Research Article https://doi.org/10.12973/eu-jer.13.4.1647

European Journal of Educational Research
Volume 13, Issue 4, 1647 - 1663.
ISSN: 2165-8714
http://www.eu-jer.com/

Combining Performance-Based and Self-Reported Measures of Executive Functions: Are Both Meaningful in Predicting Study Success in Higher Education Students?*

Diane Marcia Manuhuwa**
Saxion University of Applied Sciences, NETHERLANDS

Mirjam Snel-de-Boer
Saxion University of Applied Sciences, NETHERLANDS

Jan Willem de-Graaf
Saxion University of Applied Sciences, NETHERLANDS

Joke Fleer
University of Groningen, NETHERLANDS

Received: December 5, 2023 ▪ Revised: May 7, 2024 ▪ Accepted: April 2, 2024

Abstract: Research in higher education has revealed a significant connection between executive functions (EF) and study success. Previous investigations have typically assessed EF using either neuropsychological tasks, which provide direct and objective measures of core EF such as inhibition, working memory, and cognitive flexibility, or self-report questionnaires, which offer indirect and subjective assessments. However, studies rarely utilize both assessment methods simultaneously despite their potential to offer complementary insights into EF. This study aims to evaluate the predictive capabilities of performance-based and self-reported EF measures on study success. Employing a retrospective cohort design, 748 first-year Applied Psychology students completed performance-based and self-report questionnaires to assess EF. Maximum likelihood correlations were computed for 474 students, with data from 562-586 first-year students subsequently subjected to hierarchical regression analysis, accommodating pairwise missing values. Our results demonstrate minimal overlap between performance-based and self-reported EF measures. Additionally, the model incorporating self-reported EF accounted for 13% of the variance in study success after one year, with the inclusion of performance-based EF raising this proportion to 16%. Self-reported EF assessments modestly predict study success. However, monitoring levels of self-reported EF could offer valuable insights for students and educational institutions, given that EF play a crucial role in learning. Additionally, one in five students reports experiencing significant EF difficulties, highlighting the importance of addressing EF concerns for learning and study success.

Keywords: *Executive functions, higher education, performance-based measurements, self-reported measurements, study success.*

To cite this article: Manuhuwa, D. M., Snel-de-Boer, M., de-Graaf, J. W., & Fleer, J. (2024). Combining performance-based and self-reported measures of executive functions: Are both meaningful in predicting study success in higher education students? *European Journal of Educational Research*, *13*(4), 1647-1663. https://doi.org/10.12973/eu-jer.13.4.1647

Introduction

Executive functions (EF) are crucial in learning and predicting student study success (e.g., Baars et al., 2015; Musso et al., 2019; Ramos-Galarza et al., 2020). EF encompass a set of cognitive processes that are partially interdependent and involved in the top-down regulation of behavior, affect, and cognition (e.g., Baggetta & Alexander, 2016; Barkley, 2012; Nigg, 2017). They become active when habitual responses and routines prove inadequate, especially in situations characterized by novelty or complexity (e.g., Barkley, 2012; Diamond, 2013; Miller & Cohen, 2001), which is the case when a student learns something new. As such, EF are like the supervisory management system of the brain (e.g., personal communication, 8th of October 2022, Mc Closkey) because they orchestrate and modulate the functioning of other brain regions necessary for learning and performing (Zelazo & Carlson, 2020).

Research on EF in higher education indicates a significant association between EF and study success (e.g., Baars et al., 2015; Manuhuwa et al., 2023; Musso et al., 2019). Within this body of research, EF has been assessed through two primary methodologies: neuropsychological assessments, which directly and objectively evaluate core EF, such as inhibition, working memory, and cognitive flexibility, and self-report questionnaires, which offer an indirect and subjective measure of EF. However, it is uncommon for both assessment approaches to be utilized simultaneously despite providing distinct

* This study was presented at the 5th International Conference on Applied Research in Education on October 20-22, 2023.

** **Corresponding author:**
Diane Marcia Manuhuwa, Saxion University of Applied Sciences, School of Applied Psychology and (International) Human Resource Management, Program Employability Transition. Handelskade 75, 7417 DH Deventer, The Netherlands. ✉ d.m.manuhuwa@saxion.nl

yet complementary perspectives on EF. This investigation explores the predictive capacity of combining performance-based and self-reported EF measures concerning study success.

Literature Review

The literature distinguishes between the core and complex EF (e.g., Diamond, 2013; Miyake et al., 2000). The first of three core EF is *inhibition*, the capacity to control attention or inhibit dominant or automatic behavioral responses, thoughts, and emotions (e.g., Baggetta & Alexander, 2016). For example, you can read a text without being distracted by environmental noises or distracting thoughts. *Working memory* involves actively holding information in the mind and manipulating it (e.g., Baddeley, 2010; Diamond, 2013). For example, you are conversing with your fellow students and recalling what they just said while formulating your response in real-time. *Cognitive flexibility* encompasses the ability to literally and figuratively change perspective, ignore irrelevant data, acquire new information, change thought patterns, or modify behavior (e.g., Diamond, 2013; Nigg, 2017). For instance, you suddenly have to cover a presentation for a fellow student who has fallen ill and change your current plans. The complex EF are based on these core EF and refer to, for example, planning, organizing, and task- and self-monitoring (Diamond, 2013). Complex EF are evident during studying in various capacities, such as planning (e.g., creating schedules for reading materials, outlining plans for completing assignments, and preparing for upcoming exams). Examples of organization are organizing study materials, structuring notes, or arranging digital files. Finally, task and self-monitoring processes involve tracking your study progress, adjusting strategies as needed, and monitoring your performance while studying.

EF play a central role in learning because they are activated when students face new or complex situations, work towards achieving (learning) goals, or need to adapt. In these situations, students cannot rely on learned automatisms or responses but must overwrite old knowledge and patterns with new information or skills (Zelazo & Carlson, 2020). The demands on EF are particularly high at the onset of learning something new, gradually diminishing as knowledge and skills become more crystallized (Naglieri & Otero, 2017). Therefore, the extent to which EF are adequate influences a student's learning and study success. The relationship between EF and study success has been demonstrated in children, particularly in reading and mathematics (e.g., Alloway et al., 2010; Gathercole et al., 2004), adolescents (e.g., Effeney et al., 2013), and young adult students (e.g., Baars et al., 2015; Manuhuwa et al., 2023; Musso et al., 2019). All these studies show that better EF or fewer EF problems are correlated with obtaining more credits. Because of the relationship between EF and study success, it is helpful to identify the students at risk for poor EF so that, for instance, educational professionals can take proactive and preventive action. Knowing what measurement instruments are most effective in assessing EF is necessary.

EF are generally assessed through neuropsychological tasks involving direct and objective testing of core EF, such as working memory, inhibition, and cognitive flexibility, or through self-reported observed behaviors, which provide an indirect and subjective assessment. In their practitioner review, Toplak et al. (2013) found that only 24% of the 286 relevant correlations reported statistical significance, with an overall median correlation between the measurement types of only .19. This limited number of statistically significant correlations highlights the difficulty in establishing a reliable and consistent relationship between performance-based and self-reported measures of EF. More recent studies using both types of measurements also indicate a low degree of correlation between performance-based and self-reported measures of EF among young adult students (e.g., Follmer, 2021; Snyder et al., 2021) and adults (Bodenburg et al., 2022). These findings underscore that these two types of measurements capture distinct underlying mechanisms of EF.

