Big five personality traits and common mental disorders within a hierarchical taxonomy of psychopathology: A longitudinal study of Mexican-origin youth

SUPPLEMENTAL MATERIALS

Data Analytic Procedures.

Descriptive statistic across all waves of data collection for focal study variables are reported in Tables S1-S3. Observed sex differences in psychopathology and the Big Five across all waves of data collection are reported in Tables S4-S6. Data analytic procedures can be summarized in three steps. First, we tested for measurement invariance of a hierarchical structure of common mental disorders. Second, univariate growth curve models were used to estimate the average trajectories of Big Five personality and latent factors of psychopathology, and bivariate growth curve models were used to estimate the trajectories of residual variance in latent psychopathology factors, after accounting for the higher-order factor. Third, bivariate and multivariate growth curve models were used to estimate correlations between the growth factors of Big Five personality and latent factors of psychopathology, before and after accounting for the higher-order factor of psychopathology.

Configural Invariance.

Before estimating average growth trajectories across different levels of a hierarchical structure of psychopathology, a series of confirmatory factor analysis (CFA) models were used to test for longitudinal measurement invariance. To begin, a series of cross-sectional CFA models were used to test for configural invariance-i.e. whether symptoms-counts were indicators of the same latent constructs from age 10 to 17. Depicted in Figure 1, one-factor, two-factor, and three-factor models were fit to the data successively at each measurement occasion. In the one-factor model, all psychiatric symptoms were specified to load onto a single latent factor. In the two-factor model, symptoms of depression, anxiety, and PTSD were specified to load onto a latent internalizing factor, and symptoms of oppositional defiance disorder, conduct disorder, marijuana-use[1], and attention-deficit and hyperactivity-related symptoms of ADHD were specified to load onto an externalizing factor. In the three-factor model, attention-deficit and hyperactivity-related symptoms of ADHD were specified to load onto a third factor, while the remaining symptom-counts loaded on their respective internalizing or externalizing factors. Latent variables were scaled using unit loading identification (i.e. by fixing the factor loading of the first indicator to one), and factor variances were freely estimated. Simple structure was imposed on one-factor, two-factor, and three-factor models, whereby each symptom-count loaded onto only one factor (i.e. no cross-loadings), correlations between latent factors were freely estimated, and residual error terms were uncorrelated. Note, the correlated three-factor and higher-order factor model have the same degrees of freedom and fit statistics because the higher-order factor is just-identified (see C & D of Figure 1). For this reason, model fit statistics are reported for models “C/D” in Table S7.

Bi-factor models were also fit to the data at each measurement occasion, as bifactor models have been used in previous studies of the hierarchical structure of psychopathology. The bifactor model is more complex than the correlated three-factor model (i.e. estimates more parameters) and does not impose simple structure, as each symptom-count loads onto more than one latent factor, and correlations between latent factors are fixed to zero. Examples of bifactor models are depicted on the bottom of Figure 1. In both bifactor models, all psychiatric symptoms

[1] By age 13, less than 2% of the sample had initiated marijuana use (N = 11; see Table S1), which would not provide enough information to reliability estimate symptom thresholds from age 10 to 13. Therefore, marijuana use was only included as an ordinal indicator of psychopathology from age 14 to 17.

were specified to load onto a single common factor (i.e. the bifactor). In the first bifactor model (depicted in cell E of Figure 1), symptoms of ODD, CD, ADHD, and marijuana-use were specified to load onto an externalizing factor, while symptoms of depression, anxiety, and PTSD loaded on the bifactor only. In the second bifactor model, in addition to symptom-counts loading onto a bifactor, symptoms of ODD, CD, ADHD, and marijuana-use were specified to load onto an externalizing factor, and symptoms of depression, anxiety, and PTSD were specified to load onto a latent internalizing factor. Similar to the simple one-factor, two-factor, and three-factor models, the residual error terms were uncorrelated in the bifactor models. As noted in the body of the manuscript, there are statistical reservations with comparing bifactor models to alternative factor models. Moreover, it may be argued that the higher-order model provides an appropriate interpretative framework for capturing the general tendency for transdiagnostic dimensions of psychopathology to correlate. Nevertheless, given their prominence in previous studies of the latent structure of psychopathology, bifactor models were estimated and included in model comparisons.

As noted in the body of the manuscript, before fitting CFA models to the data, symptom-counts were transformed into ordered-categorical responses, based on the assumption that ordinal responses provide a coarsened index of an underlying continuous distribution of liability (Rijsdijk & Sham, 2002; Falconer, 1965). Ordinal symptom-counts were then specified as ordered-categorical indicators of latent factors, and models were estimated using robust weighted least squares (i.e. WLSMV; Muthén & Muthén, 2017). Consequently, discrimination parameters were estimated for symptoms-counts, which are analogous to factor loadings for polytomous indicators. Instead of intercepts, three threshold parameters were estimated for each symptom-count (one less than the number of ordered categorical responses), which indicates the location on the latent trait dimension where the probability switches from endorsing one response over another (i.e. no symptoms vs. one or two symptoms, one or two symptoms vs. three or four symptoms, and three or four symptoms vs. five or more symptoms). Consequently, although described colloquially as CFA models, measurement invariance was tested using item response theory models, specifically multivariate graded response models (Samejima, 2016).

Fit Statistics and Model Comparisons.

Measurement invariance models were compared using changes in root mean squared error of approximation (ΔRMSEA), changes in comparative fit index (ΔCFI), and changes in model chi-square ($\Delta\chi 2$) rescaled as root deterioration per restriction (*RDR;* Browne & Du Toit, 1992). This rescaling, which transforms $\Delta\chi 2$ to an RMSEA metric (RDR = $\sqrt{}(\Delta\chi^2$ - Δdf) / $\sqrt{}$[Δdf(n)]), is recommended *specifically when testing for measurement invariance* because $\Delta\chi^2$ can be sensitive to even negligible differences in estimated parameters when sample sizes are large and when comparing model that free and constrain many parameters (Hildebrandt, Wilhelm, & Robitzsch, 2009). Models in which parameter restrictions produce RDR values equal to or greater than .08 should not be considered invariant (Browne & Du Toit, 1992), and *RDR* values less than .05 “indicate that the differences in fit can be considered of only minor importance” (Hildebrandt, Wilhelm & Robitzsch, 2009, p.95). In addition to RDR, ΔRMSEA and ΔCFI were selected to compare models because simulation studies have shown that these fit statistics are generally robust when testing for measurement invariance (Cheung & Rensvold, 2002; Chen, 2007), whereby values less than or equal to .01 indicate that “the null hypothesis of invariance should not be rejected” (Cheung & Rensvold, 2002, p. 251).

The absolute fit of models was evaluated using *RMSEA* and *CFI.* We selected these fit statistics because they are commonly reported in confirmatory analyses and, more importantly, they evaluate different aspects of model fit. While both fit statistics are based on the chi-square value and the degrees of freedom for the fitted model (RMSEA = $\sqrt{((\chi^2 - df)/n-1)} / \sqrt{[df(n - 1)]}$), calculation of CFI also involves the chi-square value and degrees of freedom for the null model (CFI = $[(\chi^2\text{-df(Null Model)}) - (\chi^2\text{-df(Fitted Model)})]/\chi^2\text{-df(Null Model)}$). Markus (2019) provides a helpful analogy: Suppose you are evaluating the distance of real estate listings from your office because you'd like a short commute to work. You could compare prospective addresses using their distance from your office, and you could also compare addresses by how much shorter the distance is compared to your current address. If your current address is already close to your office, then a prospective address could score well on the first measure, because it's close to your office, but poorly on the second measure, because it is not much closer than your current address. The measure of absolute distance is analogous to RMSEA, and the measure of relative improvement is analogous to CFI. Thus, RMSEA and CFI may not always "agree" when evaluating model fit using traditional cut-offs (e.g. RMSEA < .05 & CFI > .90; Hu & Bentler, 1999), because there is nothing inconsistent about a model that fits well but not much better than a null model (Markus, 2019), which can occur when the observed covariances among manifest variables are small.

Model comparisons that tested configural invariance are reported in Table S7. All three fit statistics (ΔRMSEA, ΔCFI, and RDR) favored a bifactor model at ages 10, 14, and 15 years, but a three-factor model was preferred at ages 12 and 17 years. ΔRMSEA, ΔCFI, and RDR were discrepant at ages 11, 13, and 16 years[2], leading to an equivocal decision regarding the best-fitting model. Specifically, ΔRMSEA favored a bifactor model at ages 11 and 13 years , and a three-

---

[2] *RMSEA* favored a bifactor model at ages 11 and 13 years , and a three-factor solution at 16 years. *CFI* favored a three-factor solution at ages 11 and 13 years but favored a bi-factor solution at 16 years.

factor model at 16 years, while ΔCFI favored a three-factor model at ages 11 and 13 years but favored a bi-factor model at 16 years. Finally, RDR favored a bifactor model at ages 11, 13, and 16 years, but the bifactor model that was different than the bifactor model favored by Δ*RMSEA* and Δ*CFI*. However, the bi-factor models produced a non-positive definite residual covariance matrix, specifically a negative residual variance for one or more indicator of psychopathology at ages 10, 13, and 17 years (a.k.a. Heywood cases).

Regarding parameter estimates, there was considerable variation among the standardized factor loadings (λ) onto the latent internalizing and externalizing factors in the bifactor models (range of $\lambda = -.01$ to 1.15), and many of these parameters were estimated with low precision ($SE_{\lambda s} > \lambda s$, *p*s > .10). In contrast, the standardized factor loadings in the correlated three-factor and higher-order model were moderate to high (range of $\lambda = .43$ to .89) and statistically significant at every wave of data collection (*p*s < .001). In light of the above considerations, particularly the inadmissible estimates (a.k.a. Heywood cases), bifactor models were neither carried forward to test for longitudinal invariance, estimate growth trajectories, nor estimate correlations between growth factors of personality and latent dimensions of psychopathology. Comparing one-factor, two-factor, and three-factor models, at least two of three fit statistics were consistent with the persistence of a three-factor solution from age 11 to 17. On the other hand, at age 10 fit statistics unambiguously preferred a two-factor solution. Therefore, a three-factor model was selected as the preferred model from age 11 to 17 and carried forward to test for longitudinal invariance of the structural and measurement properties of the high-order model depicted in Figure 2.

