Supplementary Materials

Characterizing Within-Person Trajectories of Negative Affect Across Adolescence: A Longitudinal Clustering Approach

Contents

1. Inclusion and exclusion criteria

Supplementary Table 1. Overview of inclusion and exclusion criteria used in the study (Somerville et al., 2018).

Inclusion Criteria	**Exclusion Criteria**
Age 5-21 years	Premature birth
Speaks English well	Serious medical conditions (e.g., stroke, cerebral palsy)
Safe to enter MRI	Serious endocrine condition (e.g., precocious puberty, untreated growth hormone deficiency)
	Long term use of immunosuppressants or steroids
	Any history of serious head injury
	Hospitalization >2 days for certain physical or psychiatric conditions or substance use
	Treatment >12 months for psychiatric conditions
	Receiving certain special services at school
	Claustrophobia
	Pregnancy

2. Functional outcome survey descriptions

Supplementary Table 2. NIH Emotion Toolbox Social Functioning and Life Satisfaction Measure Descriptions (Gershon et al., 2013)

Functional Outcome Domain	Description
Emotional Support	The perception that people in one's social network are available to listen to one's problems with empathy, caring, and understanding
Friendship	Perceptions of the availability of friends or companions with whom to interact or affiliate
Perceived Hostility	Perceptions of hostility, e.g., measuring the perceptions of how often people argue with me, yell at me, or criticize me)
Perceived Rejection	Perceptions of rejection in daily interactions, e.g., how often people don't listen when I ask for help, or don't pay attention to me
Loneliness	The perception that one is alone, lonely, or socially isolated from others
General Life Satisfaction	Global feelings and attitudes about one's life; assessment whether the participant likes his or her life

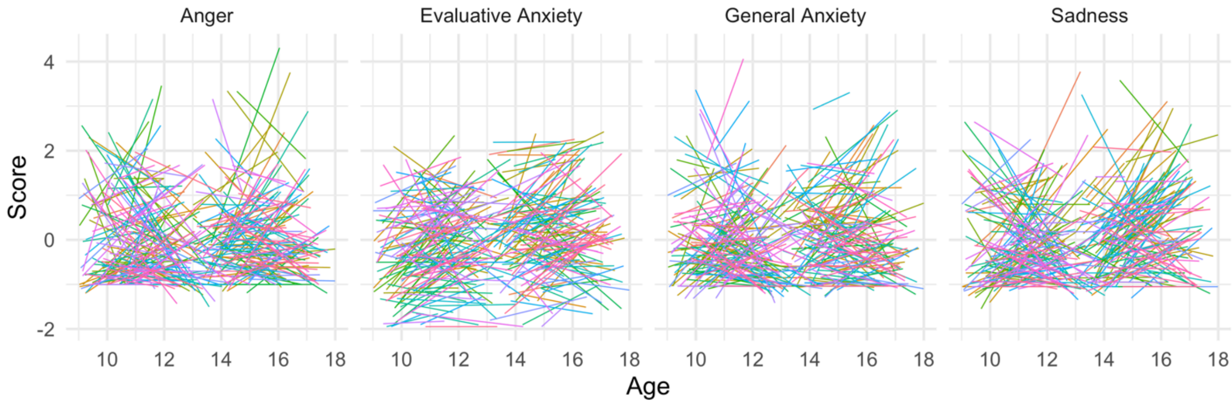
3. Analysis of COVID-related impact to analyses

Data collection for Wave 3 longitudinal participants was ongoing when the COVID-19 pandemic began and paused testing for several months, resulting in increased variability in the length of time between Waves 2 and 3. We include a brief supplemental analysis and discussion of the potential effects of COVID on our study.

Because all our analyses use precise age instead of timepoint to capture the time between visits for each participant, this variability in Wave 3 visit time did not impact our analysis. However, the pandemic had a profound effect on individuals' lives and may have influenced the emotional experiences and well-being of participants who were collected after its onset. Because of our recruitment design, post-COVID Wave 3 testing affected participants across different ages (as shown in Figure 1 in the main manuscript), thus, this confound is less likely to affect the shape of the developmental trends investigated in this study compared to a longitudinal study design tracking participants of the same age. We confirmed that the average age of participants at Wave 3 who were tested before the pandemic did not differ from the average age of participants tested after the onset of the pandemic (average Wave 3 age pre-pandemic: 14.86 years, average Wave 3 age post-pandemic: 14.90 years).

Additionally, we explored whether average negative affect scores differed between the Wave 3 pre-pandemic and Wave 3 post-pandemic participants. We found that for all negative affect types, there was no significant difference in scores between Wave 3 pre- and Wave 3 post-pandemic participants (anger: $t = 0.92$, df = 44. 76, p = .36; evaluative anxiety: $t = -0.74$, df = 55.36, p = 0.46; general anxiety: $t = 0.09$, df = 52.30, p = .46; sadness: $t = 1.37$, df = 44.45, p = .18). Despite our findings that the pandemic did not disproportionately affect a certain age group and did not significantly affecting negative affect scores, for all models conducted in the longitudinal sample we include a COVID control variable indicating whether the assessment was collected before or after the testing pause. To the extent possible, these precautions helped to isolate the changes in affect across time that are linked to normative developmental changes and reduced potential COVID-specific affect changes.

4. Individual Affect Trajectories Across Timepoints



Supplementary Figure 1. Individual trajectories across timepoints plotted for each negative affect type. Each line represents an individual. The *geom_smooth* function with method "lm" from the ggplot2 package was used for visualization.

5. Quantifying Degree of Between-Person Variability

To motivate our aim to characterize the heterogeneity in affect trajectories across age, we conducted a preliminary analysis to explore how much between-person variability exists in the data. To quantify this variability, we conducted a series of mixed-effects models using age and negative affect type to predict an individual's affect score. Because the cross-sectional age-affect relationships were previously found to be nonlinear, we fit generalized additive mixed models (GAMMs), an extension of generalized additive models incorporating random effects, with a smooth age term fit for each negative affect type, using the *gamm4* package (Wood & Scheipl, 2020) in R (R Core Team, 2023). In these models, smooth functions of covariates (i.e., age) are represented by penalized regression splines. The penalty for each smooth term is treated as a random effect, while the unpenalized component is treated as fixed. This approach balances model flexibility and parsimony by estimating the smoothness of the term through its variance parameter (Wood & Scheipl, 2020). The result is a stable, smooth curve that captures age-related changes in affect without being constrained to stereotypical polynomial shapes. To convey the complexity of the fitted curves, we report the effective degrees of freedom (EDF) of the smooth terms. For example, EDFs of 1, 2, and 3 represent approximately linear, quadratic, and cubic effects, respectively (Wood, 2017).

