Revision of the Mindfulness Manuscript

Dimensionality Test

Previous research has extensively investigated the bipolar continuum hypothesis by examining the dimensionality of *trait* self-compassion using various psychometric methods (Brenner et al. 2017; Cleare et al. 2018; Neff 2016; Neff, Whittaker, and Karl 2017; Neff et al. 2019; Petrocchi, Ottaviani, and Couyoumdjian 2014). These studies have consistently reported that models incorporating a general factor demonstrate superior fit indices, thereby providing empirical support for Neff's hypothesis.

The present study aims to extend this line of inquiry by conducting a novel psychometric evaluation of the dimensionality of *state* self-compassion. Our approach is distinctive in that it accounts for the multilevel structure inherent in our data, addressing a critical gap in the existing literature. To this end, we employed a series of multilevel confirmatory factor analysis (CFA) models, applied to the combined data from our two studies. This analytical strategy allows us to appropriately model the hierarchical nature of our data, where repeated measurements are nested within days, which are in turn nested within individuals.

We specified and compared three distinct models to elucidate the underlying structure of state self-compassion:

- In the **One-Factor Model**, we hypothesized that state self-compassion could be represented by a single latent factor, capturing variance both within and between individuals.
- The **Two-Factor Model** allowed for a conceptualization where state self-compassion is represented by two distinct latent factors—CS and UCS—operating at both the within-person and between-person levels.
- The **Bifactor Model** tested whether state self-compassion is best represented by a general self-compassion factor alongside orthogonal specific factors (CS and UCS), separating variance attributable to the general factor from variance explained by the specific factors.

The goodness-of-fit indices for each model are presented in Table 1.

Table 1. Goodness-of-Fit Indices for the One-Factor, Two-Factor, and Bifactor Models

Model	CFI	TLI	RMSEA	SRMR (Within)	SRMR (Between)	AIC	BIC
One-Factor	0.935	0.910	0.050	0.033	0.058	475881.5	476195.3
Two-Factor	0.972	0.959	0.034	0.024	0.050	474810.2	475139.6
Bifactor	0.987	0.971	0.029	0.016	0.031	474376.7	474816.0

As shown in Table 1, the goodness-of-fit indices progressively improve from the One-Factor Model to the Two-Factor Model, with the Bifactor Model demonstrating the best overall fit. The Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) both exceed the commonly accepted threshold of 0.95 for the Two-Factor and Bifactor Models, indicating strong model fit. Additionally, the Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) values decrease across models, further supporting improved fit with more complex structures.

We conducted a series of Likelihood Ratio Tests (LRTs) to formally assess improvements in model fit:

- Comparing the One-Factor Model with the Two-Factor Model yielded an LRT statistic of 1068.792 with 5 degrees of freedom ($p < 0.001$), indicating that the Two-Factor Model significantly improves fit. This suggests that conceptualizing state self-compassion as two distinct components (CS and UCS) better represents the data.
- The comparison between the Two-Factor Model and the Bifactor Model resulted in an LRT statistic of 506.656 with 17 degrees of freedom ($p < 0.001$), demonstrating that the Bifactor Model offers further improvement in fit. This suggests that, beyond distinct compassionate and uncompassionate components, a general self-compassion factor provides additional explanatory power.

In summary, from a psychometric standpoint, both the Two-Factor and Bifactor Models provide better fit than the One-Factor Model, with the Bifactor Model yielding the best statistical fit. These results suggest that self-compassion is best conceptualized as comprising both a general factor and specific components (CS and UCS). However, while the Bifactor Model offers superior statistical fit, its added complexity raises questions about the practical significance of modeling both general and specific factors.

Practical Importance

Dimensionality assessments should not rely exclusively on statistical fit indices but should also consider the substantive contributions of each factor (Rodriguez, Reise, and Haviland 2016; Pere J. Ferrando and Lorenzo-Seva 2018). While more complex models typically achieve better fit, they may overfit the data, particularly when too many factors are specified (Pere Joan Ferrando and Lorenzo-Seva 2019). A balance between model complexity and interpretability is essential for drawing meaningful conclusions.

To evaluate whether the data supports a unidimensional or multifactorial structure, we applied the following criteria:

1. **Factor Correlation**:
   In the Bifactor Model, correlations between the general factor and specific factors are constrained to zero, theoretically separating them. The lack of correlation between general and specific factors implies a distinct role for each factor, negating the need to interpret correlations in this case.

2. **Explained Common Variance (ECV)**:
   ECV estimates how much common variance is explained by the general versus specific factors. The general factor in the Bifactor Model (Gen_w) has high standardized loadings on items (e.g., SCP1 = 0.614, SCN2 = 0.579 at the within-level), suggesting it accounts for most of the variance. Specific factors (POS_w and NEG_w) show weaker loadings, indicating that the general factor dominates the variance explanation.

3. **Omega Hierarchical (H)**:
   H measures how much variance in the total score is due to the general factor. In the Bifactor Model, the general factor explains a large proportion of the variance (e.g., SCP1 = 0.614, SCN2 = 0.579), while specific factors contribute much less, supporting the predominance of the general factor.

4. **Bifactor Model Analysis**:
   The Bifactor Model confirms the dominance of the general factor, as specific factor loadings are consistently lower (e.g., POS_w on SCP3 = -0.313, NEG_b on SCN8 = 0.288), indicating that the specific factors add limited explanatory value beyond the general factor.