The literature consensuses that both types of measurements can provide a picture of the complex construct of EF. However, there are differences between them in focus and application. Firstly, core EF are often measured through performance-based measures in a neutral laboratory environment or with self-report measures, while complex EF are generally assessed with self-report measures.

Secondly, while self-reported EF provides insight into behavioral manifestations of EF in daily life (e.g., Kamradt et al., 2014), performance-based measurements focus on cognitive functions (e.g., Toplak et al., 2013; Zelazo & Müller, 2010). Therefore, the completion of self-reports demands reflective ability from the respondent, while performance-based measurements capture optimum objective performance on various tests. Thirdly, in the clinical context, the outcomes of performance-based measures of EF have been associated with brain damage and psychopathology (e.g., Kamradt et al., 2014; Rabinovici et al., 2015; Velichkovsky et al., 2023) while self-reported EF outcomes are more strongly associated with impairments in daily functioning than performance-based measures (e.g., Barkley & Fischer, 2011; Kamradt et al., 2014). In the educational context, performance-based EF have been found to mediate study success, for instance, via self-regulated learning (e.g., Musso et al., 2019; Rutherford et al., 2018), while self-reported EF influence study success directly (e.g., Baars et al., 2015; Ramos-Galarza et al., 2020). Finally, the literature often describes self-reported measurements as more 'ecologically valid' than performance-based measurements. However, this assumption does not hold, as performance-based measurements can predict functional outcomes, and 'ecologically valid' is defined differently in studies or is incorrectly used (Suchy, DesRuisseaux, et al., 2024; Suchy, Mora, et al., 2024). Instead, assuming that most self-reported measurements meet face, content, or criterion validity is likely accurate. In other words, they are a way of measuring that one sees as relevant for understanding EF, which is also reflected by items addressing daily functioning behavior. At the same time, performance-based measurements also meet the mentioned validity criteria. However, they measure a different aspect of EF: how the brain tackles and performs a specific EF task.

In summary, both measurements capture different aspects of the complex and multidimensional construct EF, providing complementary insights into EF in students or distinctive perspectives on cognitive control (Snyder et al., 2021, p. 1096). The primary aim of this study was to investigate the potential value of combining performance-based and self-reported EF measurements in predicting study success (earned credits after one academic year). Specifically, we address the following questions: (1) Do combined performance-based and self-reported EF measurements predict study success? (2) What is the relative importance of each predictor?

We hypothesize that both types of measurements contribute to the variance in study success after one academic year because they assess different EF mechanisms, potentially providing complementary and relevant information. Additionally, we expect the contribution of self-reported EF to be higher than that of performance-based EF, given its more direct influence on study success compared to the more indirect impact of performance-based EF.

Methodology

Procedure

A retrospective cohort design was used for this study (Ato et al., 2013). Performance-based and self-reported EF were measured in November 2020 and December 2021 during the module 'Diagnostic Research Part 1' (DR1), and obtained credits were collected during the academic years 2020-2021 and 2021-2022. In the module DR1, first-year Applied Psychology students learn to perform psychodiagnostic research and experience participating in research themselves by filling out questionnaires and going through neuropsychological testing.

The first part of the study, in the first week of the module, consisted of an online lecture about the aim and procedure of the study. The students were informed that participation was voluntary and confidential, that it would take approximately an hour (10-15 minutes to complete the questionnaire and 45 minutes for the full neuropsychological test battery), and that no credits would be granted. Conforming to the guidelines of the Institutional Ethical Advice Committee (SEAC-2019-14 B), informed consent was obtained at the start of the procedure. The second and third parts of the study, filling out the questionnaire and the neuropsychological testing, were executed in each academic year's second and third weeks of the module.

We collected data for this study in two consecutive academic years: 2020-2021 and 2021-2022.

Participants

Participants for this study consisted of Applied Psychology students in their first year at an Applied University in the Netherlands during the academic years 2020-2021 and 2021-2022. First-year students face a new way of learning alongside significant changes in their lifestyle, such as living independently (Carragher & McGaughey, 2016), and therefore, their EF might be extra challenged. The students were eligible to participate if they were aged between 18 and 25, a common age group during the first year of higher education. We chose 25 years as the upper boundary because EF development is assumed to be optimal due to prefrontal cortex maturation, and we would rather speak of adults than young adults (Giedd & Rapoport, 2010).

Measures

Study Success

We collected data on study success by extracting the number of credits earned from the university's database at the end of each academic year.

Self-reported Measures of EF

The *Behavior Rated Inventory of Executive Function – Adult version (BRIEF-A)* measures self-reported EF in the everyday environments of adults (Roth et al., 2005, 2013). Scholte and Noens translated the BRIEF-A into Dutch and developed norms for adults between 18 and 65 years old in the Netherlands (Roth et al., 2011). The BRIEF-A comprises nine distinct and empirical-based scales, as outlined in Table 1. These scales, consisting of 75 items on a three-point scale (1=never; 2= sometimes; 3=often), measure problematic behaviors in the past month. Higher scores are indicative of increased executive dysfunction levels. Converting the raw scores of the subscales into T-scores allows for comparison with a norm group, where a T-score of 65 or above signals severe problems in a specific executive function or cluster of EF. In clinical contexts, a cut-off T-score of ≥ 80 or ≥ 90 may offer higher specificity and a more realistic representation (Abeare et al., 2021; Shwartz et al., 2020).

Table 1. Reliability Subscales BRIEF-A and Example Items

Index (Total of Subscales) Subscale	Cronbach's α	Number of Items	Example Item
Behavioral regulation			
Inhibition	.67	8	*I am impulsive*
Shift	.72	6	*I am bad at change*
Emotional control	.89	10	*I overreact to minor problems*
Self-monitor	.70	6	*I say things without thinking*
Metacognition			
Initiate	.76	8	*I find it challenging to start working independently*
Working Memory	.76	8	*I can only concentrate for a short time*
Plan	.76	10	*I do not plan tasks.*
Task Monitor	.69	6	*I make sloppy mistakes*
Organization of materials	.82	8	*I do not clean up my mess.*

Note. Subscales and items are translated from the Dutch BRIEF-A version (Roth et al., 2011).

Furthermore, the BRIEF-A contains three validity scales for possible symptom validity: negativity, inconsistency, and infrequency. These validity scales assess if the student's response pattern is not overly negative, inconsistent, or atypical, respectively. Specifically, the cutoff score of the negativity scale is informative for exaggeration of negative symptoms within young adults (i.e., students). According to Roth et al. (2005, 2011), it is six or more items out of ten items out of ten. However, Abeare et al. (2021) and Shwartz et al. (2020) state that four or more items are a better measure and increase the specificity. We evaluated the reliability of the BRIEF-A by assessing the internal consistency of the subscales and concluded that the internal consistency reliability was acceptable (Table 1).