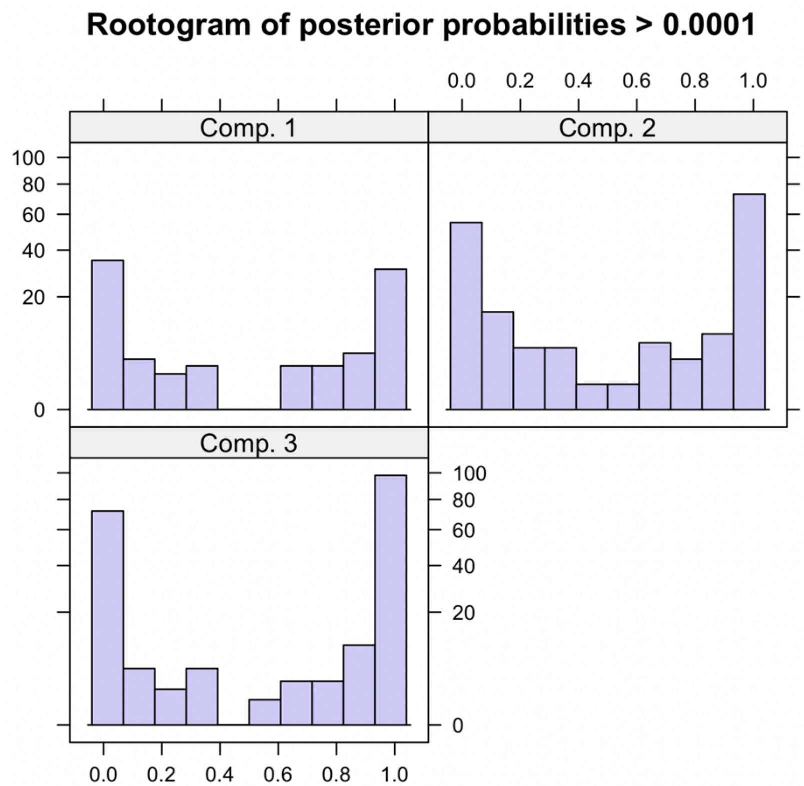
We tested five models with increasingly complex random effects structures to capture this variability, specified as follows: 1) no random effects: individual participants' intercepts and slopes are not allowed to vary, 2) random intercept (participant): individual intercepts are allowed to vary, but individual slopes are fixed, 3) random slope (age): both individual intercepts and slopes across age are allowed to vary; 4) two random slopes (age and negative affect type): individual intercepts and slopes across age are allowed to vary, and intercepts are additionally allowed to vary by negative affect type, 5) random slope interaction (age x negative affect type): individual intercepts and slopes across age are allowed to vary, and both intercepts and slopes are additionally allowed to vary by negative affect type. These random effect structures allow us to explore how much the model improves when allowing for increasing levels of individual subject variability, from only allowing their initial levels of affect to vary to allowing both their initial affect levels and affect score trajectory slopes to differ separately for each negative affect type. Exploring sex differences and the effect of the COVID-19 pandemic were not primary aims for the current analysis, so we included sex assigned at birth and a binary COVID variable in the models as controls. We compared the five models using AIC values to identify the random effect structure that best fits the data.

Visual examination of each participant's affect trajectory across visit (**Supplementary Figure 1**) reveals that within the average age curves previously identified, there is a high degree of individual variability. We found that the fifth model tested (i.e., the model in which both random intercepts and slopes were allowed to vary across different affect types) had the lowest AIC value (AIC = 6983, compared to AIC for model without random effects = 7782) and was selected for further inference. The variances of the random intercepts and slopes, indicating the degree of variability between subjects in their average levels of negative affect and their affect trajectory slopes across time, are reported in **Supplementary Table 3**. Between-subject variability in intercept and slope varied across negative affect types; 14% of the total variance in emotion score is due to individual differences in anger levels, whereas 0.2% of the variance in emotion score is due to individual differences in sadness slopes. These results indicate there is a large degree of individual variation in both initial magnitude of the negative affect score and its slope over time, and that these individual differences vary by negative affect type. This large degree of variation from averaged age curves motivates our aim to characterize the observed heterogeneity in affective trajectories.

Supplementary Table 3. Variability between individual participants' initial levels of negative affect (*random intercept variance*) and negative affect trajectories across time (*random slope variance*). Values are variances and percentage of variance explained is in parentheses.

	Random Intercept Variance (%)	Random Slope Variance (%)
Anger	0.48 (14.2%)	0.34 (7.0%)
Evaluative Anxiety	0.56 (11.6%)	0.13 (2.7%)
General Anxiety	0.29 (6.0%)	0.03 (0.6%)
Sadness	0.09 (1.9%)	0.01 (0.2%)

6. Mixture Regression Rootogram of Posterior Probabilities



Supplementary Figure 2. Rootogram generated from the *plot* method from the flexmix package. Each subplot corresponds to a cluster of individuals. Comp. 1 (Component 1) refers to Cluster 1, or the "Low-Stable" cluster; Comp. 2 (Component 2) refers to Cluster 2, or the "Moderate-Increasing" cluster; Comp. 3 (Component 3) refers to Cluster 3, or the "High-Increasing" cluster. The height of the bars corresponds to the square roots of counts to allow for low counts to be visible and peaks less emphasized. The y-axis denotes the number of observations in each bar. Because each component typically has many observations with posteriors close to zero (thus obscuring the information in other bins), all probabilities with a posterior below .0001 are ignored. Peaks near 0 and 1 indicate points clearly fitting or clearly not fitting that cluster (indicating good separation), while points in the middle of the distribution indicate a lack of separation.

7. Mixture Regression Summary Table

Supplementary Table 3. Mixture regression summary results.

	Prior	Size	Post>0	Ratio
Comp. 1	0.166	464	930	0.499
Comp. 2	0.377	1027	1924	0.534
Comp. 3	0.457	1255	2158	0.582

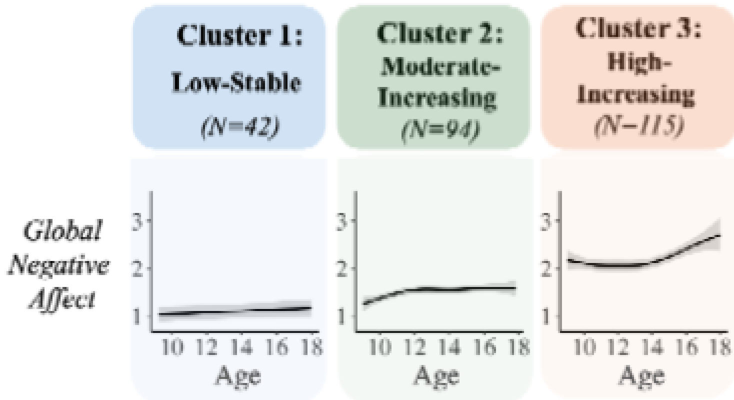
Note: Output generated by the *summary* method from the flexmix package. For each cluster, the prior probability (prior), the number of observations assigned to the corresponding cluster (size), the number of observations with a posterior probability larger than .0001 (post>0) and the ratio of the latter two numbers (ratio; indicates how separated the cluster is from the others) is provided. For example, Cluster 3 contained 2158 points with non-zero likelihood of being in that cluster, and 58.2% of those points were best fit by that cluster.

