.e., emotional support) with the supervisor(s) was the largest and most robust predictor of trainee satisfaction with academic support and overall current experience (Fig. 5, Table S12, Model 1). Respondents who reported being 'somewhat' comfortable discussing failure had 3-fold higher odds of reporting greater academic support ($\beta = 1.09$, 95% CrI [0.20, 1.97]) while being 'very' comfortable had 9-fold higher odds ($\beta = 2.23$, 95% CrI [1.22, 3.24]). Comfort level was also positively associated with current experience: trainees who reported feeling 'very' comfortable had markedly higher experience ratings (4-fold higher odds) compared to those not comfortable ($\beta =$

1.42, 95% CrI [0.45, 2.38]), while being 'somewhat' comfortable showed a smaller (1.6-fold higher odds), more uncertain positive effect ($\beta = 0.47$, 95% CrI [-0.40, 1.31]).

Trainees who showed higher mean comfort levels across waves had over 12-fold and 10-fold higher odds of reporting more satisfaction with both academic support ($\beta = 2.50$, 95% CrI [1.53, 3.47]) and current experience ($\beta = 2.33$, 95% CrI [1.28, 3.36]; Fig. 5, Table S12: Model 2). Within-person fluctuations in comfort were also positively associated with a 2.5-fold higher odds for academic support ($\beta = 0.92$, 95% CrI [0.29, 1.54]) and 1.6-fold higher odds for current experience ($\beta = 0.46$, 95% CrI [-0.13, 1.05]), with the latter estimate uncertain. Comfort in earlier waves was a strong predictor of comfort in subsequent waves (Table S12, Model 3).

3.5 RQ4: Group differences

Among the three trainee levels (i.e., MSc, PhD, postdoc), more senior roles had higher odds of providing technical support (PhD: $\beta = 0.44$, 95% CrI [0.10, 0.76], Postdoc: $\beta = 1.49$, 95% CrI [1.01, 1.98]; Table 4). Postdocs also had higher odds of providing emotional support than graduate students, but this effect was uncertain ($\beta = 0.30$, 95% CrI [-0.15, 0.77]). Consistent with previous findings, postdocs provided and reciprocated more support to graduate students relative to other professional roles (e.g., supervisors, professors, committee members; Figure S13). Postdocs had the highest odds of reciprocating emotional support ($\beta = 0.11$, 95% CrI [-0.15, 0.37]), and the second-highest odds of exchanging technical support ($\beta = 0.62$, 95% CrI [0.37, 0.87]), after supervisors. They also had the lowest odds of being considered difficult, though uncertain ($\beta = -0.53$, 95% CrI [-1.43, 0.29]).

Table 4

Mixed-effects models predicting technical support and emotional support offered by career stage (reference category: MSc students), including random effects of the respondents (id). Posterior estimates are reported alongside the 95% credible intervals (CrI). Estimates with an asterisk (*) have a 95% CrI that excludes zero.

Outcome	Predictor	Estimate	95% CrI Lower	95% CrI Upper
Technical support	Intercept 1	-0.521	-0.898	-0.150
	Intercept 2	2.471	2.054	2.902
	PhD students	0.436*	0.103	0.762
	Postdocs	1.493*	1.005	1.983
	SD (Intercept) ID	1.328	1.060	1.646
	--			
Emotional support	Intercept 1	-1.737	-2.151	-1.352
	Intercept 2	0.218	-0.157	0.582
	PhD students	-0.057	-0.368	0.264
	Postdocs	0.304	-0.151	0.770
	SD (Intercept) ID	1.313	1.028	1.653

When it came to the social norm items, group differences were modest and somewhat patterned for international trainees, women, trainees who felt underrepresented, and postdoctoral fellows (Table S14). Specifically, international students expressed lower mean support for 75% (6/8 items) of the 'ideal' and 62.5% (5/8) of the 'actual' norm items compared to domestic students, although only one of these 'ideal' effects was certain (Disinterestedness 1: $\beta = -0.99$, 95% CrI [-1.92, -0.06]). International students also had three instances where the mean 'actual' rating exceeded the mean 'ideal' rating, with one of these items showing different signs (+/-) of the estimate (Communality 2: 'ideal', $\beta = -0.68$, 95% CrI [-1.77, 0.35] and 'actual', $\beta = 0.17$, 95% CrI [-0.60, 0.91], Disinterestedness 1: 'ideal', $\beta = -0.99$, 95% CrI [-1.92, -0.06] and 'actual', $\beta = -0.22$, 95% CrI [-1.14, 0.70], Organised Skepticism 1: 'ideal', $\beta = -0.62$, 95% CrI [-1.48, 0.23] and 'actual', $\beta = -0.07$, 95% CrI [-0.92, 0.80]). International students were also more likely to indicate that 'scientists *do* consider the social attributes of the author when evaluating scientific claims' compared to their domestic peers (Universalism 2: $\beta = -0.84$, 95% CrI [-1.87, 0.16]), although the credible intervals were wide and uncertain.

On average, both women and respondents who felt underrepresented in their research field were more likely to disagree that scientists enacted Mertonian norms than their counterparts (100%; 8/8 items). Three of these effects were certain for women (Communality 2: $\beta = -1.03$, 95% CrI [-1.75, -0.35], Universalism 1: $\beta = -1.27$, 95% CrI [-2.13, -0.37], Universalism 2: $\beta = -1.03$, 95% CrI [-2.00, -0.03]), and none of these effects were certain for underrepresented trainees. Women and respondents who felt

underrepresented also disagreed, on average, that two norm items represented 'ideal' behaviour: Communality 2, that 'scientists should make findings freely accessible' (women: $\beta = -1.11$, 95% CrI [-2.12, -0.11], underrepresented: $\beta = -1.22$, 95% CrI [-2.18, -0.23]) and Disinterestedness 2, that 'no question is too controversial' (women: $\beta = -1.39$, 95% CrI [-2.39, -0.41], underrepresented: $\beta = -1.11$, 95% CrI [-2.14, -0.05]). There was considerable overlap between these groups, with nearly 80% of trainees identifying as both women and underrepresented in their research field.

Although all effects were uncertain with wide posterior distributions, postdocs consistently had lower odds than MSc students (reference group) to report that, in practice, scientists *do* adhere to Mertonian norms for 87.5% of items (7/8). They disagreed most strongly that 'scientists make findings freely accessible' (Communality 1: $\beta = -1.18$, 95% CrI [-2.37, 0.05]), and that 'scientists do not consider the social attributes of the author when evaluating scientific claims' (Universalism 2: $\beta = -1.09$, 95% CrI [-2.41, 0.25]).