5. **Change in R-squared**:
   The general factor in the Bifactor Model explains a substantial portion of variance (e.g., R-squared for SCP1 = 0.425 within-level, 0.786 between-level). In contrast, the specific factors provide only marginal additional explanation, suggesting that a unidimensional solution may be more parsimonious without significant loss of explanatory power.

In summary, the results from the multilevel CFA analysis support Neff's hypothesis that CS and UCS are related at the trait or global level, as indicated by the superior fit of the Bifactor Model. The strong correlation between CS and UCS at the latent level suggests that, across participants, individuals with high levels of CS tend to have lower levels of UCS, and vice versa.

However, it is important to emphasize that multilevel CFA examines *latent, person-level relationships*, which capture general trends over time and context. The present analysis demonstrates that CS and UCS are correlated at a global level. In contrast, other analyses focused on *momentary dynamics* may suggest that CS and UCS operate more independently within individuals over short time periods. This distinction clarifies that Neff's hypothesis may hold at a trait level but requires further exploration at the momentary level.

Modeling the Dynamic Relationship Between Compassionate and Uncompassionate Self-Responding: An Idionomic Approach

The analyses described in the previous section provided evidence supporting the "essential unidimensionality" (Reise, Bonifay, and Haviland 2013) of state self-compassion. Our multilevel framework, which accounted for repeated measurements within individuals over a three-month period, demonstrated that compassionate and uncompassionate self-responding (CS and UCS) are inversely related at the *nomothetic level* – that is, when examined across the entire sample.

Our psychometric evaluations, using purely *internal criteria* (fit indices for item scores), revealed that more complex models improved statistical fit. However, these improvements did not fundamentally challenge Neff's hypothesis of a bipolar continuum. While minor deviations and better fit indices were observed with more complex models, their practical importance was limited, and the overall pattern of findings remained aligned with the notion of CS and UCS as opposing dimensions. This conclusion was further corroborated when considering *external criteria*, including covariates such as momentary negative affect and context evaluation, as described in our two studies. These situational factors did not substantially disrupt the core relationship between CS and UCS proposed by the bipolar continuum hypothesis.

It is important to note, however, that all previous analyses were conducted at a *nomothetic* level, focusing on group-level patterns. This approach, while valuable, may obscure important differences in how self-compassion operates in the everyday life of individuals. The variability in how individuals experience and express self-compassion across different contexts might not be fully captured in these broad, averaged analyses.

To address this limitation and gain deeper insights into person-specific dynamics, we conducted an *idionomic analysis* to examine the relationship between UCS and CS at the individual level (Ciarrochi et al. 2024; Ferrari et al. 2022; Sahdra et al. 2024). This approach allows us to explore potential heterogeneity in self-compassion processes that may be masked by nomothetic analyses, providing a more nuanced understanding of how these constructs interact within individuals over time.

Idionomic Analysis of the Relationship Between UCS and CS

For the present purposes, an idionomic analysis was conducted in two stages. The first stage involved a strictly idiographic approach, focusing on the individual-level patterns. In the second stage, we applied a nomothetic approach to examine and describe group-level regularities as well as the variability in these effects across individuals (e.g., Ciarrochi et al. 2024; Sahdra et al. 2024).

Step 1: Fitting Individual-Level Hierarchical Bayesian Models. For each participant separately, we implemented a hierarchical Bayesian model using Stan to estimate the relationship between uncompassionate (UCS) and compassionate self-responding (CS). The model incorporated additional covariates, including negative affect and context evaluation, as well as lagged effects of CS from the previous measurement within the same day. The model for each participant was specified as follows:

$$\begin{aligned}\mathrm{UCS}_n \sim t_\nu\Big(&\alpha + \gamma_{\mathrm{CS}} \cdot \mathrm{CS}_n + \gamma_{\mathrm{neg\_aff}} \cdot \mathrm{neg\_aff}_n \\ &+ \gamma_{\mathrm{context}} \cdot \mathrm{context}_n + \phi \cdot \mathrm{lag\_CS}_n \\ &+ \gamma_{\mathrm{interaction}} \cdot \mathrm{CS}_n \cdot \mathrm{neg\_aff}_n, \sigma\Big),\end{aligned}$$

where:

- UCS_n represents the CS score for observation n,
- α denotes the intercept,
- γ_{CS} is the coefficient for the primary predictor, CS (CS_n),
- $\gamma_{\mathrm{neg\_aff}}$ and $\gamma_{\mathrm{context}}$ are coefficients for negative affect ($\mathrm{neg\_aff}_n$) and context evaluation ($\mathrm{context}_n$), respectively,
- ϕ represents the autoregressive coefficient for the lagged CS within the same day ($\mathrm{lag\_CS}_n$),
- σ is the scale parameter (standard deviation) of the distribution,
- ν denotes the degrees of freedom of the Student's t-distribution,
- $\gamma_{\mathrm{interaction}}$ is the coefficient for the interaction term between CS and negative affect.

This model allows for the examination of the association between UCS and CS while controlling for the effects of negative affect, context evaluation, lagged CS effects within the same day, and the CS $\times$ negative-affect interaction. We employed a Student's t-distribution (t_ν) to account for potential outliers or heavy-tailed distributions in the data (see Supplementary Information for further details).

For each participant, we examined the posterior distribution of the `gamma_CS` coefficient to estimate the proportion of posterior draws that were negative. This enabled us to evaluate whether there was evidence of a negative association between UCS and CS at the individual level, in line with Neff's bipolar continuum hypothesis. Additionally, we computed the mean posterior estimates for the `gamma_neg_aff`, `gamma_context`, and `gamma_interaction` parameters, which represent the effects of negative affect, context evaluation, and the interaction between CS and negative affect, respectively, on UCS for each participant.