Longitudinal Measurement Invariance.

Before examining growth trajectories across different levels of a hierarchical model of psychopathology, a series of models were used to test for longitudinal invariance. First, a restrictive longitudinal baseline model was fit to the data. Depicted in Figure 2, the discrimination and threshold parameters of ordinal indicators were freely estimated but constrained to equality across measurement occasion. These parameters make up the measurement portion of the hierarchical model and the equality constraints reflect longitudinal invariance. Invariance of discrimination parameters indicates that, as individuals grow older, symptom-counts are equally sound indicators of their respective factors. Invariance of symptom thresholds implies that age-related differences in how frequently symptoms are endorsed are accounted for by the effects of age on the underlying factors. In the absence of invariant discrimination and threshold parameters, interpretation of latent variables can proceed under an assumption of partial measurement invariance, that is, assuming that age-related differences in thresholds and discriminations have only a minor impact on the estimated parameters of latent factors (Byrne, Shavelson, & Muthén, 1989).

In addition to the measurement portion of the hierarchical model, the longitudinal baseline model placed equality constraints on structural parameters. Depicted in Figure 2, measurement and structural constraints at different levels of the hierarchy are highlighted in shades of gray, including the residual variances of latent internalizing, ADHD, and externalizing factors (i.e. first-order variances), the magnitude of factor loadings from first-order factors onto a general factor (i.e. higher-order factor loadings), variation in the general factor at each measurement occasion (i.e. higher-order variances), and cross-time correlations between the general factor (i.e. latent correlations). Longitudinal measurement and structural invariance were tested by comparing the baseline model to a series of less restrictive models that, for example, allowed either the discrimination or threshold parameters to vary across measurement occasion. The baseline model was also compared to models that freely estimated structural parameters at each measurement occasion, including first-order variances, higher-order factor loadings, higher-order variances, and latent correlations that capture the rank-order stability of general liability for dimensions of psychopathology to correlate. A flow chart or "roadmap" that details the models that were used to test for longitudinal invariance is depicted in Figure S1.

Fit statistics for the models that were used to test for longitudinal invariance are reported in Table S8. *RDR* < .08 indicated that first-order variances were invariant across measurement occasion. *RDR*s > .08 indicated that constraining thresholds, discrimination parameters, higher-order factor loadings, variances, and correlations to equality resulted in misfit to the data. According to ΔCFI, however, first-order variances, higher-order factor loadings, and higher-order variances were invariant across measurement occasion, but thresholds, discrimination parameters, and latent correlations were noninvariant. Then again, according to ΔRMSEA, from age 11 to 17, none of the equality constraints depicted in Figure 2 resulted in misfit to the data. Thus, model comparisons provided mixed support for measurement and structural invariance of a higher-order model from age 11 to 17.

At least two of three fit statistics provided evidence for longitudinal invariance of structural parameters, including first-order variances, higher-order factor loadings, and higher-order variances. On the other hand, only one of three fit statistics were consistent with invariance of latent correlations and measurement properties, including the thresholds and discriminations of symptom-counts. However, after freely estimating thresholds and latent correlations, placing additional equality constraints on discrimination parameters did not result in misfit to the data. Note, the same pattern of results emerged in reduced models that only included observations at measurement occasions when self-report measures of Big Five personality were available (age 12 to 17), except these models provided stronger support for the invariance of discrimination parameters (see Table S8). These reduced models were fit to the data to establish whether longitudinal invariance of the higher-order factor model held across the waves of data collection that would be used in subsequent analyses to estimate average growth trajectories of focal study constructs, as well as correlations between growth factors.

As noted in the body of the manuscript, "despite some discrepancies across fit statistics, model comparisons were largely consistent with the longitudinal invariance of a three-factor high-order model from ages 12 to 17. Although fit statistics provided strong support for the invariance of structural parameters, there was weaker support for the invariance of measurement parameters, especially thresholds for symptom-counts. This suggests that age-related differences in how frequently symptoms are endorsed may not be fully accounted for by age-related differences in the underlying factors. Nevertheless, according to root mean squared error of approximation (RMSEA = .043), a model that freely estimated latent correlations but otherwise assumed full measurement and structural invariance met traditional standards for good model fit (Hu & Bentler, 1999). Although, the comparative fit index (CFI) for this model was below these standards (CFI = .871), it may be argued that *RMSEA* should be preferred over *CFI* "in confirmatory contexts, when researchers wish to determine whether a given model fits well enough to yield interpretable parameters" (p. 378; Rigdon, 1996). In addition, small distortions from simple structure have been shown to produce misfit in incremental fit indexes, like *CFI*, but not *RMSEA* (Beauducel & Wittmann, 2005)."

Finally, when the strength of correlations between observed indicators are moderate-to-low, like correlations between symptoms of psychopathology measured 5 years apart, then CFI may fall below traditional standards for good model fit (e.g. CFI < .90) because the unique off-diagonal elements of the observed covariate matrix are relatively small and not drastically different from the off-diagonal elements of the covariate matrix for the null model. In light of the above considerations, the three-factor higher-order model was carried forward for subsequent analyses. The results of this model are reported in the body of the manuscript, including factor loadings, latent correlations, and residual variances.

Growth Models.

As described in the body of the manuscript, "A series of latent growth models were fit separately to each Big Five trait and latent internalizing, ADHD, externalizing, and higher-order factors. For each construct, an intercept-only model was fit to the data, which implies no growth." In this model, the factor loadings onto the latent intercept were fixed to one at each measurement occasion, the variance of the latent intercept was freely estimated, and factor loadings onto the latent slope were fixed to zero. This model served as a baseline model for comparing alterative solutions, including linear, quadratic, and latent-basis growth models." In the linear growth model, the factor loading at the first measurement is fixed to zero and subsequent loadings are fixed to increase by one at each measurement occasion. This way, the variance of the intercept captures interindividual differences in levels of the Big Five and psychopathology at age 12, while the mean and variance of the slope capture the average rate of linear change from age 12 to 17 and interindividual differences in rates of linear change.

In addition to examining linear change, we also examined the possibility of non-linear change using quadratic and latent-basis growth models. The quadratic growth model estimated an additional latent factor, whereby factor loadings from the first to last measurement occasion were successively fixed to zero, one, four, nine, sixteen, and twenty-five. The mean of this factor captures the average direction and rate of quadratic growth, and the variance captures interindividual differences in rates of quadratic growth. In the latent-basis growth model, the factor loadings at the first and last measurement occasion were fixed to zero and five, while the remaining factor loadings were freely estimated from the data, such that the time metric was rescaled optimally to approximate the shape of the growth curve. In all univariate growth models, correlations between intercept and slope factors were freely estimated.

The absolute fit of growth models was evaluated using *RMSEA*, with values < .08 indicating adequate fit, and values < .05 indicating good fit to the data (Browne & Cudeck, 1993). Because quadratic and latent-basis models have the same degrees of freedom when fit to six repeated-measures, when comparing these models, the one with the smaller *RMSEA* was selected as the best-fitting model. When comparing the fit of growth models, change in model degrees of freedom (Δdf) were small, relative to Δdf when comparing measurement invariance models. Therefore, we did not rely on ΔRMSEA, ΔCFI, and RDR to compare growth models, as these fit statistics are recommended specifically for comparing measurement invariance models. Instead, the best-fitting non-linear growth model was compared to linear growth and intercept-only models using the more traditional $\Delta\chi^2$. Information criteria (AIC & BIC) were also used to compare growth models of Big Five personality, as these growth models were fit to continuous observed indicators. Therefore, growth models of Big Five personality were estimated using maximum likelihood and, consequently, the fitted likelihood function permitted the calculation of these fit statistics (AIC = -2logL + 2p & BIC = 2logL + log(n)p, where L is the fitted likelihood function and p is the number of estimated parameters).

On the other hand, growth models of psychopathology were fit to latent factors with ordered-categorical indicators. Therefore, these models were estimated using robust weighted least squares, which preclude the calculation of information criteria. Developmental changes in psychopathology factors were modeled using curve of factors models (CUFFs; McArdle, 1988), whereby a latent factor at each measurement occasion was regressed on intercept and slope factors to characterize initial-levels and changes over time. To examine within-individual changes in the general tendency for latent factors of psychopathology to correlate, the higher-order factor was regressed on intercept and slope factors. This model is depicted on the bottom panel of Figure S2. Next, a CUFFs model of the higher-order factor was compared to a factor of curves model (FOCUS; McArdle, 1988), which tests whether common intercept and slope factors can account for associations among the latent intercepts and slopes of internalizing, ADHD, and externalizing factors. A path diagram of the higher-order FOCUS model is depicted on the top panel of Figure S2[3].

Finally, although small in magnitude, after accounting for the general higher-order factor, there was statistically significant residual variance in latent factors of psychopathology. Consequently, a series of bivariate curves of factors models were used to estimate the growth trajectories of the residual variance in internalizing, ADHD, and externalizing factors, while accounting for common variance explained by the general higher-order factor. As unstandardized estimates of residual variance were low (range of σ^2 = .04 to .13) and free of unsystematic measurement error, after specifying growth factors, the residual variance of internalizing, ADHD, and externalizing factors was fixed to zero. Consequently, in these models, all of the variance in latent factors of psychopathology were explained by three factors: the higher-order factor and the residual intercept and slope factors, which capture variation in initial-levels and

[3] Because the indictors of common intercept and slope factors are the intercepts and slopes of latent variables (i.e. a series of CUFFs models), a more fitting abbreviation for this model might be FOCUFFs (i.e. Factors Of CUrves oF Factors).

within-changes in latent psychopathology factors that is independent from the higher-order factor. In addition, in these models the covariances between intercepts and slopes were freely estimated within-construct (e.g. the intercept and slope factors for the residual variance in latent internalizing were allowed to correlate), and the cross-construct covariances between the growth factors for the general and subordinate factors were fixed to zero (e.g. the intercept and slope factors for the residual variance in latent internalizing were not allowed to correlate with the intercept and slope factors for the higher-order factor).