The mixture regression analysis revealed a three-cluster solution provided the best fit for the data as indicated by BIC values (BIC=7142.28). Out of the observations with a non-zero likelihood of being in a cluster, 49.9% of those observations were best fit by that cluster for Cluster 1, 53.4% for Cluster 2, and 58.2% for Cluster 3. The rootogram (**Supplementary Figure 2**) shows peaks near 1 (indicating many of the points are overwhelmingly well-represented by that cluster) and 0 (indicating points that clearly don't fit the cluster) and few points in the middle of the distribution, indicating that the clusters are well separated.

8. Cluster trajectories for global negative affect

After cluster identification, the negative affect trajectories for *global negative affect*, calculated as an average of the score for each type of negative affect, was modeled using generalized additive models consistent with prior models (i.e., global negative affect score as the dependent variable, the spline of age as a predictor, and sex and a binary covid variable as covariates of non-interest) to aid in cluster description. Global negative affect scores were calculated solely for cluster interpretation and were not used in the primary clustering analysis.

In global negative affect endorsement (**Supplementary Figure 3**), the Low-Stable cluster was characterized by low and stable levels of global negative affect (F=1.99, EDF=1, p=.162), the Moderate-Increasing cluster was characterized by moderate levels of global negative affect and a non-linear age trajectory, such that negative affect increased during childhood (~9-12 years) followed by a plateauing across adolescence (F=4.48, EDF=3.09, p=.002), and the High-Increasing cluster was characterized by stable levels of negative affect across childhood with sharp increases beginning around age 14 (F=6.71, EDF=2.88, p<.001). Thus, the Moderate-Increasing and High-Increasing clusters interestingly showed opposite non-linear trends. Given that the High-Increasing cluster showed the poorest functional outcomes, steep rises in negative affect across the adolescent years may be especially predictive of poorer well-being compared to negative affect increases that level off before adolescence.



Supplementary Figure 3. Negative affect trajectories of global negative affect for each cluster. The Y-axis scales shown correspond to an average of each negative affect score. Global negative affect was calculated and plotted for each cluster to aid in cluster interpretation and visualization but was not included in the finite mixture regression modeling.

As an additional supplementary analysis, we evaluated whether modeling distinct negative affect types provided additional explanatory value for functional outcomes beyond a global negative affect approach. We compared two clustering solutions: 1) the primary model based on four disaggregated negative affect types (anger, sadness, general anxiety, evaluative anxiety) and 2) an alternative model based on a global negative affect score. For both cluster solutions, one-way ANOVAs were conducted to examine between-cluster differences on each of the six functional outcome variables (emotional support, friendship quality, life satisfaction, loneliness, perceived hostility, perceived rejection) at the final study timepoint (V3). Eta squared (η^2) was calculated to estimate the proportion of variance in each outcome explained by cluster membership.

The disaggregated negative affect model explained more variance than the global model for four of the six functional outcomes: emotional support ($\Delta\eta^2 = .004$), perceived hostility ($\Delta\eta^2 = .021$), perceived rejection ($\Delta\eta^2 = .015$), and life satisfaction ($\Delta\eta^2 = .003$). For the remaining two outcomes, friendship quality ($\Delta\eta^2 = -.004$) and loneliness ($\Delta\eta^2 = -.028$), the global model accounted for more variance.

These exploratory findings suggest that differentiating among negative affect types may provide added value in predicting functional outcomes for specific functional outcome domains. Although the improvement in explanatory power was modest, this approach may capture outcome-specific associations that could be missed by a global negative affect approach.

9. Sensitivity analysis to evaluate effect of binary COVID variable

To evaluate the effect of the binary COVID variable that was included in all models to control for whether testing occurred before or after the onset of the COVID-19 pandemic, we conduct a sensitivity analysis and compare models with and without this term. First, the mixture regression clustering analysis was re-run without the addition of the COVID term, and hard cluster assignments (i.e., which participants fell into which clusters) remained identical for the three-cluster solution. Second, the regression models examining affect slope as a predictor of functional outcomes were re-run without the COVID variable. Models were conducted as described in the manuscript, with the six functional outcome measures collected at the latest timepoint (Wave 3) as the dependent variables, the four negative affect trajectory slopes were the predictors and controls including the baseline level of affect magnitude for all four affect types, the baseline level of the functional outcome examined, sex, and the spline of age. Multiple comparisons were controlled for using a Bonferroni correction and an adjusted alpha of .008. Results are shown in **Supplementary Table 4**.

Supplementary Table 4. Comparison of regression models evaluating effect of affect slope on functional outcome variables with and without binary COVID variable.

		COVID variable included (Results reported in main manuscript)			**COVID variable not included**		
Dependent variable	Independent variables	Coefficient	t	p-value	Coefficient	t	p-value
Emotional Support	General Anxiety Slope	-0.11	-1.10	.272	-0.22	-1.13	.260
	Anger Slope	-0.24	-2.62	.009	-0.24	-2.63	.009
	Sadness Slope	-0.20	-2.10	.037	-0.19	-1.98	.049
	Evaluative Anxiety Slope	0.015	0.17	.865	0.01	0.11	.916
Friendship	General Anxiety Slope	-0.17	-1.70	.091	-0.17	-1.72	.087
	Anger Slope	0.03	0.35	.724	0.03	0.34	.734