Network characteristics exhibited considerable between-person variance, with several smaller but all uncertain group-level differences being detected in the repeated-measures analysis (Table S15). International trainees had approximately 3-fold higher odds than domestic trainees of offering technical support within their networks ($\beta = 1.03$, 95% CrI [-0.06, 2.05]). Women had higher odds of both giving ($\beta = 1.03$, 95% CrI [-0.39, 2.36]) and receiving ($\beta = 0.89$, 95% CrI [-0.45, 2.18]) emotional support, and having more emotionally reciprocal relationships ($\beta = 0.88$, 95% CrI [-0.60, 2.28]) compared to men. Of concern, both women ($\beta = 0.62$, 95% CrI [-0.81, 2.02]) and trainees who felt underrepresented ($\beta = 0.97$, 95% CrI [-0.47, 2.36]) tended to report more difficult individuals in their networks, but credible intervals included zero. Although uncertain, postdocs had about 3-fold lower odds of having 'very close' network members ($\beta = -1.11$, 95% CrI [-2.80, 0.63]) compared to MSc students.

Finally, mostly uncertain group differences emerged in the satisfaction measures (Table S16). International students ($\beta = -1.42$, 95% CrI [-2.50, -0.34]), underrepresented trainees ($\beta = -1.16$, 95% CrI [-2.25, -0.05]), and postdocs ($\beta = -0.74$, 95% CrI [-2.14, 0.66]) were more likely to express dissatisfaction with their financial support than their respective counterparts, although the last estimate was uncertain. Underrepresented trainees ($\beta = -1.29$, 95% CrI [-2.33, -0.25]) and women ($\beta = -0.75$, 95% CrI [-1.80, 0.29]) were more likely to report lower academic support, although the latter effect was smaller and uncertain. Underrepresented trainees also showed some indication of being less satisfied with their overall current experience ($\beta = -0.99$, 95% CrI [-2.07, 0.10]), though again, this effect was moderately uncertain. Random effects for all models contributed substantially to unexplained variance.

3.6 RQ5: Social networks and satisfaction predicting 'actual' responses to Mertonian items

Network measures (i.e., technical support, emotional support, closeness, or difficult members) did not show any consistent patterns in predicting adherence to Mertonian norms (Table S17). For the satisfaction variables, academic support was positively associated with six of the eight Mertonian norm items (75%), though nearly all estimates were small and uncertain (Table S18). Likewise, satisfaction

with current experience was positively associated with six of the eight items (75%), with only one certain effect for Organized Skepticism 1 ($\beta = 0.28$, 95% CrI [0.05, 0.51]). While these broader consistencies were apparent, the estimates were often negligible.

Finally, we tested whether trainees who felt greater value congruency with peers or faculty, participated in more frequent discussions about values, or had lower concern about questionable research practices or serious research misconduct also scored higher on the Mertonian 'actual' items (Table S19). Overall, there were no large or consistent trends across these predictors, except that trainees who felt greater value congruence with faculty tended to rate seven of the eight Mertonian items more positively. Of these, only two effects were certain (Disinterestedness 1: $\beta = 0.50$, 95% CrI [0.29, 0.72] and Organized Skepticism 1: $\beta = 0.32$, 95% CrI [0.12, 0.51]).

4.0 Discussion

This study evaluated whether key patterns identified in two prior cross-sectional investigations (Kalyn Bogard & Barrett, 2026; Kalyn Bogard et al., 2026) held across time. Across four waves of data, we found broad support for the stability of previously observed associations: trainees' perceptions of social norms, their satisfaction with academic support and current experience, and relational features of their social networks showed consistent patterns, with between-person differences accounting for a substantial portion of variance across all outcomes. Concurrently, within-person fluctuations were detectable and meaningful, suggesting that trainees' experiences were not wholly fixed. Trainees apparently respond, at least modestly, to changes in their immediate material and social conditions. This distinction between stable individual differences and within-person change has direct implications for how interventions and supports might be targeted to promote trainee wellbeing.

4.1 Social norms in tension

Our data indicated that mean response patterns were relatively stable, with consistent between-person differences in the random effects across all items, aligning with past work (Kalyn Bogard & Barrett, 2026). Trainees collectively endorsed academic norms as ideals, but reported that they are not enacted to the degree they prefer. In particular, norms related to competition, status, careerism, and funding pressures showed the largest gaps from ideal to actual response. Most items showed roughly equal between- and within-person fluctuations across waves, with no systematic differences by norm type (i.e., academic versus industrial) or response type ('ideal' versus 'actual'), although within-person differences detected across the 'ideal' item responses may be partially inflated, as most responses clustered near one pole of the numeric scale, exhibiting little between-person variance. This strengthens our assessment that scientists likely view social norms differently in degree, rather than wholly in kind. Small individual-level differences in value-ranking or prioritization are appropriate and allow researchers to express their personal values within a contextualized research environment. Group-level differences, like those observed for international trainees, may reflect genuine cultural differences in how scientific practice is structured across countries, particularly in systems that place stronger emphasis on

competition (Quan et al., 2017). These cultural differences are not inherently negative but highlight opportunities for mentors to guide international trainees in navigating differing cultural expectations in research environments across national systems.

Although some between-person variation plausibly reflects legitimate value pluralism, it also may indicate structural constraints or contextual factors that condition researchers to adopt more competitive, secretive, or otherwise unethical behaviours than they prefer. Observed differences by career stage, international status, gender, and underrepresentation suggest that individuals in these groups have personal experience with cultural differences, marginalization, exclusion, or overt competition in their academic settings. For example, women and those who felt underrepresented were more likely to disagree that work *should* be openly shared amongst scientists, perhaps reflecting a protectionist instinct rooted in evidence that women perform more technical research labour on average but are less frequently credited in prominent authorship positions (Larivière et al., 2016; Macaluso et al., 2016). This differential valuation of labour may lead women to feel less compelled to share their data openly for fear of not being properly credited. Similarly, both groups disagreed that ‘no question is too controversial for study’, possibly reflecting greater sensitivity to the social implications and potential harms of pursuing controversial research (Clark et al., 2023). These patterns suggest that marginalized researchers may approach specific norms with greater reflexivity, shaped by their experience of how these norms are unevenly applied. Finally, the Governance item (i.e., that administrators control research activities on campus), which had the lowest median VPC, was also among the most uncertain responses in the initial cross-sectional survey (i.e., trainees were allowed to indicate ‘I’m not sure’ to any of the items). While our data cannot address this distinction directly, higher rates of within-person variance may represent genuine epistemic uncertainty about academic governance for trainees.

We suggest, in line with others (Porter, 1995), that a shift from informal to formalized academic organization has inadvertently weakened social norms. Historically, scientific communities were small enough that norms could be negotiated informally through mentorship, reputation, and interpersonal trust (Daston, 2023). This occurred not because scientists were more virtuous (Shapin, 2010), but because scale permitted it. As science has expanded, these informal mechanisms have been displaced by standardized systems to quantify research productivity in ways that are designed to manage larger, more heterogeneous research communities. Science may therefore be a victim of its own success: the scale enabling extraordinary discoveries also incentivizes bureaucratic tools for comparison, ranking, and evaluation that are divorced from epistemic values. As Nguyen (Nguyen, 2024) emphasizes, flourishing academic systems require space for value plurality, yet contemporary evaluation regimes increasingly outsource values to productivity metrics, citation counts, and competitive funding schemes, suppressing communally beneficial norms and amplifying inequities. Our findings are consistent with this view: perceptions of scientific norms may reflect not only cultural or individual differences, but also the structural pressures of a system that increasingly privileges measurable outputs over the more nuanced values at the heart of scientific practice.

