Supplemental content A: Simulation 1 – The potential downward bias of auto-and cross-regressive slopes

We consider the effects of no centering (NC) and within-person centering (WPC) in a multilevel VAR model, both with fixed and random specifications of the autoregressive and cross-regressive coefficients. We are mainly interested to find the extent to which the bias that has been found by Hamaker and Grasman (2015) for WPC autoregressive slopes, will also pertain to WPC cross-regressive slopes. We thus add onto the results by Hamaker and Grasman (2015) in two ways, namely (1) to replicate their results for WPC autoregressive slopes, and (2) to extend their results to multilevel VAR models that include more than a single outcome and a single lagged predictor. To avoid needless complications, we neglect the mixture framework in this small simulation, and consider a standard multilevel VAR. Because the bias of WPC and NC pertains only to the multilevel VAR part of the MMVAR, the results of this simulation based on the multilevel VAR will generalize directly to MM-VAR. We consider the effects of the following factors: (a) fixed vs. random coefficients, (b) the size of the cross-regressive slopes, and (c) the number of observations. Our simulation was performed in R (R Core Team, 2020), multilevel VAR models were estimated with the lmer function from the R-package lme4 (Bates et al., 2015) in the fashion described by Bringmann et al. (2013).[12]

Design and procedure. To generate the data, the following factors were kept constant. We generate multilevel VAR(1) models each with two outcome variables and two lagged predictors. The fixed effect within-person slopes, γ_{11}, and γ_{22}, were equal to .3 for autoregressions. The level 2 variances of the within-person mean, τ_{11} and τ_{22}, equaled 3, the grand means of both variables equaled 0. The innovation variances were set to 3. Covariances of random effects and innovations were set to zero. Each generated data set

[12]That is, multiple outcome variables are accounted for by estimating a separate multilevel equation for each outcome variable.

contains 100 individuals. The following three factors were varied in a completely crossed design with 100 replications each: (a) the standard deviations of the individual within-individual slopes, τ_{33}, τ_{44}, τ_{55}, and τ_{66}, equaled 0 or .2; (b) the fixed effect within-individual slopes for the cross-regressions, γ_{21}, and γ_{12}, equaled 0 or .2; (c) the number of observations equaled 20, 50 or 100.

All $(2 \times 2 \times 3 \times 100 =)$ 1,200 resulting data sets were analyzed with WPC and NC. For WPC, we fit two pairs of single outcome multilevel models using both within-person mean centered lagged-1 outcomes, $w_{PA\,i,\,t-1}$ and $w_{NA\,i,\,t-1}$, as predictors. Following the notation in Equations 1 and 2, the model to estimate the simulated data was:

$$y_{PA\,i\,t} = \mu^{WPC}_{PA\,i} + \phi^{WPC}_{11\,i} w_{PA\,i,\,t-1} + \phi^{WPC}_{21\,i} w_{NA\,i,\,t-1} + \varepsilon^{WPC}_{PA\,i\,t} \tag{A.9}$$

$$y_{NA\,i\,t} = \mu^{WPC}_{NA\,i} + \phi^{WPC}_{12\,i} w_{PA\,i,\,t-1} + \phi^{WPC}_{22\,i} w_{NA\,i,\,t-1} + \varepsilon^{WPC}_{NA\,i\,t}. \tag{A.10}$$

For NC, we fit two pairs of single outcome multilevel models using both lagged-1 outcomes on the original scales, $y_{PA\,i,\,t-1}$ and $y_{PA\,i,\,t-1}$, as predictors, estimating the following model:

$$y_{PA\,i\,t} = \mu^{NC}_{PA\,i} + \phi^{NC}_{11\,i} y_{PA\,i,\,t-1} + \phi^{NC}_{21\,i} y_{NA\,i,\,t-1} + \varepsilon^{NC}_{PA\,i\,t} \tag{A.11}$$

$$y_{NA\,i\,t} = \mu^{NC}_{NA\,i} + \phi^{NC}_{12\,i} y_{PA\,i,\,t-1} + \phi^{NC}_{22\,i} y_{NA\,i,\,t-1} + \varepsilon^{NC}_{NA\,i\,t} \tag{A.12}$$