	Sadness Slope	**-0.26**	**-2.76**	**.006**	-0.25	-2.62	.009
	Evaluative Anxiety Slope	-0.09	-1.04	.298	-0.10	-1.12	.263
Perceived Hostility	General Anxiety Slope	0.15	1.96	.051	0.14	1.89	.060
	Anger Slope	**0.27**	**4.00**	**<.001**	**0.27**	**3.98**	**<.001**
	Sadness Slope	0.19	2.57	.011	**0.20**	**2.83**	**.005**
	Evaluative Anxiety Slope	-0.01	-0.17	.864	-0.02	-0.26	.800
Loneliness	General Anxiety Slope	-0.04	-0.54	.589	-0.04	-0.52	.605
	Anger Slope	0.06	0.85	.396	0.06	0.85	.397
	Sadness Slope	**0.62**	**7.84**	**<.001**	**0.62**	**7.87**	**<.001**
	Evaluative Anxiety Slope	0.13	1.76	.081	0.13	1.79	.074
Perceived Rejection	General Anxiety Slope	0.14	1.52	.131	0.14	1.52	.131
	Anger Slope	**0.23**	**2.79**	**.006**	**0.23**	**2.79**	**.006**
	Sadness Slope	**0.25**	**2.84**	**.005**	**0.25**	**2.92**	**.004**
	Evaluative Anxiety Slope	-0.02	-0.19	.846	-0.02	-0.21	.832
Life Satisfaction	General Anxiety Slope	-0.05	-0.30	.769	-0.05	0.18	.770
	Anger Slope	0.06	0.34	.735	0.06	0.35	.729
	Sadness Slope	-0.40	-2.32	.023	-0.40	-2.34	.022
	Evaluative Anxiety Slope	0.13	0.86	.391	0.13	0.87	.389

Note: P-values were corrected using the Bonferroni correction (adjusted alpha = .05/6 = .008). Bold text indicates significant models. Spline age was controlled for in all models.

When the COVID variable was not included as a control in the model, the affect slope variables that significantly predicted functional outcomes were similar. As in the models reported in the manuscript with the COVID variable included, the anger trajectory slope predicted perceived hostility and perceived rejection and the sadness trajectory slope predicted perceived rejection and loneliness. However, there were two differences in significant sadness slope terms when compared to the corrected .008 alpha level: 1) with the COVID variable included, sadness slope predicted friendship (p = .006), but without the COVID variable included, sadness slope did not reach significance (p = 009), and 2) with the COVID variable included, sadness slope did not reach significance in predicting perceived hostility (p = .011), but without the COVID variable included it did (p = .005). Because the COVID variable does appear to have a small effect, we retain this variable in the primary models reported in the manuscript. However, the consistency of results suggests that the associations between affect trajectories and functional outcomes are not highly dependent on the inclusion of the COVID variable, reinforcing the overall validity of our results.

10. Identifying Negative Affect Types: Previous Cross-Sectional Analysis Methods and Results

See below for the methods and results for the exploratory factor analysis and confirmatory factor analysis to establish negative affect types, as reported in Grisanzio et al. (2023).

Methods

Exploratory Factor Analysis

The first aim of the study was to uncover the latent structure of the negative affect variables to obtain meaningful summary scores for different forms of negative affect in our sample. We implemented a data-driven approach to identify forms of negative affect rather than relying on scale summary scores, as this approach 1) allowed us to only include items measuring negative affective experience, rather than emotion-related thoughts or beliefs, and 2) allowed items capturing similar affective experiences to group together without being tied to a priori assumptions.

To achieve aim, we conducted an exploratory factor analysis (EFA) using the *fa* function from the *psych* package (version 1.8.12, Revelle, 2018) in R (R Core Team, 2020). Twenty-two items measuring a range of negative affective experiences were selected from the self-report measures. Two items were eliminated due to consistently low loadings across bootstrapped samples in a supplementary analysis and the remaining twenty items were input to the EFA. Because the items all measure negative affect and are assumed to be related, we used an oblimin rotation to achieve a non-orthogonal (oblique) solution that would allow the factors to be correlated. Additionally, the negative affect variables were all 4 or 5-point Likert scale items, and thus polychoric correlations were used in the EFA to accurately estimate the correlations between the ordinal variables. When choosing the number of factors to retain, we considered a scree plot using the elbow method, eigenvalues > 1 criteria, parallel analysis (in which a factor is considered as "significant" if its eigenvalue is larger than the 95% quantile of those obtained from a random data matrix of the same size as the original), and interpretability. Because parallel analysis is cautioned to be partially sensitive to sample size (i.e., for large samples, the eigenvalues of random factors will tend to be very small resulting a larger number of factors than using other criteria; Revelle (2015)), the parallel analysis was weighted less strongly. The resulting factor solution was compared to three additional theoretically identified and data-driven factor solutions to ensure it was the best fitting model (see "Exploratory factor analysis solution comparisons" below).

Confirmatory Factor Analysis

To evaluate the fit of the factor structure extracted in the EFA, we conducted a confirmatory factor analysis (CFA) using the *cfa* function in R's *lavaan* package (version 0.6-9, Rosseel, 2012). We calculated and report the standard measures and fit rules (Hu & Bentler, 1999) to assess how well the proposed model produced by the EFA captures the covariance between the measured items. The fit indices used include the comparative fit index (CFI, should be >= 0.95), root mean square error of approximation (RMSEA should be <= 0.05, upper CI bound <= 0.10), standardized root mean square residual (SRMR, should be <= 0.08), Tucker-Lewis index (TLI, > .90 indicates good fit), McDonald fit index (MFI, higher values indicate better fit), and the chi-squared test (a non-significant p-value suggests the model fits) was also used. While we report the chi-square for completeness, this statistic becomes more significant with larger samples and has low power at smaller samples (Gatignon, 2010); due to the large sample size in the current study, we did not rely on this statistic for determining fit. To improve model fit, we identified and removed items with non-significant loadings. These fit indices represent an upper bound on the fit that would be expected in an independent sample because the model was developed and tested using the same data.