4.2 Predicting satisfaction for trainees

Overall, current experience appeared to be shaped primarily by satisfaction-related factors (i.e., academic support, perceived research productivity, and financial support) rather than structural features of trainees' social network, which instead related primarily to perceived academic support. Academic support, research productivity, and financial support were strongly associated with current experience, and academic support continued to predict research productivity, as already demonstrated (Kalyn Bogard et al., 2026). Between-person differences were evident: individuals who reported higher average support satisfaction across waves had higher odds of reporting more positive current experiences. Encouragingly, instances of higher-than-average within-person reporting for these same variables were associated with higher current experience, albeit less so for academic support. This pattern suggests that trainees' evaluations are responsive to short-term change in material and performance-related conditions, rather than being solely a reflection of fixed dispositions between trainees. Understanding the personal and social mechanisms that drive upward shifts in satisfaction over time remains paramount to future investigation.

Social network measures played a selective role in predicting satisfaction. Indicators of network quality, such as having more 'very close' network members and fewer difficult individuals, were associated with greater satisfaction with academic support. This closely parallels previous results, where having more emotionally supportive network members was the best indicator of academic support, although emotional support and closeness are strongly correlated (Kalyn Bogard et al., 2026). To be clear, these effects were small. For example, having an additional 'very close' network member corresponded to only a 25% increase in the odds of reporting a one-unit increase in academic support. Nevertheless, these results suggest that small but consistent aspects of students' affective relationships in their academic networks play a modest yet reliable role in shaping their perceptions of academic support over time, which in turn influences their current experience. Overall, network characteristics left a substantial amount of unexplained variance in academic support, which invites future work to examine additional dimensions of support quality, timing, and effectiveness.

Time-lagged analyses, random-effects estimates, and group differences converged to demonstrate persistent between-person variation in trainees' networks, support, and satisfaction, likely shaped by personal or contextual factors that operate across the training period. Notably, several findings clustered around groups with historically marginalized positions in academia. International students, underrepresented students, and postdocs all reported lower satisfaction with financial support, a disparity likely compounded for international students by substantially higher tuition fees (Statistics Canada, 2025). Women had higher odds of providing emotional support while also reporting more difficult relationships — a precarious combination as emotional support appears instrumental for trainee development, yet receives no formal recognition as a contribution to science (Tiokhin et al., 2024). Underrepresented individuals similarly reported more difficult relationships alongside lower academic support. These patterns indicate that structural inequities manifest not only in material conditions (e.g., financial support) but also in everyday social experience within research training environments.

Career stage also played a role in social resource distribution and satisfaction. More senior roles, particularly postdocs, importantly provide both technical and emotional support to students. While the critical role of the postdoc is undeniable throughout this study and our previous work, several patterns emerged indicating the precarity of this role. Postdocs reported fewer close network members, less satisfaction with financial support, and more agreement that scientists do not live up to certain Mertonian norms related to careerism and competition. These results should be interpreted cautiously, as our postdoc sample was small, but they align with previous accounts (Blaney et al., 2020; Feldon et al., 2019). Postdocs typically navigate a fiercely competitive job market, with uncertain prospects for advancement, and frequent relocations that may limit opportunities to maintain or build strong, relational ties in their social networks. As a result, they may be particularly vulnerable to the competitive pressures of academic culture, as detected in our responses. The combination of competition for scarce opportunities, pressure to demonstrate productivity, provision of both technical and emotional support to junior trainees, and awareness of reputational hierarchies within academia may make this career stage especially attuned to behaviours that deviate from Mertonian ideals. Such perceptions may reflect the reality that postdocs operate in environments where small differences in recognition, visibility, or network position can have outsized consequences for long-term career success.

Collectively, these findings encourage reflexive mentorship that recognizes the specific circumstances of individual trainees. Between-person variation is not inherently problematic; it reflects the diversity of trainees' needs, preferences, and contexts that should be addressed on an individual basis over blanket prescription. At the same time, inequities based on immutable characteristics *are* inherently problematic. Addressing these cannot rest solely on supervisors, who vary in resource availability, but require cooperation across multiple scales, from the social to the structural. These findings highlight the importance of distinguishing stable individual differences from modifiable drivers of within-person change when evaluating and improving academic training environments.

4.3 Emotional support is a central dimension of trainee supervision

Among all the relational variables examined, trainees' comfort discussing academic failure with their supervisor(s) emerged as a robust predictor of both perceived academic support and overall current experience, consistent with our previous work (Kalyn Bogard et al, 2026). Trainees who reported greater mean comfort levels consistently evaluated their academic support more positively and reported more favourable current experiences. Notably, this association was graded rather than binary: being 'somewhat' comfortable was associated with better outcomes than being uncomfortable, while being 'very' comfortable was associated with substantially higher satisfaction ratings. This indicates that even moderate improvements in relational openness may be consequential for trainees' perception of support.

The longitudinal analyses further indicate that comfort with discussing failure is itself a relatively stable characteristic, with strong autoregressive effects across waves and substantial between-person

variability. Trainees who were more comfortable at earlier waves tended to remain so over time, suggesting that these affective relationships reflect enduring features of either supervisory relationships or individual dispositions, rather than overly transient states. However, within-person increases in comfort were associated with improvements in perceived academic support and, to a lesser extent, current experience. This finding is important, as it indicates that the relational state between trainee and supervisor is not *totally* fixed, although our data cannot speak to whether these changes arise from shifts in supervisory practices, trust building for the trainee, or both.

Together, these results highlight emotional support as a central dimension of effective supervision. Comfort discussing academic failure appears to function as a relational gateway through which academic support is perceived, likely shaping how trainees interpret their training, guidance, feedback, and supervisory expectations. The observed declines in academic support ratings across waves perhaps indicate that trainees may require more sustained emotional support over time, particularly as supervisors may assume that more established trainees require less intensive oversight (e.g., Gatfield, 2005). From a practical perspective, these findings also underscore the importance of supervisory practices that normalize difficulty, uncertainty, and failure as integral components of academic training and research more broadly. Creating environments in which trainees disclose setbacks without fear of negative judgment likely enhances perceived support and also promotes more realistic expectations of the scientific process and, ultimately, higher-quality science when mistakes are acknowledged and addressed. Of course, these affective relationships are shaped by a combination of individual, relational, and institutional factors; responsibility for fostering supportive trainee-supervisor relationships extends beyond individual supervisors to include departmental norms and broader academic cultures and expectations.

4.5 Cross-study associations are largely absent

Overall, these results provided little evidence that trainees' immediate social environments meaningfully shaped their perceptions of how Mertonian norms are enacted. Most associations were close to zero or highly uncertain, and even when trends emerged, such as for academic support or current experience, the effects were modest and largely imprecise. Rather than a failure of measurement, these results suggest a mismatch between relational experience and norm perception that we anticipated, prompting several possible explanations.

