The title

First Author1 & Ernst-August Doelle1,2

1 Wilhelm-Wundt-University

2 Konstanz Business School

Author note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

The authors made the following contributions. First Author: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Ernst-August Doelle: Writing - Review & Editing, Supervision.

Correspondence concerning this article should be addressed to First Author, Postal address. E-mail: [my@email.com](mailto:my@email.com)

Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline. Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines. One sentence clearly stating the **general problem** being addressed by this particular study. One sentence summarizing the main result (with the words “**here we show**” or their equivalent). Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge. One or two sentences to put the results into a more **general context**. Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

*Keywords:* keywords

*Word count:* X

The title

# Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

## Participants

## Material

## Procedure

## Data analysis

#### Modeling the Dynamic Relationship Between Compassionate Self and Uncompassionate Self: An Idiographic Approach.

In our study, the use of EMA data on state self-compassion enables us to test Neff’s hypothesis in an *idiographic* rather than purely *nomothetic* manner. Neff’s hypothesis posits a dynamic interaction between these two constructs, where increases in compassionate self-responding inhibit uncompassionate self-responding. This theoretical perspective suggests that CS and UCS are inversely related at the momentary level, and this relationship should be testable over time, accounting for *contextual variability*.

Following the approach proposed by Ferrari et al. (2022), we extended Neff’s framework by considering not just the global relationship between CS and UCS, but also the contextual factors that may modulate this relationship. Specifically, we considered a model that tests whether the link between CS and UCS is influenced by three key dimensions: *Time Frame* (momentary and daily fluctuations), *Current Circumstances* (momentary affective states and context evaluations), and *Individual Differences* (variability in participants’ baseline levels and responsiveness).

To achieve this, we formulated a *Bayesian hierarchical model* that incorporates these three dimensions. This model allows us to measure the connection between CS and UCS at the *individual level*, capturing how the dynamic interplay between these constructs changes over time, across different emotional and contextual conditions, and between individuals.

Our hierarchical Bayesian model posits that is influenced by a combination of fixed effects (e.g., the overall effect of CS) and random effects (e.g., individual differences in the effect of CS across participants). The Student-t distribution allows for flexibility in modeling the error structure, potentially accommodating outliers or non-normality in the data. To prevent overfitting, we utilized weakly informative regularization priors (see Supplementary Information for the full model specification).

This idiographic approach highlights the dynamic nature of the CS-UCS relationship, accounting for the fact that it varies across individuals and can be influenced by immediate emotional and situational factors. By incorporating momentary, day-level, and person-level predictors, the model provides a more precise description of how CS and UCS interact in everyday life.

Key components of our model include:

* *Compassionate (CS) and Uncompassionate (UCS) self responding*: The core variables of interest, where increases in CS are hypothesized to inhibit UCS at the individual level.
* *Momentary, Day-Level, and Person-Level Predictors*: The model includes momentary and day-level *negative affect* and *context evaluations* as predictors of the CS-UCS relationship. It also accounts for baseline differences between participants, allowing us to capture individual-specific variations in this relationship.
* *Bayesian Hierarchical Structure*: This structure allows for individual-specific estimates of how CS influences UCS, while also capturing general patterns across participants.

By including all these factors, we not only test Neff’s hypothesis at the group level but also explore how it applies *uniquely to each individual*, making it a truly idiographic analysis, while accounting for the different levels of centering and scaling in both negative affect and context evaluation.

# Results

The key parameters of the Bayesian hierarchical model are , , and . While the fixed-effect analysis supported Neff’s hypothesis, an idiographic analysis revealed significant individual heterogeneity in the relationship between UCS and CS. The posterior mean of was -0.380 (89% CI [-0.406, -0.352]), indicating a general negative association at the nomothetic level. However, a large of 0.324, 89% CI [0.304, 0.345], suggests substantial individual variability in this relationship. Indeed, only 67% of individuals exhibited the expected negative association between UCS and CS, highlighting the importance of considering individual-level variation when evaluating Neff’s hypothesis (see Figure XX). This result is consistent with hypothesis 3 (“Individual Differences Moderate the Structure of Self-compassion”) of Ferrari, Ciarrochi, Yap, Sahdra, and Hayes (2022).