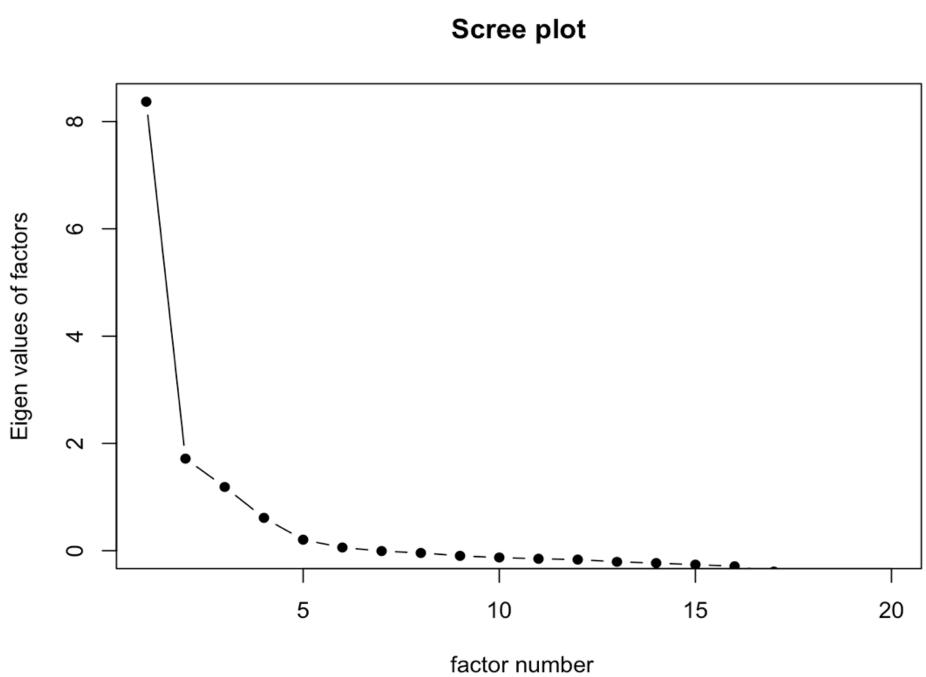
Results

Exploratory Factor Analysis

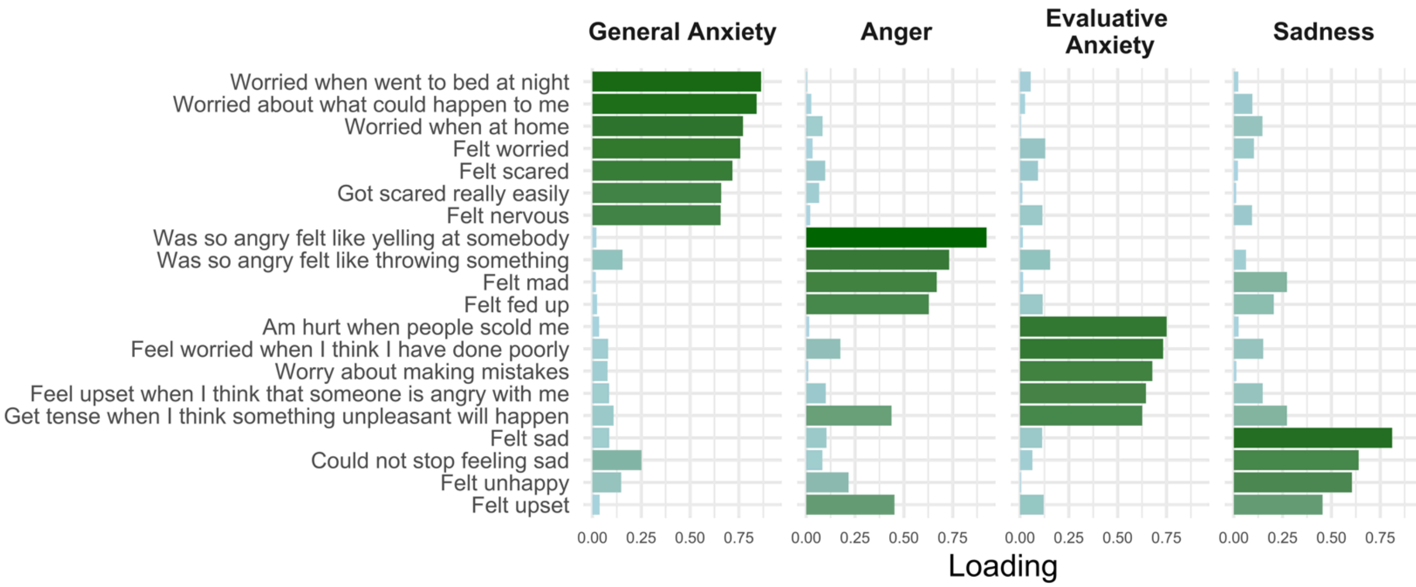
The scree plot (see **Supplementary Figure 4**) indicated that 2 or 5 factors may be optimal, the eigenvalues > 1 criteria indicated 3 factors, and the parallel analysis indicated 5 factors. Two, 3, 4, 5-factor solutions were extracted to compare interpretability. A 2-factor solution (52.3% of variance explained) consisted of a factor with items relating to evaluative anxiety (i.e., items centered around anxiety about making mistakes or being negatively evaluated) and a factor with anger, general anxiety (i.e., general feelings of worry, fear, or nervousness), and sadness items. A 3-factor solution (59.6% of variance explained) consisted of a general anxiety factor, an anger factor, and an evaluative anxiety factor, with sadness items loading weakly on the anxiety and anger factors. A 4-factor solution (64.5% of variance explained) resulted in a general anxiety factor, an anger factor, an evaluative anxiety factor, and a sadness factor. A 5-factor solution (68.5% of variance explained) was consistent with the 4-factor solution with an added factor of one item, "I usually get very tense when I think something unpleasant is going to happen". Due to common conceptualization of anger, anxiety, and sadness as distinctly experienced negative emotions (Russell, 1980) we eliminated the 2-factor solution. Due to the weak loadings of the sadness items, we eliminated the 3-factor solution. Finally, due to the difficulty of interpreting the 5th factor distinctly from the general anxiety and evaluative anxiety factors of the 4-factor solution and the limited utility in a 1-item factor, we chose the 4-factor solution as the optimal solution (**Supplementary Figure 5**).

Due to the data-driven nature of our approach, we selected the following names to describe the affective state expressed in each group of items without direct reference to the original scales or to clinical terminology: *general anxiety*, *anger*, *evaluative anxiety,* and *sadness*. *General anxiety* contained items originally from the NIH Toolbox Fear Survey, *anger* contained items from the NIH Toolbox Anger Survey, *sadness* contained items from the NIH Toolbox Sadness and Anger Surveys, and *evaluative anxiety* contained items originally from the BIS scale.

The correlations between factors were as follows: *general anxiety* and *anger* ($r(768) = .61, p < .001$), *general anxiety* and *evaluative anxiety* ($r(768) = .31, p < .001$), *general anxiety* and *sadness* ($r(768) = .64, p < .001$), *anger* and *evaluative anxiety* ($r(768) = .30, p < .001$), *anger* and *sadness* ($r(768) = .45, p < .001$), and *evaluative anxiety* and *sadness* ($r(768) = .18, p < .001$).



Supplementary Figure 4. Scree plot for 4-factor solution. Eigenvalues for one through twenty factors are plotted.



Supplementary Figure 5. EFA loading plot. The absolute values of the loadings are plotted, sorted by loading strength. Higher loading strengths are depicted by larger values on the x-axis and a darker green color. Items submitted to affect factor analysis are paraphrased on the left; the full items are available in the "Negative affect items and data-driven factor assignment" section below.