Step 2: Aggregate Analysis Using a Hierarchical Model. After the idiographic analysis, we used a hierarchical model (e.g., Ciarrochi et al. 2024) in the `brms` package (R) to summarize the proportion of negative estimates for the `gamma_CS` parameter across participants. This model employed a binomial distribution, with the total number of posterior samples as the denominator and the proportion of negative estimates as the response. A random intercept for participants was included to account for individual variability.

Hierarchical models were also applied to the mean posterior estimates of the `gamma_neg_aff`, `gamma_context`, and `gamma_interaction` parameters, representing the influences of negative affect, context evaluation, and the interaction between CS and negative affect on UCS. Each model included a fixed effect (intercept) and random intercepts for participants, capturing individual differences. A Student-t likelihood was used to account for potential outliers and accommodate the heavy-tailed nature of the effect distributions, providing robust aggregate estimates of each parameter and the heterogeneity of their relationships with UCS (see Supplementary Information for further details).

Results

The analysis of the `gamma_CS` parameter across participants indicated that 81.0% (89% CI [0.796, 0.823]) of the posterior estimates for the association between uncompassionate self-responding (UCS) and compassionate self-responding (CS) were negative. This supports Neff's bipolar continuum hypothesis, which posits an inverse relationship between UCS and CS. However, the variability in individual effects (`sd(Intercept)` = 1.25, corresponding to 0.196 on the probability scale) suggests moderate heterogeneity in the strength of this relationship across participants.

For the influence of negative affect on UCS, represented by the `gamma_neg_aff` parameter, the analysis revealed a positive overall effect (intercept = 0.38, 89% CI [0.34, 0.42]), indicating that higher levels of negative affect are generally associated with increased UCS. The variability between individuals (`sd(Intercept)` = 0.19, corresponding to 0.547 on the probability scale) suggests substantial heterogeneity in how strongly negative affect influences UCS.

The analysis of context evaluation showed a small but credible negative effect on UCS (mean = -0.04, 95% CI [-0.07, -0.01]), suggesting that higher context evaluation scores are associated with a slight reduction in UCS. The variability in individual responses (`sd(Intercept)` = 0.17, corresponding to 0.543 on the probability scale) also indicates substantial heterogeneity in the relationship between context evaluation and UCS across participants.

Finally, the interaction between CS and negative affect was negligible, with the 89% credible interval spanning zero (-0.03, 0.00), suggesting no meaningful interaction effect between these variables on UCS.

Discussion

Our idionomic analysis provides insight into the complex, momentary relationship between CS and UCS. Although the group-level results generally support Neff's bipolar continuum hypothesis, substantial heterogeneity emerged in the strength and direction of the CS-UCS association across individuals. While most participants displayed the anticipated inverse relationship between CS and UCS, a notable proportion showed no clear association, and a smaller subset exhibited positive correlations—indicating that, for some individuals, CS and UCS may vary together rather than oppositely, which contradicts the bipolar continuum hypothesis [Ferrari et al., 2023; Ullrich et al., 2020]. This variability suggests that Neff's hypothesis may not be universally applicable and underscores the importance of idiographic approaches in revealing the nuanced patterns often obscured in aggregate, group-level analyses.

Additionally, our analysis revealed that negative affect consistently increased UCS across participants, indicating that elevated negative emotional states tend to intensify uncompassionate self-responding. However, the variability observed across individuals highlights that the influence of negative affect on UCS is not uniform. Contrary to expectations, we found no credible evidence that negative affect moderated the CS-UCS relationship, suggesting that momentary emotional states do not substantially alter how CS influences UCS within individuals (Dejonckheere et al., 2021).

Supplementary Material Description

Dimensionality Test

Mplus script for the One-Factor Model

```
TITLE: Multilevel One-Factor Model with covariates at moment, day, and person level;

DATA:
  FILE = "data.dat";

VARIABLE:
  NAMES = user_id day time_window scs_pos_1 scs_neg_2 scs_pos_3 scs_neg_4
    scs_neg_5 scs_pos_6 scs_pos_7 scs_neg_8 neg_aff_Moment neg_aff_Day neg_aff_Person;
  USEVARIABLES = scs_pos_1 scs_neg_2 scs_pos_3 scs_neg_4 scs_neg_5 scs_pos_6
    scs_pos_7 scs_neg_8 neg_aff_Moment neg_aff_Day neg_aff_Person;
  CLUSTER = user_id;  ! Clustering at the subject level
  WITHIN = neg_aff_Moment neg_aff_Day;  ! Covariates at the moment and day level
  BETWEEN = neg_aff_Person;  ! Covariate at the person level

MISSING = .;

ANALYSIS:
  TYPE = TWOLEVEL;  ! Multilevel model with user_id as the cluster
  ESTIMATOR = ML;  ! Maximum likelihood estimation

MODEL:
  %WITHIN%  ! Level 1: Within-subject and within-day variation
    SelfCompassion_w BY scs_pos_1 scs_neg_2 scs_pos_3 scs_neg_4 scs_neg_5
      scs_pos_6 scs_pos_7 scs_neg_8;
    SelfCompassion_w ON neg_aff_Moment neg_aff_Day;
    ! Regress self-compassion factor on momentary and daily negative affect

  %BETWEEN%  ! Level 2: Between-subject variation
    SelfCompassion_b BY scs_pos_1 scs_neg_2 scs_pos_3 scs_neg_4 scs_neg_5
      scs_pos_6 scs_pos_7 scs_neg_8;
    SelfCompassion_b ON neg_aff_Person;
    ! Regress between-level self-compassion factor on person-level negative affect

OUTPUT:
  TECH1 TECH8 STANDARDIZED MODINDICES;
```