First, the Mertonian norms represent a diffuse and multidimensional set of behaviours that manifest across multiple levels, from individual interactions to departmental cultures and broader disciplinary conventions. In our social norms survey, respondents were instructed to "think about the scientists and students in Canada you have interacted with in your department(s) or research field throughout your academic training". This wording was intended to anchor trainees' judgements in their immediate research environments rather than in abstract notions of science or sensational cases of misconduct. However, across two surveys and four waves, we have ample evidence that not all respondents attended closely to instructions. Thus, while the social network data capture genuine features of respondents'

local environments, their evaluations of enacted Mertonian norms may reflect a more generalized sense of “how science works” rather than their day-to-day interactions with specific people.

Second, the Mertonian items showed only weak inter-item correlation and do not represent a single coherent construct. We clearly cast our analytical net too wide by expecting variables from the immediate social environment to shape perceptions across all eight norms similarly. Future work should focus on specific norms that are most plausibly influenced by interpersonal dynamics, such as communality or disinterestedness, and recruit comparative samples of scientists who may and may not be directly affected by exclusion, competition, or gatekeeping.

Other possibilities may include ceiling effects, restricted variance in trainees’ norm ratings, and low sample size. While we detected provisional group differences in response, the lack of more extreme variability in our sample and a larger sample may constrain the ability to detect stronger or more certain associations. However, a trainee’s social environment may also simply be more strongly tied to well-being, satisfaction, and productivity rather than evaluations of scientific norms. In this view, the mismatch lies not in the data but in our theoretical assumption that relational experiences would translate into judgements about epistemic norms. Perhaps perceptions of Mertonian norms are shaped more by structural or institutional features, such as funding models, lab size, disciplinary conventions, departmental politics, or broader academic incentives, rather than by the micro-level composition of a trainee’s social network. Trainees may observe norms as emergent properties of their field or department, not as functions of the individuals with whom they interact most closely. Our future work intends to use qualitative analysis to investigate these perceptions in more detail.

5.0 Limitations

Numerous limitations should be considered when evaluating this study. First, sample size and composition constrain the precision and generalizability of our estimates. Although the Bayesian framework allowed us to model uncertainty explicitly and to estimate effects under relatively sparse data conditions, the overall sample’s limited statistical power for detecting more certain effects, cross-level interactions, and heterogeneous trajectories in slope, particularly for certain subgroups. This limitation is especially salient for postdoctoral fellows, who were underrepresented in our sample, even though the patterns observed were compelling and consistent (e.g., Gewin, 2025; Woolston, 2020). Results should be considered exploratory and hypothesis-generating, not confirmatory.

Second, the sampling strategy and context limit generalizability. Participants self-selected into a study focused on scientific norms, values, and academic experiences, which may have attracted individuals who are more reflective or attuned to these issues. In addition, the sample is limited to the three largest institutions in Alberta, Canada, which should not be taken to represent ‘all’ Canadian contexts, let alone other national systems that differ in funding structures, training pathways, and evaluation regimes. In fact, generalizability should not be the primary aim of work such as this, as we believe the drivers of variation include personal, contextual, and structural factors that inherently differ within academic

environments. Disciplinary representation was limited to the Natural Sciences, and some norms may be more salient or differently enacted across research fields, potentially attenuating associations when analyses are pooled.

Third, measuring personal social networks captures an incomplete representation of trainees' social environments. The ego-centric network approach focuses on perceived relationships from the ego. It does not capture departmental hierarchies, resource asymmetries, lab structures, or departmental cultures that may strongly shape both support and norm enactment. Further, although we distinguished affective qualities such as closeness and difficulty, there are other social dimensions (e.g., informal gatekeeping) within the broader network that could be captured more effectively with a whole-network approach.

Fourth, temporal scope and causal inference are limited by the four-wave design of this study. While our longitudinal analyses here highlight the value of using multilevel approaches and variance decomposition, our data did not permit modelling of longer-term developmental trajectories or critical transition points in academic training. Also, reciprocal relationships between our variables remain plausible. For example, greater satisfaction may facilitate stronger social networks, and vice versa.

Finally, these studies relied on self-reported measures for social norms, social networks, support and productivity. Even though we attempted to account for this by using hierarchical analytical frameworks, the possibility of shared method bias and social desirability bias can still be inflating associations among closely related constructs (Podsakoff et al., 2012). At the same time, subjective evaluations are central to our research questions, particularly given that *perceptions* of support, comfort, and norm enactment are likely to shape behaviour regardless of objective conditions. Nonetheless, future work may benefit from integrating complementary data from qualitative accounts (Gardner, 2008) or institutional indicators, even as these can be weak indicators of graduate student success (Miller & Stassun, 2014; Moneta-Koehler et al., 2017).

Declarations

Competing Interest Statement

The authors declare that there are no competing interests.

Funding Information

This research was supported by a University of Lethbridge Graduate Student Association Research Bursary Award and the Social Sciences and Humanities Research Council of Canada through a Doctoral Award (#752-2023-2790) to HKB and an Explore Grant (#41840-4200-8001) to LB.

Author Contribution

Author ContributionsHKB: Conceptualization, Investigation, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Validation, Visualization, Software, Writing – original draft, review & editing. JP and CL: Software, Data curationLB: Conceptualization, Supervision, Methodology, Funding acquisition, Writing – review & editing.

Acknowledgement

We warmly thank Drs. Tyler Bonnell and Gert Stulp for input on our study design.

Data Availability

Data and code underlying this study are openly available on GitHub at <https://github.com/hkaly/Longitudinal-investigation-of-norms-and-networks.git>. This code will be published on Zenodo under a Creative Commons Attribution 4.0 International License (CC BY 4.0) after review.

References

1. Anderson, M. S. (2000). Normative orientations of university faculty and doctoral students. *Science and Engineering Ethics*, 6(4), 443–461. <https://doi.org/10.1007/s11948-000-0002-6>
2. Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007). The perverse effects of competition on scientists' work and relationships. *Science and Engineering Ethics*, 13(4), 437–461. <https://doi.org/10.1007/s11948-007-9042-5>
3. Bieliński, J., & Tomczyńska, A. (2019). The ethos of science in contemporary Poland. *Minerva*, 57(2), 151–173. <https://doi.org/10.1007/s11024-018-9365-1>
4. Blaney, J. M., Kang, J., Wofford, A. M., & Feldon, D. F. (2020). Mentoring relationships between doctoral students and postdocs in the lab sciences. *Studies in Graduate and Postdoctoral Education*, 11(3), 263–279. <https://doi.org/10.1108/SGPE-08-2019-0071>
5. Bray, D., & Von Storch, H. (2017). The normative orientations of climate scientists. *Science and Engineering Ethics*, 23(5), 1351–1367. <https://doi.org/10.1007/s11948-014-9605-1>
6. Bright, L. K., & Heesen, R. (2023). To be scientific is to be communist. *Social Epistemology*, 37(3), 249–258. <https://doi.org/10.1080/02691728.2022.2156308>
7. Brooker, R., & Allum, N. (2024). Investigating the links between questionable research practices, scientific norms and organizational culture. *Research Integrity and Peer Review*, 9(1), 12. <https://doi.org/10.1186/s41073-024-00151-x>

