Confirmatory Factor Analysis of the BRIEF-A: Exploring Executive Function Structures in a Norwegian University Student Sample

Candidate x

UiT The Arctic University of Norway

PSY-9140

25 January 2024

Introduction

Executive functions (EF) encompass a range of high-level cognitive processes that facilitate goal-directed behavior and adaptive responses to complex situations. These processes are associated with the coordinated activity of multiple brain regions, with the prefrontal cortex playing a central integrative role. According to Miyake and Friedman's unity and diversity paradigm, EF comprises three core factors: updating, shifting, and inhibiting. These factors have been reliably measured in laboratory settings and are foundational concepts in introductory psychology books. However, the ecological validity of EF tests has been questioned due to their low predictive validity in real-life situations, which is concerning given the wide use of teste tests (Spooner & Pachana, 2006).

The Behavior Rating Inventory of Executive Functioning adult (BRIEF-A-SR), from now referred to just as BRIEF, is a self-report questionnaire designed to assess executive functioning in everyday contexts, and tries add ecological validity to the EF construct (Roth & Gioia, 2005). It has gained popularity in neuropsychology for its practical relevance, and is one more the most used questionnaires by Norwegian psychologists (Ryder, 2021a). Unfortunately, BRIEF's validation in Norway is limited, more specifically, it relies on U.S. norm data that use a two-factor structure to divide the 9 subscales into behavioral regulation and metacognition, but this is not the only way to divide these subscales (Roth & Gioia, 2005; Ryder, 2021b). Alternative models, including one, three and four-factor structures, have been proposed in the literature (Roth et al., 2013). Norwegian clinicians are informing clinical decisions based on these factors, which is based on a model that might not fit a Norwegian population (Egeland & Fallmyr, 2010).

The current study aims to determine the most appropriate factor model for BRIEF-A in a Norwegian context. Four theoretical models will be tested using confirmatory factor analysis (CFA): a one-factor model representing EF as a unitary construct, the original two-factor model from BRIEF-A's creators, a three-factor model that divides behavioral regulation into emotional and cognitive components, and a four-factor model that further divides metacognition. CFA is a theory-driven approach suitable for testing expected structures and comparing model fit, and is also the method used to test the original two-factor model (Roth & Gioia, 2005). This study will employ methodological choices consistent with previous BRIEF research to ensure valid comparisons.

Methods

Participants and procedure

This study aimed to recruit 315 university students from campuses in northern Norway. Participants were part of a larger study investigating the effects of mood on executive functions, which also included assessments of depression, anxiety, personality, and sleep, in addition to the self-report version of BRIEF. All participants completed a pen-and-paper version of these questionnaires, taking approximately 20 minutes per person. The current paper focuses solely on the factor structure of BRIEF. Inclusion criteria required participants to be healthy students without neurodevelopmental disorders, a history of drug/alcohol abuse, psychiatric diagnoses, or traumatic brain injury.

Statistical analysis

Confirmatory factor analysis (CFA) was performed on the 9 BRIEF subscales. Given that each subsequent model added a factor to the previous one, they were considered nested models and compared using chi-square tests. For all statistical analysis RStudio version 4.3.2 were used together with the packages Lavaan (0.6-17) and semPlot (1.1.6). The scripts and data used for these analyses are available in a public GitHub repository for transparency and reproducibility purposes. The repository can be accessed at https://github.com/holsteirik/fl_project

BRIEF subscales

BRIEF consist of 75 items where each is belonging to a subscale. Each item can be scored from one (never), two (often) or three (always). These items are summed up into each subscale. Subscales will be analyzed to check reliability by Cronbachs Alpha and McDonalds Omega. Previous studies have used only alpha values, but omega might be prudent considering the subfactors are ordinal values. Pearson's R correlations for all nine subfactors will also be investigated. These subscales are then the subject of the CFA.

Estimation

Prior research on BRIEF has utilized maximum likelihood (ML) and weighted least squares means and variance (WLSMV) estimation methods (Albuquerque et al., 2023; Egeland & Fallmyr, 2010; Mohammadnia et al., 2022; Skogan et al., 2016). Given the slight positive skew of BRIEF scores in healthy samples, a robust estimation method was deemed appropriate. Although the skewness was minor, it could potentially bias the results. WLSMV is made to analyze ordinal data, since BRIEF subscales are summed scores, treating them as scale variables is justified, and robust maximum likelihood (MLR) may provide more accurate fit statistics. The primary analysis employed MLR, but ML and WLSMV also examined for comparison.

Models

Four models were compared:

1. A one-factor model with all nine subscales loading onto a single EF construct.
2. A two-factor model with behavioral regulation (Impulse regulation, Flexibility, Emotional control, and Self-monitoring) and Metacognition (Initiation, Working memory, Plan/Organizing, Task monitoring, and Organization of materials).