Mplus script for the Two-Factor Model

```
TITLE: Multilevel Two-Factor Model with covariates at moment, day, and person level;

DATA:
  FILE = "data.dat";

VARIABLE:
  NAMES = user_id day time_window scs_pos_1 scs_neg_2 scs_pos_3 scs_neg_4
    scs_neg_5 scs_pos_6 scs_pos_7 scs_neg_8 neg_aff_Moment neg_aff_Day neg_aff_Person;
  USEVARIABLES = scs_pos_1 scs_neg_2 scs_pos_3 scs_neg_4 scs_neg_5 scs_pos_6
    scs_pos_7 scs_neg_8 neg_aff_Moment neg_aff_Day neg_aff_Person;
  CLUSTER = user_id;  ! Clustering at the subject level
  WITHIN = neg_aff_Moment neg_aff_Day;  ! Covariates at the moment and day level
  BETWEEN = neg_aff_Person;  ! Covariate at the person level

MISSING = .;

ANALYSIS:
  TYPE = TWOLEVEL;  ! Multilevel model with user_id as the cluster
  ESTIMATOR = ML;  ! Maximum likelihood estimation

MODEL:
  %WITHIN%  ! Level 1: Within-subject and within-day variation
    PositiveSelfCompassion_w BY scs_pos_1 scs_pos_3 scs_pos_6 scs_pos_7;
    ! Positive self-compassion factor on within level
    NegativeSelfCompassion_w BY scs_neg_2 scs_neg_4 scs_neg_5 scs_neg_8;
    ! Negative self-compassion factor on within level
    PositiveSelfCompassion_w ON neg_aff_Moment neg_aff_Day;
    ! Regress positive factor on momentary and daily negative affect
    NegativeSelfCompassion_w ON neg_aff_Moment neg_aff_Day;
    ! Regress negative factor on momentary and daily negative affect

  %BETWEEN%  ! Level 2: Between-subject variation
    PositiveSelfCompassion_b BY scs_pos_1 scs_pos_3 scs_pos_6 scs_pos_7;
    ! Positive self-compassion factor on between level
    NegativeSelfCompassion_b BY scs_neg_2 scs_neg_4 scs_neg_5 scs_neg_8;
    ! Negative self-compassion factor on between level
    PositiveSelfCompassion_b ON neg_aff_Person;
    ! Regress positive factor on person-level negative affect
    NegativeSelfCompassion_b ON neg_aff_Person;
```

```
    ! Regress negative factor on person-level negative affect

OUTPUT:
  TECH1 TECH8 STANDARDIZED MODINDICES;
```

Mplus script for the Bifactor Model

```
TITLE: Multilevel Bifactor Model with covariates at moment, day, and person level;

DATA:
  FILE = "data.dat";

VARIABLE:
  NAMES = uid day tw scp1 scn2 scp3 scn4 scn5 scp6 scp7 scn8 na_mom na_day na_per;
  USEVARIABLES = scp1 scn2 scp3 scn4 scn5 scp6 scp7 scn8 na_mom na_day na_per;
  CLUSTER = uid;
  WITHIN = na_mom na_day;
  BETWEEN = na_per;

  MISSING = .;

ANALYSIS:
  TYPE = TWOLEVEL;
  ESTIMATOR = ML;

MODEL:
  %WITHIN%
    Gen_w BY scp1* scn2 scp3 scn4 scn5 scp6 scp7 scn8;
    Pos_w BY scp1* scp3 scp6 scp7;
    Neg_w BY scn2* scn4 scn5 scn8;

    ! Fix variances to help with model identification
    Gen_w@1;  ! Fix variance of the general factor
    Pos_w@1;  ! Constrain variance of the positive factor
    Neg_w@1;  ! Constrain variance of the negative factor

    ! Ensure no covariance between factors
    Pos_w WITH Gen_w@0;
    Neg_w WITH Gen_w@0;
    Pos_w WITH Neg_w@0;

    ! Regress factors on covariates
    Gen_w ON na_mom na_day;
    Pos_w ON na_mom na_day;
    Neg_w ON na_mom na_day;
```

```
  %BETWEEN%
    Gen_b BY scp1* scn2 scp3 scn4 scn5 scp6 scp7 scn8;
    Pos_b BY scp1* scp3 scp6 scp7;
    Neg_b BY scn2* scn4 scn5 scn8;

    ! Fix variances on the between level as well
    Gen_b@1;  ! Fix variance of the general factor
    Pos_b@1;  ! Constrain variance of the positive factor
    Neg_b@1;  ! Constrain variance of the negative factor

    ! Ensure no covariance between factors
    Pos_b WITH Gen_b@0;
    Neg_b WITH Gen_b@0;
    Pos_b WITH Neg_b@0;

    ! Regress factors on covariates
    Gen_b ON na_per;
    Pos_b ON na_per;
    Neg_b ON na_per;

OUTPUT:
  TECH1 TECH8 STANDARDIZED MODINDICES(ALL);
```

Idiographic Analysis

This section details the specification of the model fitted separately to individual data from both experiments, constituting the first step of our idiographic analysis.