Subsequently, per condition, the bias was estimated for the autoregressive and cross-regressive slopes separately. In the following we take WPC as example, but the following applies equally to the NC estimates. The parameters of interest are the fixed effects $\gamma^{WPC}_{11}, \gamma^{WPC}_{12}, \gamma^{WPC}_{21}, \gamma^{WPC}_{22}$ that correspond to the random slopes $\phi^{WPC}_{11\,i}, \phi^{WPC}_{12\,i}, \phi^{WPC}_{21\,i}, \phi^{WPC}_{22\,i}$. Bias was estimated by averaging across the 100 replications within a simulation condition (e.g., $\widehat{\text{Bias}}(\gamma^{WPC}_{11}) = \frac{1}{100}\sum_{r=1}^{100} \hat{\gamma}^{WPC}_{11r} - \gamma_{11}$). Because the two single outcome multilevel models represented in Equations A.9 and A.10 are symmetric, we denote the average of $\widehat{\text{Bias}}(\gamma_{11}^{WPC})$ and $\widehat{\text{Bias}}(\gamma_{22}^{WPC})$ as the bias estimate for the *autoregressive* coefficients. Similarly, we denote the average of $\widehat{\text{Bias}}(\gamma_{12}^{WPC})$ and $\widehat{\text{Bias}}(\gamma_{21}^{WPC})$ as the bias estimate for the *cross-regressive* coefficients. To quantify the uncertainty of the bias estimates due to the limited number of replications, we calculated Monte Carlo errors using equation 7 in Koehler et al. (2009). Just like the autoregressive and cross-regressive estimates, the Monte Carlo errors are displayed as averages across the two variables *PA* and *NA*. All results are multiplied by 1,000 and included in Table A1.

Results. The following conclusions can be drawn from Table A1.

1. Similar to Hamaker and Grasman (2015), WPC yields autoregressive coefficients that are underestimated across different conditions; cross-regressive coefficients seem biased for WPC only when there is variation for the random slopes.

2. For NC, autoregressive coefficients seem unbiased; in contrast, cross-regressive coefficients are underestimated when there is variation for the random slopes.

3. The bias of cross-regressive coefficients are in general more severe for NC than for WPC; the bias tends to increase with the variance of the random slopes for NC and WPC.

4. For WPC and NC, the biases discussed above for autoregressive and cross-regressive coefficients reduce when the number of observations increases.

Table A1

Simulation results: downward bias in autoregressive coefficients for WPC.

	Bias (Monte Carlo Error) ×1000					
	Autoregressive			Cross-regressive		
Estimator	$T = 20$	$T = 50$	$T = 100$	$T = 20$	$T = 50$	$T = 100$
$\tau_{33} = \tau_{44} = \tau_{55} = \tau_{66} = 0$, $\gamma_{21} = \gamma_{12} = 0$						
NC	−0.025(2.335)	−0.969(1.385)	−0.667(0.865)	0.502(2.246)	−2.347(1.320)	1.265(0.932)
WPC	−75.997(2.172)	−28.412(1.355)	−13.966(0.855)	0.104(2.423)	−2.440(1.379)	1.210(0.940)
$\tau_{33} = \tau_{44} = \tau_{55} = \tau_{66} = 0$, $\gamma_{21} = \gamma_{12} = .2$						
NC	7.591(3.335)	−2.331(2.481)	−2.444(2.073)	0.040(2.886)	0.032(2.164)	−0.497(2.121)
WPC	−70.022(3.007)	−28.485(2.407)	−14.528(2.052)	0.308(3.121)	0.039(2.229)	−0.510(2.153)
$\tau_{33} = \tau_{44} = \tau_{55} = \tau_{66} = .2$, $\gamma_{21} = \gamma_{12} = 0$						
NC	−6.131(2.370)	−2.984(1.555)	−1.939(1.059)	−32.912(2.096)	−15.129(1.292)	−6.999(0.873)
WPC	−77.226(2.200)	−28.481(1.510)	−14.229(1.048)	−15.436(2.226)	−6.436(1.316)	−2.488(0.881)
$\tau_{33} = \tau_{44} = \tau_{55} = \tau_{66} = .2$, $\gamma_{21} = \gamma_{12} = .2$						
NC	−6.267(3.168)	−4.586(2.441)	−3.381(2.102)	−29.804(2.645)	−13.820(2.297)	−7.925(2.185)
WPC	−78.731(2.927)	−29.158(2.394)	−14.603(2.084)	−14.420(2.822)	−6.430(2.340)	−4.240(2.225)