3. A three-factor model dividing behavioral regulation into Emotional regulation (Emotional control and Flexibility) and Cognitive regulation (Impulse control and Self-monitoring), with the Metacognition factor unchanged.
4. A four-factor model maintaining the Emotional and Cognitive regulation factors and dividing Metacognition into External metacognition (Initiation, Working memory, and Plan/Organizing) and Internal metacognition (Task monitoring and Organization of materials).

All factors were allowed to correlate in each model.

Fit

Model fit was assessed using several indices with rules of thumb from Hu & Bentler (1999). The Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) values close to .95 or higher are indicative of a good fit. The Root Mean Square Error of Approximation (RMSEA) values below .06 also suggest a good fit. Additionally, a Standardized Root Mean Square Residual (SRMR) value below .08 is considered acceptable. Chi-square tests were used to compare the models.

Missing

Missing data on individual items are scored as one, and up to two items per subscale may be counted this way, as per the BRIEF manual (Roth & Gioia, 2005). The items for each scale are then summed up, each forming one of the 9 subscales. If there are more than two missing items per scale, the subscale cannot be calculated and will be estimated using Full Information Maximum Likelihood (FIML), under the assumption that data were missing at random.

Results

Participants

A total of 304 university students were included in the final sample, with a mean age of 23.5 years (SD = 4.63, range = 18-52). The majority (70%) were women. The sample representing a diverse range of academic disciplines. Data collection spanned approximately three months.

BRIEF scales

The nine BRIEF subscales demonstrated acceptable to good internal consistency, with Cronbach's alpha values ranging from .68 to .89. McDonald's omega values were comparable. All subscale correlations were highly significant, correlation coefficients ranging from .19 to .68, as reported in Table 1. None of the subscales missed more then two items, so no missing data in these scales had to be estimated.

Table 1

Reliability and Pearson's Correlations between subscales

Scale	M	SD	α	ω	1	2	3	4	5	6	7	8
1. Inhibit	12.7	2.67	.68	.69								
2. Shift	9.78	2.36	.70	.70	.43							
3. Emotional control	15.5	4.42	.89	.89	.48	.50						
4. Self-monitoring	8.69	2.12	.69	.69	.59	.32	.29					
5. Initiation	14.2	3.26	.77	.77	.45	.47	.31	.28				
6. Working memory	13.6	3.06	.78	.78	.39	.29	.19	.26	.30			
7. Plan/organizing	16.8	3.61	.77	.77	.57	.50	.38	.46	.68	.42		
8. Task monitoring	10.7	2.05	.68	.69	.55	.46	.36	.45	.59	.36	.68	
9. Organization [a]	12.9	3.37	.74	.74	.40	.32	.31	.37	.44	.30	.48	.46

Note. M = mean, SD = standard deviation, α = Cronbach’s alpha, ω = McDonald’s Omega. All correlations were significant < .001.[a] = Organization of materials.

Model fit

The one-factor model exhibited a poor fit, with fit indices below acceptable thresholds and accounting for an average of 48.4% of the variance in the subscales. For all models, the SRMR values were below 0.08, while the RMSEA scores were high, exceeding 0.06. The two-factor model showed improved fit, but indices remained suboptimal. The behavioral regulation index accounted for 43.6% of the variance, while metacognition accounted for 57.2%. A chi-square difference test confirmed that the two-factor model fit significantly better than the one-factor model.

The three-factor model yielded the best fit, with acceptable CFI and borderline TLI scores. This model significantly outperformed the two-factor model, with behavioral regulation, emotional regulation, and metacognition factors explaining 61%, 51%, and 57% of the variance in the scales, respectively (see Figure 1).

The four-factor model fit worse than the three factor, but not significantly so, but was significantly superior to the two-factor model, $\Delta\chi^2(5) = 39.136$, $p < .001$. The factors of behavioral regulation, emotional regulation, external metacognition, and internal metacognition accounted for 62%, 50.1%, 63.8%, and 48% of the variance, respectively.

WLSMV estimation yielded perfect CFI and TLI scores for all models, suggesting it may not be stringent enough for this data. ML estimation did not alter any measures, indicating minimal deviation from normal distribution. Therefore, non-robust RMSEA values are reported. Model comparisons are detailed in Table 2.