Curve of Factors vs. Factor of Curves.

Two modeling approaches were used to account for developmental changes in the strength of correlations among internalizing, ADHD, and externalizing factors: A CUFFs model of a higher-order factor and a FOCUS model of second-order growth factors (see Figure S2). According to RMSEA, both approaches showed good fit to the data (range of RMSEA for CUFFs models =.042 to .044; range of *RMSEA* for FOCUS models = .037 - .039), but nested comparisons ($\Delta\chi 2$s > 370.00, Δdf = 13, *p*s < .001) indicated that FOCUS models showed better fit to the data. With respect to estimating intercept-intercept and slope-slope associations between personality and psychopathology, the two modeling approaches produced nearly identical results. Of the Big Five, only initial-levels of conscientiousness, agreeableness, and neuroticism were associated with initial-levels of general liability for psychopathology, and only interindividual differences in within-individual changes in extraversion and neuroticism were significantly associated with interindividual differences in within-individual change in higher-order psychopathology. Further, the point estimates and corresponding standard errors were nearly identical across the two modeling approaches. A comprehensive comparison of estimated covariances across the two modeling approaches can be found in Table S11.

Finally, to ensure that the correlations between the growth factors of personality and psychopathology were not confounded by sex differences, self-reported biological sex was specified as an exogenous covariate of growth factors. The effects of biological sex on growth factors of Big Five personality and psychopathology are reported in Table 4. Parameter estimates from growth models and CUFFs models are reported in Table 1. The correlations between the growth factors of personality and the curves of latent factors (i.e. initial-levels and slopes of latent psychopathology factors) are reported in Table 2 and Table 3.

Sensitivity Analyses.

At the request of a reviewer, additional models were fit to the data that freely estimated correlations among the residual variances of symptoms of common mental disorders after loading onto latent psychopathology factors, such that residual variance in symptoms of generalized anxiety, for example, were allowed to correlate directly over time. These correlations capture non-factor-related reasons why symptoms of each common mental disorder may be related over time. These parameter estimates are post-hoc in the sense that these correlations were neither included in initial data analytic procedures nor predicted *a priori* before models were fit to the data. The quadratic factor of curves model of the higher-order factor of psychopathology that included freely estimated correlations among the residual variances of symptom-counts showed good fit to the data (χ^2 = 2182.021, *df* =. 1318, *RMSEA* = .032, *CFI* = .948), but produced a Heywood case (e.g. a correlation greater than 1 between the residual variances of major depressive disorder) as the standardized residual variances in symptoms of depression at each measurement occasion were low (range of $\sigma 2$ = .14 to .19). Therefore, to ensure that intercept-intercept and slope-slope associations between Big Five personality and psychopathology were not biased due to the absence of these residual correlations, additional bivariate and multivariate models were fit to the data that included correlations among the residuals of latent psychopathology factors, excluding correlations between major depressive disorder, which were fixed to zero. These models showed good fit to the data (RMEAs < .05, CFIs > .90) and the size and precision of intercept-intercept and slope-slope correlations between the Big Five and psychopathology, as well as the results of null hypothesis significance tests remained largely unchanged. These results are reported in Tables S14 & S15. One exception was the intercept-intercept correlation between neuroticism and the residual variance in latent internalizing psychopathology, after accounting for the p-factor (r = .134, *SE* = .059, p = .023), which met a conventional but not a conservative threshold for statistical significance but was smaller and no longer statistically significant after including correlations among the residual variances of indicators of latent psychopathology factors (r = .079, *SE* = . .068, p = .245). As the present study adopted a conservative threshold for interpreting effects as statistically significant (α = .005), this change in results does not alter the conclusions that were drawn from the present study.

References

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 136−162). Newbury Park, CA: Sage.

Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling*, *12*(1), 41-75.

Browne, M. W., & Du Toit, S. H. (1992). Automated fitting of nonstandard models. *Multivariate Behavioral Research*, *27*(2), 269-300.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. Psychological Bulletin, 105, 456–466

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464-504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255.

Falconer, D. S. (1965). The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Annals of Human Genetics*, *29*(1), 51-76.

Hildebrandt, A., Wilhelm, O., & Robitzsch, A. (2009). Complementary and competing factor analytic approaches for the investigation of measurement invariance. *Review of Psychology*, *16*(2), 87-102.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Markus, K.A. (2019) Re: Configural Invariance - bad model fit. Email Correspondence from *SEMNET* on Dec 13, 2019, 9:35 AM.

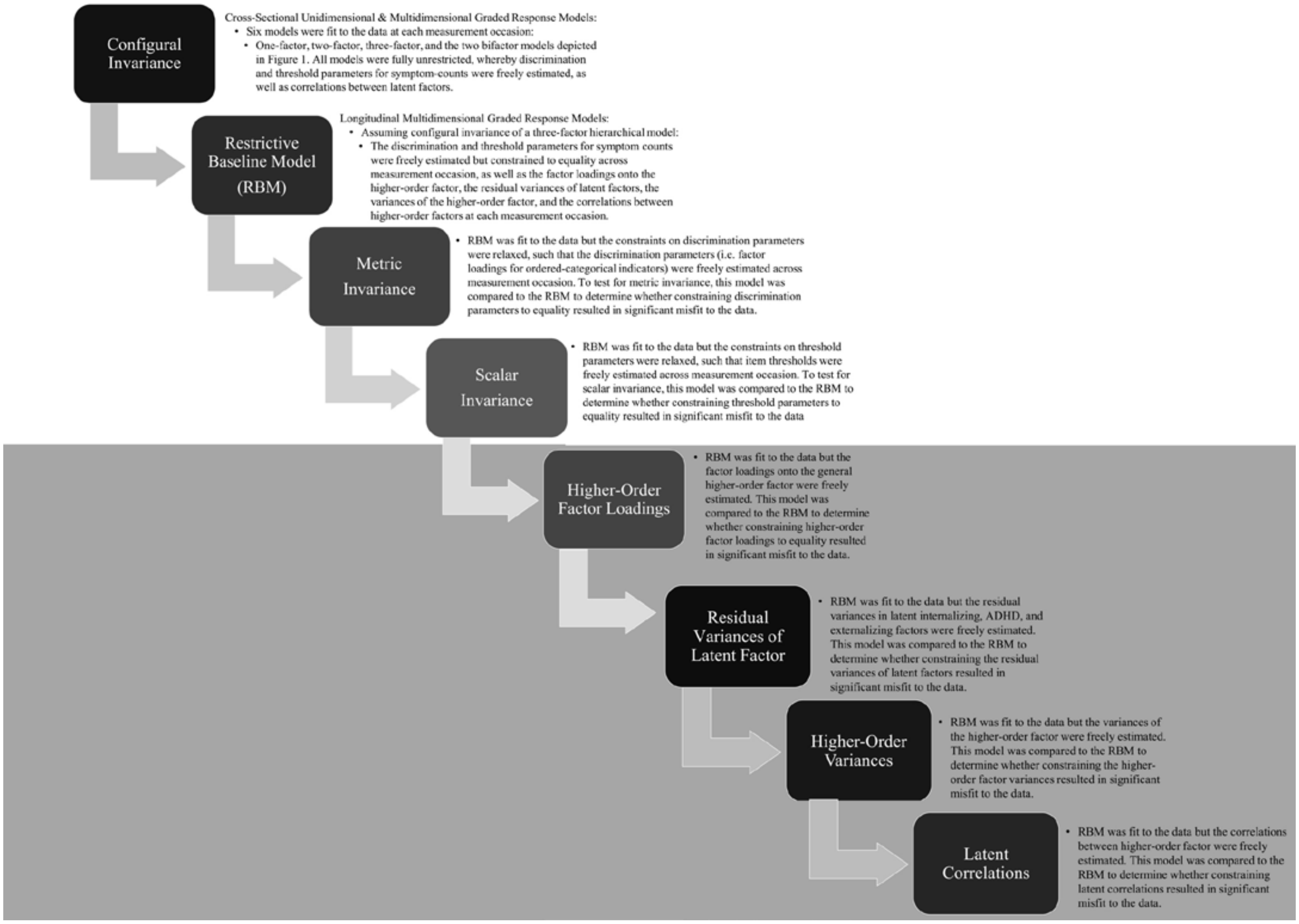
McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In *Handbook of multivariate experimental psychology* (pp. 561-614). Springer, Boston, MA.

Muthén, L. K., & Muthén, B. (1998-2017). Mplus. *The comprehensive modelling program for applied researchers: User's guide, 5*.

Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(4), 369-379.

