SUMMATIVE ASSESSMENT COVERSHEET

Undergraduate Courses

How to use this coversheet:

- Complete all shaded areas in section 1
- Read the statement in section 2
- Complete the tick box in section 3 if appropriate
- Complete the checklist in section 4
- Insert this coversheet as the first page of your assessment submission; you should add it before converting the file to PDF so that you are only submitting one document to Moodle.

Please be aware that these coversheets are unique to each assessment and should not be reused.

1. BASIC DETAILS

Course: PB230 - Intermediate Statistics and Research Methods for Psychological and Behavioural Science
Assessment: Secondary Data Analysis
Summative Component: This assessment counts for 50% of the overall mark in PB230

5 digit candidate number for 2021/22 *This can be found on LSE for You.* *Please note that candidate numbers change each year.*	2	6	1	4	5
Tasks Completed	1, 2, 3				
Word Count:	3,500				

2. PLAGIARISM

All work submitted as part of the requirements for any assessment of the school (e.g. exams, essays, dissertations, and any other work, including computer programs) must be expressed in your own words and incorporate your own ideas and judgements. Plagiarism must be avoided in all such work. Plagiarism can involve the presentation of another person's thoughts or words as if they were your own. However, please note that plagiarism also includes self-plagiarism, which is where you, as the author re-use your own previously submitted work or data in a 'new' written piece of work without letting the reader know that this material has appeared elsewhere. The definition of 'your own work; also includes work produced by collaboration or group-work unless expressly permitted by the Department. For the avoidance of doubt, it is acceptable to reuse work that was submitted as part of a 'draft' formative assessment. Plagiarism does not need to be deliberate or intentional for it to constitute an assessment offence.

Having read and understood the LSE Regulations on Assessment Offences: Plagiarism I hereby confirm that the work contained in this submission is my own apart properly referenced quotations. I understand that my work will be stored in the Turnitin Repository for the purposes of plagiarism detection.

3. SHARING YOUR WORK WITH FUTURE COHORTS

We would like to share excellent examples of work with future cohorts.
If you are happy for us to do this please check this box. ☒

4. SUBMISSION CHECKLIST

☒ I have added this coversheet as the first page of my submission.
☒ I have added my 2021/22 candidate number to the box in Section 1.
☒ I have noted the tasks I have completed in Section 1.
☒ I have included an accurate word count for my write up in Section 1.
☒ I have read and understood the declaration I am making in Section 2.
☒ I will save my work as a PDF File.
☒ I will name my PDF File using my candidate number ONLY.
☒ I will remove any information which may identify me, or my fellow group members (including any 'Author' information in the metadata).

The role of Self-Compassion in mediating the relationship between Openness to Experience and college student's Subjective Well-being

Research Questions (RQ) and Hypotheses

RQ_1: Does the hypothesized mediation model fit the data?

RQ_2: Is the relationship between openness and subjective well-being (SWB) fully mediated by self-compassion in college students?

Null Hypothesis: The relationship between openness and SWB is not fully mediated by self-compassion in college students.

Alternate Hypothesis: The relationship between openness and SWB is fully mediated by self-compassion in college students.

Data Analysis Method and Justification

A mediation model was hypothesized by the researchers (*Figure 1*) as they were interested in exploring whether self-compassion fully mediated the positive relationship between openness to experience and SWB. They are multiple ways to test such mediation models including multiple regression models and path analysis.

Firstly, the present hypothesized model was a full mediation model, meaning the relationship between openness and SWB depended entirely on the mediator, as opposed to a partial mediation model, where they would also be a direct relationship between openness and SWB. Secondly, this implies the model needed to test two simultaneous outcomes predicted by two different independent variables: whether openness predicted outcome 1 (self-compassion), and whether self-compassion predicted outcome 2 (SWB).

Hence, because of these two features, path analysis was the best option as it i) allows for parsimony and the use of Occam's Razor by leaving the direct relationship between openness and SWB out and adjudicating the fit of overidentified models (Blumer et al., 1987); ii) is a multivariate analysis thus can include multiple outcomes and test the paths between all outcomes simultaneously (Keith, 2019).

Therefore, as the hypothesized relationships were specified between all variables, path analysis traced in conformity with Wright's rules (1934) enabled to infer potential causality between variables and subsequently, inform theory. Then, certain features of path analysis were specified to enhance the validity and reliability of the analysis.

Maximum likelihood estimation with 5,000 bootstrapped resamples decreased the Type 1 Error probability (Preacher & Hayes, 2008) and enabled for removal of the normality of sampling distribution condition (Nevitt & Hancock, 2001).

CFI, SRMR, RMSEA and TLI were used to assess model fit according to the following metrics: CFI$\geq$.90, SRMR $\leq$.10, RMSEA $\leq$.10, and TLI$>$.90 (Marsh et al., 2004); the latter correcting for parsimony by providing estimates independent of sample size.

Analyses were conducted in R using the *laavan* package (Rosseel, 2012).

Preliminary Analyses

Data Screening.

Prior to running the primary analysis, the data were screened to check the conformity of the values within the range of the Likert Scale (1-7). Values in *swb1* and *open2* were outside the appropriate range, thus replaced with missing values.

Missing Data.

Then, the data were screened for missing values. There were 157 complete cases and 54 cases with incomplete data. The missing data were missing completely at random (MCAR): the probability of the pattern of missing values diverging from randomness was greater than .05 ($p = .06$) (Jamshidian et al., 2014).

Hence, the appropriate strategy for the small amount of missing data (i.e., 2.5%) and highly correlated items was applied: listwise deletion of participants with more than 10% of missing data (Cole, 2008) followed by Scale Mean Imputation (SMI) (replacing the missing values using the mean of each participant's available non-missing items from the relevant subscale) (Schafer and Graham, 2002). This resulted in the removal of 2 participants.

Outliers.

Participants were identified as univariates outliers if they had standardized z-scores larger than 3.29 ($p < .001$); and as multivariate outliers when Mahalanobi's distances were greater than $\chi 2$ (3) = 16.27 ($p < .001$) (Tabachnick & Fidell, 2007). Two more participants were removed (2 univariate and 0 multivariate outliers).

Descriptive Statistics.

This yielded a final sample of 207 participants, whose score distributions on the three variables of interests (openness, self-compassion and SWB) can be seen in *Table 1* and *Figures 2-4*. These data were approximately univariate normal (average absolute skew = -.25; average absolute kurtosis = -.61).

Model Assumptions.

Lastly, normality and independence of residuals, equality of variance, absence of influential observations multicollinearity and homoscedasticity were verified using the *performance* R package (Lüdecke et al., 2021). The literature on the subject matter enabled validation of the theoretical assumptions of the model (Neff et al., 2007; Saricaoglu & Arslan, 2013).

Primary Analyses

Correlations.

All the variables were significantly positively correlated ($p<.001$) (*Table 2*): openness and self-compassion had the largest correlation (.62), while SWB and self-compassion had a moderate correlation (.38) and openness and SWB had a small correlation (.29) (Cohen, 1988; 1992).

Over-identified path analysis.

The path model displayed adequate fit to the data (CFI = 1.00, TLI = 1.01, RMSEA = .00, SRMR = .02) (*Table 3*).

The path coefficient between self-compassion and SWB was significant (b = .44, β = .34, 95% bootstrap CI: .25, .63). Likewise, the effect between openness and self-compassion was significant (b = .40, β = .61, 95% bootstrap CI: .33, .48). Determining whether the indirect mediating effect of openness on SWB through self-compassion (ab = .18) was statistically significant, the 95% confidence interval derived from 5,000 bootstrapped resamples did not include zero (95% CI: .10, .26) (*Table 4*).

Examination of R2 revealed approximately 11% of the variance in SWB and 37% of the variance in self-compassion was explained by the path model.

A path analysis plot visualizes all these effects (*Figure 5*).

Interpretations

Answering RQ_1, the hypothesized mediation model suitably fitted the data (CFI = 1.00, TLI = 1.01, RMSEA = .00, SRMR = .02) (*Table 3*).

Answering RQ_2, consistent with the alternate hypothesis, the positive relationship between openness and SWB was fully mediated by self-compassion in the finalized sample of 207 students (ab = .18, β = .21, 95% CI: .10, .26) (*Table 4*). The 5,000 bootstrapped confidence intervals did not include 0 thus generalizability of this causal relationship to the population of college students, in general, can be inferred while accounting for potential limitations of the study design.

As such, students that scored higher on self-reported openness to experience reported higher SWB via their higher self-compassion.

References

Blumer, A., Ehrenfeucht, A., Haussler, D., & Warmuth, M. K. (1987). Occam's Razor. *Information Processing Letters*, *24*(6), 377–380. https://doi.org/10.1016/0020-0190(87)90114-1

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Routledge. ISBN 978-1-134-74270-7.

Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, *1*(3), 98-101.