Model Structure

We employed a hierarchical Bayesian model with the following characteristics:

1. **Response Variable**: Uncompassionate Self-Responding (UCS).
2. **Primary Predictor**: Compassionate Self-Responding (CS).
3. **Covariates**:
    - `neg_aff_Moment`: Momentary negative affect.
    - `context_Moment`: Context evaluation.
    - `lag_CS_same_day`: Lagged CS from the same day.
4. **Distribution**: Student's t-distribution to account for potential outliers or heavy-tailed response distributions
5. **Temporal Dependency**: Incorporated an AR(1) component to account for autoregressive effects of UCS.

Model Parameters

- `gamma_CS`: Coefficient representing the association between CS and UCS.
- `gamma_neg_aff`: Slope for negative affect (`neg_aff_Moment`).
- `gamma_context`: Slope for context evaluation (`context_Moment`).
- `phi`: Autoregressive coefficient for the lagged CS effect.
- `nu`: Degrees of freedom for the Student's t-distribution (estimated by the model).
- `alpha`: Intercept term.
- `sigma`: Scale parameter of the Student's t-distribution.

Stan Model Specification

```
data {
  int<lower=1> N;
  vector[N] CS;
  vector[N] UCS;
  vector[N] neg_aff_Moment;
  vector[N] context_Moment;
  vector[N] lag_CS_same_day;
```

```
}

parameters {
  real alpha_raw;
  real gamma_CS_raw;
  real gamma_neg_aff_raw;
  real gamma_context_raw;
  real phi_raw;
  real<lower=0> sigma;
  real<lower=2> nu;
}

transformed parameters {
  real alpha = alpha_raw;
  real gamma_CS = gamma_CS_raw;
  real gamma_neg_aff = gamma_neg_aff_raw;
  real gamma_context = gamma_context_raw;
  real phi = phi_raw * 0.5;  // Constraining phi to [-0.5, 0.5] for stability
}

model {
  // Priors
  alpha_raw ~ normal(0, 1);
  gamma_CS_raw ~ normal(0, 1);
  gamma_neg_aff_raw ~ normal(0, 1);
  gamma_context_raw ~ normal(0, 1);
  phi_raw ~ normal(0, 1);
  sigma ~ cauchy(0, 2.5);
  nu ~ gamma(2, 0.1);

  // Likelihood
  UCS ~ student_t(nu, alpha + gamma_CS * CS + gamma_neg_aff * neg_aff_Moment +
                  gamma_context * context_Moment + phi * lag_CS_same_day, sigma);
}
```

This Stan code implements the hierarchical Bayesian model described above. The model uses weakly informative priors for all parameters and constrains the autoregressive coefficient `phi` to the interval [-0.5, 0.5] to ensure model stability. The degrees of freedom parameter `nu` is constrained to be greater than 2 to ensure finite variance of the Student's t-distribution.

Hierarchical Model Analysis

In the third step of our idiographic analysis, we employed a hierarchical model approach, analogous to a meta-analysis (Ciarrochi et al. 2024). This method allowed us to synthesize the individual-level results across all participants, providing both an aggregate estimate of the effects and a measure of heterogeneity among participants.

Model Specification

We fitted a hierarchical binomial model to the proportion of negative posterior draws for the `gamma_CS` parameter. In this framework, each participant was treated as a unique "study," enabling us to estimate:

1. The aggregate proportion of negative effects across all participants.
2. The degree of heterogeneity in these effects among participants.

The model was implemented using the `brms` package in R.

R Code for Model Fitting

```r
fit_gamma_cs_binom <- brm(
  bf(n_negative | trials(n_total) ~ 1 + (1 | id), family = "binomial"),
  data = beta_cs_df,
  prior = c(
    prior(normal(0, 1), class = "Intercept"),
    prior(normal(0, 1), class = "sd")
  ),
  iter = 100000, warmup = 2000, chains = 4, cores = 4,
  backend = "cmdstanr",
  control = list(adapt_delta = 0.99, max_treedepth = 15)
)
```

Model Components

- **Formula**: `n_negative | trials(n_total) ~ 1 + (1 | id)`
  - `n_negative`: Count of negative posterior draws.
  - `n_total`: Total number of posterior draws
  - `(1 | id)`: Random intercept for each participant.
- **Family**: Binomial, appropriate for modeling proportions.

- **Priors**:
  - `normal(0, 1)` for the intercept and standard deviation of random effects.
  - These priors are weakly informative, allowing the data to drive the results while providing some regularization.
- **MCMC Settings**:
  - 100,000 iterations with 2,000 warmup.
  - 4 chains run in parallel.
  - Increased `adapt_delta` and `max_treedepth` for improved MCMC convergence.

Extension to Other Parameters

We applied a similar hierarchical modeling approach to analyze the posterior distributions of other key parameters from the idiographic analyses:

- `gamma_neg_aff`: Coefficient for negative affect.
- `gamma_context`: Coefficient for context evaluation.
- `gamma_interaction`: Coefficient for potential interaction effects.

Posterior Predictive Checks

In the following sections, we present the results of the Posterior Predictive Checks for the four models used to assess the aggregate effects and heterogeneity in our idiographic analysis. These models evaluate the linear association between uncompassionate self-responding (UCS) and compassionate self-responding (CS), the effect of negative affect, the influence of contextual evaluation, and the interaction between CS and negative affect.