Supplemental content B: Simulation 2 – Performance of MMVAR model estimation

Recovery of random effect variances of within-person means. Figure B1a shows the mean absolute deviations (MAD) between estimated and true random effect variances of within-person means per simulation condition. Figure B2a displays the average estimates for the random effect variances per simulation condition; the value that was used to generate the data is represented by a dotted line. The figures indicate that overall the random effect variances were recovered well across simulation conditions with the exception of a few outliers. Table B1 displays the ANOVA, the effects were all non-significant suggesting there are none to only small differences in recovery across simulation conditions.

Recovery of random effect variances of VAR coefficients. Figure B1b displays the MADs between estimated and true random effect variances of VAR coefficients per simulation condition. Figure B2b shows the average random effect variances of the VAR coefficients per simulation condition. These figures suggest that the variances are recovered well across all simulation conditions, with estimated variances very close to the true data generating value. The variances are, however, consistently underestimated, as can be seen in Figure B2b.

Table B1 shows the ANOVA results, the effects were all non-significant suggesting there are none to only small differences in recovery across simulation conditions.

Recovery of mixing proportions. Figure B3 displays the MADs per simulation condition. The effect sizes in Table B2 suggest that a high number of observations ($\hat{\eta}_p^2 = .112$) and large distances between components ($\hat{\eta}_p^2 = .253$) aided recovery of mixing proportions. This indicates there were medium differences across simulation conditions. On average, MADs indicate mixing proportions were retrieved well ($M = .017$, $SD = .010$).

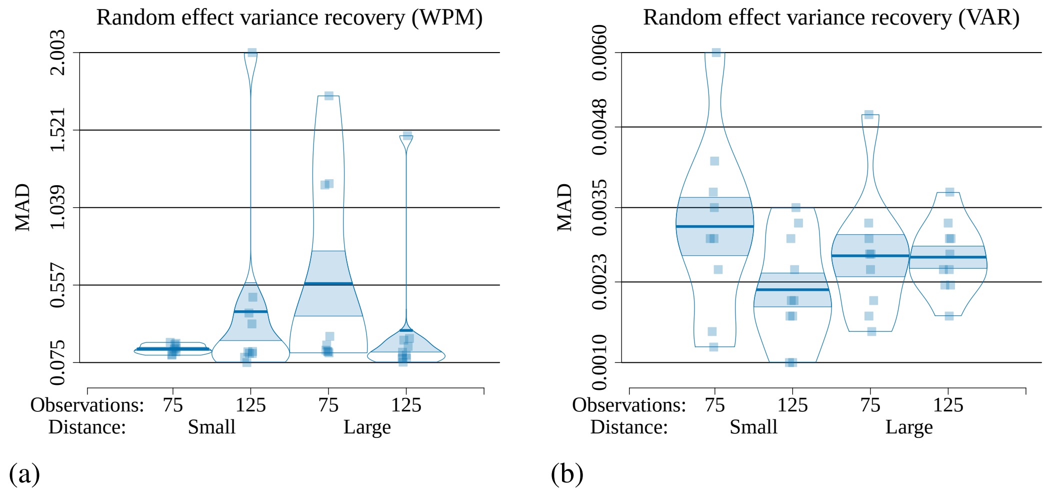


Figure B1. MADs between estimated and true random effect variances indicating the parameter recovery for the random effect variances across simulation conditions. A line indicates the mean MAD per condition, a colored band the region within one standard error of the mean.