In our model, we included an interaction term to test whether momentary negative affect influences the relationship between CS and UCS. The estimated posterior mean for this interaction term () was -0.0304, with an 89% credibility interval ranging from -0.0380 to -0.0230. This suggests that, during moments of higher negative affect, the inverse relationship between CS and UCS becomes more pronounced. This result aligns with the hypothesis that emotional distress amplifies the dynamic opposition between self-compassion and self-criticism, and provides evidence for the context-dependency of this relationship. Specifically, higher levels of negative affect strengthen the extent to which CS leads to reductions in UCS. This result is consistent with Ferrari et al. (2022) and Dejonckheere et al. (2021).

# Discussion

We used R [Version 4.4.1; R Core Team (2024)] for all our analyses.

# Supplementary Material Description

Here, we provide a detailed description of the Bayesian hierarchical model used to test the relationship between compassionate (CS) and uncompassionate (UCS) self-responding in the context of Neff’s bipolar hypothesis. The model’s structure incorporates multiple levels of variation, capturing *momentary* and *daily changes* in emotional and contextual factors, as well as *individual differences* in baseline CS and UCS levels.

The core structure of the model can be described as follows:

1. **Data Structure**:
   * *CS and UCS* are measured repeatedly over time via Ecological Momentary Assessment (EMA), with participants reporting their CS and UCS scores multiple times a day.
   * The model also includes momentary and day-level predictors such as *negative affect* and *context evaluation*.
2. **Model Parameters**:
   * *Fixed Effects*: The model includes an overall effect of CS on UCS (beta\_cs), reflecting the hypothesized *inverse relationship* between these two constructs. A negative beta\_cs would support Neff’s hypothesis that *higher CS predicts lower UCS*.
   * **Random Effects**: The model allows for *individual-level deviations* in this relationship, captured by the *random slopes* for CS (z\_participant\_slope\_cs), as well as individual differences in baseline UCS levels (z\_participant).
   * **Interaction Terms**: The model includes an interaction between *CS and momentary negative affect* (beta\_interaction), allowing us to test whether the inverse CS-UCS relationship is modulated by emotional states.
   * **Contextual Effects**: To account for fluctuations across time and context, the model includes three levels of predictors for *negative affect* and *context evaluation*:
     + *Momentary (within-day)*: Measures the effect of immediate negative affect and context on UCS.
     + *Daily (between-days)*: Measures the effect of average negative affect and context across a given day.
     + *Person-level (between-participants)*: Captures individual differences in baseline negative affect and context evaluations.
3. **Bayesian Structure**:
   * The model is implemented using Bayesian inference, which allows for posterior distributions of all parameters to be computed. This provides a full probabilistic representation of uncertainty around the estimated effects.
   * *Hierarchical Priors*: The hierarchical structure of the model allows for individual-specific estimates of the relationship between CS and UCS. Random slopes are modeled with normal distributions centered on the group-level fixed effects, capturing the variability across participants.
4. **Generated Quantities**:
   * The model includes a *generated quantities* block to compute *posterior predictions* for UCS (pred\_UCS) at each measurement occasion, based on the estimated relationships between CS, negative affect, and context.
   * *Log-likelihoods* (log\_lik) for each observation are computed, enabling posterior predictive checks and model comparison using *leave-one-out cross-validation (LOO)*.
5. **Sampling**:
   * The model was estimated using Markov Chain Monte Carlo (MCMC) sampling. Four chains were run with 2000 iterations each for burn-in and 2000 iterations for sampling. This approach provided efficient approximation of the posterior distribution.

#### Key Parameters and Interpretation.

* **beta\_cs**: Represents the overall effect of *CS on UCS*. A *negative value* would indicate that, on average, increases in CS are associated with decreases in UCS, supporting Neff’s hypothesis.
* **beta\_interaction**: Tests the interaction between *CS and negative affect*, indicating whether the CS-UCS relationship becomes stronger or weaker depending on momentary emotional states.
* **sigma\_participant\_slope\_cs**: The standard deviation of the *random slopes* for CS, indicating how much the effect of CS on UCS varies across individuals.
* **Random Intercepts**: Variability in UCS across participants (sigma\_participant), days (sigma\_day), and measurement occasions (sigma\_measurement), capturing how uncompassionate responses fluctuate due to time or context.
* **Posterior Predictive Checks and Model Fit Assessment**: To evaluate model fit, we employed posterior predictive checks and leave-one-out cross-validation (LOO). Pareto statistics were calculated to quantify how well each data point was predicted by the model. There were no divergent transitions. Maximum Rhat: 1.0150; mean Rhat: 1.0006.