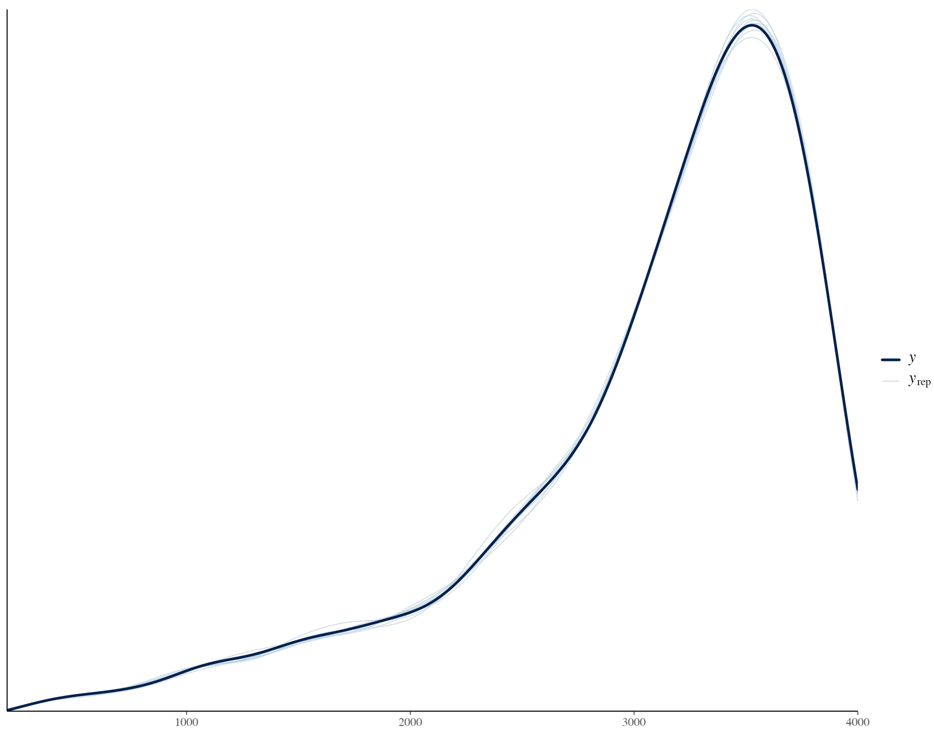

Figure 1: Idiographic analysis. Posterior Predictive Check for the hierarchical model fitted to the proportion of posterior draws of the `gamma_CS`parameter that were negative.

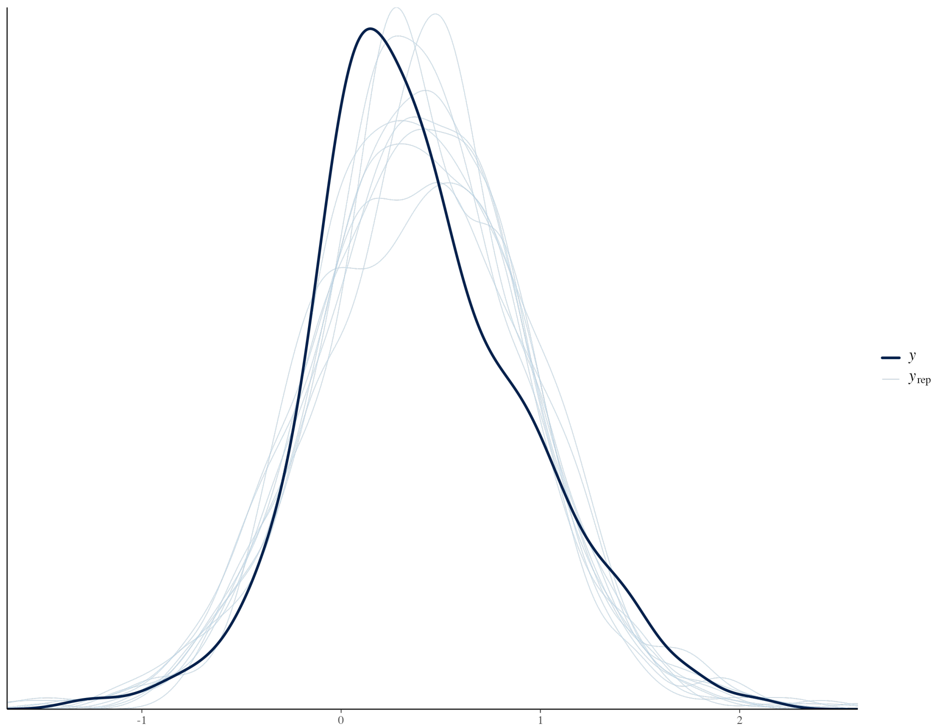

Figure 2: Idiographic analysis. Posterior Predictive Check for the hierarchical model fitted to the mean of the posterior draws of the `gamma_neg_aff` parameter.