Performance-Based Measures of EF

The Central Nervous System Vital Signs (CNSVS) is a computerized neurocognitive test battery designed for regular clinical screening of cognitive problems (CNS Vital Signs, 2023; Gualtieri & Johnson, 2006). It is a valid and reliable instrument (e.g., Campman et al., 2017; Gualtieri & Johnson, 2006; Plourde et al., 2018) suitable as a screening tool or a serial assessment measurement.

The CNSVS is a neuropsychological test battery with seven well-known and well-established tests (Gualtieri & Johnson, 2006). From this test battery, we selected the Stroop Test, the four-part continuous performance test including a working memory test and a test for shifting attention. These three tests measure the core EF: inhibition, working memory, and cognitive flexibility.

The Stroop Test (ST) measures *Inhibition* and consists of three parts: the first part measures reaction time; the student must press the space bar as fast as possible when they see a word (e.g., blue or green). The second part becomes more complex than the first part by adding a command: the student must press the word only if it is displayed in its corresponding color. For instance, the word 'red' must be pressed only if it is typed in red. Finally, the third part requires inhibition. The student must press the space bar if the word does not appear in the corresponding color. So, for example, when the word red is typed in blue, it means to push the space bar; red in red does not. The average reaction times (in ms) of parts two and three are taken as a measure of reaction time to measure inhibition. In addition, we used the number of commission errors (i.e., the stimuli the student missed and should have pressed) to assess the performance's adequacy.

The Four-Part Continuous Performance Test (FPCPT) measures the visual part of *working memory* and is administered in four parts to measure sustained attention and visual working memory (CNSVS manual). Only part four is administered. Part four is a "two-back" CPT, and the student must press the space bar if the symbol shown is the same as the two symbols back. Here, we use the number of correct responses, namely, the incorrect responses subtracted from the total number of correct responses.

The Shifting Attention Test (SAT) measures *Cognitive flexibility* by assessing whether students can adequately switch from one instruction to another. The student is shown a symbol at the top of the screen, in red or blue. Two symbols appear at the bottom, left and right. The student is instructed to choose 'shape' or 'color'. For example, suppose the symbol at the top is a red-filled circle, and the symbols on the left and right are a red rectangle and a blue circle, respectively. In the case of the command, choose 'shape,' the student must now choose the blue circle (right Shift key). In the case of the command, choose 'color' the left red rectangle (left Shift key). The symbols and assignments change constantly, and the challenge for the student is to be fast and make as few mistakes as possible. This study tests cognitive flexibility by measuring the number of correct responses (total correct responses - total errors) and the average response time (in milliseconds).

Statistical Plan

The first step was identifying the measurement model for self-reported EF with confirmatory factor analyses (CFA) in Amos 26.0, using maximum likelihood (ML) estimation, which is robust with non-normal distributed data (Gaskin & Happell, 2014).

Subsequently, because of high multicollinearity and inadequate model fit measures of the CFA ($x2/df$ = 2.39, $p < .001$; CFI = .73), we conducted an exploratory factor analysis (EFA) per theoretical construct. We calculated the factor loadings of the BRIEF-A via the Unweighted Least Squares (ULS) method, which has several advantages (Watkins, 2018). First, the ULS method works with both normal and non-normal distributed data. Second, the ULS should consider the ordinality of the BRIEF-A answers, limited to three response choices. Third, the ULS diminishes the attenuation that works better against multicollinearity, which is high within the BRIEF scales (Lorenzo-Seva & Ferrando, 2021). Consequently, the latent self-reported EF variables were established, and no constructs were excluded. However, from the observable measures of the latent variables, some items were removed due to cross-loadings or low factor loadings (< .3) (an overview of the included and excluded items are reported in supplements 1 and 2).

Subsequently, the correlations between the BRIEF-A and CNSVS subscale scores were calculated using the Maximum Likelihood method. For this analysis, we used a sample of 474 students who completed the questionnaire and the tests without missing values or invalid responses. The correlations are interpreted according to Gignac and Szodorai (2016), where correlations of ≥ .10 are considered relatively small, ≥ .20 as typical (medium), and ≥ 0.30 as relatively large. In addition, we excluded correlations below .20, as we expected them to be too small to be meaningful for the educational context.

Again, a CFA was conducted using ML estimation, now without significant multicollinearity and acceptable model fit measures ($\chi2/df$ = 1.76; SRMR = .05; RMSEA = .04; CFI = .93). The assumptions of linearity and homoscedasticity for hierarchical regression were met. Also, the errors of the dependent variable were normally distributed (skewness < 2; kurtosis < 2), and there was no multicollinearity among predictors (VIF < 5) (Hair et al., 2022).

Finally, we conducted a hierarchical regression analysis of the data. This analysis allows us to examine the incremental contribution of EF measures in predicting variations in study success while controlling for relevant covariates, such as age, gender, or prior academic achievement. We used pairwise missing values for this analysis because not every student conducted every test, and many missing values within tests are present. Therefore, we use the maximal sample possible, ranging from 562-586. We conducted the hierarchical regression analysis with the variable “Academic year” added as a covariate because the groups differed significantly in “Obtained credits” ($t(624) = 3.85$, $p < .001$) with an average of 47.19 (*SD* = 17.10) for cohort 2020-2021 and an average of 41.82 (*SD* = 17.27) for cohort 2021-2022. Incorporating the study year as a covariate can decrease the dependent variable (credits) variance attributable to the study year. This can make the model more robust by reducing the noise caused by different years of study, allowing a more accurate assessment of the effect of other variables (VanLunen, 2020).

Results

Descriptive Statistics

Over two academic years, 748 students completed questionnaires and conducted neuropsychological performances. Of these, 626 signed the informed consent (33.2% male, 66.0% female, and 0.8% defined themselves differently; the mean age was 19.7 years (*SD* = 1.70)).

Of these 626, 562 students validly conducted the three CNSVS neuropsychological performances, and 586 completed the questionnaire (BRIEF-A). The student’s total BRIEF-A raw scores ranged from 72-182; the mean score was 118.74 (*SD* = 18.65). More specifically, 21.5% scored in the clinical range (*t*-score ≥ 65), meaning these students report severe EF problems. Finally, the average obtained credits after one academic year was 45.57 (*SD* = 17.82), ranging from 0 to 60, with 60 points being the maximum possible.

Data analysis

Figure 1 illustrates the sample selection procedure and provides an overview of the chosen analyses.