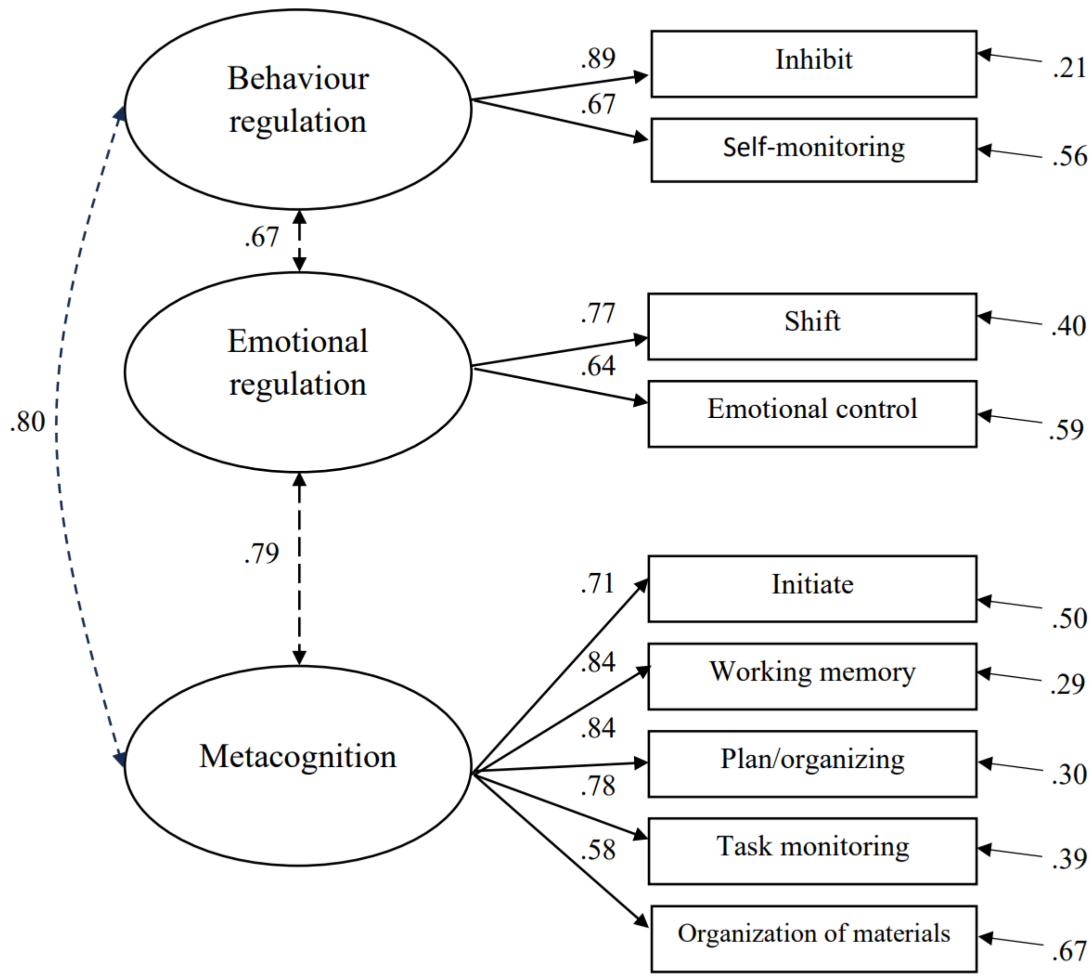
Table 2 fit indices

Model	df	X^2	RMSEA (90% CI)	SRMR	TIL	CFI
1. One-factor	26	128.8***	.111 (.092-.131)	.048	.894	.921
Model 1 versus 2	1	19.0***				
2. Two-factor	26	109.8***	.103 (.084-.123)	.045	.910	.935
Model 2 versus 3	2	37.6 ***				
3. Three-factor	24	72.2***	.081 (.060-.103)	.034	.944	.963
Model 3 versus 4	3	1.59				
4. Four-factor	21	70.6***	.088 (.066-.111)	.034	.934	.961

Note: df = model degrees of freedom, RMSEA = root mean squared error of approximation, 90% CI = 90% confidence interval for RMSEA, SRMR = standardized root mean residual, TLI = Tucker–Lewis Index, CFI = comparative fit index. ***p < .001

Figure 1

Confirmatory factor analysis for the three-factor model



Note. All path coefficients for latent variables are presented as standardized estimates. Latent factor variances are fixed at 1. Standardized variances for observed variables reflect the unique variance of each indicator. Correlations between latent variables are depicted by dotted lines.

Discussion

The findings from this study indicate that a three-factor model is the best fit for a Norwegian population The results suggest that EF, at least as measured by BRIEF, is not a unitary construct, but rather consist of separate but correlated factors. This fits with Miyake and Friedmans models. This is also in line with the authors of BRIEFs study which also found support for a three-factor model (Roth et al., 2013). Both studies were done on healthy participants, but this study was done on a Norwegian population, suggesting that a three-factor model is most appropriate. This goes against the currently used model, which only report a two-factor model, which according to our data, does not fit a Norwegian population.