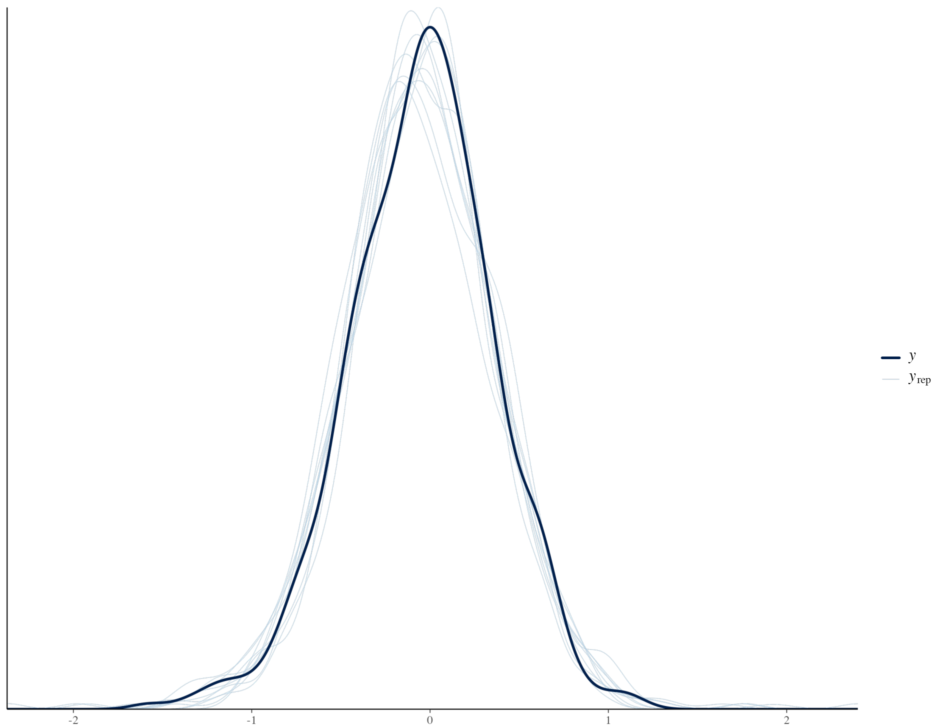

Figure 3: Idiographic analysis. Posterior Predictive Check for the hierarchical model fitted to the mean of the posterior draws of the `gamma_context` parameter.

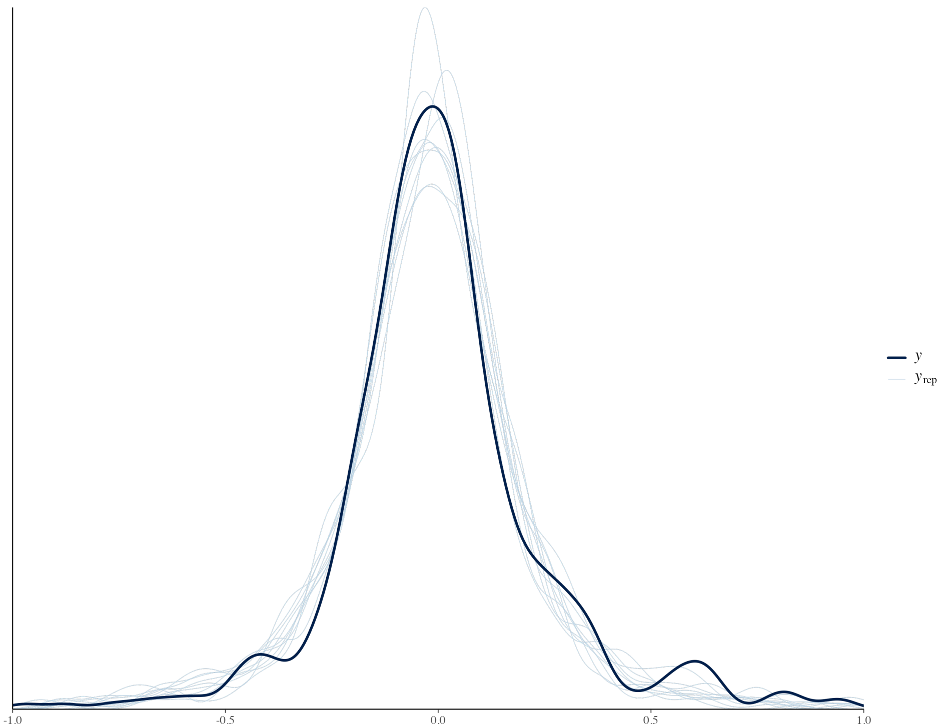

Figure 4: Idiographic analysis. Posterior Predictive Check for the hierarchical model fitted to the mean of the posterior draws of the `gamma_interaction` parameter.

Response to Reviewer Comments

Reviewer 1

- Point 2: "The introduction provides a reasonable rationale for testing the bipolar continuum hypothesis using EMA. However, I wonder if the authors could provide a brief review of recent self-compassion research using EMA (e.g., Mey et al., 2023; DOI: 10.1007/s12671-022-02050-y, which was in the reference list but was not cited) and explain how previous research might inform hypotheses."
  - *Thank you for this helpful suggestion. We have added a brief review of recent work on self-compassion using EMA to the introduction of the manuscript, drawing particularly on Mey et al. (2023). This work highlights that, while self-compassion (SC) has traditionally been viewed as a stable trait, there is growing evidence that SC fluctuates within individuals over time and that these fluctuations impact well-being. For instance, individuals report healthier eating behaviors, better body image, less perceived stress, and a lower likelihood of clinical impairment due to eating habits on days when they report higher SC than usual.*
  - Mey et al. (2023) utilized EMA to investigate the implications of momentary self-compassion (SC) for affective dynamics and well-being in daily life, finding that higher recent SC was related to higher momentary positive affect and lower momentary negative affect. Moreover, they found that higher recent SC was associated with lower stress reactivity to daily hassles. The study importantly shows that SC may not only be a beneficial trait but also a helpful state that can be cultivated in daily life to enhance well-being and mitigate the negative impact of stress. This is significant because it highlights the potential for individuals to experience the benefits of SC regardless of their general level of self-compassion. Mey et al. (2023) also discovered that SC fluctuates over time within the same individual, challenging the traditional view of SC as a static trait. This dynamic nature of SC supports our present design, which uses EMA to examine how the CS and UCS covary across different contexts to evaluate the hypothesis of a bipolar continuum. Furthermore, the study's findings suggest that repeatedly increasing one's momentary SC could potentially lead to increases in trait levels of SC, offering more enduring well-being benefits. This aligns with research on mindfulness, which has shown that repeatedly increasing state mindfulness in brief meditations over several weeks led to heightened trait mindfulness at the end of the intervention. The sources also note a "paradox" within the SC concept. While SC aims to alleviate suffering, it also emphasizes acceptance of the current moment, creating a potential tension between these two aspects. The findings of Mey et al. (2023) thus demonstrate the value of EMA as a method to explore the dynamic nature of SC and its implications for well-being. This method captures the fluctuation of SC within individuals, which might not be revealed by static, trait-based measures.