The full Stan code for the model is provided below, along with additional posterior predictive checks, and the full posterior parameter estimates.

data {  
 int<lower=0> N; // Total number of observations  
 int<lower=0> P; // Number of participants  
 int<lower=0> D; // Number of days  
 int<lower=0> M; // Number of measurements per day per participant  
 array[N] int<lower=1, upper=P> participant; // Participant index for each observation  
 array[N] int<lower=1, upper=D> day; // Day index for each observation  
 array[N] int<lower=1, upper=M> measurement; // Measurement index for each observation  
 array[N] real CS; // Compassionate Self measures  
 array[N] real UCS; // Uncompassionate Self measures  
  
 // Scaled and centered negative affect variables  
 array[N] real neg\_aff\_Moment; // Negative affect moment-centered  
 array[N] real neg\_aff\_Day; // Negative affect day-centered  
 array[N] real neg\_aff\_Person; // Negative affect person-centered  
  
 // Scaled and centered context evaluation variables  
 array[N] real context\_eval\_Moment; // Context evaluation moment-centered  
 array[N] real context\_eval\_Day; // Context evaluation day-centered  
 array[N] real context\_eval\_Person; // Context evaluation person-centered  
}  
  
parameters {  
 // Fixed effects  
 real alpha\_ucs; // Intercept for UCS  
 real beta\_cs; // Overall effect of CS on UCS  
 real beta\_interaction; // Effect of interaction between CS and NA  
  
 // Coefficients for the three negative affect components  
 real beta\_neg\_aff\_Moment;  
 real beta\_neg\_aff\_Day;  
 real beta\_neg\_aff\_Person;  
  
 // Coefficients for the three context evaluation components  
 real beta\_context\_eval\_Moment;  
 real beta\_context\_eval\_Day;  
 real beta\_context\_eval\_Person;  
  
 // Random intercepts  
 vector[P] z\_participant; // Random intercept for participants  
 vector[D] z\_day; // Random intercept for days  
 vector[M] z\_measurement; // Random intercept for measurements  
  
 // Random slopes for CS at the participant level  
 vector[P] z\_participant\_slope\_cs;  
  
 // Random slopes for NA at the participant level  
 vector[P] z\_participant\_slope\_na;  
  
 // Variance parameters  
 real<lower=0> sigma\_participant; // SD of participant intercepts  
 real<lower=0> sigma\_day; // SD of day intercepts  
 real<lower=0> sigma\_measurement; // SD of measurement intercepts  
 real<lower=0> sigma\_participant\_slope\_cs; // SD of participant slopes for CS  
 real<lower=0> sigma\_participant\_slope\_na; // SD of participant slopes for NA  
 real<lower=0> sigma\_ucs; // Error term for UCS model  
  
 real<lower=0> nu; // Degrees of freedom for t-distribution  
}  
  
model {  
 // Priors for fixed effects  
 alpha\_ucs ~ normal(0, 1);  
 beta\_cs ~ normal(0, 1);  
 beta\_interaction ~ normal(0, 1); // Prior for interaction term  
 beta\_neg\_aff\_Moment ~ normal(0, 1);  
 beta\_neg\_aff\_Day ~ normal(0, 1);  
 beta\_neg\_aff\_Person ~ normal(0, 1);  
 beta\_context\_eval\_Moment ~ normal(0, 1);  
 beta\_context\_eval\_Day ~ normal(0, 1);  
 beta\_context\_eval\_Person ~ normal(0, 1);  
  
 // Priors for random effects (latent variables)  
 z\_participant ~ normal(0, 1);  
 z\_day ~ normal(0, 1);  
 z\_measurement ~ normal(0, 1);  
 z\_participant\_slope\_cs ~ normal(0, 1);  
 z\_participant\_slope\_na ~ normal(0, 1);  
  
 // Priors for variances  
 sigma\_participant ~ exponential(1);  
 sigma\_day ~ exponential(1);  
 sigma\_measurement ~ exponential(1);  
 sigma\_participant\_slope\_cs ~ exponential(1);  
 sigma\_participant\_slope\_na ~ exponential(1);  
 sigma\_ucs ~ exponential(1);  
 nu ~ gamma(2, 0.1);  
  
 // Likelihood for UCS using a t-distribution  
 for (n in 1:N) {  
 UCS[n] ~ student\_t(  
 nu,  
 // Random intercepts  
 alpha\_ucs +  
 sigma\_participant \* z\_participant[participant[n]] +  
 sigma\_day \* z\_day[day[n]] +  
 sigma\_measurement \* z\_measurement[measurement[n]] +  
   
 // Main effect of CS with individual random slope  
 (beta\_cs + sigma\_participant\_slope\_cs \* z\_participant\_slope\_cs[participant[n]]) \* CS[n] +  
   