8. Bürkner, P.-C. (2018). Advanced Bayesian multilevel modelling with the R package brms. *The R Journal*, 10(1), 395. <https://doi.org/10.32614/RJ-2018-017>
9. Clark, C. J., Graso, M., Redstone, I., & Tetlock, P. E. (2023). Harm hypervigilance in public reactions to scientific evidence. *Psychological Science*, 34(7), 834–848. <https://doi.org/10.1177/09567976231168777>
10. Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62(1), 583–619. <https://doi.org/10.1146/annurev.psych.093008.100356>
11. Daston, L. (2023). *Rivals: How Scientists Learned to Cooperate*. Columbia Global Reports. <https://doi.org/10.2307/jj.15552483>
12. Deanna, R., Merkle, B. G., Chun, K. P., Navarro-Rosenblatt, D., Baxter, I., Oleas, N., Bortolus, A., Geesink, P., Diele-Viegas, L., Aschero, V., de Leone, M. J., Oliferuk, S., Zuo, R., Cosacov, A., Grossi, M., Knapp, S., Lopez-Mendez, A., Welchen, E., Ribone, P., & Auge, G. (2022). Community voices: The importance of diverse networks in academic mentoring. *Nature Communications*, 13(1), 1681. <https://doi.org/10.1038/s41467-022-28667-0>
13. Feldon, D. F., Litson, K., Jeong, S., Blaney, J. M., Kang, J., Miller, C., Griffin, K., & Roksa, J. (2019). Postdocs' lab engagement predicts trajectories of PhD students' skill development. *Proceedings of the National Academy of Sciences of the United States of America*, 116(42), 20910–20916. <https://doi.org/10.1073/pnas.1912488116>
14. Gardner, S. K. (2008). "What's too much and what's too little?": The process of becoming an independent researcher in doctoral education. *The Journal of Higher Education*, 79(3), 326–350. <https://doi.org/10.1353/jhe.0.0007>
15. Gatfield, T. (2005). An investigation into PhD supervisory management styles: Development of a dynamic conceptual model and its managerial implications. *Journal of Higher Education Policy and Management*, 27(3), 311–325. <https://doi.org/10.1080/13600800500283585>
16. Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511790942>
17. Gewin, V. (2025, August 12). Postdoc depression and anxiety rates are rising, finds survey of 872 researchers. *Nature*. <https://doi.org/10.1038/d41586-025-02450-9>
18. Goldstein, H., Browne, W., & Rasbash, J. (2002). Partitioning variation in multilevel models. *Understanding Statistics*, 1(4), 223–231. https://doi.org/10.1207/S15328031US0104_02
19. Grilli, L., & Rampichini, C. (2007). Multilevel factor models for ordinal variables. *Structural Equation Modelling: A Multidisciplinary Journal*, 14(1), 1–25. <https://doi.org/10.1080/10705510709336734>
20. Heffernan, T. (2021). Academic networks and career trajectory: 'There's no career in academia without networks.' *Higher Education Research & Development*, 40(5), 981–994. <https://doi.org/10.1080/07294360.2020.1799948>
21. Hoffman, L., & Stawski, R. S. (2009). Persons as contexts: Evaluating between-person and within-person effects in longitudinal analysis. *Research in Human Development*, 6(2–3), 97–120.

<https://doi.org/10.1080/15427600902911189>

22. Kalyn Bogard, H. J., & Barrett, L. (2026). From Merton to market: Normative tensions in Canadian science. *Quantitative Science Studies*, in press. <https://doi.org/10.1162/QSS.a.480>
23. Kalyn Bogard, H. J., Lloyd, C., Petluk, J., & Barrett, L. (2026). It takes a village to raise a scientist: Social network predictors of satisfaction for academic trainees in Canada. *Quantitative Science Studies*, in press. <https://doi.org/10.1162/QSS.a.499>
24. Kim, S. Y., & Kim, Y. (2018). The ethos of science and its correlates: An empirical analysis of scientists' endorsement of Mertonian norms. *Science, Technology and Society*, 23(1), 1–24. <https://doi.org/10.1177/0971721817744438>
25. Kristman, V., Manno, M., & Côté, P. (2004). Loss to follow-up in cohort studies: How much is too much? *European Journal of Epidemiology*, 19(8), 751–760. <https://doi.org/10.1023/B:EJEP.0000036568.02655.f8>
26. Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., & Sugimoto, C. R. (2016). Contributorship and division of labor in knowledge production. *Social Studies of Science*, 46(3), 417–435. <https://doi.org/10.1177/0306312716650046>
27. Li, W., Zhang, S., Zheng, Z., Cranmer, S. J., & Clauset, A. (2022). Untangling the network effects of productivity and prominence among scientists. *Nature Communications*, 13(1), 4907. <https://doi.org/10.1038/s41467-022-32604-6>
28. Liu, X., Zeng, A., & Cui, X. (2026). Breaking through the constraints of academic social capital: Insights from top scientists' collaboration strategies. *Journal of Informetrics*, 20(2), 101798. <https://doi.org/10.1016/j.joi.2026.101798>
29. Longino, H. E. (1990). *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton University Press. <https://doi.org/10.2307/j.ctvx5wbfbz>
30. Macaluso, B., Larivière, V., Sugimoto, T., & Sugimoto, C. R. (2016). Is science built on the shoulders of women? A study of gender differences in contributorship. *Academic Medicine*, 91(8), 1136–1142. <https://doi.org/10.1097/ACM.0000000000001261>
31. McIntyre, L. C. (2019). *The Scientific Attitude: Defending Science from Denial, Fraud, and Pseudoscience*. MIT press. <https://doi.org/DOI:10.7551/mitpress/12203.001.0001>
32. Miller, C., & Stassun, K. (2014). A test that fails. *Nature*, 510(7504), 303–304. <https://doi.org/10.1038/nj7504-303a>
33. Moneta-Koehler, L., Brown, A. M., Petrie, K. A., Evans, B. J., & Chalkley, R. (2017). The limitations of the GRE in predicting success in biomedical graduate school. *PLOS ONE*, 12(1), e0166742. <https://doi.org/10.1371/journal.pone.0166742>
34. Nguyen, C. (2024). Value capture. *Journal of Ethics and Social Philosophy*, 27(3). <https://doi.org/10.26556/jesp.v27i3.3048>
35. Oreskes, N. (2021). *Why Trust Science? With a new preface by the author*. Princeton University Press. <https://doi.org/10.1515/9780691222370>

36. Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63(1), 539–569. <https://doi.org/10.1146/annurev-psych-120710-100452>
37. Porter, T. (1995). *Trust in Numbers*. Princeton University Press.
38. Quan, W., Chen, B., & Shu, F. (2017). Publish or impoverish: An investigation of the monetary reward system of science in China (1999–2016). *Aslib Journal of Information Management*, 69(5), 486–502. <https://doi.org/10.1108/AJIM-01-2017-0014>
39. Shapin, S. (2010). *Never Pure: Historical Studies of Science as if it Was Produced by People with Bodies, Situated in Time, Space, Culture, and Society, and Struggling for Credibility and Authority*. Johns Hopkins University Press.
40. Statistics Canada. (2025, September 10). Tuition in Canada: Modest increases and widening gaps, 2025/2026. *The Daily*. <https://www150.statcan.gc.ca/n1/daily-quotidien/250910/dq250910d-eng.htm>
41. Tiokhin, L., Panchanathan, K., Smaldino, P. E., & Lakens, D. (2024). Shifting the level of selection in science. *Perspectives on Psychological Science*, 19(6), 908–920. <https://doi.org/10.1177/17456916231182568>
42. Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
43. Woolston, C. (2020). Postdocs under pressure: ‘Can I even do this any more?’ *Nature*, 587(7835), 689–692. <https://doi.org/10.1038/d41586-020-03235-y>