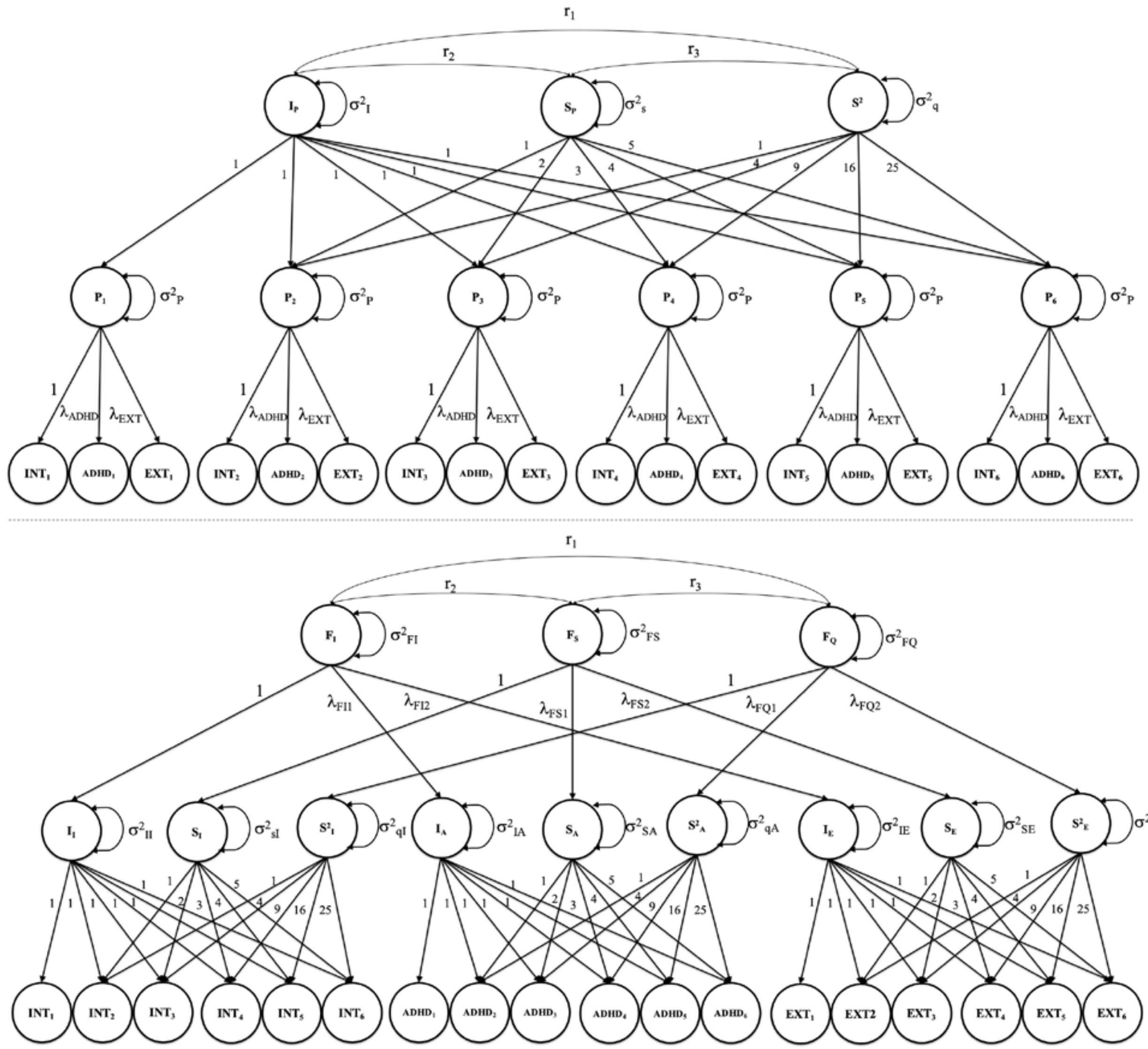
Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in Bioinformatics*, *3*(2), 119-133.

Figure S1. Roadmap for Testing Longitudinal Measurement Invariance of a Hierarchical Model of Common Mental Disorders



N**ote**. The top panel describes the analytic procedures that were used to evaluate the longitudinal invariance of the measurement properties of the higher-order model. The bottom panel describes the analytic procedures that were used to evaluate the longitudinal invariance of the structural components of the higher-order model.

Figure S2. Path Diagrams of Higher-Order Quadratic Curve of Factors and Factor of Curves Models



Note. Observed indicators, thresholds, discrimination parameters, intercepts, means, and the residual variances of first-order factors were omitted from the path diagrams to ease visualization

Table S1. Descriptive Statistics for Internalizing Symptoms & Marijuana Use

	Age (years)							
Symptoms	10	11	12	13	14	15	16	17
Depression								
M	5.50	3.72	3.99	3.95	4.08	3.79	3.36	2.96
σ2	16.99	13.98	14.64	14.65	17.79	15.07	14.16	11.59
N	643	564	575	590	597	588	599	598
(0)	58	113	101	105	146	131	163	182
(1 - 2)	122	167	155	173	132	153	169	163
(3 - 4)	133	97	118	97	105	113	92	100
(≥ 5)	330	187	201	215	214	191	175	153
(MDD) Dx	1.3%	0.4%	0.7%	1.5%	1.6%	1.5%	0.4%	1.3%
Generalized Anxiety								
M	3.81	2.58	2.37	1.97	1.96	1.62	1.40	1.23
σ2	4.80	3.28	3.13	2.95	3.40	2.85	2.46	2.32
N	644	562	574	590	589	578	596	596
(0)	28	71	82	127	146	179	211	250
(1 - 2)	162	221	245	271	250	261	261	247
(3 - 4)	245	197	189	143	142	98	90	73
(≥ 5)	209	73	58	49	51	40	34	26
(GAD) Dx	3.7%	0.6%	0.6%	0.3%	0.7%	1.3%	1.2%	0.3%
Post-Traumatic Stress								
M	1.17	0.55	0.52	0.36	0.28	0.44	0.30	0.20
σ2	10.33	4.66	3.77	2.67	2.34	3.56	2.04	1.19
N	663	565	575	590	603	590	600	599
(0)	572	520	526	557	577	550	566	574
(1 - 2)	3	7	5	4	4	5	7	5
(3 - 4)	7	8	14	4	6	9	7	7
(≥ 5)	81	30	30	25	16	26	20	13
(PTSD) Dx	1.9%	0.4%	0.1%	0%	0.1%	0.7%	0.2%	0.0%
Marijuana Use								
M	0.00	0.01	0.02	0.08	0.20	0.28	0.40	0.41
σ2	0.00	0.04	0.09	0.41	0.89	1.35	2.20	2.27
N	663	565	575	590	603	590	600	599
(0)	663	564	572	579	569	543	540	541
(1 - 2)	0	0	2	3	15	24	26	19
(3 - 4)	0	0	0	4	8	6	10	14
(≥ 5)	0	1	1	4	11	17	24	25
(Dependence) Dx	0%	0.1%	0.1%	0.4%	1.8%	2.1%	3.4%	3.3%

Notes. Means and variances of continuous symptom counts, and distributions of cut variables are reported across eight waves of data collection for depression, anxiety, PTSD, and marijuana use. Dx = percent of youth meeting diagnostic criteria for the respective psychiatric disorder.

Table S2. Descriptive Statistics for ADHD and Externalizing Symptoms

	Age (years)							
Symptoms	10	11	12	13	14	15	16	17
Attention Deficit								
M	1.21	0.87	1.01	1.16	1.07	1.16	0.93	0.79
σ2	2.22	1.50	1.74	2.13	2.01	2.01	1.97	1.50
N	640	557	574	588	433	543	584	598
(0)	269	290	268	258	205	259	314	344
(1 - 2)	264	211	243	241	169	212	205	196
(3 - 4)	85	44	48	68	42	51	44	44
(≥ 5)	22	12	15	21	17	21	21	14
(ADHD) Dx	0.3%	0.0%	0.0%	0.0%	0.0%	0.3%	0.0%	0.0%
Hyperactivity								
M	1.03	0.86	1.01	1.02	0.97	0.88	0.71	0.50
σ2	1.92	1.85	2.22	2.23	2.00	2.15	1.82	1.15
N	642	564	574	588	429	533	581	596
(0)	298	315	280	300	225	313	393	435
(1 - 2)	273	191	231	212	153	160	127	126
(3 - 4)	49	44	40	51	35	41	44	27
(≥ 5)	22	14	23	25	16	19	17	8
(ADHD) Dx	0.3%	0.0%	0.0%	0.0%	0.0%	3.0%	0.0%	0.0%
Defiance								
M	0.98	0.94	1.11	1.30	1.19	1.11	0.76	0.68
σ2	2.12	2.14	2.37	2.74	2.61	2.26	1.56	1.42
N	638	557	575	589	594	533	581	596
(0)	338	318	298	275	298	255	351	382
(1 - 2)	205	164	178	189	187	197	173	168
(3 - 4)	68	51	71	88	71	56	41	34
(≥ 5)	27	24	28	37	38	25	16	12
(ODD) Dx	0.9%	1.8%	1.9%	2.7%	2.5%	2.2%	1.3%	1.2%
Emotion Dysregulation								
M	1.18	1.00	1.01	1.08	1.10	1.11	0.80	0.77
σ2	1.52	1.56	1.65	1.72	1.82	1.90	1.37	1.35
N	638	557	575	589	594	533	581	596
(0)	251	271	296	286	302	270	346	369
(1 - 2)	277	201	185	194	167	155	166	158
(3 - 4)	110	85	94	109	125	108	69	69
(≥ 5)	0	0	0	0	0	0	0	0
(ODD) Dx	0.9%	1.8%	1.9%	2.7	2.5%	2.2%	1.3%	1.2%
Conduct Disorder								
M	0.39	0.31	0.47	0.69	0.76	0.74	0.69	0.69
σ2	1.09	0.86	0.99	1.58	1.67	1.55	1.31	1.23
N	641	564	574	590	596	579	597	597
(0)	505	471	440	418	409	406	437	434
(1 - 2)	119	77	106	118	135	117	114	116
(3 - 4)	11	14	24	41	40	46	41	43
(≥ 5)	6	2	4	13	12	10	5	4
(CD) Dx	2.5%	2.4%	5.2%	8.5%	5.9%	5.6%	4.9%	3.0%

Notes. Means and variances of continuous symptom counts, and distributions of cut variables are reported across eight waves of data collection for CD and facets of ADHD and ODD. Dx = percent of youth meeting diagnostic criteria for the respective psychiatric disorder.

Table S3. Descriptive Statistics for Big Five Personality

		Age (years)					
Symptoms		12	13	14	15	16	17
Openness							
	M	3.75	3.73	3.85	3.79	3.77	3.89
	σ2	0.49	0.46	0.41	0.40	0.39	0.34
	N	313	588	604	589	600	600
Conscientiousness							
	M	3.82	3.75	3.78	3.82	3.72	3.89
	σ2	0.42	0.45	0.40	0.39	0.40	0.38
	N	313	590	604	588	600	600
Extraversion							
	M	3.31	3.40	3.52	3.42	3.39	3.46
	σ2	0.63	0.58	0.57	0.55	0.48	0.59
	N	312	589	603	589	600	600
Agreeableness							
	M	3.47	3.62	3.63	3.72	3.62	3.90
	σ2	0.46	0.49	0.43	0.38	0.40	0.37
	N	310	585	604	589	600	600
Neuroticism							
	M	2.48	2.47	2.45	2.37	2.40	2.33
	σ2	0.52	0.56	0.46	0.49	0.50	0.54
	N	314	589	604	589	600	600

Notes. Means, variances, and sample sizes are reported across six waves of data collection.