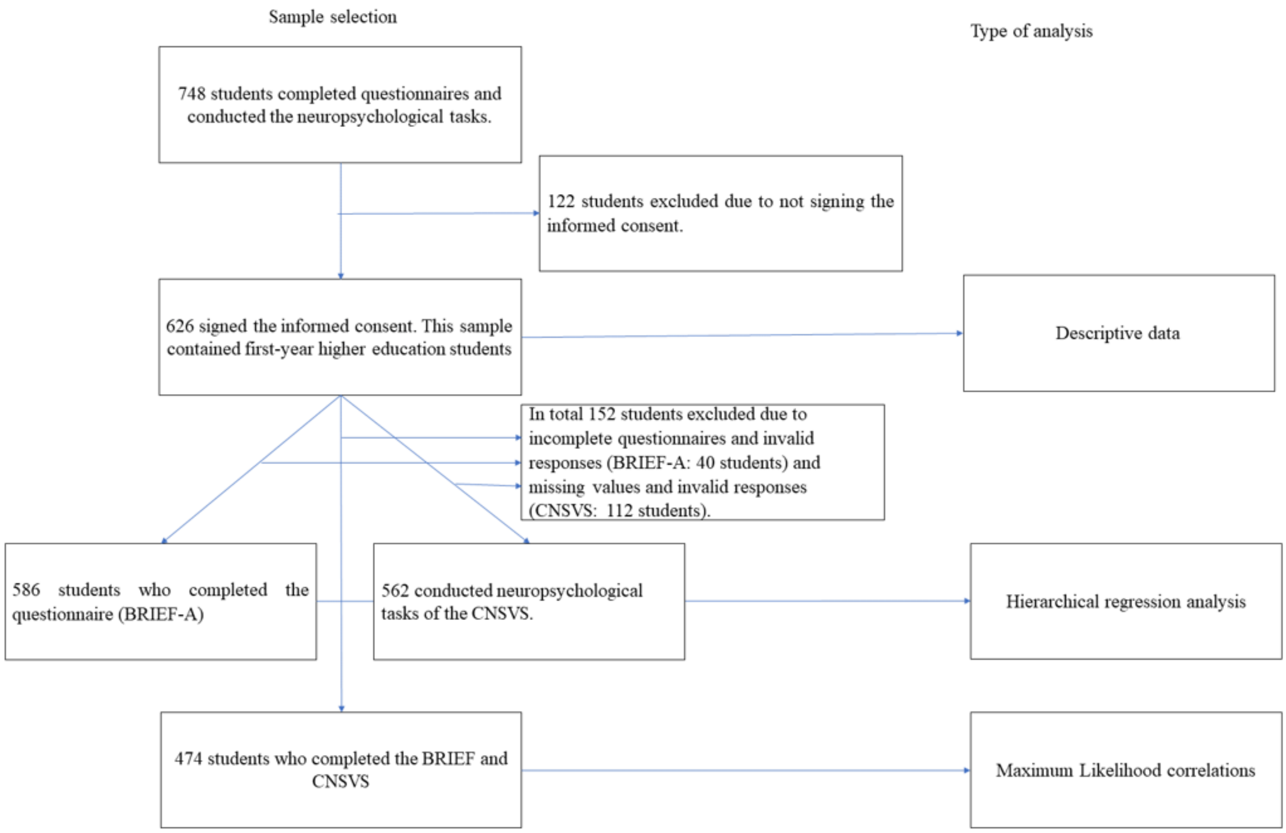

Figure 1. Flowchart of Sample Selection and Corresponding Analysis

Table 2 describes the relationships between self-reported EF (measures with the BRIEF), performance-based EF (measured with the CNSVS), and study success. The performance-based EF and self-reported EF showed no correlations higher than .20, demonstrating no meaningful relationships. We found weak to strong correlations (range r = .11 to r = -.90) among the self-reported EF, meaning that problems with one executive function often coincide with problems in another. The self-reported EF 'initiate' (-.25) and 'plan/organize' (-.23) are medium and statistically significantly correlated with the obtained credits after one academic year, whereby more EF problems come with fewer credits.

Subsequently, there were a few statistically significant correlations within the performance-based EF. Namely, 'flexibility-reaction time' correlates with 'inhibition-reaction time' (.59; i.e., being fast on a flexibility test comes with being fast on an inhibition test), and the 'flexibility-errors' correlate with 'inhibition-reaction time' (-.38) and 'flexibility-reaction time' (-.65), i.e., in these cases being fast comes with making many errors. No correlations were found between performance-based EF and study success, suggesting that there is not a (direct) relationship between them.

Predictive Power of Self-reported EF and Performance-based EF to Study Success

The causal model incorporating self-reported EF and performance-based EF as predictors of study success and academic year as a covariate exhibited acceptable model fit measures ($\chi 2/df$ = 1.63; SRMR = .04; RMSEA = .04; CFI = .93). The factor loadings of the measured variables ranged from at least .51 to .89 (see Appendix 2).

Table 3 summarizes the results of the hierarchical regression analysis for EF variables predicting study success for the two academic years. For the hierarchical regression analysis, we used the observed variables of self-reported EF after the confirmatory and exploratory analyses.

In step 1, we added the covariate 'Academic year' to correct for differences in study success, R = .15, F(1, 492) = 11.69, p < .001. In step 2, we added self-reported EF into the equation, R = .36, F(12, 481) = 5.98, p < .001. Self-reported EF accounted statistically significantly for thirteen percent of the variance in study success. In step 3, performance-based EF were forced into the equation, R = .40, F (17, 476) = 5.19 p < .001. Sixteen percent of the variance was accounted for after step 2; the F-change was statistically significant (p = .012), indicating that adding performance-based EF does improve the model prediction.

Table 2. Maximum Likelihood Correlations Between Self-Reported EF and Performance-Based EF (n=474)

EF Self-Reported Subscales[1]	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Study Success (0-60 max)	44.82	17.53	1														
2. Inhibit[1] (6-18)	13.81	2.90	-.20***	1													
3. Shift (6-18)	13.05	3.26	-.05	.19***	1												
4. Emotional Control (6-30)	15.90	4.53	.07	.20***	.61***	1											
5. Self-Monitor (6-18)	8.63	2.06	-.12*	.83***	.17**	.17**	1										
6. Initiate (8-24)	14.50	3.42	-.24***	.58***	.35***	.15**	.26***	1									
7. Working Memory (8-24)	14.24	3.29	-.19***	.75***	.43***	.23***	.38***	.76***	1								
8. Plan/Organize (10-30)	16.60	3.70	-.23***	.65***	.35***	.14**	.41***	.87***	.82***	1							
9. Task Monitor (6-18)	10.90	2.22	-.20***	.72***	.29***	.13*	.45***	.86***	.81***	.90***	1						
EF task-based subscales																	
Inhibition																	
11. Inhibition-reaction time[3] (ms)	653.33	76.15	-.04	-.09	-.00	-.03	.01	-.11*	.01	-.03	-.11*	-.04	1				
12. Inhibition-errors[2]	1.43	1.26	-.01	.08	-.02	-.08	.01	.14**	.03	.13*	.13*	.11*	-.16***	1			
Working memory																	
13. Working memory-errors[2]	12.01	2.99	.10*	.06	-.14**	-.05	.03	.08	-.01	.01	.05	.04	-.23***	-.04	1		
Flexibility																	
14. Flexibility-reaction time[3] (ms)	1000.45	129.74	-.01	-.12*	.06	.02	-.07	-.06	.05	-.05	-.08	.00	-.59***	-.11*	-.28***	1	
15. Flexibility-errors[2]	45.37	9.53	.02	.04	-.02	.00	.05	-.04	-.07	.01	.03	-.03	-.38***	-.13**	.19***	-.65***	1