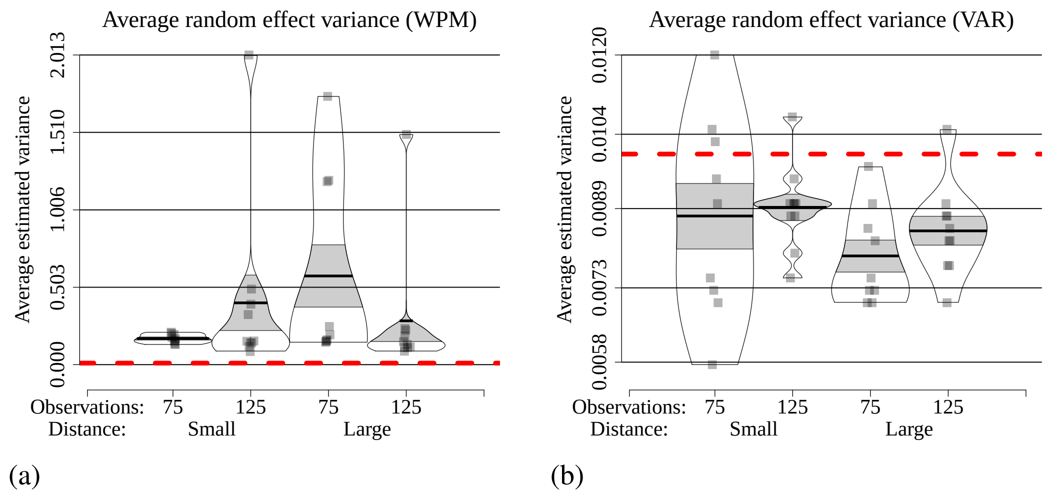


Figure B2. Average estimates for the random effect variances across simulation conditions. A line indicates the mean estimated variance, a colored band the region within one standard error of the mean. A dotted line shows the true random variance value used in data generation.

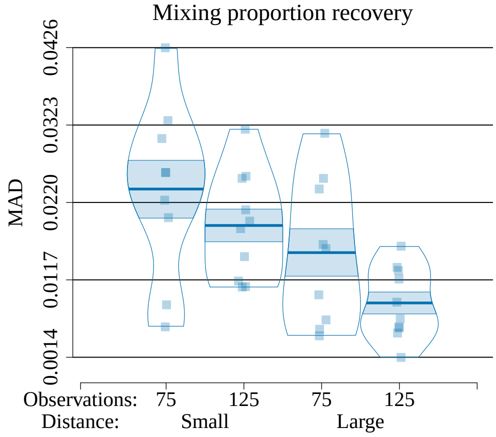


Figure B3. MADs between estimated and true mixing proportions indicating the recovery of mixing proportions across simulation conditions. A dark line indicates the mean MAD, a band the standard error.

Table B1

ANOVA results on MADs between estimated and true random effect variances of within-person means (WPM), and on MADs between estimated and true random effect variances of VAR coefficients. Any effect size above .1 is highlighted in bold.

	WPM				VAR coefficients			
	df	*F*	*p*	$\hat{\eta}_p^2$	*df*	*F*	*p*	$\hat{\eta}_p^2$
Observations	1	.037	.849	.001	1	2.500	.123	.068
Distance	1	.712	.405	.021	1	.026	.874	.001
Observations × Distance	1	2.830	.102	.077	1	2.290	.140	.063

Table B2

ANOVA results on MADs between estimated and true mixing proportions. Any effect size above .1 is highlighted in bold.

	df	*F*	*p*	$\hat{\eta}_p^2$
Observations	1	4.290	.046	**.112**
Distance	1	11.500	.002	**.253**
Observations × Distance	1	.110	.742	.003

Supplemental content C: Application COGITO study

Results: The 2×2 component solution

The estimated MMVAR parameters of the 2×2 component model are presented in Table C1. The fixed effect within-person means shown in this table, which we interpret as average trait scores, are graphically presented in Figure C1a, the fixed effect VAR coefficients in Figure C1b. The confidence intervals (CIs) of the fixed effects, displayed in Table C1, all exclude zero, with the exception of the CI for the PA to NA cross-regressive coefficient for Components 3 and 4.

Average trait scores appear to be separated into 'Unhappy components' (Components 1 and 2) and 'Happy components' (Components 3 and 4). The 'Unhappy components' are characterized by low PA and high NA, as is evident from the fixed effect within-person means displayed in Figure C1a and in Table C1. The reversed pattern, high PA and low NA, can be seen for the 'Happy components'. The 'Happy' and the 'Unhappy' components are further split into 'High inertia' and 'Low inertia' components (Components 1 and 3, and Components 2 and 4 respectively). Differences in inertia are particularly pronounced for NA as can be seen in Figure C1b and in Table C1.