Why p = 0.89?

We generated 89% Credibility Intervals instead of the conventional 95% to avoid encouraging implicit hypothesis testing, as suggested by McElreath (2020). Bayesian inference prioritizes estimation over hypothesis testing, making alternative intervals, such as 89% or even 50%, equally valid and informative (e.g., Burger, Ralph-Nearman, and Levinson 2022).

Bibliography

Brenner, Rachel E, Patrick J Heath, David L Vogel, and Marcus Credé. 2017. "Two Is More Valid Than One: Examining the Factor Structure of the Self-Compassion Scale (SCS)." *Journal of Counseling Psychology* 64 (6): 696–707.

Burger, Julian, Christina Ralph-Nearman, and Cheri A Levinson. 2022. "Integrating Clinician and Patient Case Conceptualization with Momentary Assessment Data to Construct Idiographic Networks: Moving Toward Personalized Treatment for Eating Disorders." *Behaviour Research and Therapy* 159: 104221.

Ciarrochi, Joseph, Baljinder Sahdra, Madeleine I Fraser, Steven C Hayes, Keong Yap, and Andrew T Gloster. 2024. "The Compassion Connection: Experience Sampling Insights into Romantic Attraction." *Journal of Contextual Behavioral Science* 32: 100749.

Cleare, Seonaid, Andrew Gumley, Chris J Cleare, and Rory C O'Connor. 2018. "An Investigation of the Factor Structure of the Self-Compassion Scale." *Mindfulness* 9: 618–28.

Ferrando, Pere J, and Urbano Lorenzo-Seva. 2018. "Assessing the Quality and Appropriateness of Factor Solutions and Factor Score Estimates in Exploratory Item Factor Analysis." *Educational and Osychological Measurement* 78 (5): 762–80.

Ferrando, Pere Joan, and Urbano Lorenzo-Seva. 2019. "An External Validity Approach for Assessing Essential Unidimensionality in Correlated-Factor Models." *Educational and Psychological Measurement* 79 (3): 437–61.

Ferrari, Madeleine, Joseph Ciarrochi, Keong Yap, Baljinder Sahdra, and Steven C Hayes. 2022. "Embracing the Complexity of Our Inner Worlds: Understanding the Dynamics of Self-Compassion and Self-Criticism." *Mindfulness* 13 (7): 1652–61.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. 2nd ed. Chapman; Hall/CRC.

Mey, Lara Kristin, Mario Wenzel, Karolina Morello, Zarah Rowland, Thomas Kubiak, and Oliver Tüscher. 2023. "Be Kind to Yourself: The Implications of Momentary Self-Compassion for Affective Dynamics and Well-Being in Daily Life." *Mindfulness* 14 (3): 622–36.

Neff, Kristin D. 2016. "The Self-Compassion Scale Is a Valid and Theoretically Coherent Measure of Self-Compassion." *Mindfulness* 7: 264–74.

Neff, Kristin D, István Tóth-Király, Lisa M Yarnell, Kohki Arimitsu, Paula Castilho, Nima Ghorbani, Hailan Xiaoxia Guo, et al. 2019. "Examining the Factor Structure of the Self-Compassion Scale in 20 Diverse Samples: Support for Use of a Total Score and Six Subscale Scores." *Psychological Assessment* 31 (1): 27–45.

Neff, Kristin D, Tiffany A Whittaker, and Anke Karl. 2017. "Examining the Factor Structure of the Self-Compassion Scale in Four Distinct Populations: Is the Use of a Total Scale Score Justified?" *Journal of Personality Assessment* 99 (6): 596–607.

Petrocchi, Nicola, Cristina Ottaviani, and Alessandro Couyoumdjian. 2014. "Dimensionality of Self-Compassion: Translation and Construct Validation of the Self-Compassion Scale in an Italian Sample." *Journal of Mental Health* 23 (2): 72–77.

Reise, Steven P, Wes E Bonifay, and Mark G Haviland. 2013. "Scoring and Modeling Psychological Measures in the Presence of Multidimensionality." *Journal of Personality Assess-*

ment 95 (2): 129–40.

Rodriguez, Anthony, Steven P Reise, and Mark G Haviland. 2016. "Evaluating Bifactor Models: Calculating and Interpreting Statistical Indices." *Psychological Methods* 21 (2): 137–50.

Sahdra, Baljinder K, Joseph Ciarrochi, Korena S Klimczak, Jennifer Krafft, Steven C Hayes, and Michael Levin. 2024. "Testing the Applicability of Idionomic Statistics in Longitudinal Studies: The Example of 'Doing What Matters'." *Journal of Contextual Behavioral Science*, 100728.