 // Main effect of NA (including random slope for NA)  
 (beta\_neg\_aff\_Moment + sigma\_participant\_slope\_na \* z\_participant\_slope\_na[participant[n]]) \* neg\_aff\_Moment[n] +  
 beta\_neg\_aff\_Day \* neg\_aff\_Day[n] +  
 beta\_neg\_aff\_Person \* neg\_aff\_Person[n] +  
   
 // Interaction term between CS and NA (Moment-level)  
 beta\_interaction \* CS[n] \* neg\_aff\_Moment[n] +  
   
 // Main effect of context evaluation components  
 beta\_context\_eval\_Moment \* context\_eval\_Moment[n] +  
 beta\_context\_eval\_Day \* context\_eval\_Day[n] +  
 beta\_context\_eval\_Person \* context\_eval\_Person[n],  
   
 // Error term  
 sigma\_ucs  
 );  
 }  
}  
  
generated quantities {  
 array[N] real pred\_UCS; // Posterior predictions for UCS  
 array[N] real log\_lik; // Log-likelihood for UCS  
  
 for (n in 1:N) {  
 pred\_UCS[n] = student\_t\_rng(  
 nu,  
 alpha\_ucs +  
 sigma\_participant \* z\_participant[participant[n]] +  
 sigma\_day \* z\_day[day[n]] +  
 sigma\_measurement \* z\_measurement[measurement[n]] +  
 (beta\_cs + sigma\_participant\_slope\_cs \* z\_participant\_slope\_cs[participant[n]]) \* CS[n] +  
 (beta\_neg\_aff\_Moment + sigma\_participant\_slope\_na \* z\_participant\_slope\_na[participant[n]]) \* neg\_aff\_Moment[n] +  
 beta\_neg\_aff\_Day \* neg\_aff\_Day[n] +  
 beta\_neg\_aff\_Person \* neg\_aff\_Person[n] +  
 beta\_interaction \* CS[n] \* neg\_aff\_Moment[n] +  
 beta\_context\_eval\_Moment \* context\_eval\_Moment[n] +  
 beta\_context\_eval\_Day \* context\_eval\_Day[n] +  
 beta\_context\_eval\_Person \* context\_eval\_Person[n],  
 sigma\_ucs  
 );  
  
 log\_lik[n] = student\_t\_lpdf(  
 UCS[n] |  
 nu,  
 alpha\_ucs +  
 sigma\_participant \* z\_participant[participant[n]] +  
 sigma\_day \* z\_day[day[n]] +  
 sigma\_measurement \* z\_measurement[measurement[n]] +  
 (beta\_cs + sigma\_participant\_slope\_cs \* z\_participant\_slope\_cs[participant[n]]) \* CS[n] +  
 (beta\_neg\_aff\_Moment + sigma\_participant\_slope\_na \* z\_participant\_slope\_na[participant[n]]) \* neg\_aff\_Moment[n] +  
 beta\_neg\_aff\_Day \* neg\_aff\_Day[n] +  
 beta\_neg\_aff\_Person \* neg\_aff\_Person[n] +  
 beta\_interaction \* CS[n] \* neg\_aff\_Moment[n] +  
 beta\_context\_eval\_Moment \* context\_eval\_Moment[n] +  
 beta\_context\_eval\_Day \* context\_eval\_Day[n] +  
 beta\_context\_eval\_Person \* context\_eval\_Person[n],  
 sigma\_ucs  
 );  
 }  
}

In summary, the idiographic approach enabled by this model allows for a deeper exploration of Neff’s bipolar hypothesis, testing whether the CS-UCS relationship holds not only at the group level but also at the *individual level*, across different time frames and contexts.

## Additional Model Results

In this section, we provide a comprehensive overview of the additional results from the Bayesian hierarchical model, focusing on the relationship between **Compassionate Self (CS)** and **Uncompassionate Self (UCS)**, with further details on the **random effects**, **interaction terms**, and other key parameters.

#### 1. **Posterior Estimates of Fixed Effects**.

The core fixed effects in the model include and , which have been discussed in the main manuscript.

#### 2. **Random Effects and Variability Across Individuals**.

One of the strengths of this hierarchical model is its ability to capture **individual variability** in the relationship between **CS** and **UCS**.