We compared the resulting four components on external variables that have not been included in the model estimation. In the following, we assign individuals into components based on their modal component membership probabilities. Table C2 lists age and gender of the components' members, and also the estimated mixing proportions of the components. The mixing proportions in Table C2 show that more individuals were classified into the 'Happy' rather than the 'Unhappy' components. Whether the 'High inertia' or 'Low inertia' components were more frequent depended on the component membership for the within-person means: Individuals in the 'Happy' components were more likely to be classified into the 'Low inertia' rather than the 'High inertia' component (i.e., Component 4 over

Component 3). Individuals in the 'Unhappy' components were more likely to be classified into the 'High inertia' rather than the 'Low inertia' component (i.e., Component 1 over Component 2).

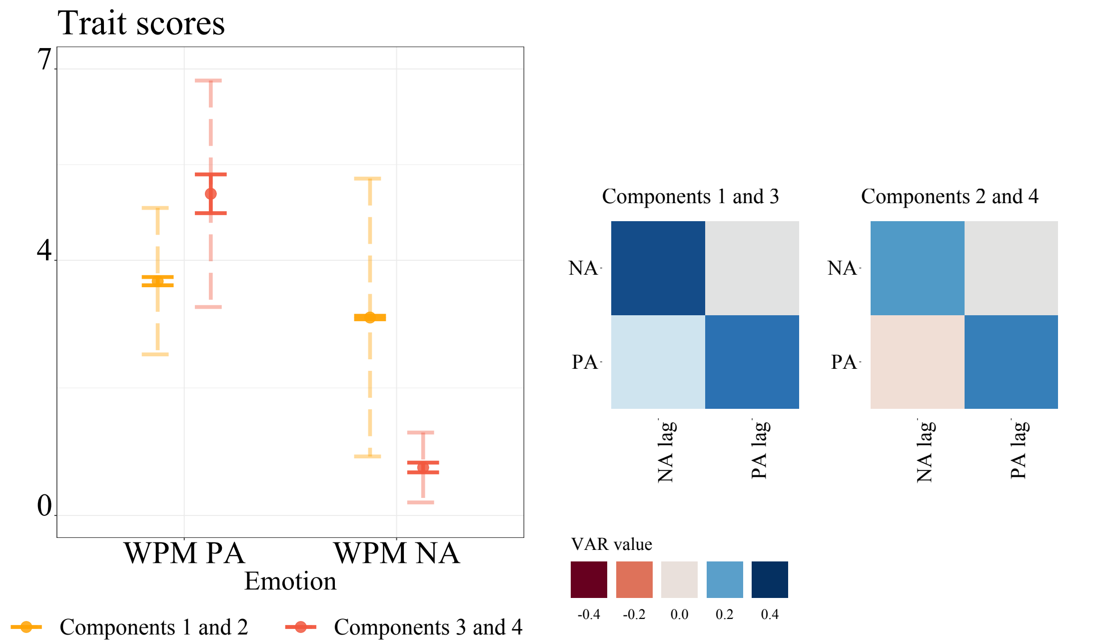
Components are clearly ordered by age, as can be seen in Table C2: Individuals in the 'Happy' components are older than the 'Unhappy' components, while the individuals in the 'High inertia' components have a lower mean age than the 'Low inertia' components. The differences in gender proportions across components, shown in Table C2, seem negligible. Figure C2 depicts the distribution of age within the four components. Figure C2 shows that, like in the 3 component solution, two of the MMVAR components correspond largely to the two different age groups. Figure C3 compares the members of the four components on their PANAS scores before the ESM study took place (left) and two years after the ESM study took place (right), the corresponding values are listed in Table C3. Table C3 shows also the Neuroticism, Openness, and Extraversion scores of the components' members before the ESM study. Looking at external variables that were not included in the model estimation (see Table C2 and C3), components seem most different on values of NA. Differences in NA are observable even 2 years after the study (see Table C3). Components seem to differ in Neuroticism but are comparable on Extraversion and Openness (see Table C3).

We conducted a MANOVA to investigate the component differences that are displayed in Table C2 and C3. All continuous variables that were assessed before the ESM study were used as outcome variables (i.e., PA, NA, Neuroticism, Extraversion, Openness and age); Modal component membership was used as independent variable. The MANOVA was significant ($p < .001$). To test subsequently which components differed and on which variables, we performed a paired t-test for each of the six variables. Each of these six post-hoc tests was Bonferroni-corrected for multiple testing across the combinations of components. Because we have to account also for multiple testing with regard to the number of post-hoc tests, we present in Table C4 only those differences for which the adjusted

p-values are below $\alpha = .05/6$. Table C4 shows that for age, Neuroticism and NA there are significant differences between components.