Figures

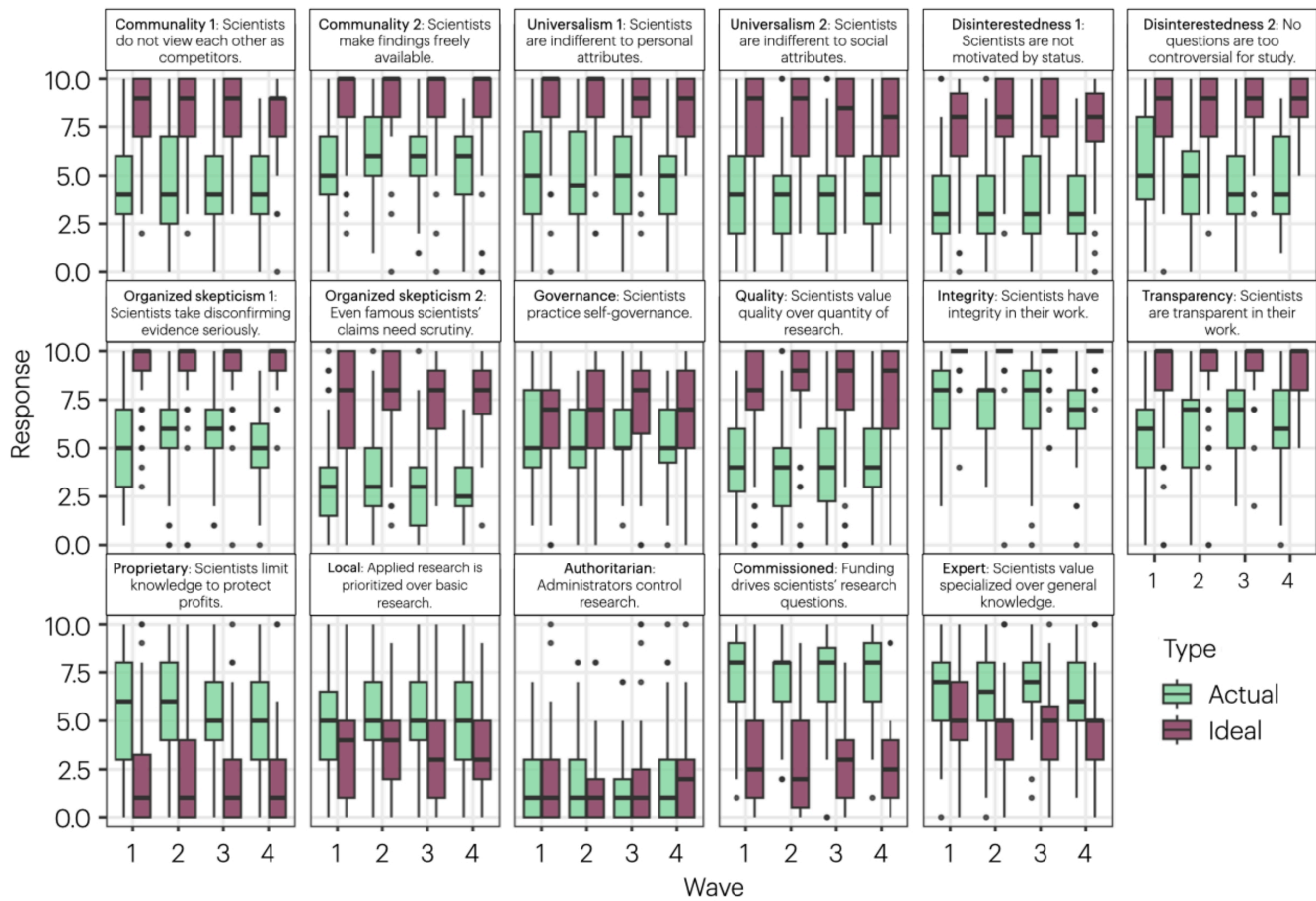


Figure 1

Observed 'ideal' and 'actual' ratings of 17 social norms items across four waves. Boxplots show the median (centre line), interquartile range (box), and spread of responses (whiskers), with points indicating outlying responses.