Confirmatory Factor Analysis

The CFA indicated that all items had significant loadings (range of standardized loadings: .623 - .877), so all items were retained for the analysis. The fit statistics of the final four-factor solution were as follows: chi-squared test statistic = 523.26 (df = 164, $p < 0.001$), CFI = .992, RMSEA = .053 (CI = .048, .059), SRMR = .051, TLI = .991, MFI = .792. The CFI, SRMR, and TLI are all within the recommended range to suggest a well-fitting model (CFI >= .95, SRMR <= .08, TLI > .90). The RMSEA is slightly higher than the recommended value (<= .05), however, the upper CI bound is within the recommended range (<= .10). Therefore, taken together, the CFA fit indices suggest the four-factor solution is a well-fitting model.

11. Negative affect items and data-driven factor assignment

Supplementary Table 5. Final selection of negative affect items, the original measure they were selected from, and their factor assignment.

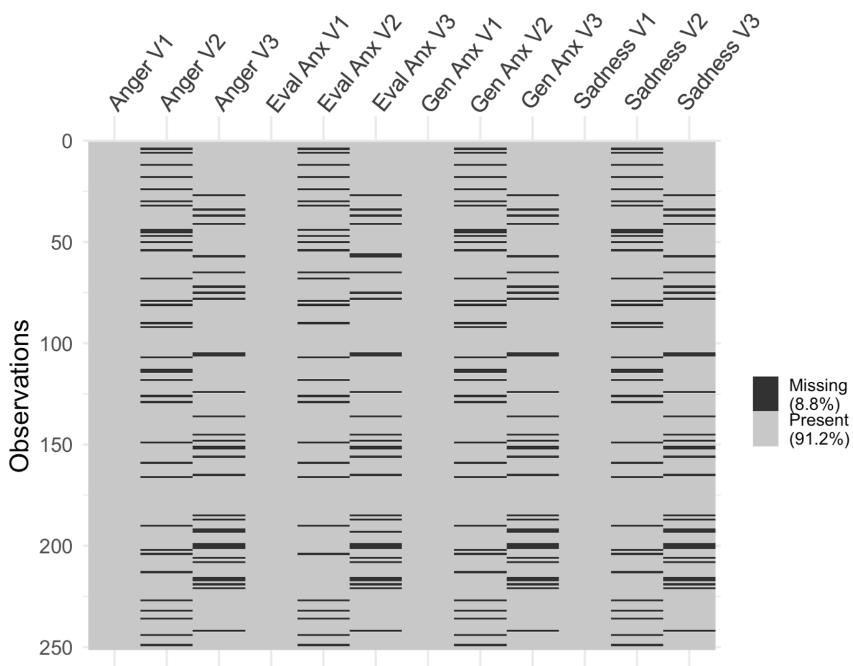
Full Item	**Abbreviation**	**Measure**	**Factor Assignment**
I felt mad	PedRepAng13	NIH Toolbox Anger Subscale	Anger
I was so angry I felt like yelling at somebody	PedRepAng14	NIH Toolbox Anger Subscale	Anger
I felt fed up	PedRepAng16	NIH Toolbox Anger Subscale	Anger
I was so angry I felt like throwing something	PedRepAng17	NIH Toolbox Anger Subscale	Anger
I felt upset	PedRepAng18	NIH Toolbox Anger Subscale	Sadness
I felt scared	PedRepAnx42	NIH Toolbox Fear Subscale	General anxiety
I worried about what could happen to me	PedRepAnx43	NIH Toolbox Fear Subscale	General anxiety
I felt worried	PedRepAnx44	NIH Toolbox Fear Subscale	General anxiety
I worried when I went to bed at night	PedRepAnx46	NIH Toolbox Fear Subscale	General anxiety
I felt nervous	PedRepAnx48	NIH Toolbox Fear Subscale	General anxiety
I worried when I was at home	PedRepAnx50	NIH Toolbox Fear Subscale	General anxiety
I got scared really easily	PedRepAnx51	NIH Toolbox Fear Subscale	General anxiety

I felt unhappy	PedRepDep36	NIH Toolbox Sadness Subscale	Sadness
I felt sad	PedRepDep38	NIH Toolbox Sadness Subscale	Sadness
I could not stop feeling sad	PedRepDep41	NIH Toolbox Sadness Subscale	Sadness
I usually get very tense when I think something unpleasant is going to happen	bisbas1	BIS/BAS Scale	Evaluative anxiety
I worry about making mistakes	bisbas2	BIS/BAS Scale	Evaluative anxiety
I am hurt when people scold me or tell me that I do something wrong	bisbas3	BIS/BAS Scale	Evaluative anxiety
I feel pretty upset when I think that someone is angry with me	bisbas4	BIS/BAS Scale	Evaluative anxiety
I do not become fearful or nervous, even when something bad happens to me	bisbas5	BIS/BAS Scale	N/A
I feel worried when I think I have done poorly at something	bisbas6	BIS/BAS Scale	Evaluative anxiety
I am very fearful compared to my friends	bisbas7	BIS/BAS Scale	N/A

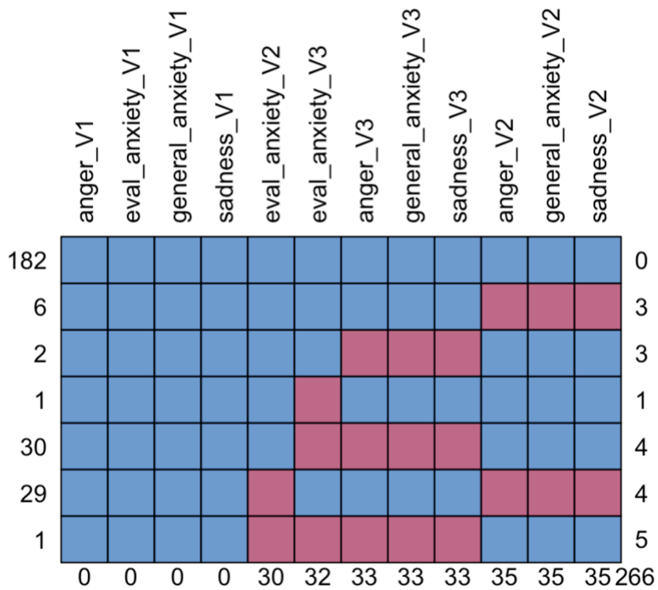
12. Evaluation of missingness: Little's MCAR

To assess the missingness pattern of our negative affect variables of interest (i.e., mean score for each negative affect type at each of the three study waves), we conducted Little's MCAR test using the *naniar* package (Tierney & Cook, 2023) in R. This test evaluates whether the means of observed variables differ across distinct patterns of missingness, based on the observed data. Little's MCAR test was significant ($\chi^2(52) = 87.0$, $p = .002$), indicating that the assumption of Missing Completely at Random (MCAR) does not hold. This suggests that missingness is systematically related to observed or unobserved variables rather than occurring purely at random. Missingness was visualized in **Supplementary Figures 6 and 7** using the *visdat* (Tierney, 2017) and *mice* (van Buuren & Groothuis-Oudshoorn, 2011) packages, showing an increase in missingness in visits 2 and 3 compared to visit 1, with some missingness patterns aligning with visit.