Table C1
The final MMAVR model parameters: fixed effects (random effect variances) [95% confidence intervals for the fixed effect].

	Component 1	Component 2	Component 3	Component 4
$\hat{\gamma}_{PA\,k}$ (WPM PA)	3.67 (1.32) [3.61; 3.73]		5.04 (3.15) [4.74; 5.34]	
$\hat{\gamma}_{NA\,k}$ (WPM NA)	3.1 (4.74) [3.07; 3.13]		.75 (.3) [.67; .83]	
$\hat{\gamma}_{11\,k}$ (PA to PA)	.29 (.01) [.24; .34]	.26 (.01) [.23; .29]	.29 (.01) [.24; .34]	.26 (.01) [.23; .29]
$\hat{\gamma}_{21\,k}$ (NA to PA)	.05 (.01) [.02; .08]	$-.01$ ($<.01$) [$-.02; -.002$]	.05 (.01) [.02; .08]	$-.01$ ($<.01$) [$-.02; -.002$]
$\hat{\gamma}_{12\,k}$ (PA to NA)	.01 ($<.01$) [$-.01; .03$]	.01 ($<.01$) [$-.002; .02$]	.01 ($<.01$) [$-.01; .03$]	.01 ($<.01$) [$-.002; .02$]
$\hat{\gamma}_{22\,k}$ (NA to NA)	.35 (.01) [.30; .40]	.21 ($<.01$) [.18; .24]	.35 (.01) [.30; .40]	.21 ($<.01$) [.18; .24]



(a) The fixed effect within-person means. Solid error bars represent the 95% CIs. Dashed oblique error bars show one standard deviation estimates of the corresponding random effects (i.e., $\sqrt{\hat{\tau}_{11\,k}}$ or $\sqrt{\hat{\tau}_{22\,k}}$) above and below the fixed effect.

(b) The fixed effect VAR coefficients.

Figure C1

Table C2

Means (SDs) on age, and gender proportions and mixing proportions for each component. Individuals were assigned into components based on their modal component membership probabilities.

	Age	Proportion female	Mixing proportion
Component 1	29.55 (14.44)	.55	.20
Component 2	39.07 (20.67)	.62	.14
Component 3	42.39 (22.60)	.47	.29
Component 4	65.95 (15.00)	.46	.37

Table C3

Means (SDs) of positive affect (PA) and negative affect (NA) either before the start of the intensive longitudinal data collection, or 2 years after the intensive longitudinal data collection, and Means (SDs) on the NEO before the start of the intensive longitudinal data collection. PA and NA were measured with the PANAS on a continuous scale from 0 to 7. The NEO was assessed on a continuous scale from 0 to 4. Individuals were assigned into components based on modal component membership probabilities.

	PANAS				NEO		
	Before study PA	Before study NA	Follow-up PA	Follow-up NA	Neuroticism	Extraversion	Openness
Component 1	4.54 (.98)	3.43 (.90)	4.29 (.85)	2.88 (1.26)	2.04 (.45)	2.42 (.43)	2.63 (.35)
Component 2	4.92 (.98)	2.86 (1.36)	4.51 (1.26)	2.24 (1.34)	1.75 (.43)	2.21 (.42)	2.59 (.36)
Component 3	4.37 (.97)	2.91 (1.23)	4.40 (1.00)	2.44 (1.10)	1.88 (.40)	2.15 (.36)	2.45 (.33)
Component 4	4.75 (.81)	2.01 (.94)	4.67 (.98)	1.53 (.97)	1.50 (.37)	2.16 (.40)	2.44 (.29)

Table C4

All significant comparisons of the Bonferroni-adjusted post-hoc paired t-tests with modal component memberships as the independent variable. Different components were compared on their mean values of age, PA, NA, Neuroticism, Extraversion, and Openness. Because six of these paired t-tests were conducted, we show only the results where the adjusted p-values are below $.05/6$.