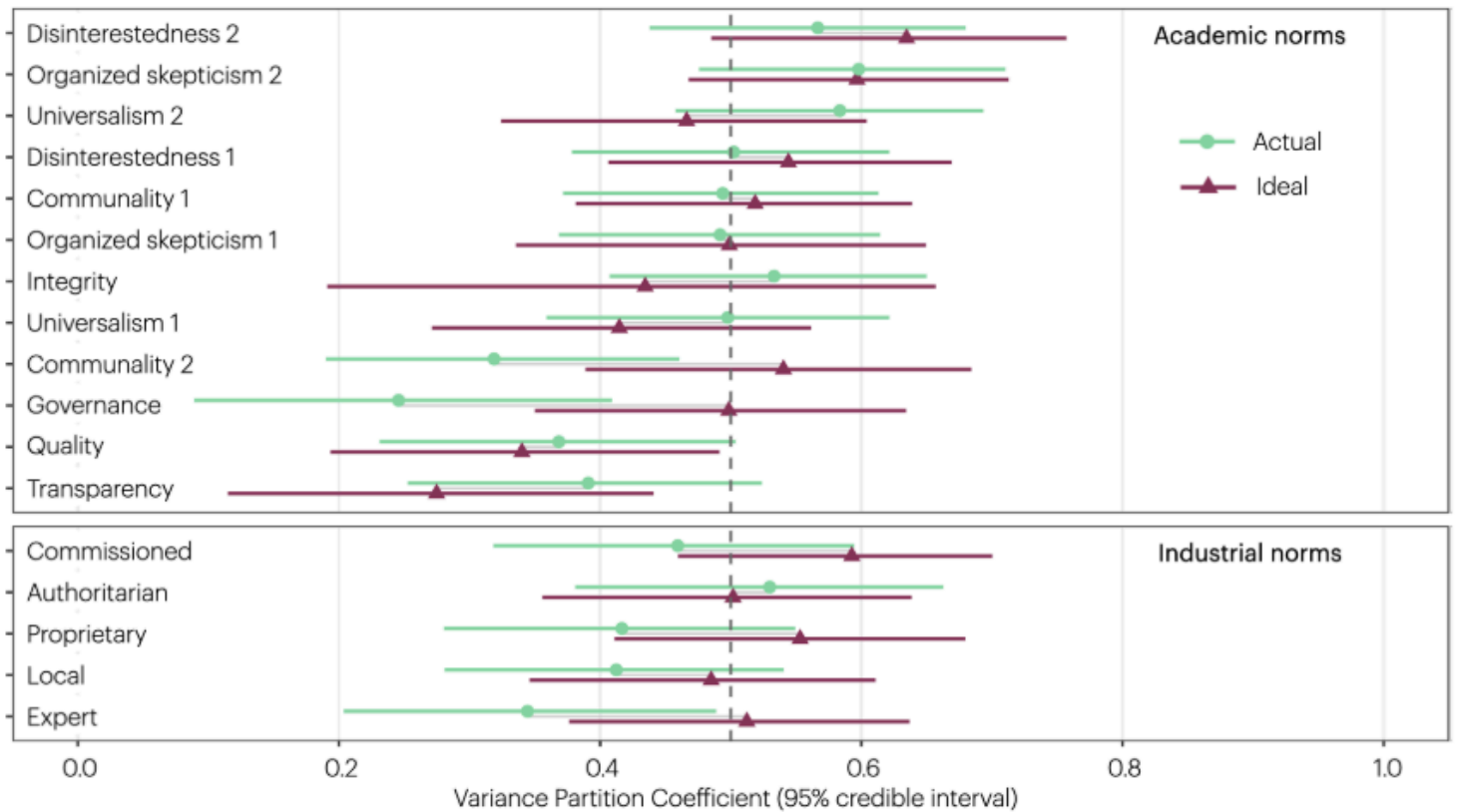


Figure 2

Median variance partition coefficients (VPC) with 95% credible intervals (horizontal lines) for 'actual' (circle) and 'ideal' (triangle) rating across academic (top) and industrial (bottom) norm items. The dashed vertical line at 0.5 indicates equal partitioning of within- and between-person variance. Values above 0.5 indicate more between-person variance, while those below 0.5 indicate more within-person variance.

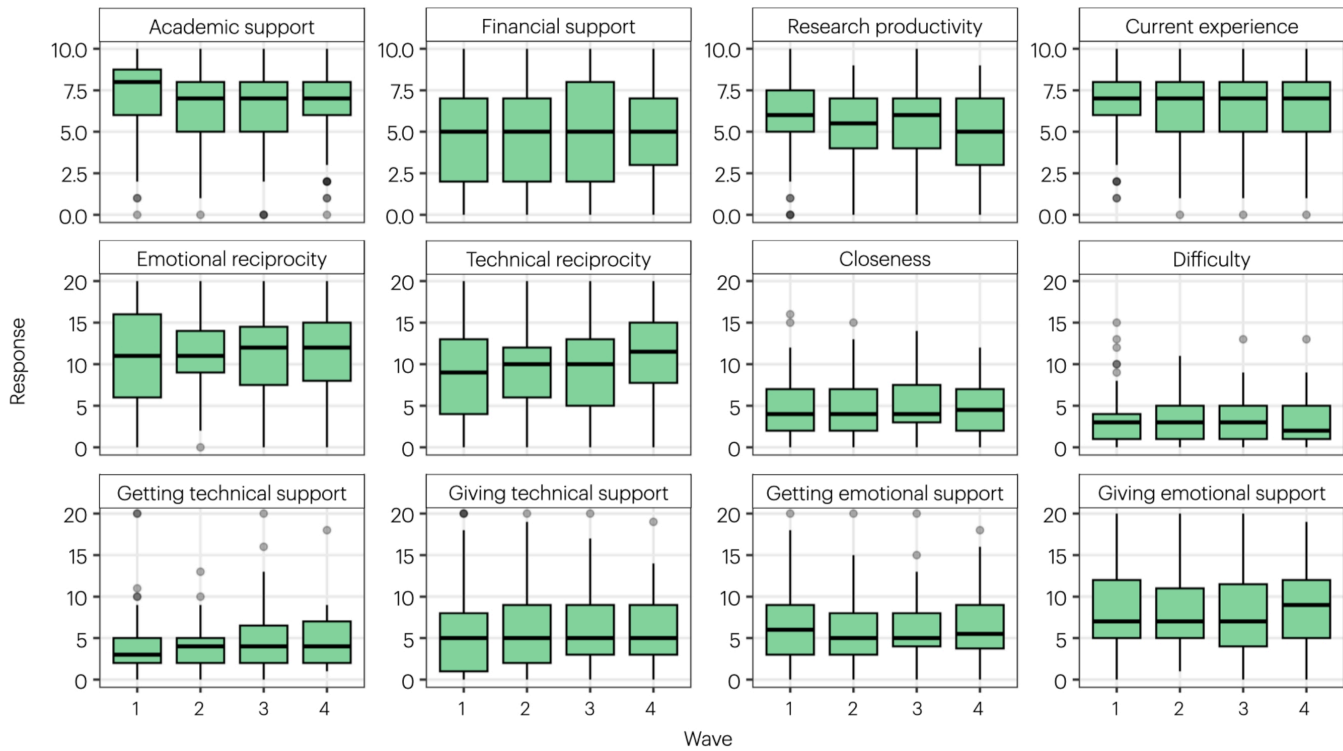


Figure 3

Observed responses across four waves for measures of satisfaction, each rated from 0 (not at all satisfied) to 10 (very satisfied): academic support, financial support, research productivity, and current experience (top row). Remaining panels show network characteristics, measured as the number of network members (0 to 20): exchanging emotional and technical reciprocity; considered 'very' close or 'difficult' to interact with; providing or receiving 'a lot' of technical support; and providing or receiving emotional support that the respondent is 'very' comfortable with. Boxplots show the median (centre line), interquartile range (box), and spread of responses (whiskers), with points indicating outlying responses.