Cole, J. (2008). How to deal with missing data conceptual overview and details for implementing two modern methods. In Osborne, J. (Ed.), *Best practices in quantitative methods* (pp. 214-238). SAGE Publications, Inc., https://www.doi.org/10.4135/9781412995627

Jamshidian, M., Jalal, S., & Jansen, C. (2014). MissMech: An R Package for Testing Homoscedasticity, Multivariate Normality, and Missing Completely at Random (MCAR). *Journal of Statistical Software, 56*(6), 1 - 31. doi:http://dx.doi.org/10.18637/jss.v056.i06

Keith, T. (2019). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*.

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P. & Makowski, D. (2021). Assessment, Testing and Comparison of Statistical Models using R. *Journal of Open Source Software*, 6(59), 3112. doi: 10.31234/osf.io/vtq8f.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320-341.

Neff, K. D., Rude, S. S., & Kirkpatrick, K. L. (2007). An examination of self-compassion in relation to positive psychological functioning and personality traits. *Journal of research in personality*, 41(4), 908-916.

Nevitt, J., & Hancock, G. (2001). Performance of Bootstrapping Approaches to Model Test Statistics and Parameter Standard Error Estimation in Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 353– 377. doi: 10.1207/s15328007sem0803_2

Osborne, J. W. (2013*). Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data*. Sage. Chicago

Preacher, K. J., & Hayes, A. F. (2008). *Assessing mediation in communication research* (pp. 13-54). London: The Sage sourcebook of advanced data analysis methods for communication research. Chicago

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, *Foundation for Open Access Statistics,* vol. 48(i02). doi: http://hdl.handle.net/10.18637/jss.v048.i02

Saricaoglu, H., & Arslan, C. (2013). An Investigation into Psychological Well-Being Levels of Higher Education Students with Respect to Personality Traits and Self-Compassion. *Educational Sciences: Theory and Practice*, 13(4), 2097-2104. Chicago

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147–177. https://doi.org/10.1037/1082- 989X.7.2.147

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Boston: Pearson Education.

Wright, S. (1934). The Method of Path Coefficients. *The Annals of Mathematical Statistics*, *5*(3), 161–215. https://doi.org/10.1214/aoms/1177732676

Tables.

Table 1.

Descriptive Statistics for Openness, Self-Compassion and SWB within the finalized sample population of college students (n=207)

	Min-Max	*M*	*SD*	*Med.*	*Skewness*	*Kurtosis*
Openness	1.75-7.00	5.34	1.21	5.50	-0.68	-0.10
Self-Compassion	3.50-7.00	5.57	0.80	5.75	-0.43	-0.33
SWB	2.00-7.00	5.40	1.05	5.33	-0.18	-0.48

Note. M, SD and *Med.* are used to represent mean, standard deviation and median, respectively.

Table 2.

Summary of Pearson correlations between Openness, Self-compassion and SWB within the finalized sample population of college students (n=207)

	Openness	Self-Compassion	SWB
Openness	-		
Self-Compassion	0.62 **	-	
SWB	0.29 **	0.38 **	-

Note. ** indicates $p < .001$.

Table 3.

Fit Indices for the Over-Identified Full Mediation Model between Openness and SWB via Self-Compassion within the finalized sample population of college students (n=207)

	Estimate	90% CI [LL, UL]
Comparative Fit Index (CFI)	1.00	-
Tucker Lewis Index (TLI)	1.01	-
Root Mean Square Error of Approximation (RMSEA)	0.00	[0.00, 0.16]
Standardized Root Mean Square Residual (SRMR)	0.02	-

Note. LL and UL represent the lower limit and upper limit for the 90% confidence intervals.

Table 4.

Path coefficients for the Over-Identified Full Mediation Model between Openness and SWB via Self-Compassion within the finalized sample population of college students (n=207)

	Label	Model Estimates		Standardized Estimates	*SE*
		Estimate	95% CI [LL, UL]		
SWB ~ Self-Compassion	b	0.44**	[0.25, 0.62]	0.34	0.10
Self-Compassion ~ Openness	a	0.40**	[0.33, 0.48]	0.61	0.04
Indirect effect (a*b)	ab	0.18**	[0.10, 0.26]	0.21	0.04

Note. LL and UL represent the lower limit and upper limit for the 95% confidence intervals; *SE* represents the standard error;
* indicates $p < .05$. ** indicates $p < .001$.

Figures.

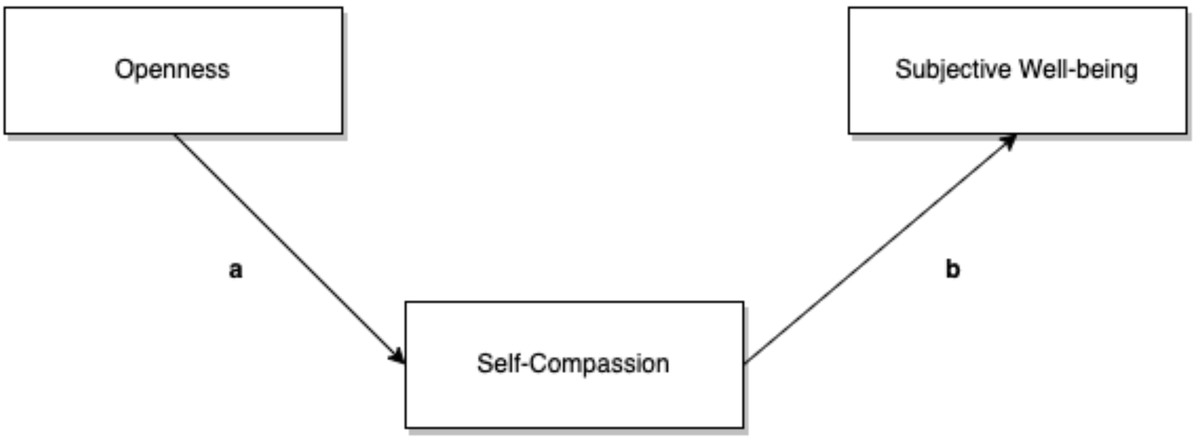

Figure 1.
Hypothesised Over-Identified Full Mediation Model between Openness and SWB via Self-Compassion for college students

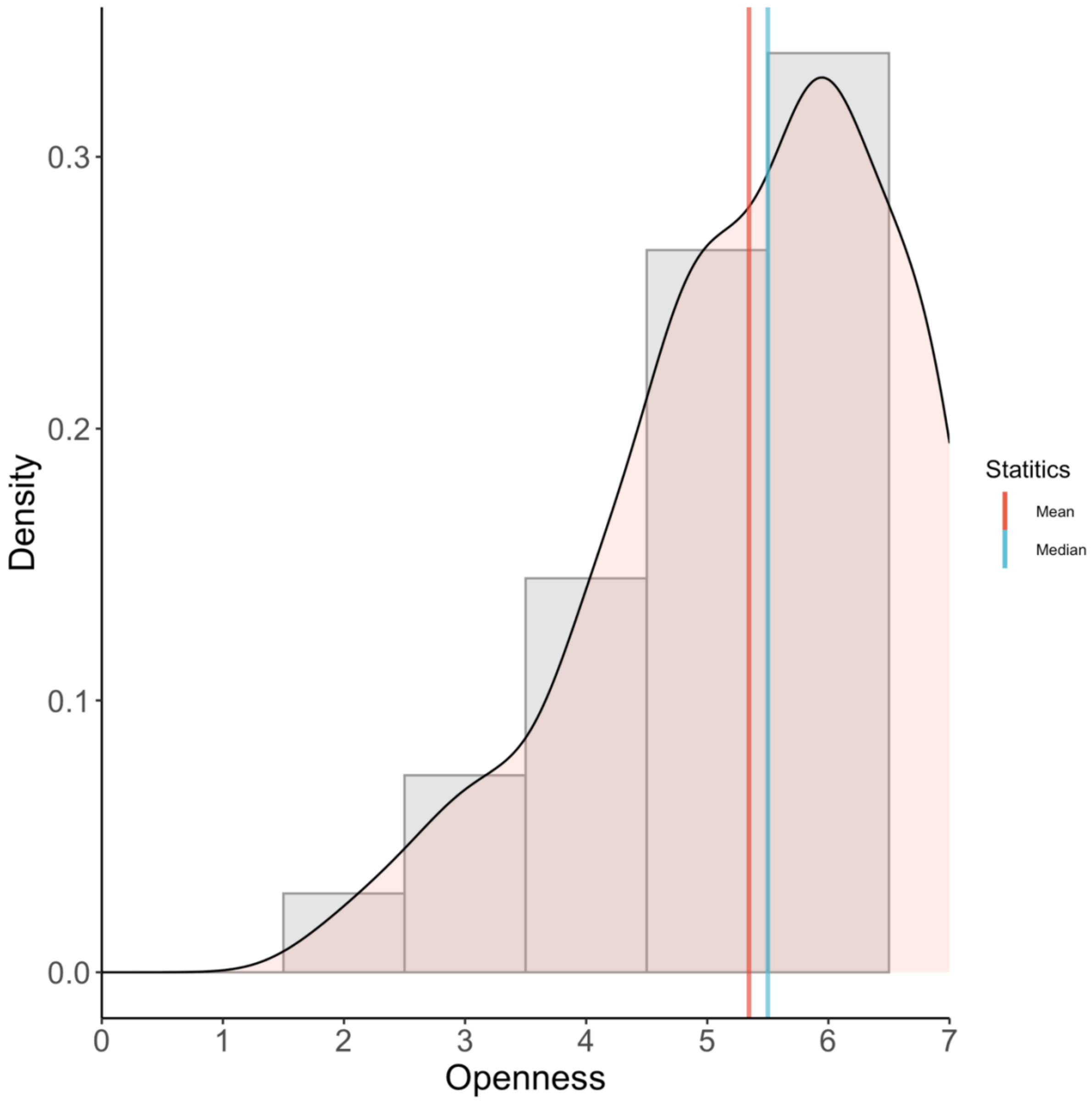

Figure 2.
Distribution of Openness within the finalized sample population of college students (n=207)