	Components	t	df	adjusted p
Age	1 vs. 3	−3.45	96.73	.005
Age	1 vs. 4	−12.73	82.10	$< .001$
Age	2 vs. 4	−6.39	39.79	$< .001$
Age	3 vs. 4	−6.91	95.80	$< .0010$
Before study NA	1 vs. 4	7.97	82.79	$< .001$
Before study NA	3 vs. 4	4.67	106.34	$< .001$
Neuroticism	1 vs. 4	6.43	65.76	$< .001$
Neuroticism	3 vs. 4	5.66	117.89	$< .001$

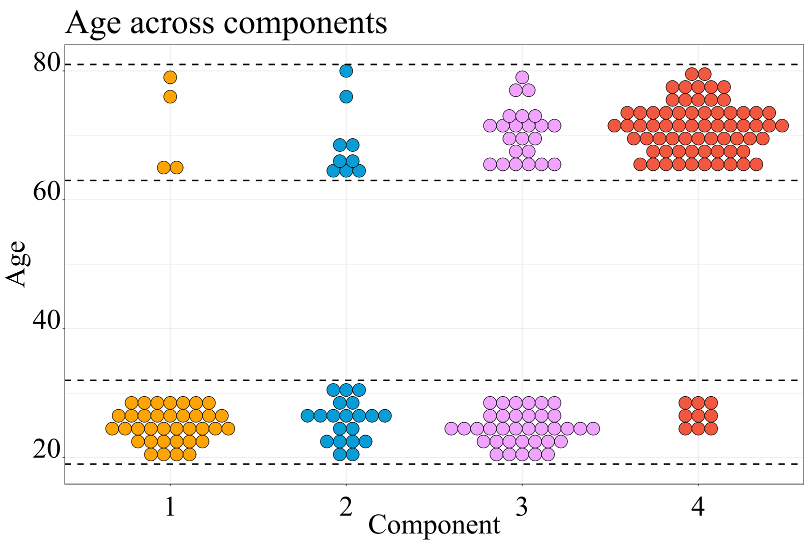


Figure C2. Age distributions within components. Dashed lines indicate the boundaries of the age groups in this sample. Individuals are assigned into components based on their modal component membership probability.

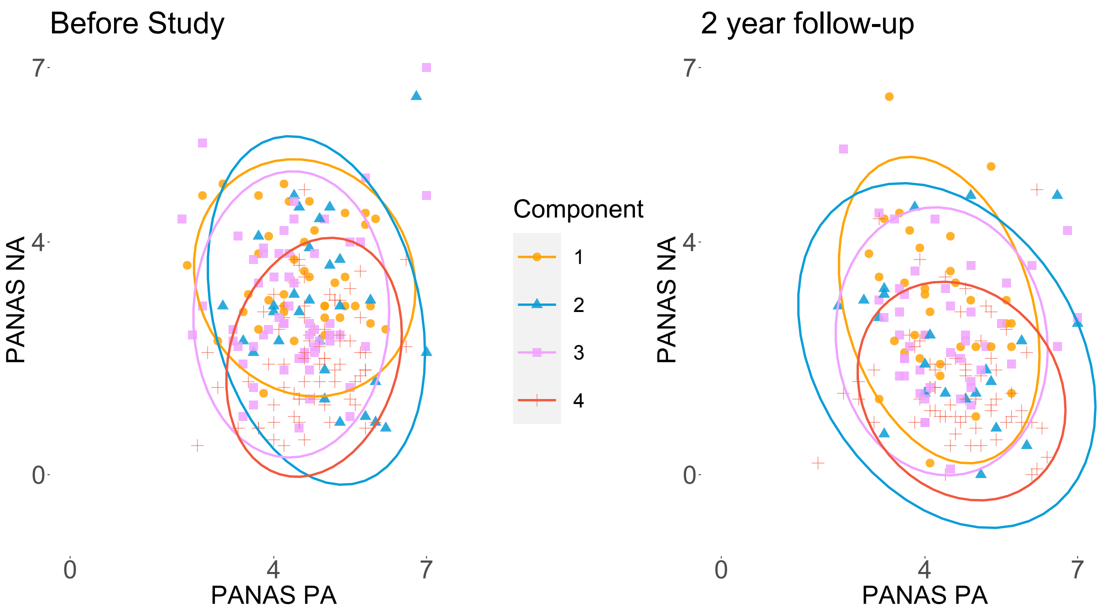


Figure C3. PANAS scores before the ESM study took place (left) and two years after the ESM study took place (right), split by modal component membership of the individual. Ellipses show the multivariate t-distribution at the 95% confidence level.