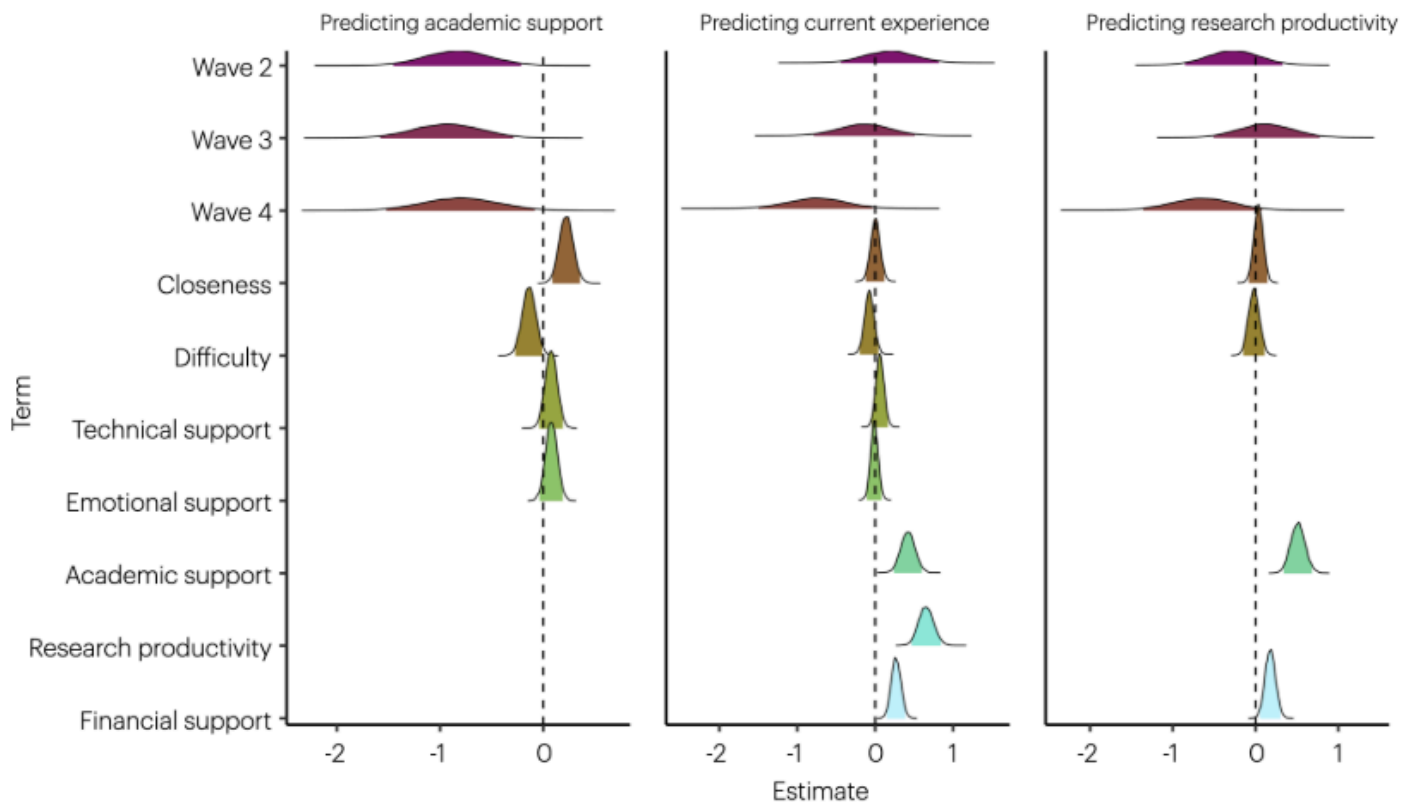


Figure 4

Posterior distributions of jointly estimated Bayesian mixed-effects model estimates predicting academic support, current experience, and research productivity (left to right). Coloured density curves show the 95% credible interval (CrI) of the estimated effect of each predictor (rows) with wave as a fixed effect (ref. category: Wave 1). The dashed vertical line at 0 indicates no effect; posterior distributions whose 95% CrI includes zero are uncertain in the direction of effect.

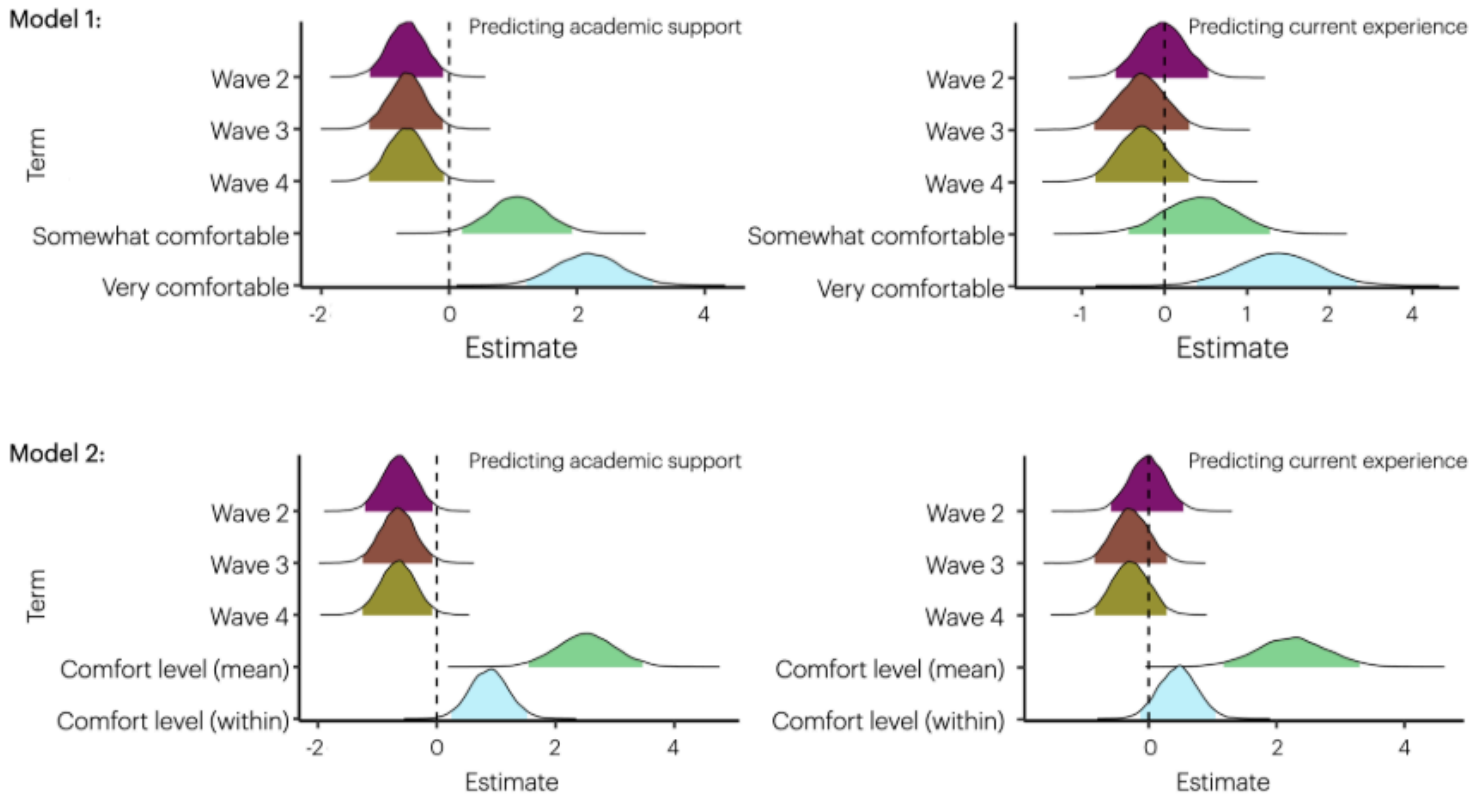


Figure 5

Posterior distributions of jointly estimated Bayesian mixed-effects model estimates predicting academic support (left) and current experience (right), for Model 1 (top) and Model 2 (bottom). In Model 1, comfort level is a categorical predictor (reference category: not comfortable). In Model 2, comfort level is decomposed into between-person (mean) and within-person (within-person deviations) components, treated as fixed effects. Wave is also treated as a fixed effect (ref. category: Wave 1). The dashed vertical line at 0 indicates no effect; posterior distributions whose 95% CrI includes zero are uncertain in the direction of effect.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Ch.3supplementarymaterials.docx](#)