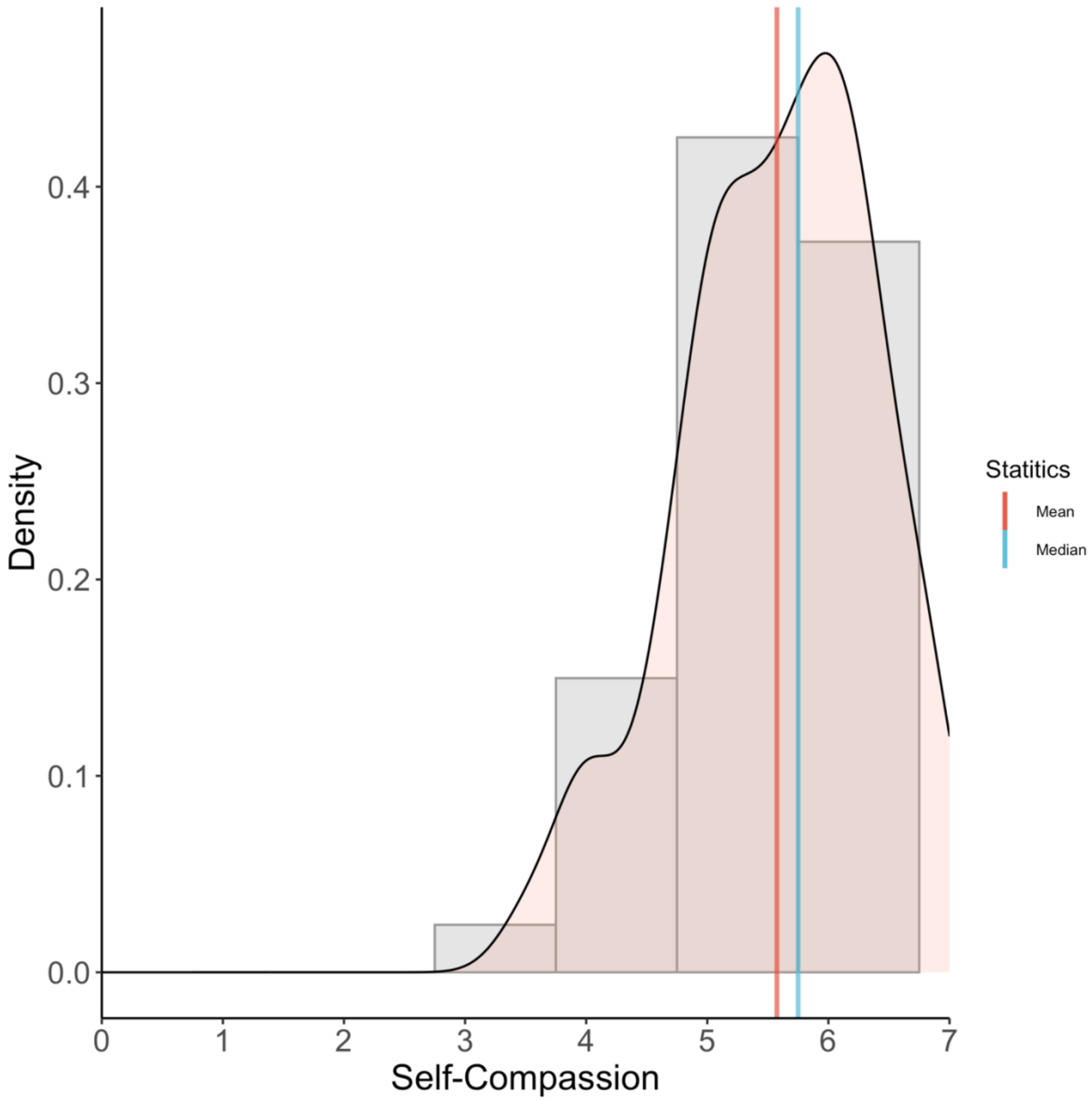

Figure 3.
Distribution of Self-Compassion within the finalized sample population of college students (n=207)

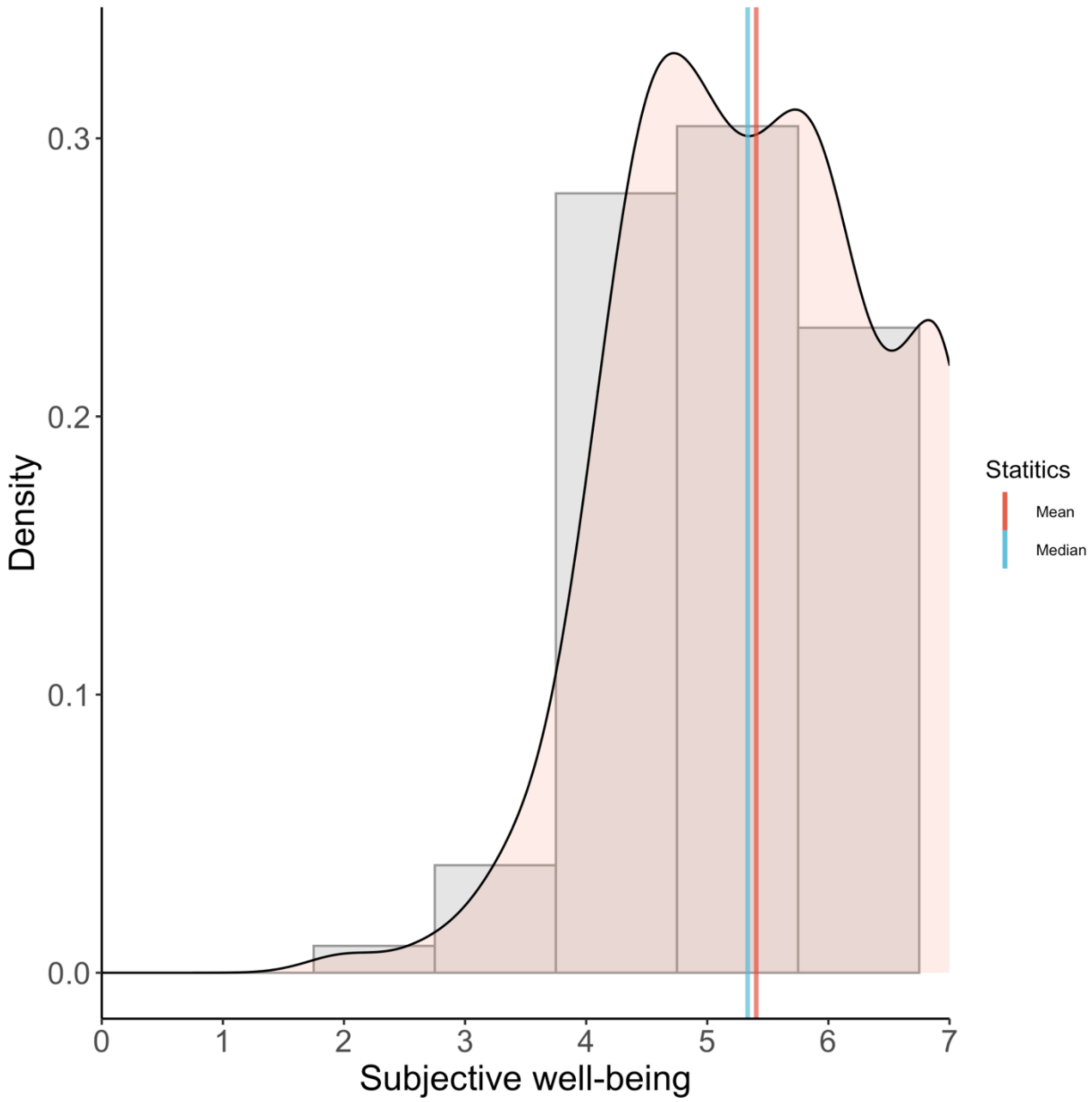

Figure 4.
Distribution of Subjective Well-being within the finalized sample population of college students (n=207)