Table S4. Observed Sex Differences for Internalizing Symptoms & Marijuana Use

	Age (years)							
Symptoms	10	11	12	13	14	15	16	17
Depression								
female mean	5.59	3.86	4.16	4.88	5.22	4.80	4.23	3.71
male mean	5.42	3.58	3.80	3.00	2.92	2.76	2.48	2.19
t	0.53	0.87	1.13	6.17	6.93	6.61	5.83	5.61
df	641	562	573	588	595	685	597	596
p	.596	.386	.256	< .001	< .001	< .001	< .001	< .001
Cohen's *d*	0.04	0.07	0.09	0.51	0.57	0.54	.48	0.46
Anxiety								
female mean	4.14	2.65	2.55	2.33	2.30	1.93	1.66	1.46
male mean	3.49	2.50	2.18	1.61	1.61	1.31	1.14	1.00
t	3.81	0.98	2.47	5.13	4.58	4.49	4.16	3.78
df	642	560	572	588	587	576	594	594
p	< .001	.326	.013	< .001	< .001	< .001	< .001	< .001
Cohen's *d*	0.30	0.08	0.21	0.42	0.37	0.37	0.34	0.31
PTSD								
female mean	1.25	0.52	0.68	0.48	0.49	0.75	0.39	0.25
male mean	1.09	0.57	0.35	0.24	0.07	0.12	0.20	0.15
t	0.64	0.31	2.01	1.76	3.39	4.08	1.64	1.20
df	661	563	573	588	601	588	598	597
p	.524	.764	.045	.078	< .001	< .001	.100	.228
Cohen's *d*	0.05	0.03	0.17	0.14	0.28	0.34	0.13	0.10
Marijuana Use								
female mean	0	0.00	0.02	0.08	0.21	0.28	0.43	0.38
male mean	0	0.02	0.01	0.08	0.19	0.27	0.38	0.44
t	0	0.99	0.52	0.03	0.24	0.13	0.39	0.53
df	661	563	573	588	601	588	598	597
p	NA	.319	.601	.976	.808	.893	.696	.593
Cohen's *d*	NA	0.08	0.04	0.00	0.02	0.01	.03	.04

Notes. t = t-test of biological sex difference. df = degrees of freedom. p = two-tailed probability of type-I error. Cohen's d = absolute value of standardized mean difference. NA = not applicable.

Table S5. Observed Sex Differences for ADHD and Externalizing Symptoms

	Age (years)							
Symptoms	10	11	12	13	14	15	16	17
Attention (ADHD)								
female mean	1.23	0.93	0.96	1.37	1.14	1.18	0.97	0.79
male mean	1.19	0.81	1.06	0.95	1.00	0.93	0.88	0.79
t	0.30	1.13	0.95	3.51	1.05	2.07	0.78	0.04
df	638	555	572	586	431	541	582	596
p	.765	.259	.341	< .001	.295	.038	.433	.965
Cohen's *d*	0.02	0.10	0.08	0.29	0.10	0.18	0.06	0.00
Hyperactivity (ADHD)								
female mean	1.01	.914	0.95	1.17	1.06	1.00	0.74	0.48
male mean	1.04	.805	1.06	0.88	0.89	0.77	0.68	0.52
t	0.26	0.96	0.87	2.37	1.23	1.82	0.53	0.38
df	640	562	572	586	427	531	579	594
p	.789	.338	.384	.018	.218	.069	.598	.703
Cohen's *d*	0.02	0.08	0.07	0.19	0.12	0.16	0.04	0.03
Defiance (ODD)								
female mean	0.99	0.98	1.25	1.62	1.46	1.26	0.83	0.79
male mean	0.96	0.90	0.98	0.97	0.92	0.95	0.69	0.56
t	0.13	0.66	2.14	4.87	4.10	2.33	1.40	2.44
df	636	555	573	587	592	531	579	594
p	.895	.509	.032	< .001	< .001	.019	.161	.015
Cohen's *d*	0.01	0.06	0.18	0.40	0.34	0.20	0.12	0.20
Emotion Dys. (ODD)								
female mean	1.19	1.13	1.11	1.46	1.41	1.37	1.02	0.99
male mean	1.18	0.86	0.90	0.70	0.78	0.84	0.57	0.54
t	0.09	2.52	1.94	7.32	5.81	4.53	4.71	4.84
df	636	555	573	587	592	531	579	594
p	.903	.012	.052	< .001	< .001	< .001	< .001	< .001
Cohen's *d*	0.01	0.21	0.16	0.60	0.48	0.40	0.40	0.40
Conduct Disorder								
female mean	0.35	0.28	0.43	0.67	0.75	0.74	0.69	0.70
male mean	0.42	0.35	0.52	0.70	0.76	0.74	0.68	0.68
t	0.86	0.84	1.14	0.29	0.17	0.01	0.14	0.25
df	639	562	572	588	594	577	595	595
p	.388	.401	.253	.765	.861	.994	.888	.805
Cohen's *d*	0.07	0.07	0.09	0.02	0.01	0.00	0.01	0.02

Notes. t = t-test of biological sex difference. df = degrees of freedom. p = two-tailed probability of type-I error. Cohen's d = absolute value of standardized mean difference.

Table S6. Observed Sex Differences for Big Five Personality

	Age (years)					
Symptoms	12	13	14	15	16	17
Openness						
female mean	3.77	3.73	3.91	3.81	3.78	3.90
male mean	3.73	3.72	3.78	3.77	3.77	3.87
t	0.50	0.28	2.51	0.71	0.09	0.81
df	311	586	602	587	598	598
p	.614	.781	.012	.478	.924	.418
Cohen's *d*	0.06	0.02	0.20	0.06	0.01	0.07
Conscientiousness						
female mean	3.90	3.84	3.86	3.87	3.69	3.94
male mean	3.74	3.66	3.70	3.76	3.75	3.84
t	2.24	3.17	3.16	2.26	1.19	2.04
df	311	588	602	586	598	598
p	.025	.002	.002	.024	.233	.042
Cohen's *d*	0.25	0.26	0.26	0.19	0.10	0.17
Extraversion						
female mean	3.29	3.37	3.52	3.35	3.36	3.43
male mean	3.33	3.44	3.51	3.49	3.40	3.48
t	0.55	1.00	0.15	2.32	0.63	0.83
df	310	587	601	587	598	598
p	.580	.318	.877	.020	.525	.408
Cohen's *d*	0.06	0.08	0.01	0.19	0.05	0.07
Agreeableness						
female mean	3.51	3.67	3.70	3.82	3.72	3.96
male mean	3.42	3.56	3.56	3.62	3.53	3.84
t	1.16	1.89	2.57	4.08	3.67	2.42
df	308	583	602	587	598	598
p	.245	.058	.010	< .001	< .001	.015
Cohen's *d*	0.13	0.16	0.21	0.33	0.30	0.20
Neuroticism						
female mean	2.54	2.56	2.51	2.41	2.50	2.43
male mean	2.42	2.37	2.38	2.33	2.29	2.21
t	1.48	3.15	2.53	1.34	3.64	3.63
df	312	587	602	587	598	598
p	.141	.002	.011	.180	< .001	< .001
Cohen's *d*	0.17	0.26	0.21	0.11	0.30	0.30

Notes. t = t-test of biological sex difference. df = degrees of freedom. p = two-tailed probability of type-I error. Cohen's d = absolute value of standardized mean difference.

Table S7. Model Fit Statistics for Cross-Sectional CFA Models of a Hierarchical Structure of Common Mental Disorders from Age 10 to 17