[1] The self-reported scales are reversed; a higher score indicates more perceived EF problems.
[2] Errors are reversed; a higher score refers to more errors, which implies more EF problems.
[3] Time is reversed because lower means the subject is faster, which is preferable unless it comes at the expense of quality (i.e., many errors are made).
* Significant at the 0.05 level (2-tailed)
** Significant at the 0.01 level (2-tailed)
*** Significant at the 0.001 level (2-tailed)

Table 3. Summary of Hierarchical Regression Analysis for EF Variables Predicting Study Success With Cohort as a Covariate 2020-2021 and 2021-2022 (N=562-586)

Predictor	β	R^2	ΔR^2	*F*
Step 1		.02	.02	11.69
Cohorts	-.15*			
Step 2		.13**	.11	5.98
Cohorts	-.15**			
Self-reported EF				
Physical Turmoil	-.06			
Trouble Accepting Change	.08			
React Impulsively	.02			
Trouble Remembering	-.04			
Trouble Initiating	-.13*			
Anger outbursts	-.13**			
React emotionally	.11*			
Trouble planning	-.10*			
Make sloppy mistakes	.04			
Trouble keeping environment organized	-.03			
Step 3		.16***	.03	5.19
Cohorts	-.16***			
Self-reported EF				
Physical Turmoil	-.07			
Trouble Accepting Change	.09			
React Impulsively	.02			
Trouble Remembering	-.03			
Trouble Initiating	-.15**			
Anger outbursts	-.11*'			
React easily emotionally	.10			
Trouble planning	-.10*			
Make sloppy mistakes	.03			
Trouble keeping environment organized	-.04			
EF performance-based				
Inhibition-reaction time	-.06			
Inhibition-errors	.04			
Working memory-errors	.15***			
Cognitive flexibility-reaction time	.09			
Cognitive flexibility-errors	.00			

* Significant at the .05 level
** Significant at the .01 level
*** Significant at the .001 level

In the model of step 3, precisely, self-reported EF 'Trouble initiating' (β = .-15**), 'Anger outbursts' (β = -.11*), and 'Plan' (β = -.10*) statistically significantly predict study success, where more problems lead to fewer obtained credits. Additionally, performance-based EF 'working memory' predicts a portion of study success (β = .15***), with a better achievement predicting more study success.

Discussion

In this study, we investigated how both performance-based and self-reported EF explain variances in students' study success. Our data revealed that approximately one in five students reported significant EF difficulties. We found no substantial correlations between performance-based EF measures and self-reported EF assessments, indicating that these two measurement approaches tap into different underlying aspects of EF. When considered combined, self-reported and performance-based EF collectively accounted for 16% of the variation in study success, measured by obtained credits. Notably, self-reported EF contributed a more significant portion (13%) to this explanation, with performance-based EF adding 3%. Based on these findings, we conclude that self-reported EF contributes to study success. It might be relevant for educational institutions to consider and support EF in students.

Our finding that self-reported EF impacts students' study success is consistent with prior limited research on the predictive role of EF in academic performance. For instance, Baars et al. (2015) demonstrated that students' self-reported levels of attention, planning, self-control, and self-monitoring predicted their academic progress after one year by 9%. Another example is Knouse et al. (2014), who highlighted that self-reported EF -particularly self-motivation- explained 17% of the variance in study success after one semester in a sample of 229 female students. Finally, Ramos-Galarza et al. (2020) demonstrated an explained variance of 31% of self-reported EF, significantly higher than the percentage we found. However, unlike our analyses, they added the statistically significant associated variables to the logical ordinal regression equation. By doing so, the EF aspects that did not contribute to study success did not attenuate their results. Despite the different percentages, these studies give the same indication as ours, namely that self-reported EF is a significant variable in learning and study success.

This study also revealed a lack of significant correlation between self-reported EF and performance-based EF, a trend observed in several similar studies (e.g., Barkley & Fischer, 2011; Follmer, 2021; Toplak et al., 2013). Follmer (2021) notes that such low correlations are frequently encountered in research, occurring in comparisons of both direct measurements with each other and of indirect and direct ones, thereby complicating the interpretation of results. His study additionally reveals a disparity between self-reported EF scores and actual performance on a neuropsychological task among overconfident students. Nonetheless, he asserts that despite these challenges, the findings keep value and warrant interpretation, albeit cautiously, considering the constraints inherent in the different testing methodologies.

A reason for the varied measurements could be substantive. The difference may arise from the separate underlying mechanisms that these measures evaluate: overt behavior, which entails executive control and is assessed in unstructured environments through self-reports, and covert cognitive abilities, which involve supervisory processes and are evaluated in structured environments using neuropsychological tests (Toplak et al., 2016).

Self-report questionnaires like the BRIEF provide insights into individuals' behaviors in daily situations. For instance, a statement such as "I have trouble sitting still" may indicate difficulties with self-control (inhibition problems), reflecting past behavior. In contrast, performance-based measures entail administering specific tests or tasks to evaluate cognitive abilities at the moment. For example, the Stroop test assesses impulse control, with success indicating adequate inhibition abilities. Both methods shed light on EF, like inhibition, albeit through different lenses, showcasing varied approaches to understanding EF performance. This overarching distinction between overt behavior and covert cognitive abilities could explain the low convergent validity between the two types of measurements, as they show weak associations. The lack of correlations may stem from differences in EF operationalization between the questionnaires and the tests (Snyder et al., 2021). However, both measurements conceptually pertain to EF, indicating sufficient content validity and providing valuable insights into EF. The divergence in operationalizations of the EF construct may stem from its complex and multidimensional nature.

Additionally, performance-based measures primarily evaluate the efficiency of cognitive processes, while self-reports provide insights into success in goal pursuit. Furthermore, the distinction aligns with the psychometric differentiation between typical and optimal performance situations, with self-reports assessing typical performance in unstructured environments and performance-based measures reflecting optimal performance under highly organized conditions (Toplak et al., 2016).

Finally, task impurity presents a challenge for performance-based tests in assessing isolated EF tests, as each test typically measures more than one executive function and non-EF aspects, such as processing visual or auditory information (Miyake et al., 2000). Similar considerations may extend to self-report questionnaires, as they evaluate daily-life behavior influenced by various factors beyond EF alone. Consequently, both measurements introduce 'noise,' albeit of different kinds. In conclusion, accurately measuring EF proves challenging; however, employing multiple tests provides the most comprehensive assessment of a student's EF levels.

Our findings suggest that self-reported EF have a more significant impact, at least directly on study success, than performance-based EF measures. This implies that students' acquired behaviors and strategies play a more significant role in determining their study success than the efficacy of their cognitive functions. For instance, even if students exhibit weak performance on cognitive tests assessing working memory, they may still excel academically due to their utilization of learned behaviors and strategies, as evidenced by their self-reported EF.

This notion appears particularly applicable to young adults, in line with research on EF development (e.g., Barkley, 2012; Huizinga et al., 2006). Core EF abilities commence development in childhood but remain relatively rudimentary. As individuals progress through childhood and adolescence, they become more adept at managing complex tasks and refining complex EF skills. Consequently, young adults demonstrate enhanced capabilities for learning independent living as they can effectively navigate intricate tasks.