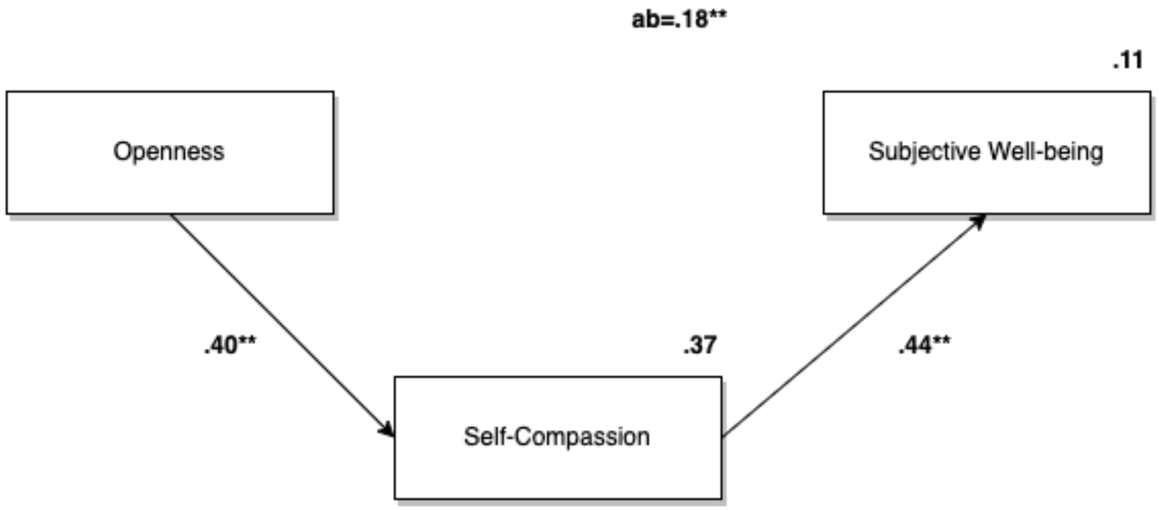

Figure 5.
Path Coefficients and variance explained (R2) by the Over-Identified Full Mediation Model between Openness and SWB via Self-Compassion fitted within the finalized sample population of college students (n=207)
Note. Fit Metrics: CFI = 1.00; TLI = 1.01; SRMR = .02; RMSEA = .00 (90%CI: .00 to .16);
** indicates p < .001.

The role of Rumination in mediating the relationship between academic's Exhaustion and Perfectionism

Research Questions and Hypotheses

RQ_1: Does the hypothesized mediation model fit the data?

RQ_2: What is the size, direction, and statistical significance of the paths in the hypothesized mediation model?

RQ_3: Does rumination fully mediate the relationship between perfectionism and exhaustion?

Null Hypothesis: Rumination doesn’t mediate fully the relationship between perfectionism and exhaustion.

Alternate Hypothesis: Rumination fully mediates the relationship between perfectionism and exhaustion.

Data Analysis Method and Justification

A mediation model was hypothesized by the researchers as they were interested in exploring whether rumination fully mediated the positive relationship between perfectionism and exhaustion in academics. The three latent variables were measured through several corresponding self-report items.

They were three relevant components to this analysis: i) testing the fit of the underlying rumination, perfectionism, and exhaustion unobserved constructs with the measurement items used, ii) estimating the relationship between these constructs while controlling for potential confounding variables, and iii) inferring causality between all three constructs simultaneously.

Structural equation modelling (SEM) combined all the required components of i) confirmatory factor analysis (CFA), ii) multiple regression and iii) path analysis, using the variables' variances and covariances (Loehlin 2017).

A full latent variable SEM (*Figure 2*) includes the outer model measuring adequacy of the factor loading and the inner structural model exploring the relationships between those factors; inner model and outer model will be tested subsequently (Anderson & Gerbin, 1988).

The full SEM's main asset is testing the relationships in the inner model in the absence of any measurement error, absence which is assumed for any regression. This is essential in psychology and for the testing of constructs such as rumination, exhaustion and perfectionism which are unobservable and not directly measurable (Flora and Flake, 2017).

Then, certain features of SEM were specified to enhance the validity and reliability of the analysis.

Maximum likelihood estimation with 5,000 bootstrapped resamples and the use of Wright's parsimonious tracing rules (1934) warranted finding the best fitting parameters for the data.

Strong cut-off factor loadings > .60 and composite reliability >.70 were the metrics used for the CFA (Garson, 2013; Nunnally, 1978).

CFI, SRMR, RMSEA and TLI were used to assess inner and full model fit according to the following metrics: CFI≥.90, SRMR ≤.10, RMSEA ≤.10, and TLI>.90 (Marsh et al., 2004).

Analyses were conducted in R using *lavaan* package (Rosseel, 2012) fitting the ordinal structure of the data measured on Likert scales.

Preliminary Analyses

Data Screening.

Preliminary analyses started by screening conformity of the range of values within the Likert Scale used for surveying (1-7). All the values were within the appropriate range.

Missing Data.

Screening for missing values followed. There were 208 complete cases and 6 cases with incomplete data. Jamshidian and colleagues' (2014) test of the pattern of missing data revealed it was MCAR ($p = .06$).

Hence, appropriately for the little amount of missing data (i.e., 0.4%) and the highly correlated latent constructs items would load on (perfectionism $\alpha = .81$; exhaustion $\alpha = .86$; rumination $\alpha = .87$); participants with more than 10% of missing data were deleted listwise ($n=1$) (Cole, 2008), and the missing values were replaced using SMI (Schafer and Graham, 2002).

Outliers.

No outliers were identified on the item-level: neither univariate outliers with standardized z-scores larger than 3.29 ($p < .001$); nor multivariate outliers with Mahalanobi's distances greater than $\chi2$ (10) = 29.59 ($p < .001$) (Tabachnick & Fidell, 2007).

Identifying outliers on the construct level was not deemed relevant as items will subsequently inform the latent variables.

Descriptive Statistics.

The descriptive statistics of the items for the final sample of 213 participants can be visualized in *Table 1*. Overall, an approximately univariate normal distribution was observed (average absolute skew = .53; average absolute kurtosis = .10).

Model Assumptions.

Finally, independence of residuals, equality of variance, absence of influential observations multicollinearity and homoscedasticity were verified using the *performance* R package (Lüdecke et al., 2021). A moderate violation of the assumption of multivariate normality was observed, but analyses nonetheless persisted as CFA carried out with maximum likelihood estimation overrides this issue (Yang and Liang, 2013).

Moreover, the literature on the subject matter enabled validation of the theoretical assumptions of the model (Flaxman et al., 2012).

Primary Analyses

CFA.

The measurement model consisting of three latent variables exhibited an acceptable fit to the data: $\chi^2 = 59.26$ (32), $p < .05$; TLI = .96; CFI = .97; SRMR = .05; RMSEA = .06 (90% CI = .04 to .09) (*Table 2*). Three perfectionism items, three rumination items and four exhaustion items were respectively used as the measured variables for the perfectionism, rumination and exhaustion factors.

Standardized factor loadings and Composite Reliability

All standardised factor loadings for the measured variables on their latent factors were significant (perfectionism β range = 0.71 to 0.80; rumination β range = .83 to .90; exhaustion β range = .72 to .81). Furthermore, each of these factors demonstrated satisfactory composite reliability (perfectionism $\rho = .92$; rumination $\rho = .89$; exhaustion $\rho = .85$).

Error Free Correlations

All error-free correlations between latent factors were positive, statistically significant, and ranged in magnitude from moderate to large according to conventional effect size criteria (i.e., small $\geq .10$, moderate $\geq .30$, large $\geq .50$) (Cohen, 1988; 1992) (*Table 3*).

These results allowed advancement to the full structural equation model.

SEM.

A full mediation model including one indirect path from perfectionism to exhaustion via rumination was chosen based on fit indexes suggesting this model possessed an acceptable fit to the data: $\chi 2$= 60.17 (33); TLI = .96; CFI = .97; SRMR = .05; RMSEA = .06 [90%CI: .04, .09] (*Table 4*).

Perfectionism positively predicted rumination (b = .55, β = .57, 95% CI = .37,.75). In turn, rumination positively predicted exhaustion (b = .21, β = .32, 95% CI = .09,.35). Both path coefficients were statistically significant.

This model successfully explained 33% of the variance in rumination and 10% of the variance in exhaustion.

The indirect pathway from perfectionism to exhaustion via rumination was also significant as its 95% percentile confidence intervals derived from 5,000 bootstrap iterations did not include 0 (ab = .12, β = .18, 95% CI= .04, .21) (*Table 5*).

The path analysis plot conjures up these results (*Figure 3*).

Interpretations

Answering RQ_1, the CFA revealed the measurement model was adequate: $\chi^2 = 59.26$ (32), $p < .05$; TLI = .96; CFI = .97; SRMR = .05; RMSEA = .06 (90% CI = .04 to .09) (*Table 2)*; and the full SEM exhibited the hypothesized mediation model suitably fitted the data $\chi 2=$ 60.17 (33); TLI = .96; CFI = .97; SRMR = .05; RMSEA = .06 [90%CI: .04, .09] (*Table 4*).