		Absolute Fit					Model Comparisons						
Age	Models	χ2	df	p	RMSEA	CFI	Model	ΔRMSEA[1]	ΔCFI[2]	Δχ2	Δdf	p	RDR[3]
10	A: 1-Factor	82.936	20	< .001	.069	.964	AvsB	-.014	.014	20.530	1	< .001	.172
	B: 2-Factor	57.559	19	< .001	.055	.978	BvsC	.003	.000	2.366	2	.306	.016
	C/D: 3-Factor	55.358	17	< .001	.058	.978	CvsD	-.026	.016	26.627	2	< .001	.136
	E: 1-Bi-Factor[1,2,3]	25.153	15	.048	.032	.994	DvsE	-.004	.002	6.599	3	.085	.042
	F: 2-Bi-Factor*	18.321	12	.106	.028	.996							
11	A: 1-Factor	64.770	20	< .001	.063	.978	AvsB	-.015	.009	17.137	1	< .001	.169
	B: 2- Factor	44.169	19	< .001	.048	.987	BvsC	-.018	.009	16.497	2	< .001	.113
	C/D: 3-Factor[2]	25.383	17	.086	.030	.996	CvsD	-.011	.003	6.753	2	.034	.065
	E: 1-Bi-Factor[3]	17.932	15	.266	.019	.999	DvsE	-.010	.001	4.880	3	.181	.033
	F: 2-Bi-Factor[1]	12.495	12	.407	.009	1.00							
12	A: 1-Factor	72.121	20	< .001	.067	.974	AvsB	-.002	.002	6.818	1	.009	.100
	B: 2-Factor	65.823	19	< .001	.065	.976	BvsC	-.014	.011	20.127	2	< .001	.125
	C/D: 3-Factor[1,2,3]	42.056	17	< .001	.051	.987	CvsD	.002	.002	2.868	2	.238	.027
	E: 1-Bi-Factor	39.419	15	< .001	.053	.988	DvsE	.011	-.003	0.568	3	.904	NA
	F: 2-Bi-Factor	40.639	12	< .001	.064	.985							
13	A: 1-Factor	111.551	20	< .001	.088	.964	AvsB	-.006	.006	15.668	1	< .001	.157
	B: 2-Factor	94.567	19	< .001	.082	.970	BvsC	-.024	.016	27.358	2	< .001	.146
	C/D: 3-Factor[2]	60.438	17	< .001	.066	.983	CvsD	-.010	.006	15.426	2	< .001	.106
	E: 1-Bi-Factor*[1]	43.052	15	< .001	.056	.989	DvsE	-.008	.004	14.349	3	.002	.080
	F: 2-Bi-Factor*[3]	28.532	12	.005	.048	.993							
14	A: 1-Factor	157.451	27	< .001	.090	.950	AvsB	-.012	.013	30.187	1	< .001	.222
	B: 2-Factor	121.629	26	< .001	.078	.963	BvsC	-.003	.003	17.114	2	< .001	.113
	C/D: 3-Factor	105.073	24	< .001	.075	.969	CvsD	-.034	.023	56.300	3	< .001	.173
	E: 1-Bi-Factor[1,2,3]	42.190	21	.004	.041	.992	DvsE	.001	.001	5.736	3	.125	.039
	F: 2-Bi-Factor	37.159	18	.005	.042	.993							
15	A: 1-Factor	204.693	27	< .001	.106	.910	AvsB	-.009	.017	27.389	1	< .001	.208
	B: 2-Factor	171.572	26	< .001	.097	.927	BvsC	-.006	.014	26.961	2	< .001	.144
	C/D: 3-Factor	141.855	24	< .001	.091	.941	CvsD	-.023	.030	66.289	3	< .001	.187
	E: 1-Bi-Factor[1,2,3]	77.775	21	< .001	.068	.971	DvsE	.004	.001	7.289	3	.063	.051
	F: 2-Bi-Factor	73.005	18	< .001	.072	.972							
16	A: 1-Factor	153.396	27	< .001	.088	.918	AvsB	-.008	.017	24.540	1	< .001	.198
	B: 2-Factor	126.035	26	< .001	.080	.935	BvsC	-.010	.019	28.569	2	< .001	.149
	C/D: 3-Factor[1]	94.434	24	< .001	.070	.954	CvsD	-.006	.013	23.641	3	< .001	.107
	E: 1-Bi-Factor[2]	71.811	21	< .001	.064	.967	DvsE	.003	.002	7.843	3	.049	.052
	F: 2-Bi-Factor[3]	66.538	18	< .001	.067	.969							
17	A: 1-Factor	142.876	27	< .001	.085	.933	AvsB	-.006	.011	18.707	1	< .001	.172
	B: 2-Factor	122.989	26	< .001	.079	.944	BvsC	-.018	.025	38.240	2	< .001	.174
	C/D: 3-Factor[1,2,3]	76.936	24	< .001	.061	.969	CvsD	.007	-.003	3.719	3	.293	.020
	E: 1-Bi-Factor	78.638	21	< .001	.068	.966	DvsE	.002	.003	8.749	3	.032	.056
	F: 2-Bi-Factor*	71.170	18	< .001	.070	.969							

Notes. χ2 = model chi-squared. df = model degrees of freedom. p = p-value for model chi-squared. RMSEA = root mean squared error of approximation. CFI = comparative fit index. Δχ2 = change in model chi-squared. Δdf = change in model degrees of freedom. RDR = root deterioration per restriction. Asterisks denote models the produced a non-positive definite residual covariance matrix.

Table S8. Model Fit Statistics for Longitudinal Measurement Invariance Models of a Hierarchical Structure of Common Mental Disorders from Age 11 to 17 (Top Panel) and from Age 12 to 17 (Bottom Panel)

	Absolute Fit					Model Comparisons						
Models (Age 11 to 17)	$\chi 2$	df	p	RMSEA	CFI	Model	Δ RMSEA[1]	ΔCFI[2]	$\Delta\chi 2$	Δdf	p	RDR
1. Fully Restricted	4578.12	1093	< .001	.046	.852							
2. Free Discriminations	4853.37	1870	< .001	.049	.835	1 v. **3**	.003	.017	92.51	33	< .001	.05
3. Free Thresholds	4056.96	1757	< .001	.045	.872	1 v. **2**	-.001	.020	1055.01	146	< .001	.10
4. Free Higher-Order λ	4735.47	1891	< .001	.048	.843	**1** v. 5	.002	-.009	36.39	12	< .001	.06
5. Free Residual σ^2	4576.37	1885	< .001	.046	.851	**1** v. 4	.000	-.001	19.34	18	.371	.01
6. Free Higher-Order σ^2	4601.53	1898	< .001	.046	.851	**1** v. 6	.000	-.001	20.00	6	.002	.06
7. Free Correlations	4222.95	1883	< .001	.043	.871	1 v. **7**	-.003	.019	200.93	20	< .001	.12
8. Free #7 & #3	3671.30	1737	< .001	.041	.893	7 v. **8**	-.002	.022	1055.09	16	< .001	.10
9. Free #7 & #2	4437.95	1850	< .001	.046	.857	**7** v. 9	.003	-.014	58.30	33	.004	.03
10. Free #7 & #2 & #3	3846.55	1704	< .001	.044	.882	**8** v.10	.003	-.011	58.30	33	.004	.03

Models (Age 12 to 17)	$\chi 2$	df	p	RMSEA	CFI	Model	Δ RMSEA[1]	ΔCFI[2]	$\Delta\chi 2$	Δdf	p	RDR
1. Fully Restricted	3923.28	1436	< .001	.052	.850							
2. Free Thresholds	3472.69	1313	< .001	.050	.870	1 v. **2**	-.002	.020	865.98	123	< .001	.10
3. Free Discriminations	4191.25	1408	< .001	.055	.832	**1** v. 3	.003	-.029	75.35	28	< .001	.05
4. Free Lower-Order σ^2	3924.51	1421	< .001	.052	.849	**1** v. 4	.000	-.012	19.36	15	.197	.02
5. Free Higher-Order λ	4040.47	1426	< .001	.053	.842	**1** v. 5	.001	-.019	35.62	10	< .001	.06
6. Free Higher-Order σ^2	3947.29	1431	< .001	.052	.848	**1** v. 6	.000	-.013	19.72	5	.001	.07
7. Free Correlations	3707.88	1422	< .001	.050	.862	1 v. **7**	-.002	.012	148.63	14	< .001	.12
8. Free #7 & #3	3238.42	1299	< .001	.048	.883	7 v. **8**	-.002	.021	865.98	123	< .001	.10
9. Free #7 & #2	3911.43	1394	< .001	.053	.848	**7** v. 9	.003	-.014	54.49	28	.002	.04
10. Free #7 & #2 & #3	3411.76	1271	< .001	.051	.871	**8** v.10	.003	-.012	54.49	28	.002	.04

Notes. $\chi 2$ = model chi-squared. df = model degrees of freedom. p = p-value for model chi-squared. RMS = root mean squared error of approximation. CFI = comparative fit index. Δ = change. RDR = root deterioration per restriction. RMS decreasing more than or equal to .01 would indicate that equality constraints result in significant misfit to the data and, therefore, the more restrictive model should be rejected. CFI increasing more than or equal to .01 would indicate that equality constraints result in misfit to the data. RDR values equal to or greater than .05 indicate that equality constraints result in misfit to the data. Preferred models are printed in bold.

Table S9. Model Fit Statistics for Factor of Curves Models for the Higher-Order Factor and Internalizing, Attention-Hyperactivity, and Externalizing Factors of Psychopathology

		Model Fit				Model Comparisons			
Factor	Model	$\chi2$	df	RMSEA	CFI	Model	$\Delta\chi2$	Δdf	p
P	1. Intercept	3923.28	1436	.052	.850				
	2. Linear	3670.91	1433	.049	.865	1 vs. 2	75.55	3	< .001
	3. **Quadratic**	3514.82	1429	**.048**	**.874**	2 vs. 3	75.62	4	< .001
	4. Latent-Basis	3648.63	1429	.049	.866	2 vs. 4	31.05	4	< .001
INT	1. Intercept	809.39	194	.070	.868				
	2. Linear	505.68	191	.051	.932	1 vs. 2	150.69	3	< .001
	3. **Quadratic**	450.58	187	**.047**	**.943**	2 vs. 3	46.75	4	< .001
	4. Latent-Basis	506.00	187	.051	.932	2 vs. 4	2.08	4	.719
ADHD	1. Intercept	395.08	93	.071	.899				
	2. Linear	272.89	90	.056	.939	1 vs. 2	67.36	3	< .001
	3. **Quadratic**	243.13	86	**.053**	**.947**	2 vs. 3	25.98	4	< .001
	4. Latent-Basis	259.89	86	.056	.942	2 vs. 4	15.35	4	.004
EXT	1. Intercept	1442.89	275	.081	.844				
	2. Linear	1321.50	272	.077	.860	1 vs. 2	69.78	3	< .001
	3. Quadratic	1230.88	268	**.075**	**.871**	2 vs. 3	70.16	4	< .001
	4. Latent-Basis	1306.90	268	.077	.861	2 vs. 4	25.02	4	< .001
INT	1. Intercept	3272.65	1431	.045	.889				
(Residual	2. Linear	3232.03	1428	.044	.891	1 vs. 2	49.623	3	< .001
Variance)	**3. Quadratic**	3222.06	1424	.044	.892	2 vs. 3	18.612	4	< .001
	4. Latent-Basis	3223.76	1424	.044	.891	2 vs. 4	14.664	4	.006
ADHD	1. Intercept	3439.11	1429	.047	.879				
(Residual	**2. Linear**	3434.98	1426	.047	.879	1 vs. 2	11.176	3	.011
Variance)	3. Quadratic	3431.32	1422	.047	.879	2 vs. 3	8.935	4	.063
	4. Latent-Basis	3432.33	1422	.047	.879	2 vs. 4	8.951	4	.062
EXT	1. Intercept	3072.66	1429	.042	.901				
(Residual	**2. Linear**	3039.36	1426	.042	.903	1 vs. 2	33.752	3	< .001
Variance)	3. Quadratic*	3025.99	1422	.042	.903	2 vs. 3	22.222	4	< .001
	4. Latent-Basis	3032.97	1422	.042	.903	2 vs. 4	15.092	4	.004

Notes. $\chi2$ = model chi-squares were all statistically significant at p < .001. df = model degrees of freedom. p = probability. RMS = root mean squared error of approximation. $\Delta\chi2$ = chi-squared difference. Δdf = chang in model degrees of freedom. Best-fitting models are printed in bold font. P = General psychopathology or "P Factor". INT = Internalizing. ADHD = Attention Deficit Hyperactivity Disorder. EXT = Externalizing Asterisks denote models the produced a non-positive definite residual covariance matrix.