To examine whether missingness was related to demographic or COVID-related variables, we additionally conducted a logistic mixed-effects regression predicting whether a data point was missing as a function of age, sex, and whether the assessment occurred post-COVID (**Supplementary Table 6**).



Supplementary Figure 6. **Visual overview of missingness across negative affect variables.** Each participant is shown as a horizontal line, and each variable of interest is plotted as a column. Eval Anx: evaluative anxiety, Gen Anx: general anxiety, V1: visit 1, V2: visit 2, V3: visit 3.



Supplementary Figure 7. **Patterns of missingness.** Negative affect variables across visits are shown as columns, with rows corresponding to a missing data pattern. Blue squares indicate observed data, red squares represent missing data. Rows and columns are sorted by increasing amounts of missing information. Row labels on the left refer to the number of observations with a particular missing data pattern. Row values on the right refer to the number of missing data points. Column values on the bottom represent the number of participants with missing data for each negative affect variable.

Supplementary Table 6. Logistic mixed-effects regression predicting missingness

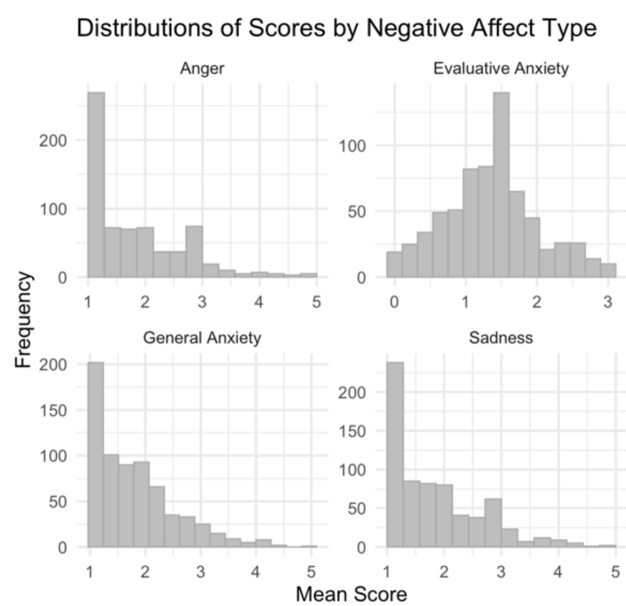
Predictor	**Estimate**	**Std. Error**	**z value**	**p-value**
Intercept	-16.118	7.318	-2.203	0.028
Age	-0.079	0.564	-0.140	0.888
Sex (M vs. F)	0.491	2.358	0.208	0.835
Post-COVID Status	0.533	3.363	0.159	0.874

Note. Missingness was coded as 1 (missing) vs. 0 (observed).

The results indicate that age, sex, and post-COVID status did not significantly predict missingness when accounting for within-subject variation. However, the rejection of MCAR in

Little's test implies that missingness may be influenced by other, unmeasured factors. While the inclusion of age, sex, and post-COVID status as control variables in our main analyses helps to mitigate concerns about systematic bias and visualizations show that missingness may align to some degree with visit, it is important to acknowledge that missing data patterns could still introduce some bias if they are related to other variables not included in our models. Thus, while missingness does not appear to be strongly associated with key demographic factors, the results caution against assuming missing data are completely random.

13. Variable distributions



Supplementary Figure 8. Distributions of scores for each negative affect type. Mean score is shown on the x-axis, count is shown on the y-axis.

Supplementary Table 7. Bivariate correlations between affect variables and functional outcomes across visits.