* : Also this parameter has been discussed in the main text.
* **Random Intercepts**:
  + : The posterior mean of this parameter is **0.458**, with an **89% CI** of **[0.437, 0.480]**, indicating substantial variation in baseline **UCS** levels across individuals.
  + : The posterior mean of this parameter is **0.089**, with an **89% CI** of **[0.076, 0.102]**, capturing daily fluctuations in **UCS**.
  + : The posterior mean of this parameter is **0.067**, with an **89% CI** of **[0.059, 0.075]**, capturing within-day, moment-to-moment variability in **UCS**.

#### 3. **Contextual Effects: Negative Affect and Context Evaluation**.

To account for contextual variability, the model includes **momentary**, **daily**, and **person-level** predictors for both **negative affect** and **context evaluation**.

* **Negative Affect (NA)**:
  + : The momentary effect of negative affect on **UCS** was estimated to have a posterior mean of **0.092**, with an **89% CI** of **[0.071, 0.113]**, indicating that higher momentary negative affect is associated with increased levels of **UCS**.
  + : The day-level effect of negative affect on **UCS** was estimated to have a posterior mean of **0.047**, with an **89% CI** of **[0.026, 0.068]**, indicating that higher average negative affect across the day contributes to increased **UCS**.
  + : The person-level effect of negative affect was estimated to have a posterior mean of **0.067**, with an **89% CI** of **[0.045, 0.089]**, capturing how individuals with higher baseline levels of negative affect tend to have higher levels of **UCS**.
* **Context Evaluation**:
  + : The momentary effect of context evaluation on **UCS** was estimated to have a posterior mean of **-0.061**, with an **89% CI** of **[-0.082, -0.040]**, suggesting that more positive context evaluations at a given moment are associated with lower levels of **UCS**.
  + : The day-level effect of context evaluation on **UCS** was estimated to have a posterior mean of **-0.042**, with an **89% CI** of **[-0.064, -0.021]**, indicating that days characterized by more positive evaluations tend to be associated with lower levels of **UCS**.
  + : The person-level effect of context evaluation was estimated to have a posterior mean of **-0.056**, with an **89% CI** of **[-0.078, -0.034]**, indicating that individuals with more positive overall evaluations of their context tend to experience lower levels of **UCS**.

#### 4. **Posterior Predictive Checks and Model Fit**.

To assess the fit of the model, we performed **posterior predictive checks** using **density overlays** and **histogram comparisons** between the observed data and the posterior predictions. The results indicated that the model accurately captures the distribution of **UCS** scores, providing confidence in the model’s predictive performance.

We also computed **leave-one-out cross-validation (LOO)** for model comparison and fit assessment. The **Pareto statistics** indicated that all observations were well predicted, with no problematic points (i.e., for all data points). The maximum value across all parameters was **1.0150**, with a mean value of **1.0006**, suggesting good convergence.

#### 5. **Discussion of Model Results**.

The results of the model support **Neff’s hypothesis** that **CS and UCS** are dynamically related, with **increases in CS leading to reductions in UCS**. However, the large variability across individuals, as captured by , suggests that this relationship is not uniform. While most participants exhibit the expected inverse relationship, a significant minority either show a weaker association or even a positive relationship, suggesting that the interaction between **self-compassion** and **self-criticism** may be influenced by individual traits or momentary emotional states.

Additionally, the significant interaction between **CS and momentary negative affect** () highlights the importance of **context-dependent fluctuations** in the relationship between these two constructs. When individuals are experiencing higher levels of negative affect, the protective effects of **self-compassion** on **self-criticism** appear to be stronger.

These findings align with the theoretical proposal that **individual differences** and **emotional context** play key roles in shaping the structure of **self-compassion** and **self-criticism**, as discussed by Ferrari et al. (2022) and De Jonckheere et al. (2021).

#### **Conclusion**.

This idiographic approach, facilitated by our Bayesian hierarchical model, allows for a deeper exploration of the dynamic relationship between **CS and UCS** across multiple levels of contextual and individual variation. Our findings provide empirical support for Neff’s hypothesis but also highlight the complexity and variability of this relationship across individuals and emotional contexts.

# References

Dejonckheere, E., Mestdagh, M., Verdonck, S., Lafit, G., Ceulemans, E., Bastian, B., & Kalokerinos, E. K. (2021). The relation between positive and negative affect becomes more negative in response to personally relevant events. *Emotion*, *21*(2), 326–336.

Ferrari, M., Ciarrochi, J., Yap, K., Sahdra, B., & Hayes, S. C. (2022). Embracing the complexity of our inner worlds: Understanding the dynamics of self-compassion and self-criticism. *Mindfulness*, *13*(7), 1652–1661.

R Core Team. (2024). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>