Table S10. Model Fit Statistics for Growth Models of Big Five Personality

								Model Comparisons			
Trait	Model	AIC	BIC	χ2	df	CFI	RMSEA	Model	Δχ2	Δdf	p
O	1. Intercept	6077.17	**6112.92**	55.72	19	.897	.055				
	2. Linear	**6064.55**	6113.71	37.11	16	.941	.045	1 vs. 2	18.61	3	< .001
	3. Quadratic	6068.08	6135.11	32.63	12	.942	.052	2 vs. 3	4.48	4	.345
	4. Latent-Basis*	6056.24	6123.28	20.80	12	.975	.034	2 vs. 4	16.31	4	.002
C	1. Intercept	5846.45	5882.21	105.19	19	.865	.084				
	2. Linear	5818.59	5867.75	71.33	16	.913	.073	1 vs. 2	33.86	3	< .001
	3. Quadratic	**5792.43**	**5859.47**	37.17	12	.961	.057	2 vs. 3	34.16	4	< .001
	4. Latent-Basis*	5801.48	5868.52	46.23	12	.946	.067	2 vs. 4	25.10	4	< .001
E	1. Intercept	6716.37	6752.12	98.75	19	.902	.081				
	2. Linear	6678.79	**6727.96**	55.17	16	.952	.062	1 vs. 2	43.58	3	< .001
	3. Quadratic	**6677.97**	6745.01	46.35	12	.958	.067	2 vs. 3	8.82	4	.065
	4. Latent-Basis	6681.28	6748.32	49.67	12	.954	.070	2 vs. 4	5.50	4	.239
A	1. Intercept	6301.53	6337.28	167.73	19	.534	.110				
	2. Linear	6207.03	6256.19	67.23	16	.839	.070	1 vs. 2	100.50	3	< .001
	3. Quadratic	6210.43	6277.40	62.64	12	.841	.081	2 vs. 3	4.59	4	.332
	4. Latent-Basis	**6157.59**	**6224.63**	9.79	12	1.00	.000	2 vs. 4	57.44	4	< .001
N	1. Intercept	6587.31	6623.06	65.80	19	.921	.062				
	2. Linear	**6547.74**	**6596.90**	20.23	16	.993	.020	1 vs. 2	45.57	3	< .001
	3. Quadratic	6553.09	6620.13	17.58	12	.991	.027	2 vs. 3	2.65	4	.618
	4. Latent-Basis	6554.43	6621.47	18.92	12	.988	.030	2 vs. 4	1.31	4	.859

Notes. Repeated-measures were specified as continuous indicators for growth models of Big Five personality, which were estimated using maximum likelihood. AIC = Akaike Information Criteria. BIC = Bayesian Information Criteria. χ2 = model chi-square. df = model degrees of freedom. RMSEA = root mean squared error of approximation. Δχ2 = chi-square difference. Δdf = change in model degrees of freedom. RDR = root deterioration per restriction. Asterisks denote models that produced a Haywood case. Best-fitting models are printed in bold font.

Table S11. Comparing Covariances Between the Growth Factors of Big Five Personality and the Growth Factors of Higher-Order Factor of Psychopathology Estimated using CUFFS and FOCUS Models

Construct:	P-Factor					
Big Five / Model:	CUFFS	FOCUS	CUFFS	FOCUS	CUFFS	FOCUS
Personality: / Covariance:	**Intercept-Intercept**	**Intercept-Intercept**	**Slope1-Linear**	**Slope1-Linear**	**Slope1/Quadratic-Quadratic**	**Slope1/Quadratic-Quadratic**
Openness/Intellect	-.003 (.010)	-.003 (.010)	-.003 (.002)	-.003 (.002)	.001 (.000)	.000 (.000)
Conscientiousness	-.041*(.010)	-.043*(.010)	.000 (.001)	.000 (.001)	.000 (.000)	.000 (.000)
Extraversion	.017 (.012)	.019 (.013)	-.007*(.002)	-.008*(.002)	.001 (.000)	.000 (.000)
Agreeableness	-.040*(.010)	-.042*(.010)	.005 (.010)	.004 (.010)	.000 (.002)	.000 (.002)
Neuroticism	.101*(.012)	.105*(.103)	.006*(.002)	.007*(.002)	-.001 (.000)	-.001 (.000)

Notes. Standard errors are reported in parentheses. Asterisks denote estimates that were statistically significant at $p_{(two\text{-}tailed)} < .005$. CUFFS = Curve of Factors Model; FOCUS = Factor of Curves Model.

S12. Correlations Between the Growth Factors of Big Five Personality and the Curves of Latent Psychopathology Factors Before & After Controlling for Self-Report Biological Sex

Construct:	P-Factor					
Big Five Personality: Model:	Zero-Order	Partial	Zero-Order	Partial	Zero-Order	Partial
Covariance:	**Intercept-Intercept**	**Intercept-Intercept**	**Slope-Linear**	**Slope-Linear**	**Slope-Quadratic**	**Slope-Quadratic**
Openness/Intellect	-.018 (.052)	-.029 (.050)	-.241 (.149)	-.252 (.152)	.224 (.161)	.241 (.167)
Conscientiousness	-.162*(041)	-.199*(.042)	-.003 (.023)	-.004 (.023)	-.045 (.030)	-.041 (.030)
Extraversion	.068 (.047)	.078 (.046)	-.290*(.092)	-.291*(.094)	.189 (.100)	.191 (.103)
Agreeableness	-.206*(.050)	-.275*(.047)	.052(.113)	.065 (.080)	-.027 (.123)	-.080 (.090)
Neuroticism	.458*(.045)	.435*(.044)	.288*(.100)	.251 (.099)	-.322*(.109)	-.283 (.109)

Construct:	Internalizing					
Big Five Personality: Model:	Zero-Order	Partial	Zero-Order	Partial	Zero-Order	Partial
Covariance:	**Intercept-Intercept**	**Intercept-Intercept**	**Slope1-Linear**	**Slope1-Linear**	**Slope1-Quadratic**	**Slope1-Quadratic**
Openness/Intellect	-.039 (.053)	-.057 (.053)	-.081 (.163)	-.081 (.999)	-.006 (.188)	-.012 (.192)
Conscientiousness	-.085 (.042)	-.131*(.043)	-.021 (.026)	-.027 (.027)	-.017 (.034)	-.006 (.034)
Extraversion	-.041 (.049)	-.031 (.049)	-.226 (.104)	-.219 (.107)	.061 (.123)	.046 (.127)
Agreeableness	-.151*(.053)	-.204*(.050)	.220 (.144)	.112 (.088)	-.222 (.163)	-.140 (.103)
Neuroticism	.430*(.049)	.400*(.047)	.160 (.112)	.112 (.110)	-.124 (.129)	-.075 (.127)

Construct:	ADHD					
Big Five Personality: Model:	Zero-Order	Partial	Zero-Order	Partial	Zero-Order	Partial
Covariance:	**Intercept-Intercept**	**Intercept-Intercept**	**Slope1-Linear**	**Slope1-Linear**	**Slope1-Quadratic**	**Slope1-Quadratic**
Openness/Intellect	.007 (.054)	.004 (.054)	-.408 (.232)	-.453 (.255)	.482 (.273)	.551 (.314)
Conscientiousness	-.246*(.045)	-.261*(.047)	.030 (.029)	.026 (.031)	-.089 (.042)	-.086 (.046)
Extraversion	.080 (.052)	.082 (.052)	-.249 (.133)	-.256 (.143)	.188 (.150)	.193 (.164)
Agreeableness	-.199*(.051)	-.206*(.050)	.103 (.155)	-.018 (.102)	-.035 (.174)	.052 (.124)
Neuroticism	.371*(.050)	.365*(.049)	.363 (.146)	.318 (.149)	-.436 (.176)	-.389 (.185)

Construct:	Externalizing					
Big Five Personality: Model:	Zero-Order	Partial	Zero-Order	Partial	Zero-Order	Partial
Covariance:	**Intercept-Intercept**	**Intercept-Intercept**	**Slope1-Linear**	**Slope1-Linear**	**Slope1-Quadratic**	**Slope1-Quadratic**
Openness/Intellect	-.012 (.051)	-.022 (.050)	-.212 (.144)	-.208 (.144)	.214 (.167)	.216 (.170)
Conscientiousness	-.124*(.040)	-.158*(.041)	-.011 (.023)	-.005 (.023)	-.025 (.032)	-.031 (.033)
Extraversion	.127 (.047)	.137 (.046)	-.283*(.092)	-.285*(.091)	.241 (.108)	.245 (.109)
Agreeableness	-.208*(.052)	-.293*(.049)	-.115 (.128)	.026 (.073)	.151 (.155)	-.035 (.089)
Neuroticism	.424*(.047)	.404*(.047)	.253* (.095)	.233 (.095)	-.328*(.111)	-.304(.112)

Notes. Partial correlations control for biological sex. Standard errors are reported in parentheses. Asterisks denote estimates that were significantly different than zero at $p_{(two\text{-}tailed)} < .005$.