An Iranian study have recently found support for the original two-factor model in a clinically depressed sample and argue that their population does not separate between emotional and behavioral regulation (Mohammadnia et al., 2022). Extrapolating from this logic would mean that Norwegian students do separate behavior and emotion, at least in a healthy sample. This raises the possibility that depression may obscure the differentiation between emotional and behavioral regulation.

The preference for a three-factor model is also supported by neuroscientific research linking EF to specific brain functions (Denckla, 2002; Egeland & Fallmyr, 2010). Zelazo and Carlson argue that there is "hot" and "cool" EF, where hot functions are necessary in situations where emotions and motivations are active, and cool EF, which operates in emotionally neutral contexts (2020). These implications are important and could help inform interventions. Zelazo & Carlson point out that EF can in certain cases be a better predictor of school performance than IQ. This differentiation has important implications for interventions, as it allows for targeted strategies.

In conclusion, the three-factor model emerged as the most suitable representation of EF for our Norwegian sample. There is a pressing need for further research, including larger studies in Norway, to confirm these findings and establish new factor norms. Relying on the current two-factor model may obscure critical clinical insights, underscoring the importance of updating our understanding and measurement of EF.

References

Albuquerque, C. P., Moura, O., Vilar, M., Seabra-Santos, M. J., Pereira, M., Major, S., Moreira, A., Lopes, A. F., Gaudêncio, A., & Simões, M. R. (2023). BRIEF-A: Factor structure and measurement invariance across self-report and informant forms. *Applied Neuropsychology: Adult*, *0*(0), 1–11. https://doi.org/10.1080/23279095.2023.2283080

Denckla, M. B. (2002). The Behavior Rating Inventory of Executive Function: Commentary. *Child Neuropsychology*, *8*(4), 304–306. https://doi.org/10.1076/chin.8.4.304.13512

Egeland, J., & Fallmyr, Ø. (2010). Confirmatory Factor Analysis of the Behavior Rating Inventory of Executive Function (BRIEF): Support for a distinction between Emotional and Behavioral Regulation. *Child Neuropsychology*, *16*(4), 326–337. https://doi.org/10.1080/09297041003601462

Hu, L., & 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. https://doi.org/10.1080/10705519909540118

Mohammadnia, S., Bigdeli, I., Mashhadi, A., Ghanaei Chamanabad, A., & Roth, R. M. (2022). Behavior Rating Inventory of Executive Function – adult version (BRIEF-A) in Iranian University students: Factor structure and relationship to depressive symptom severity. *Applied Neuropsychology: Adult*, *29*(4), 786–792. https://doi.org/10.1080/23279095.2020.1810689

Roth, R. M., & Gioia, G. A. (2005). *Behavior rating inventory of executive function–adult version*. Psychological Assessment Resources Lutz, FL.

Roth, R. M., Lance, C. E., Isquith, P. K., Fischer, A. S., & Giancola, P. R. (2013). Confirmatory Factor Analysis of the Behavior Rating Inventory of Executive Function-Adult

Version in Healthy Adults and Application to Attention-Deficit/Hyperactivity Disorder. *Archives of Clinical Neuropsychology*, *28*(5), 425–434. https://doi.org/10.1093/arclin/act031

Ryder, T. M. (2021a). Testkvalitetsprosjektet - del 1: Norske psykologers testholdninger og testbruk. *Tidsskrift for Norsk psykologforening*, *58*(1). https://psykologtidsskriftet.no/rapport/2021/01/testkvalitetsprosjektet-del-1-norske-psykologers-testholdninger-og-testbruk

Ryder, T. M. (2021b). Testkvalitetsprosjektet – del 2: Tester med behov for kvalitetstiltak. *Tidsskrift for Norsk psykologforening*, *58*(2). https://psykologtidsskriftet.no/rapport/2021/02/testkvalitetsprosjektet-del-2-tester-med-behov-kvalitetstiltak

Skogan, A. H., Egeland, J., Zeiner, P., Øvergaard, K. R., Oerbeck, B., Reichborn-Kjennerud, T., & Aase, H. (2016). Factor structure of the Behavior Rating Inventory of Executive Functions (BRIEF-P) at age three years. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *22*(4), 472–492. https://doi.org/10.1080/09297049.2014.992401

Spooner, D., & Pachana, N. (2006). Ecological validity in neuropsychological assessment: A case for greater consideration in research with neurologically intact populations. *Archives of Clinical Neuropsychology*, *21*(4), 327–337. https://doi.org/10.1016/j.acn.2006.04.004

Zelazo, P. D., & Carlson, S. M. (2020). The neurodevelopment of executive function skills: Implications for academic achievement gaps. *Psychology & Neuroscience*, *13*(3), 273.