Subsequently, the inclusion of performance-based EF resulted in a statistically marginal increase in predictive value, amounting to only 3%, primarily attributed to working memory. However, the contribution of performance-based EF remains minimal; extensive research has consistently demonstrated that performance-based EF measures, particularly those assessing working memory, exhibit a direct predictive capacity for study success (e.g., Gathercole et al., 2004; Zorza et al., 2016). Conversely, this relationship appears particularly robust in children, as evidenced by a review and meta-analysis conducted by Cortés Pascual et al. (2019), which demonstrated the predictive weight for working memory with an effect size of $r = 0.370$. Thus, although the incremental value of performance-based EF measures is almost negligible, this trend is consistent with findings observed in studies involving children.

In summary, during the early stages of development, the correlation between core EF, particularly working memory, and learning and performance is likely more straightforward. However, as adolescents transition into young adulthood, the influence of core EF abilities assessed through performance measures may become less direct as they have acquired a repertoire of EF strategies and can employ complex EF skills or compensate for weaker EF.

Educational Implications

Educational professionals and students can assess their EF levels using a self-report measurement instrument. While understanding students' EF offers limited predictive insight into their study success after a year, monitoring students' EF can still help identify those struggling with studying and ultimately achieving study success. Since EF are integral to thinking and learning processes, guiding students to optimize their EF can enhance their learning efficacy. As an example of how measuring self-reported EF can be implemented in education, we illustrate how we approach this within our educational environment. Our teaching practice administers online questionnaires as part of a year-long personal and professional development module. These questionnaires are completed by students following psychoeducation sessions on EF. Subsequently, students receive their results via email and can access an instructional video to interpret them. Following this, students participate in individual coaching sessions with their mentors to discuss the results and develop personalized plans to address any identified weaknesses in EF through practice strategies. Additionally, class activities involve small group discussions on EF strategies facilitated by coaching cards containing prompting questions to encourage reflection and sharing experiences.

Performance-based measurements hardly add value and should not be standard implemented in educational practice. Perhaps consideration could be given to measuring performance-based EF in specific cases where the self-reported information is insufficient to help a student move forward. Alternatively, in cases where performance-based measurements provide information on how cognitions function differently because of a brain disorder or a developmental gap such as ADHD. However, unlike self-reported EF measurements, its integration into education is not practically effortless. They are labor-intensive neuropsychological tests, mainly at a cost, and interpretation of the results requires specific diagnostic and neuropsychological knowledge (Naglieri & Goldstein, 2014).

Conclusion

Self-reported and performance-based EF together account for 16% of the variance in study success after an academic year, with self-reported EF contributing the most. Despite the relatively modest predictive value, gaining insights into students' EF levels is valuable, as EF are critical for learning, and approximately one in five students report significant EF difficulties. Therefore, further investigation into the implementation of EF measurements in educational practice may yield a better understanding of students' learning and study success.

Recommendations

Future research could focus on identifying the most suitable measurement instruments for integration into educational practice to predict study success, taking into account factors like age or clinical conditions. Furthermore, additional research could address questions such as "What indicators identify students at risk?", "What type of information provided by these instruments is beneficial for students?" and "What are the appropriate cut-off scores for the student population?"

If freely available instruments, which are quick and easy to administer and provide feedback to students, were accessible, applied research could explore strategies for delivering these results to students effectively. For instance, how could student coaches or mentors engage in discussions with students regarding the results? Subsequently, the disparities between performance-based and self-reported EF could be further examined to determine their impact on learning and study success. For example, to what extent does it pose a challenge if a student reports high EF levels, indicating perceived competence, but performs poorly on neuropsychological tests, suggesting suboptimal functioning of core EF skills? Lastly, exploring the underlying mechanisms of EF in the educational context could yield valuable insights. For instance, do EF serve as mediators for other concepts aside from self-regulated learning, such as procrastination?

Educational professionals and students can utilize self-report measurement instruments to gain insights into their EF levels. By monitoring students' EF, educational institutions can identify at-risk students and implement interventions, such as psychoeducation, to help them comprehend their EF profile.

Limitations

A strength of this study is that it combines two types of EF measurements, thus measuring the concept of EF in multiple ways, contributing to better construct validity in general (McDonald, 2008). In addition, the measurement instruments measure multiple variables, allowing multiple correlations (i.e., correlations) to be calculated, which is desirable within social science research to make statements about the hypothetical constructs from different perspectives (Samuel & Okey, 2015).

A limitation of the CNSVS is that it is not intended to replace formal neuropsychological testing (Gualtieri & Johnson, 2006). Additionally, the various abilities measured by the CNSVS are inter-correlated, meaning that individual cognitive functions cannot be assessed independently (Plourde et al., 2018). As a result, the CNSVS should not be used to make determinations about neuropsychological functioning in the context of clinical conditions such as brain disorders or neurobiological developmental disorders like ADHD, nor should educational professionals conclude "malfunctioning EF." Furthermore, the tests within the CNSVS do not measure EF in isolation.

Although the CNSVS is not designed for accurate neuropsychological diagnosis, it can screen for mild cognitive impairments quickly and cost-effectively. Moreover, the administrative burden of administering and scoring the tests is minimal (Bauer et al., 2012). Additionally, changes in cognitive function can be tracked over time, allowing for assessment of level and monitoring of progress (e.g., Bauer et al., 2012; Brooks & Sherman, 2012). Consequently, the CNSVS represents a valuable tool for research purposes and may hold promise for application in educational settings.

Ethics Statements

Ellis van Dooren-Wissebom, chairperson of the LLM Saxion Ethical Advice Committee (SEAC), Saxion University of Applied Sciences, reviewed and approved the studies involving human participants. The participants provided their written informed consent for participation in this study.

Conflict of Interest

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

This study was supported by a grant from Saxion University of Applied Sciences.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors and is available via: https://doi.org/10.17026/dans-24g-bc5a

Authorship Contribution Statement

Manuhuwa: Concept and design, data acquisition, data analysis and interpretation, manuscript drafting, writing, securing funding. Snel-De Boer: Data acquisition, data analysis and interpretation, revision Method and Result section. De Graaf: Editing/reviewing, supervision, final approval. Fleer: Editing/reviewing, supervision, final approval

References

Abeare, K., Razvi, P., Sirianni, C. D., Giromini, L., Holcomb, M., Cutler, L., Kuzmenka, P., & Erdodi, L. A. (2021). Introducing alternative validity cutoffs to improve the detection of non-credible symptom report on the BRIEF. *Psychological Injury and Law, 14,* 2-16. https://doi.org/10.1007/s12207-021-09402-4

Alloway, T. P., Banner, G. E., & Smith, P. (2010). Working memory and cognitive styles in adolescents' attainment. *British Journal of Educational Psychology*, *80*(4), 567-581. https://doi.org/10.1348/000709910X494566

Ato, M., López-García, J. J., & Benavente, A. (2013). A classification system for research designs in psychology. *Anales de Psicología, 29*(3), 1038-1059. https://doi.org/10.6018/analesps.29.3.178511

Baars, M. A. E., Nije Bijvank, M., Tonnaer, G. H., & Jolles, J. (2015). Self-report measures of executive functioning are a determinant of academic performance in first-year students at a University of Applied Sciences. *Frontiers in Psychology*, *6,* Article 1131. https://doi.org/10.3389/fpsyg.2015.01131

Baddeley, A. (2010). Working memory. *Current Biology, 20*(4), 136-140. https://doi.org/10.1016/j.cub.2009.12.014

Baggetta, P., & Alexander, P. A. (2016). Conceptualization and operationalization of executive function. *Mind, Brain, and Education, 10*(1), 10-33. https://doi.org/10.1111/mbe.12100

Barkley, R. A. (2012). *Executive functions. What they are, how they work and why they evolved.* Guilford Press.