	Ang V1	Emo sup V1	Evl anx V1	Frnd V1	Gen anx V1	Life sat V1	Lon V1	Perc hos V1	Perc rej V1	Sad V1	Ang V2	Evl anx V2	Gen anx V2	Sad V2	Ang V3	Emo sup V3	Evl anx V3	Frnd V3	Gen anx V3	Lon V3	Perc hos V3	Perc rej V3	Sad V3	Life sat V3
Ang V1	1.0	-.25	.16	-.27	.40	-.25	.38	.45	.42	.59	.32	.12	.20	.29	.19	-.06	.07	-.13	.08	.17	.23	.16	.11	-.17
Emo sup V1	-.25	1.0	-.06	.50	-.24	.41	-.44	-.27	-.32	-.32	-.13	-.08	-.19	-.19	-.15	.32	-.06	.25	-.12	-.23	-.18	-.21	-.17	.29
Evl anx V1	.16	-.06	1.0	-.15	.35	.04	.25	.15	.27	.18	.15	.52	.24	.31	.25	-.08	.56	-.17	.40	.34	.24	.27	.28	-.15
Frnd V1	-.27	.50	-.15	1.0	-.25	.35	-.61	-.21	-.48	-.45	-.10	-.05	-.26	-.25	-.16	.21	-.10	.41	-.08	-.27	-.18	-.15	-.15	.26
Gen anx V1	.40	-.24	.35	-.25	1.0	-.19	.33	.36	.34	.53	.16	.29	.46	.41	.12	-.07	.28	-.18	.31	.17	.13	.07	.21	-.10
Life sat V1	-.25	.41	.04	.35	-.19	1.0	-.40	-.23	-.33	-.40	.01	.13	-.03	-.16	-.07	.14	.07	.18	-.10	-.21	-.16	-.08	-.17	.25
Lon V1	.38	-.44	.25	-.61	.33	-.40	1.0	.44	.70	.63	.13	.19	.26	.29	.21	-.19	.23	-.29	.21	.35	.27	.26	.29	-.33
Perc hos V1	.45	-.27	.15	-.21	.36	-.23	.44	1.0	.55	.43	.22	.14	.09	.15	.15	-.12	.08	-.14	.10	.05	.31	.13	.13	-.19
Perc rej V1	.42	-.32	.27	-.48	.34	-.33	.70	.55	1.0	.51	.19	.17	.26	.31	.15	-.16	.16	-.23	.16	.25	.32	.32	.22	-.34
Sad V1	.59	-.32	.18	-.45	.53	-.40	.63	.43	.51	1.0	.24	.22	.30	.40	.20	-.12	.20	-.21	.19	.23	.20	.12	.23	-.26
Ang V2	.32	-.13	.15	-.10	.16	.01	.13	.22	.19	.24	1.0	.27	.34	.55	.48	-.08	.15	-.13	.15	.16	.30	.09	.31	-.15
Evl anx V2	.12	-.08	.52	-.05	.29	.13	.19	.14	.17	.22	.27	1.0	.38	.41	.18	-.08	.57	-.06	.34	.20	.06	.13	.28	-.12
Gen anx V2	.20	-.19	.24	-.26	.46	-.03	.26	.09	.26	.30	.34	.38	1.0	.58	.22	-.11	.27	-.10	.31	.27	.24	.21	.22	-.13
Sad V2	.29	-.19	.31	-.25	.41	-.16	.29	.15	.31	.40	.55	.41	.58	1.0	.35	-.17	.27	-.22	.31	.41	.26	.25	.41	-.17
Ang V3	.19	-.15	.25	-.16	.12	-.07	.21	.15	.15	.20	.48	.18	.22	.35	1.0	-.40	.38	-.31	.62	.51	.58	.50	.67	-.28
Emo sup V3	-.06	.32	-.08	.21	-.07	.14	-.19	-.12	-.16	-.12	-.08	-.08	-.11	-.17	-.40	1.0	-.10	.53	-.31	-.59	-.37	-.50	-.37	.60
Evl anx V3	.07	-.06	.56	-.10	.28	.07	.23	.08	.16	.20	.15	.57	.27	.27	.38	-.10	1.0	-.23	.55	.38	.27	.27	.43	-.12
Frnd V3	-.13	.25	-.17	.41	-.18	.18	-.29	-.14	-.23	-.21	-.13	-.06	-.10	-.22	-.31	.53	-.23	1.0	-.34	-.55	-.30	-.34	-.38	.52
Gen anx V3	.08	-.12	.40	-.08	.31	-.10	.21	.10	.16	.19	.15	.34	.31	.31	.62	-.31	.55	-.34	1.0	.52	.49	.45	.68	-.38
Lon V3	.17	-.23	.34	-.27	.17	-.21	.35	.05	.25	.23	.16	.20	.27	.41	.51	-.59	.38	-.55	.52	1.0	.47	.65	.69	-.48
Perc hos V3	.23	-.18	.24	-.18	.13	-.16	.27	.31	.32	.20	.30	.06	.24	.26	.58	-.37	.27	-.30	.49	.47	1.0	.58	.52	-.35
Perc rej V3	.16	-.21	.27	-.15	.07	-.08	.26	.13	.32	.12	.09	.13	.21	.25	.50	-.50	.27	-.34	.45	.65	.58	1.0	.50	-.46
Sad V3	.11	-.17	.28	-.15	.21	-.17	.29	.13	.22	.23	.31	.28	.22	.41	.67	-.37	.43	-.38	.68	.69	.52	.50	1.0	-.38
Life sat V3	-.17	.29	-.15	.26	-.10	.25	-.33	-.19	-.34	-.26	-.15	-.12	-.13	-.17	-.28	.60	-.12	.52	-.38	-.48	-.35	-.46	-.38	1.0

Note: Values correspond to Pearson correlation coefficients. Ang: anger, Emo sup: emotional support, Evl anx: evaluative anxiety, Frnd: friendship, Gen anx: general anxiety, Life sat: life satisfaction, Lon: loneliness, Perc hos: perceived hostility, Perc rej: perceived rejection, Sad: sadness. Negative affect variables across all study visits (V1, V2, V3) and functional outcomes across the first and last study visit (V1, V3) are shown to correspond to variables used for primary analyses.

14. Comparison of four separate negative affect types vs. global negative affect in predicting outcomes

Because primary clustering results revealed similar patterns across negative affect types, we conducted an additional post-hoc analysis to evaluate whether the negative affect types had differential effects on the social functioning and life satisfaction outcomes. Specifically, we examine the differential effect of negative affect slope, as an extension of our third analysis (i.e., "*Exploring Affect Slope as a Continuous Predictor of Functional Outcomes*").

We fit two regression models: one that is reported in the main manuscript, in which the coefficients were allowed to differ for each of the negative affect factors, and one in which the coefficients were constrained to be equal across all negative affect factors. The two models were then compared using Akaike Information Criterion (AIC) values. The difference in AIC values between the two models (Δ AIC) is reported.

The unconstrained models, where the coefficients were allowed to differ for each of the negative affect factors, had a lower AIC value than the constrained models for loneliness (Δ AIC = 23.66), perceived hostility (Δ AIC = 3.87), and perceived rejection (Δ AIC = 1.02), indicating that modeling the negative affect types separately provided a better fit to the data for these outcomes. In contrast, for life satisfaction (Δ AIC = 0.95), friendship (Δ AIC = 1.61), and emotional support (Δ AIC = 0.78), the constrained model had a lower AIC value than the unconstrained model, suggesting that for these outcomes, the slopes of the negative affect factors did not substantially differ.

For the emotional support and life satisfaction outcomes, none of the negative affect slopes emerged as significant in the primary analysis (main manuscript **Table 2**). Thus, the lack of improvement in model fit when allowing slopes to vary across negative affect types likely reflects a general lack of association between negative affect slope and these particular outcomes, rather than evidence that the effects of different negative affect types can be meaningfully collapsed into a single, common effect. In other words, the negative affect trajectories, whether considered globally or by type, may not play a substantial role in predicting life satisfaction and emotional support in this sample.

For the friendship outcome, the sadness slope emerged as a significant predictor in the primary analysis. However, the unconstrained model, in which the effects of the negative affect slopes were allowed to vary by type, did not provide a better fit than the constrained model. This pattern suggests that while the trajectory of sadness specifically may hold unique relevance for predicting friendship quality, the overall model comparison did not provide strong evidence that differentiating between negative affect types improved the model's explanatory power for this outcome.

Taken together, these results suggest that disaggregating negative affect into its constituent types provides added predictive value for certain functioning outcomes (loneliness, perceived hostility, and perceived rejection), but may offer less advantage when predicting friendship, emotional support, and overall life satisfaction.

References

Stef van Buuren, Karin Groothuis-Oudshoorn (2011). mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1-67. DOI 10.18637/jss.v045.i03.

Tierney N, Cook D (2023). "Expanding Tidy Data Principles to Facilitate Missing Data Exploration, Visualization and Assessment of Imputations." _Journal of Statistical Software_, **105**(7), 1-31. doi:10.18637/jss.v105.i07 <https://doi.org/10.18637/jss.v105.i07>.

Tierney N (2017). "visdat: Visualising Whole Data Frames." _JOSS_, **2**(16), 355. doi:10.21105/joss.00355 <https://doi.org/10.21105/joss.00355>, <http://dx.doi.org/10.21105/joss.00355>.