S13. Correlations Between the Growth Factors of Big Five Personality and the Residual Variance in the Curves of Latent Psychopathology Factors Controlling for the P-Factor and Self-Reported Biological Sex

Big Five Personality:	Construct: Model: Correlation:	Residual Variance in Internalizing					
		Zero-Order	Partial	Zero-Order	Partial	Zero-Order	Partial
		Intercept-Intercept	**Intercept-Intercept**	**Slope-Slope**	**Slope-Slope**	**Slope-Quadratic**	**Slope-Quadratic**
Openness/Intellect		-.065 (.064)	-.082 (.061)	.242 (.239)	.222 (.223)	-.426 (.333)	-.387 (.290)
Conscientiousness		.086 (.052)	.023 (.051)	-.039 (.040)	.177 (.144)	.047 (.057)	.424 (.245)
Extraversion		-.221*(.061)	-.204*(.060)	.002 (.151)	.008 (.140)	-.199 (.209)	-.188 (.183)
Agreeableness		.035 (.062)	-.012 (.058)	.390 (.217)	.307 (.185)	-.460 (.287)	-.375 (.228)
Neuroticism		.134*(.059)	.117*(.054)	-.124 (.165)	-.137 (.150)	.286 (.234)	.260 (.194)

Big Five Personality:	Construct: Model: Correlation:	Residual Variance in ADHD					
		Zero-Order	Partial	Zero-Order	Partial	Zero-Order	Partial
		Intercept-Intercept	**Intercept-Intercept**	**Slope-Slope**	**Slope-Slope**	**Slope-Quadratic**	**Slope-Quadratic**
Openness/Intellect		.009 (.056)	.027 (.059)	.166 (.164)	.178 (.174)	-	-
Conscientiousness		-.238*(.051)	-.210*(.052)	.158 (.117)	.162 (.111)	-.185 (.132)	-.189 (.119)
Extraversion		.042 (.060)	.029 (.062)	.138 (.109)	.152 (.116)	-	-
Agreeableness		-.054 (.056)	.006 (.057)	.170 (.130)	.172 (.141)	-	-
Neuroticism		-.009 (.051)	-.005 (.052)	-.074 (.096)	-.064 (.103)	-	-

Big Five Personality:	Construct: Model: Correlation:	Residual Variance in Externalizing					
		Zero-Order	Partial	Zero-Order	Partial	Zero-Order	Partial
		Intercept-Intercept	**Intercept-Intercept**	**Slope-Slope**	**Slope-Slope**	**Slope-Quadratic**	**Slope-Quadratic**
Openness/Intellect		.003 (.048)	-.001 (.048)	-.074 (.131)	-.057 (.129)	-	-
Conscientiousness		.016 (.039)	.000 (.039)	-.140 (.084)	-.105 (.081)	.135 (.092)	.102 (.088)
Extraversion		.155*(.048)	.158*(.048)	-.041 (.087)	-.048 (.085)	-	-
Agreeableness		-.145*(.049)	-.173*(.049)	-.131 (.102)	-.068 (.099)	-	-
Neuroticism		.223*(.047)	.217*(.047)	-.043 (.084)	-.028 (.083)	-	-

Notes. Partial correlations control for biological sex. Standard errors are reported in parentheses. Asterisks denote estimates that were significantly different than zero at $p_{(two-tailed)} < .005$.

S14. Correlations Between the Growth Factors of Big Five Personality and the Curves of Latent Psychopathology Factors Before & After Estimating Correlations Between the Residual Variances of Indicators of Psychopathology

		P-Factor					
Big Five Personality:	Construct: Model: Covariance:	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations
		Intercept-Intercept	**Intercept-Intercept**	**Slope-Linear**	**Slope-Linear**	**Slope-Quadratic**	**Slope-Quadratic**
Openness/Intellect		-.018 (.052)	-.019 (.055)	-.241 (.149)	-.272 (.167)	.224 (.161)	.243 (.176)
Conscientiousness		-.162*(041)	-.173*(.043)	-.003 (.023)	-.004 (.025)	-.045 (.030)	-.049 (.032)
Extraversion		.068 (.047)	.068 (.050)	-.290*(.092)	-.323*(.105)	.189 (.100)	.201 (.110)
Agreeableness		-.206*(.050)	-.218*(.053)	.052 (.113)	.067 (.127)	-.027 (.123)	-.039 (.134)
Neuroticism		.458*(.045)	.488*(.047)	.288*(.100)	.327*(.113)	-.322*(.109)	-.351*(.119)
		Internalizing					
Big Five Personality:	Construct: Model: Covariance:	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations
		Intercept-Intercept	**Intercept-Intercept**	**Slope1-Linear**	**Slope1-Linear**	**Slope1-Quadratic**	**Slope1-Quadratic**
Openness/Intellect		-.039 (.053)	-.047 (.058)	-.081 (.163)	-.096 (.175)	-.006 (.188)	.005 (.193)
Conscientiousness		-.085 (.042)	-.100 (.045)	-.021 (.026)	-.022 (.028)	-.017 (.034)	-.017 (.0350
Extraversion		-.041 (.049)	-.046 (.053)	-.226 (.104)	-.254 (.113)	.061 (.123)	.074 (.126)
Agreeableness		-.151*(.053)	-.162*(.057)	.220 (.144)	.230 (.156)	-.222 (.163)	-.223 (.168)
Neuroticism		.430*(.049)	.462*(.052)	.160 (.112)	.194 (.121)	-.124 (.129)	-.153 (.132)
		ADHD					
Big Five Personality:	Construct: Model: Covariance:	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations
		Intercept-Intercept	**Intercept-Intercept**	**Slope1-Linear**	**Slope1-Linear**	**Slope1-Quadratic**	**Slope1-Quadratic**
Openness/Intellect		.007 (.054)	.020 (.061)	-.408 (.232)	-.835 (.748)	.482 (.273)	.953 (.891)
Conscientiousness		-.246*(.045)	-.269*(.051)	.030 (.029)	.061 (.065)	-.089 (.042)	-.171 (.143)
Extraversion		.080 (.052)	.112 (.059)	-.249 (.133)	-.413 (.389)	.188 (.150)	.281 (.353)
Agreeableness		-.199*(.051)	-.227*(.058)	.103 (.155)	.209 (.328)	-.035 (.174)	-.068 (.329)
Neuroticism		.371*(.050)	.417*(.058)	.363 (.146)	.665 (.505)	-.436 (.176)	-.772 (.625)
		Externalizing					
Big Five Personality:	Construct: Model: Covariance:	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations
		Intercept-Intercept	**Intercept-Intercept**	**Slope1-Linear**	**Slope1-Linear**	**Slope1-Quadratic**	**Slope1-Quadratic**
Openness/Intellect		-.012 (.051)	-.013 (.056)	-.212 (.144)	-.270 (.182)	.214 (.167)	.274 (.216)
Conscientiousness		-.124*(.040)	-.138*(.044)	-.011 (.023)	-.014 (.029)	-.025 (.032)	-.034 (.041)
Extraversion		.127 (.047)	.144 (.052)	-.283*(.092)	-.346*(.122)	.241 (.108)	.290 (.145)
Agreeableness		-.208*(.052)	-.236*(.058)	-.115 (.128)	-.144 (.162)	.151 (.155)	.191 (.198)
Neuroticism		.424*(.047)	.479*(.052)	.253* (.095)	.321 (.120)	-.328*(.111)	-.411*(.143)

Notes. Models titled "Residual Correlations" include cross-time correlations between the residuals of indicators of latent psychopathology factors. Standard errors are reported in parentheses. Asterisks denote estimates that were significantly different than zero at $p_{(two\text{-}tailed)} < .005$.

S15. Correlations Between the Growth Factors of Big Five Personality and the Curves of Latent Psychopathology Factors Before & After Estimating Correlations Between the Residual Variances of Indicators of Psychopathology

Big Five Personality:	Construct: Model: Correlation:	Residual Variance in Internalizing					
		Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations
		Intercept-Intercept	**Intercept-Intercept**	**Slope-Slope**	**Slope-Slope**	**Slope-Quadratic**	**Slope-Quadratic**
Openness/Intellect		-.065 (.064)	-.081 (.074)	.242 (.239)	.237 (.200)	-.426 (.333)	-.354 (.240)
Conscientiousness		.086 (.052)	.124 (.063)	-.039 (.040)	-.032 (.032)	.047 (.057)	.043 (.042)
Extraversion		-.221*(.061)	-.275*(.075)	.002 (.151)	.030 (.128)	-.199 (.209)	-.166 (.154)
Agreeableness		.035 (.062)	.078 (.073)	.390 (.217)	.313 (.166)	-.460 (.287)	-.340 (.188)
Neuroticism		.134 (.059)	.079 (.068)	-.124 (.165)	-.135 (.141)	.286 (.234)	.250 (.171)

Big Five Personality:	Construct: Model: Correlation:	Residual Variance in ADHD					
		Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations
		Intercept-Intercept	**Intercept-Intercept**	**Slope-Slope**	**Slope-Slope**	**Slope-Quadratic**	**Slope-Quadratic**
Openness/Intellect		.009 (.056)	.016 (.071)	.166 (.164)	.188 (.162)	-	-
Conscientiousness		-.238*(.051)	-.280*(.068)	.158 (.117)	.158 (.114)	-.185 (.132)	-.174 (.128)
Extraversion		.042 (.060)	.051 (.075)	.138 (.109)	.180 (.110)	-	-
Agreeableness		-.054 (.056)	-.042 (.071)	.170 (.130)	.160 (.127)	-	-
Neuroticism		-.009 (.051)	-.090 (.066)	-.074 (.096)	-.092 (.095)	-	-

Big Five Personality:	Construct: Model: Correlation:	Residual Variance in Externalizing					
		Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations	Zero-Order	Residuals Correlations
		Intercept-Intercept	**Intercept-Intercept**	**Slope-Slope**	**Slope-Slope**	**Slope-Quadratic**	**Slope-Quadratic**
Openness/Intellect		.003 (.048)	.012 (.056)	-.074 (.131)	-.089 (.151)	-	-
Conscientiousness		.016 (.039)	.044 (.047)	-.140 (.084)	-.173 (.099)	.135 (.092)	.170 (.108)
Extraversion		.155*(.048)	.184*(.056)	-.041 (.087)	-.030 (.101)	-	-
Agreeableness		-.145*(.049)	-.153*(.057)	-.131 (.102)	-.165 (.118)	-	-
Neuroticism		.223*(.047)	.228*(.055)	-.043 (.084)	-.043 (.096)	-	-

Notes. Models titled "Residual Correlations" include cross-time correlations between the residuals of indicators of latent psychopathology factors. Standard errors are reported in parentheses. Asterisks denote estimates that were significantly different than zero at $p_{(two\text{-}tailed)} < .005$.