Answering RQ_2, perfectionism largely positively predicted rumination ($b = .55$, $\beta = .57$, 95% CI = .37,.75). In turn, rumination positively predicted exhaustion in a small manner ($b = .21$, $\beta = .32$, 95% CI = .09,.35). Moreover, the indirect relationship from perfectionism to exhaustion via rumination was also positive and small ($ab = .12$, $\beta = .18$, 95% CI= .04, .21) (*Table 5*). All three paths' confidence intervals did not include 0 indicating statistical significance.

This latter magnitude and significance of the indirect path allowed to answer RQ_3: rumination significantly mediated the relationship between perfectionism and exhaustion in the finalized sample of 213 academics. The null hypothesis was rejected and the maximum likelihood estimation simulating the general population of academics allowed for generalizability within the limitations of the study design.

British academics with higher perfectionism scores will report higher exhaustion because of their tendencies to ruminate more.

References

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, *103*(3), 411.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Routledge. ISBN 978-1-134-74270-7.

Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, *1*(3), 98-101.

Cole, J. (2008). How to deal with missing data conceptual overview and details for implementing two modern methods. In Osborne, J. (Ed.), *Best practices in quantitative methods* (pp. 214-238). SAGE Publications, Inc., https://www.doi.org/10.4135/9781412995627

Flaxman, P. E., Ménard, J., Bond, F. W., & Kinman, G. (2012). Academics' experiences of a respite from work: Effects of self-critical perfectionism and perseverative cognition on postrespite well-being. *Journal of Applied Psychology, 97*(4), 854–865. https://doi.org/10.1037/a0028055

Flora, D. B. & Flake, J. K. (2017). The Purpose and Practice of Exploratory and Confirmatory Factor Analysis in Psychological Research. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement, 49* (2), 78-88. doi: 10.1037/cbs0000069.

Garson, G. D. (2013). *Factor Analysis*. Asheboro, NC: Statistical Associates Publishers.

Jamshidian, M., Jalal, S., & Jansen, C. (2014). MissMech: An R Package for Testing Homoscedasticity, Multivariate Normality, and Missing Completely at Random (MCAR). *Journal of Statistical Software, 56*(6), 1 - 31. doi:http://dx.doi.org/10.18637/jss.v056.i06

Loehlin, John C. (2017) Latent variable models : an introduction to factor, path, and structural equation analysis, null Routledge

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P. & Makowski, D. (2021). Assessment, Testing and Comparison of Statistical Models using R. *Journal of Open Source Software*, 6(59), 3112. doi: 10.31234/osf.io/vtq8f.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320-341.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed. ed.). New York: McGraw-Hill.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. In *Journal of Statistical Software* (Vol. 48, Issue 2, pp. 1–36). http://www.jstatsoft.org/v48/i02/

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147–177. https://doi.org/10.1037/1082- 989X.7.2.147

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Boston: Pearson Education.

Wright, S. (1934). The Method of Path Coefficients. *The Annals of Mathematical Statistics*, *5*(3), 161–215. https://doi.org/10.1214/aoms/1177732676

Yang, Y. & Liang, X. (2013). Confirmatory factor analysis under violations of distributional and structural assumptions. *Int. J. Quantitative Research in Education*, Vol. 1, No. 1, pp.61–84. doi:10.1504/ijqre.2013.055642

Tables.

Table 1.

Descriptive Statistics of the measured items underlying the respective constructs of Perfectionism, Rumination and Exhaustion within the finalized sample population of British academics (n=213)

	Min-Max	*M*	*SD*	*Med.*	*Skewness*	*Kurtosis*
Perfectionism						
p1	1.00-7.00	2.99	1.62	3.00	0.53	-0.62
p2	1.00-7.00	2.51	1.49	2.00	0.91	-0.01
p3	1.00-7.00	2.46	1.54	2.00	0.89	-0.08
Rumination						
r1	1.00-7.00	2.62	1.35	2.00	0.92	0.56
r2	1.00-6.00	2.81	1.33	3.00	0.31	-0.86
r3	1.00-7.00	3.15	1.32	3.00	0.21	-0.71
Exhaustion						
e1	1.00-5.00	2.02	1.05	2.00	0.95	0.19
e2	1.00-5.00	1.97	1.02	2.00	0.81	-0.14
e3	1.00-5.00	2.15	1.11	2.00	0.73	-0.22
e4	1.00-5.00	2.00	1.09	2.00	0.92	0.07

Note. M, SD and *Med.* are used to represent mean, standard deviation and median, respectively.

Table 2.

Fit Indices for the Confirmatory Factor Analysis within the finalized sample population of British academics (n=213)

	Estimate	90% CI [LL, UL]
Comparative Fit Index (CFI)	0.97	-
Tucker Lewis Index (TLI)	0.96	-
Root Mean Square Error of Approximation (RMSEA)	0.06	[0.03, 0.09]
Standardized Root Mean Square Residual (SRMR)	0.05	-

Note. LL and UL represent the lower limit and upper limit for the 90% confidence intervals.

Table 3.

Error-free Correlations and Composite Reliabilities for Latent Variables within the finalized sample population of British academics (n=213)

	Perfectionism	Rumination	Exhaustion
Perfectionism	-		
Rumination	0.58**	-	
Exhaustion	0.12	0.33**	-
Composite reliability (ϱ)	0.85	0.89	0.91

Note. * indicates $p < .05$. ** indicates $p < .001$.

Table 4.

Fit Indices for the Full Latent Variable Structural Equation Model between Perfectionism and Exhaustion via Rumination within the finalized sample population of British academics (n=213)

	Estimate	90% CI [LL, UL]
Comparative Fit Index (CFI)	0.97	-
Tucker Lewis Index (TLI)	0.96	-
Root Mean Square Error of Approximation (RMSEA)	0.06	[0.03, 0.09]
Standardized Root Mean Square Residual (SRMR)	0.05	-

Note. LL and UL represent the lower limit and upper limit for the 90% confidence intervals.

Table 5.

Path coefficients for the Full Latent Variable Structural Equation Model between Perfectionism and Exhaustion via Rumination within the finalized sample population of British academics (n=213)

	Label	Model Estimates		Standardized Estimates	*SE*
		Estimate	95% CI [LL, UL]		
Exhaustion ~ Rumination	b	0.21*	[0.09, 0.35]	0.32	0.07
Rumination ~ Perfectionism	a	0.55**	[0.37, 0.76]	0.57	0.10
Indirect effect (a*b)	ab	0.12*	[0.04, 0.21]	0.18	0.04

Note. LL and UL represent the lower limit and upper limit for the 95% confidence intervals; *SE* represents the standard error;
* indicates $p < .05$. ** indicates $p < .001$.

Figures.

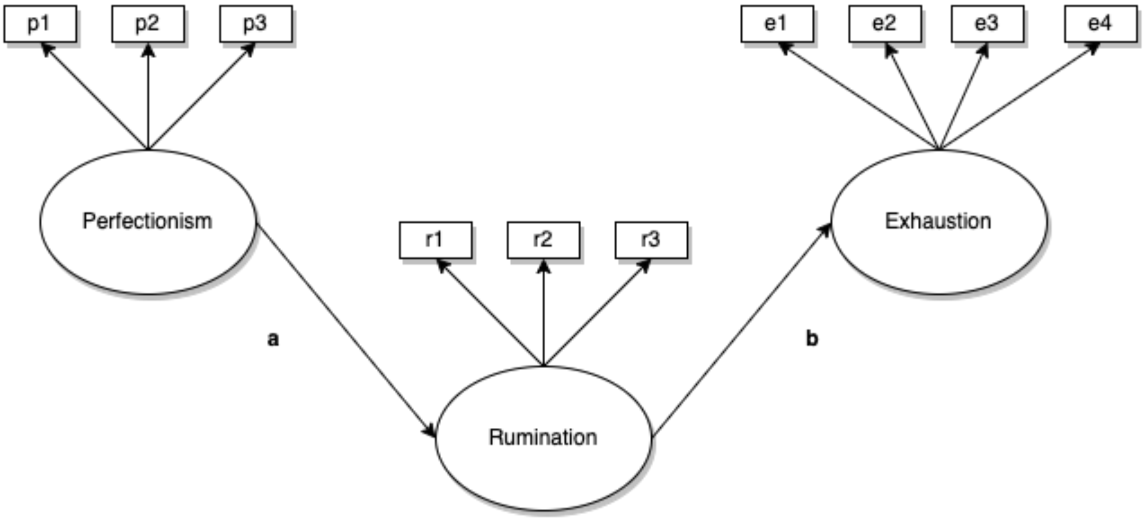

Figure 1.
Hypothesised Full Structural Equation Model between Perfectionism and Exhaustion via Rumination for British academics