Barkley, R. A., & Fischer, M. (2011). Predicting impairment in major life activities and occupational functioning in hyperactive children as adults: Self-reported executive function (EF) deficits versus EF tests. *Developmental Neuropsychology, 36*(2), 137-161. https://doi.org/10.1080/87565641.2010.549877

Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized neuropsychological assessment devices: Joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *The Clinical Neuropsychologist, 26*(2), 177-196. https://doi.org/10.1080/13854046.2012.663001

Bodenburg, S., Wendiggensen, J., & Kasten, E. (2022). Scores in self-report questionnaires assessing adult ADHD can be influenced by negative response bias but are unrelated to performance on executive function and attention tests. *Psychological Injury and Law, 15*, 189-199. https://doi.org/10.1007/s12207-022-09448-y

Brooks, B. L., & Sherman, E. M. S. (2012). Computerized neuropsychological testing to rapidly evaluate cognition in pediatric patients with neurologic disorders. *Journal of Child Neurology, 27*(8), 982-991. http://doi.org/10.1177/0883073811430863

Campman, C., Van Ranst, D., Meijer, J. W., & Sitskoorn, M. (2017). Computerized screening for cognitive impairment in patients with COPD. *International Journal of Chronic Obstructive Pulmonary Disease, 12,* 3075-3083. https://doi.org/10.2147/COPD.S142871

Carragher, J., & McGaughey, J. (2016). The effectiveness of peer mentoring in promoting a positive transition to higher education for first-year undergraduate students: A mixed methods systematic review protocol. *Systematic Reviews, 5,* Article 68. https://doi.org/10.1186/s13643-016-0245-1

CNS Vital Signs. (2023). *CNS Vital Signs interpretation guide* https://bit.ly/3XbicLj

Cortés Pascual, A., Moyano Muñoz, N., & Quilez Robres, A. (2019). The relationship between executive functions and academic performance in primary education: Review and meta-analysis. *Frontiers in Psychology, 10,* Article 1582. https://doi.org/10.3389/fpsyg.2019.01582

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Effeney, G., Carroll, A., & Bahr, N. (2013). Self-regulated learning and executive function: Exploring the relationships in a sample of adolescent males. *Educational Psychology, 33*(7), 773-796. https://doi.org/10.1080/01443410.2013.785054

Follmer, D. J. (2021). Examining the role of calibration of executive function performance in college learners' regulation. *Applied Cognitive Psychology, 35*(3), 646-658. https://doi.org/10.1002/acp.3787

Gaskin, C. J., & Happell, B. (2014). On exploratory factor analysis: A review of recent evidence, an assessment of current practice, and recommendations for future use. *International Journal of Nursing Studies, 51*(3), 511-521. https://doi.org/10.1016/j.ijnurstu.2013.10.005

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*(1), 1-16. https://doi.org/10.1002/acp.934

Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of pediatric brain development: What have we learned and where are we going? *Neuron, 67*(5), 728-734. https://doi.org/10.1016/j.neuron.2010.08.040

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual researchers. *Personality and Individual Differences, 102,* 74-78. https://doi.org/10.1016/j.paid.2016.06.069

Gualtieri, C. T., & Johnson, L. G. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Archives of Clinical Neuropsychology, 21*(7), 623-643. https://doi.org/10.1016/j.acn.2006.05.007

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). SAGE Publications.

Huizinga, M., Dolan, C. V., & Van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*(11), 2017-2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Kamradt, J. M., Ullsperger, J. M., & Nikolas, M. A. (2014). Executive function assessment and adult attention-deficit/hyperactivity disorder: Tasks versus ratings on the Barkley Deficits in Executive Functioning Scale. *Psychological Assessment, 26*(4), 1095-1105. https://doi.org/10.1037/pas0000006

Knouse, L. E., Feldman, G., & Blevins, E. J. (2014). Executive functioning difficulties as predictors of academic performance: Examining the role of grade goals. *Learning and Individual Differences, 36,* 19-26. https://doi.org/10.1016/j.lindif.2014.07.001

Lorenzo-Seva, U., & Ferrando, P. J. (2021). Not positive definite correlation matrices in exploratory item factor analysis: Causes, consequences and a proposed solution. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(1), 138-147. https://doi.org/10.1080/10705511.2020.1735393

Manuhuwa, D. M., Snel-De Boer, M., Jaarsma, A. D. C., Fleer, J., & De Graaf, J. W. (2023). Executive functions and self-regulated learning profiles in higher education students and their relationship with study success. In *INTED2023*

Proceedings: 17th International Technology, Education and Development Conference (INTED) (pp. 2507-2516). IATED. https://doi.org/10.21125/inted.2023

McDonald, J. D. (2008). Measuring personality constructs: The advantages and disadvantages of self-reports, informant reports and behavioural assessments. *Enquire, 1*(1), 75-94. https://bit.ly/3xo9TAW

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167-202. https://doi.org/10.1146/annurev.neuro.24.1.167

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. https://doi.org/10.1006/cogp.1999.0734

Musso, M. F., Boekaerts, M., Segers, M., & Cascallar, E. C. (2019). Individual differences in basic cognitive processes and self-regulated learning: Their interaction effects on math performance. *Learning and Individual Differences, 71*, 58-70. https://doi.org/10.1016/j.lindif.2019.03.003

Naglieri, J. A., & Goldstein, S. (2014). Assessment of executive function using rated scales: Psychometric considerations. In S. Goldstein, & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 159-170). Springer. https://doi.org/10.1007/978-1-4614-8106-5_10

Naglieri, J. A., & Otero, T. M. (2017). *Essentials of CAS2 Assessment*. John Wiley & Sons, Inc.