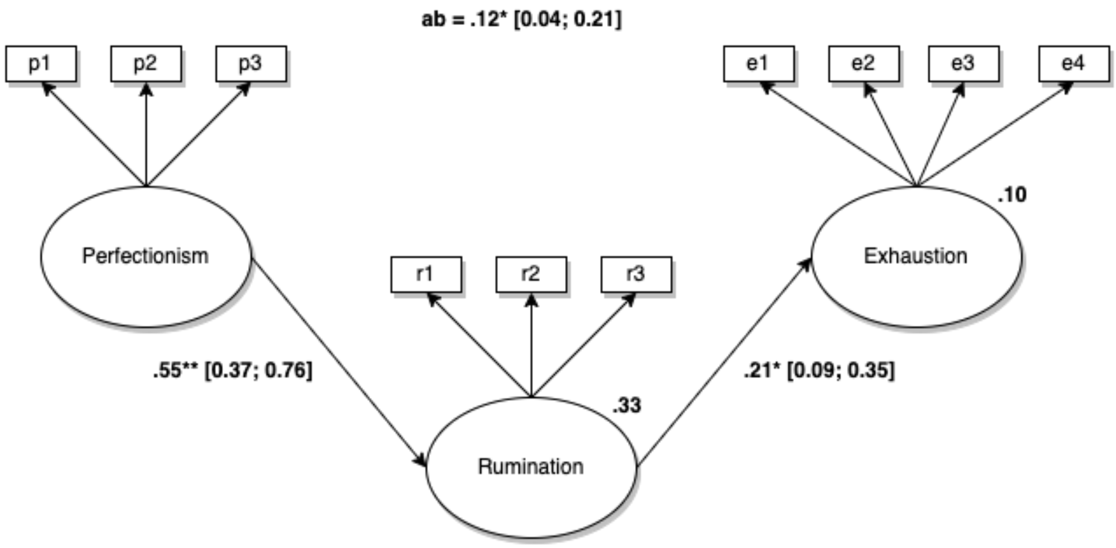

Figure 2.
Path Coefficients and variance explained (R2) by the Full Latent Variable Structural Equation Model between Perfectionism and Exhaustion via Rumination within the finalized sample population of British academics (n=213)
Note. Fit Metrics: $\chi 2$ = 60.17 (33); CFI = 0.97; TLI = .96; SRMR = .05; RMSEA = .06 (90%CI: .04, .09) ; * indicates p < .05. ** indicates p < .001.

The role of Parental Neglect on the trajectory of Aggression among children in care

Research Questions and Hypotheses

RQ_1: Is child aggression changing over time among children in care?

Null $Hypothesis_1$: There will be no change in child aggression over time among children in care.

Alternate $Hypothesis_1$: There will be a change in child aggression over time among children in care.

RQ_2: Do levels of parental neglect (PN) explain variance in the change trajectories of aggression among children in care?

Null $Hypothesis_2$: Levels of PN do not explain variance in the trajectories of aggression among children in care.

Alternate $Hypothesis_2$: Levels of PN explain variance in the trajectories of aggression among children in care.

Data Analysis Method and Justification

Researchers wanted to explore the aggression change trajectories of children in care over four years while accounting for the potential effect of levels of PN on those change trajectories.

Growth analysis of longitudinal data can be effectuated with a type of SEM (a latent growth curve model [LGC]) or with a multilevel model (MLM) using time as a Level 1 within-person predictor and adding subsequent levels of predictors (Hayes, 2006). This latter approach's popularity is growing for the analysis of repeated measurements (Steele, 2008) and has been chosen for this study adding PN as a Level 2 between-person predictor (*Figure 1*).

Even if a LGC is sometime seen as a more sophisticated way to study longitudinal data, allowing for more complex relationships among residuals and more versatility in the inclusion of variables; a MLM was here chosen in line with the KISS principle 'Keep it simple, stupid' (Axelrod, 1997). The MLM allows for an optimal control of the dependencies, here individuality of the child, that are by essence present in longitudinal data. Indeed, all measurements are nested within people (Ram et al., 2012); thus violate the assumption of independence of the general linear model (Nezlek, 2008).

Analyses were conducted in R using *lme4* package (Bates et al., 2015). Moreover, 95% confidence intervals from 5,000 bootstrapped resamples were requested to inspect statistical significance of the results (Nevitt & Hancock, 2001).

Preliminary Analyses

Grand Mean Centring.

The hypothesized predictor of within- and between-person variance, PN, was grand mean centred so that the unit of measurement represented a deviation from the sample mean of PN (Peugh, 2010).

Empty Model and Intraclass Correlation Coefficient.

Visualisation of the change trajectories of aggression in a subset of children in care revealed variance in those trajectories (*Figure 2*). An empty model was fitted to provide information about how much of the total variance in aggression was between-person variance and how much variance was yet to be explored (Field et al., 2012).

The intraclass correlation for aggression was calculated to determine whether the Level 1 outcome (i.e. aggression) had substantial person-to-person variation (*Table 1*). The intraclass correlation coefficient of aggression was .30, indicating substantial person to person variance in aggression.

The residual section representing the variability left unexplained in the empty model was 2.39. This variance could be partly explained by variability in aggression from year to year irrespective of person (i.e., time) or by other factors.

Then, the variance was partitioned between predictors which revealed that 11% of the variance (i.e., .30 / (.30 +2.39)) in aggression was due to between-person differences. This is substantial, but also means a great proportion of the variance is left unexplained, a gap which the fitted model attempted to fill.

Primary Analyses

Level 1: Within-person Predictor.

First, fixed and random effects for the intercepts and time slopes of aggression were added to the empty model.

The expected aggression score at the first point of measurement was 2.70 (95%CI: 2.60, 2.79). The fixed effect of time on the slope of aggression was significant (b = 1.12; 95%CI = 1.07, 1.16) (*Table 2*).

There was also significant between-person variance in both the intercepts (*variance* = 0.57, *SD* = 0.76; 95%CI: 0.68, 0.83) and time slopes (*variance* = 0.07, *SD* = 0.27; 95%CI: 0.22, 0.32) of aggression (*Table 1*; *Figure 3*). As such, intercepts and slopes were permitted to vary randomly in tests of PN to aggression over time. These results imply that while adding time as a predictor reduced the intercept variance, there was still a great proportion of variance to be explained in the starting aggression score and a little but significant proportion in the trajectory of aggression.

At last, the significant correlation between the random intercept and random slope (*covariance* = 0.29, 95%CI: 0.09, 0.50) indicated that children with higher starting points in aggression had an increased likelihood to exhibit rising aggression over time.

Level 2: Between-person predictor.

Subsequently, the cross-level interaction of time (Level 1; within-person) and PN (Level 2; between person) was added to the model.

For every 1 unit increase in PN away from the group mean, there was a corresponding .05 deviation in aggression from the group mean: meaning on average, people with higher PN tended to experience higher aggression (b = 0.05; 95%CI: 0.01, 0.09).

More pertinently, the relationship between time and the change trajectory differed depending on the level of a between-person moderator: deviations from the grand mean in PN just significantly moderated the relationship between time and aggression (b = .02, *SD*=.01; 95%CI: .004,.04). The positive sign of this interaction between time and PN is consistent with the interpretation that higher PN is associated with a steeper incline trajectory of the aggression slopes across the period of study.

While adding between-person differences in PN as a predictor accounted for the decrease of variance in the intercepts and slopes of aggression, random effects disclosed there is still substantial unexplained variance in both starting points (*variance* = 0.56 *SD* = 0.75; 95%CI: 0.65, 0.82) and change trajectories (*variance* = 0.07, *SD* = 0.27; 95% CI: 0.23, 0.31) of aggression.

The conditional mean of the trajectory slopes and 95% confidence bands for aggression were plotted across the range of PN using the Johnson-Neyman technique (1936) (*Figure 4*). The trajectory slopes for aggression were significant across the study period for all values of PN (i.e., confidence bands did not include zero), but the increase was steeper the higher the PN was.

Interpretations

Answering RQ_1, the fixed effect of time on the slope of aggression was significant (b = 1.12; 95%CI = 1.07, 1.16) (*Table 2*). Meaning for every unit increase in time, so every year, aggression increased by on average 1.12. The null hypothesis was rejected: child aggression changes over time among children in care.

Answering RQ_2, after analysing the variance around the change trajectories of aggression thanks to the within-person predictor of time and between-person variation (*variance* = 0.07407, *SD* = 0.2722; 95% CI: 0.23, 0.31); adding levels of PN as a between-person predictor of change trajectories of aggression very slightly reduced the variance around those change trajectories (*variance* = 0.07159, *SD* = 0.2676; 95%CI: 0.22, 0.32) (*Table 1*). Thus, the null hypothesis was rejected: levels of PN explained some variance in the change trajectories of aggression in the sample of 405 children in care.

While the change in trajectory slopes for aggression were significant across the study period for all values of PN (b = .02, *SD*=.01; 95%CI: .004,.04), the very small nature of the variance explained by between-person differences in PN (i.e., .00248) stipulate taking major caution in terms of generalizability of the results to the general population of children in care, as they might be subject to other confounding variables.

References

Axelrod, R. (1997). Advancing the Art of Simulation in the Social Sciences. In R. Conte, R. Hegselmann, & P. Terna (Eds.), *Simulating Social Phenomena* (Vol. 456, pp. 21–40). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-03366-1_2

Bates, D., Mächler, M., Bolker, B. &Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi: 10.18637/jss.v067.i01.

Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage.

Hayes, A. F. (2006). A primer on multilevel modeling. *Human communication research*, *32*(4), 385-410.

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs*, *1*, 57–93.

Nevitt, J., & Hancock, G. (2001). Performance of Bootstrapping Approaches to Model Test Statistics and Parameter Standard Error Estimation in Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 353– 377. doi: 10.1207/s15328007sem0803_2

Nezlek, J. B. (2008). An introduction to multilevel modeling for social and personality psychology. *Social and Personality Psychology Compass*, 2(2), 842-860. Chicago

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of school psychology*, 48(1), 85-112. Chicago

Ram, N., Conroy, D. E., Pincus, A. L., Hyde, A. L., & Molloy, L. E. (2012). Tethering theory to method: Using measures of intraindividual variability to operationalize individuals' dynamic characteristics. In G. Hancock & J. Harring (Eds.), Advances in longitudinal methods in the social and behavioral sciences (pp. 81-110). New York: Information Age.

Steele, Fiona (2008) Multilevel models for longitudinal data. Journal of the Royal Statistical Society: series A (statistics in society), 171 (1). pp. 5-19. ISSN 0964-1998;DOI: 10.1111/j.1467-985X.2007.00509.x

Tables.

Table 1.

Variance Components (Random Effects) for the change trajectory slopes of aggression models within the finalized sample population of children in care (n=405)

	Empty Model		1-Level Model		2-Level Model	
	Variance	*SD*	Variance	*SD* [95%CI: LL; UL]	Variance	*SD* [95%CI: LL; UL]
Id (Intercept)	0.30	0.55	0.57	0.75 [0.66; 0.83]	0.56	0.75 [0.65; 0.82]
Time	-	-	0.07	0.27 [0.23; 0.32]	0.07	0.27 [0.23; 0.31]
Residual	2.39	1.55	0.35	0.59 [0.56; 0.62]	0.35	0.59 [0.56; 0.62]
Autocorrelation	-	-	-	0.29 [0.09; 0.50]		0.26 [0.05; 0.46]

Note. LL and UL represent the lower limit and upper limit for the 95% confidence intervals; *SD* represents the standard deviation.

Table 2.

Regression coefficients (Fixed Effects) for the change trajectory slopes of aggression models within the finalized sample population of children in care (n=405)

	Empty Model		1-Level Model		2-Level Model	
	Estimate	*SE*	Estimate [95%CI: LL; UL]	*SE*	Estimate [95%CI: LL; UL]	*SE*
Intercept	4.11	0.05	2.70 [2.60; 2.79]	0.05	2.70 [2.61; 2.79]	0.04
Time	-	-	1.11 [1.08; 1.16]	0.02	1.12 [1.07; 1.16]	0.02
PN	-	-	-	-	0.05 [0.02; 0.09]	0.02
Interaction (Time:PN)	-	-	-	-	0.02 [0.00; 0.04]	0.01

Note. LL and UL represent the lower limit and upper limit for the 95% confidence intervals; *SE* represents the standard error.

Figures.

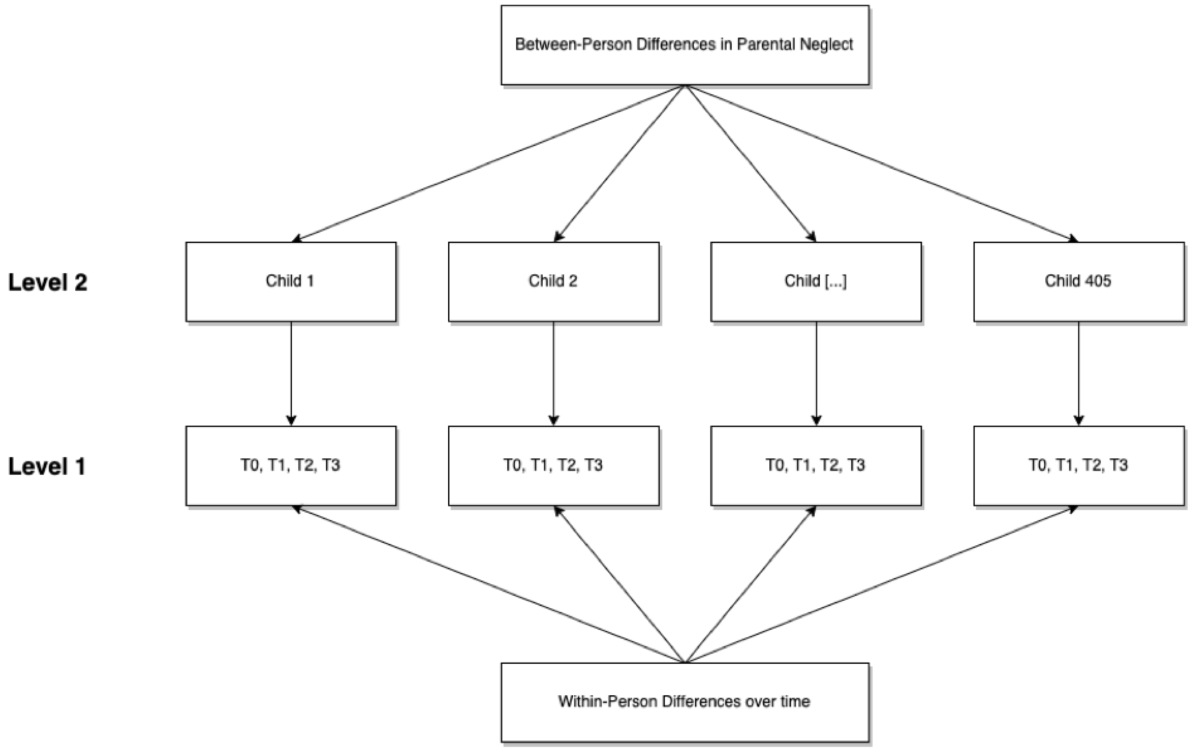

Figure 1.
Two-Level Nested Longitudinal Data Structure for the change trajectory slopes of aggression within children in care

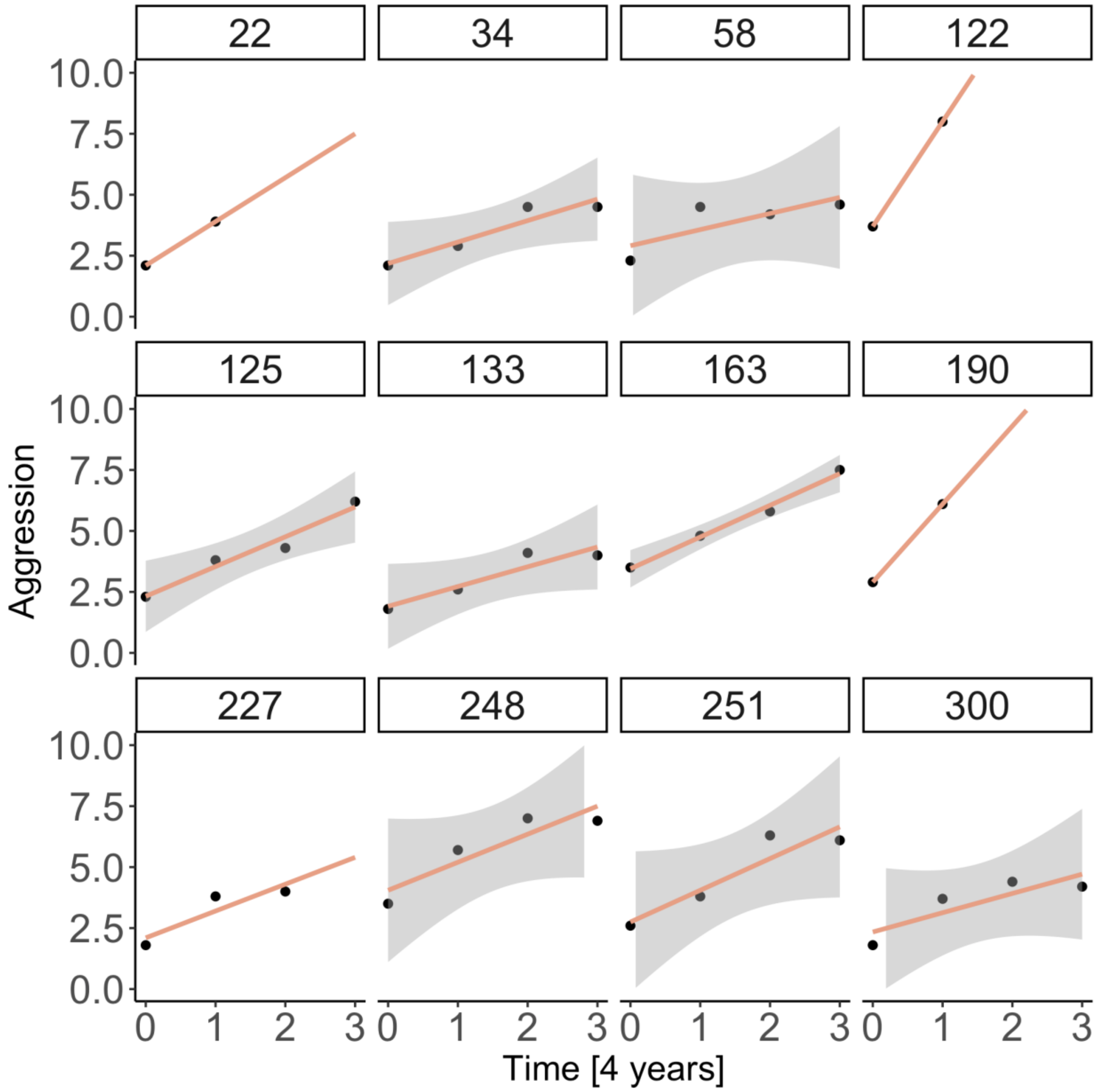

Figure 2.
Within-Person Change Trajectories of Aggression over time [4 years] in a subset of the sampled children in care (n=405)

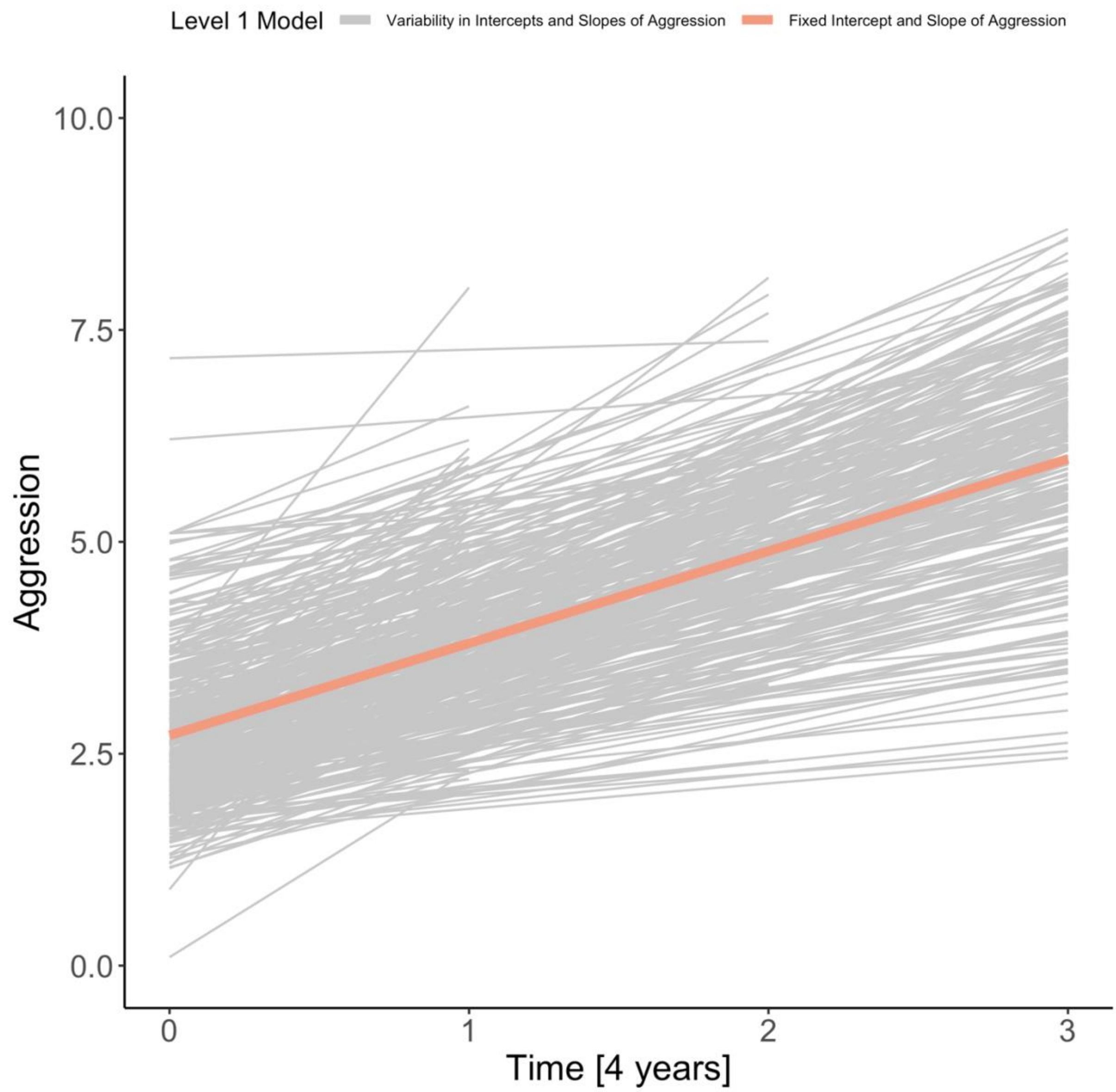

Figure 3.

Within-Person fixed effects and variability in Change Trajectories of Aggression over time [4 years] in the finalized sample of children in care (n=405)

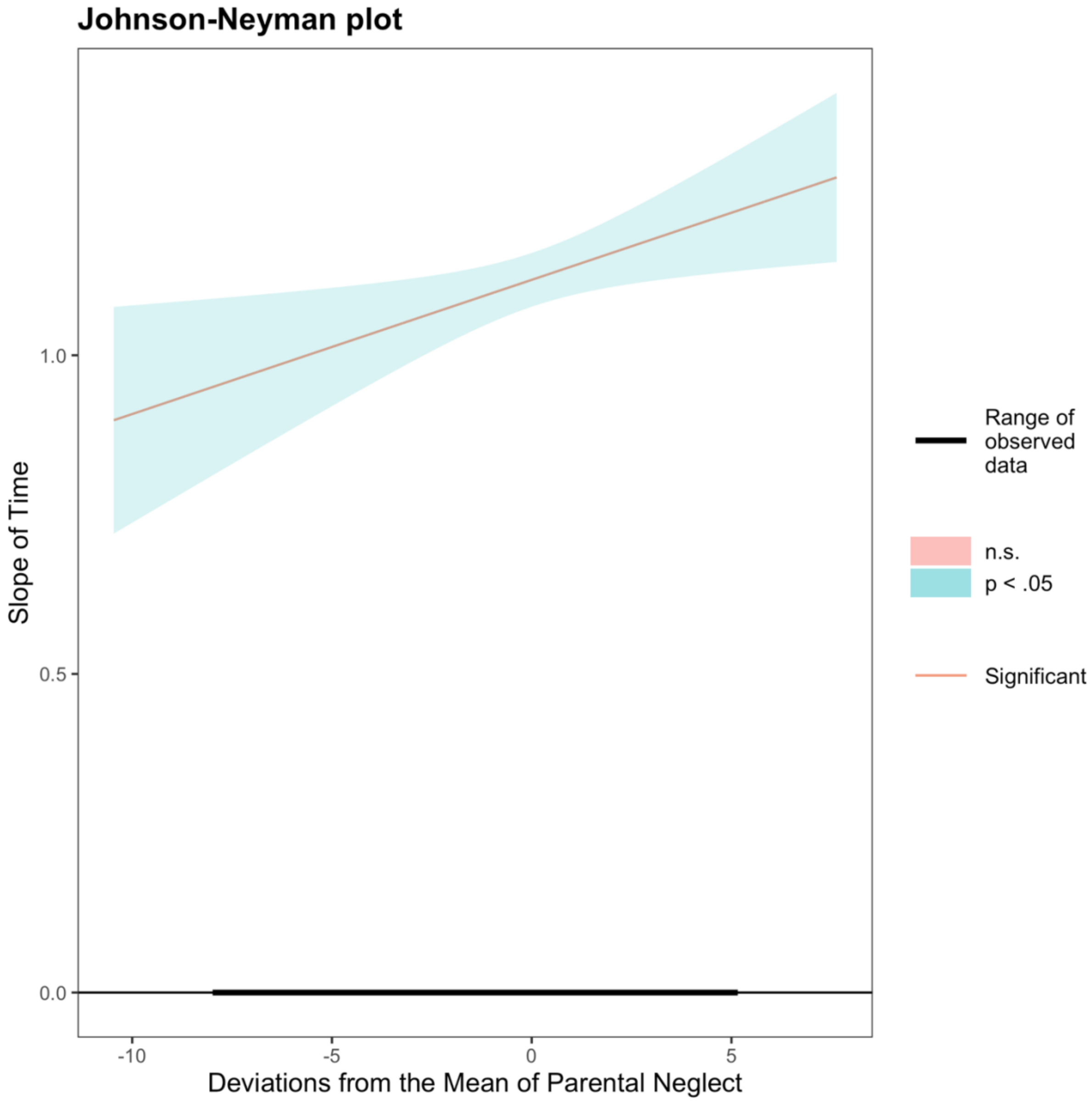

Figure 4.
Johnson-Neyman plot inspecting the significance of the effect of Parental Neglect on the Change Trajectories of Aggression over time [4 years] in the finalized sample of children in care (n=405)