Nigg, J. T. (2017). Annual research review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychology. *Journal of Child Psychology and Psychiatry, 58*(4), 361-383. https://doi.org/10.1111/jcpp.12675

Plourde, V., Hrabok, M., Sherman, E. M. S., & Brooks, B. L. (2018). Validity of a computerized cognitive battery in children and adolescents with neurological diagnoses. *Archives of Clinical Neuropsychology, 33*(2), 247-253. https://doi.org/10.1093/arclin/acx067

Rabinovici, G. D., Stephens, M. L., & Possin, K. L. (2015). Executive dysfunction. *Continuum, 21*(3), 646-659. https://doi.org/10.1212/01.CON.0000466658.05156.54

Ramos-Galarza, C., Acosta-Rodas, P., Bolaňos-Pasquel, M., & Lepe-Martínez, N. (2020). The role of executive functions in academic performance and behaviour of university students. *Journal of Applied Research in Higher Education, 12*(3), 444-455. https://doi.org/10.1108/JARHE-10-2018-0221

Roth, R. M., Isquith, P. K., & Gioia, G. A. (2005). *Behavior Rating Inventory of Executive Function®--Adult Version* (BRIEF-A) [Database record]. APA PsycTests. https://doi.org/10.1037/t86244-000

Roth, R. M., Isquith, P. K., & Gioia, G. A. (2011). *BRIEF-A: Vragenlijst executieve functies voor volwassenen* (E. Scholte, & I. Noens, Trans.) HogrEFe. (Original work published 2005). https://bit.ly/4bYNXeO

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

Rutherford, T., Buschkuehl, M., Jaeggi, S. M., & Farkas, G. (2018). Links between achievement, executive functions, and self-regulated learning. *Applied Cognitive Psychology, 32*(6), 763-774. https://doi.org/10.1002/acp.3462

Samuel, M., & Okey, L. E. (2015). The relevance and significance of correlation in social science research. *International Journal of Sociology and Anthropology Research, 1*(3), 22-28. https://bit.ly/3Kxq2HG

Shwartz, S. K., Roper, B. L., Arentsen, T. J., Crouse, E. M., & Adler, M. C. (2020). The Behavior Rating Inventory of Executive Function®-Adult Version is related to emotional distress, not executive dysfunction, in a veteran sample. *Archives of Clinical Neuropsychology, 35*(6), 701-716. https://doi.org/10.1093/arclin/acaa024

Snyder, H. R., Friedman, N. P., & Hankin, B. L. (2021). Associations between task performance and self-report measures of cognitive control: Shared versus distinct abilities. *Assessment*, *28*(4), 1080-1096. https://doi.org/10.1177/1073191120965694

Suchy, Y., DesRuisseaux, L. A., Mora, M. G., Brothers, S. L., & Niermeyer, M. A. (2024). Conceptualization of the term "ecological validity" in neuropsychological research on executive function assessment: A systematic review and call to action. *Journal of the International Neuropsychological Society, 30*(5), 499-522. https://doi.org/10.1017/S1355617723000735

Suchy, Y., Mora, M. G., DesRuisseaux, L. A., Niermeyer, M. A., & Brothers, S. L. (2024). Pitfalls in research on ecological validity of novel executive function tests: A systematic review and a call to action. *Psychological Assessment*, *36*(4), 243-261. https://doi.org/10.1037/pas0001297

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, 54*(2), 131-143. https://doi.org/10.1111/jcpp.12001

Toplak, M. E., West, R. F., & Stanovich, K. E. (2016). The assessment of executive functions in Attention-Deficit/Hyperactivity Disorder: Performance-based measures versus ratings of behavior. In J. Hoskyn (Ed.), *Executive functions in children's everyday lives: A handbook for professionals in applied psychology* (pp. 157-174). Oxford Academic. https://doi.org/10.1093/acprof:oso/9780199980864.003.0011

VanLunen, D. (2020, March 15). *Get a grip! When to add covariates in a linear regression. A guide to accurately and precisely measuring effects!* Towards Data Science. https://bit.ly/4amWRSO

Velichkovsky, B. B., Razvaliaeva, A. Y., Khlebnikova, A. A., Manukyan, P. A., Kasatkin, V. N., & Barmin, A. V. (2023). Systematic Review and Meta-Analysis of Clinically Relevant Executive Functions Tests Performance after COVID-19. *Behavioural Neurology, 2023*, Article 1094267. https://doi.org/10.1155/2023/1094267

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology, 44*(3), 219-246. https://doi.org/10.1177/0095798418771807

Zelazo, P. D., & Carlson, S. M. (2020). The neurodevelopment of executive function skills: Implications for academic achievement gaps. *Psychology and Neuroscience, 13*(3), 273-298. https://doi.org/10.1037/pne0000208

Zelazo, P. D., & Müller, U. (2010). Executive function in typical and atypical development. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 574-603). Wiley-Blackwell. https://doi.org/10.1002/9781444325485.ch22

Zorza, J. P., Marino, J., & Mesas, A. A. (2016). Executive functions as predictors of school performance and social relationships: Primary and secondary school students. *The Spanish Journal of Psychology, 19,* Article E23. https://doi.org/10.1017/sjp.2016.23

Appendices

Table A1. Appendix 1: Subscales and items BRIEF-A

Subscale	Items	Dutch translation
Inhibit (INH)	5. I drum my fingers or wiggle my legs.	Ik trommel met mijn vingers of wiebel met mijn benen.
	16. I have difficulty staying still.	Ik heb moeite om stil te blijven zitten.
	29. I find it hard to wait my turn.	Ik vind het moeilijk om op mijn beurt te wachten.
	36. I make inappropriate sexual comments.	Ik maak ongepaste seksueel getinte opmerkingen.
	43. I make decisions that get me into trouble (legally, financially, socially).	Ik neem beslissingen die me in de problemen brengen (wettelijk, financieel, sociaal).
	55. People say I am easily distracted.	Men zegt dat ik snel afgeleid ben.
	58. I rattle things off.	Ik raffel dingen af.
	73. I am impulsive.	Ik ben impulsief.
Shift (SH)	8. I have difficulty making the transition from one activity or task to another.	Ik heb moeite om de overgang van de ene activiteit of taak naar de andere te maken.
	22. I have difficulty accepting that there are different ways to solve problems at work, with friends, or with tasks.	Ik heb moeite te accepteren dat er verschillende manieren zijn om problemen met werk, met vrienden of met taken op te lossen.
	32. I struggle to think of another solution to a problem when I am stuck.	Ik heb moeite om een andere oplossing voor een probleem te bedenken als ik ben vastgelopen.
	44. I have a hard time coping with change.	Ik kan slecht tegen veranderingen.
Emotional control (EC)	61. I get upset at unexpected changes in my daily routine.	Ik raak van slag bij onverwachte veranderingen in mijn dagelijkse routine.
	67. I don't get over problems easily.	Ik zet me niet gemakkelijk over problemen heen.
	1. I have outbursts of anger.	Ik heb woede-uitbarstingen.
	12. I overreact emotionally.	Ik reageer overdreven emotioneel.
	19. I can react very emotionally to small things.	Ik kan om kleine dingen zeer emotioneel reageren.
	28. I react more emotionally to situations than my friends.	Ik reageer meer emotioneel op situaties dan mijn vrienden.
	33. I overreact to minor problems.	Ik reageer overdreven op kleine problemen.
	42. I get emotionally upset quickly.	Ik raak emotioneel snel overstuur.
	51. My angry outbursts are intense but quickly over.	Mijn woede-uitbarstingen zijn hevig, maar snel weer voorbij.
	57. People say I get emotional too quickly.	Men zegt dat ik te snel emotioneel ben.
	69. My temper turns quickly.	Mijn humeur slaat snel om.
	72. Little things upset me quickly or easily.	Ik raak door kleine dingen snel of gemakkelijk van streek.
Self-monitor (SM)	13. If I cause others to feel bad or get angry, I don't notice until it's already too late.	Als ik ervoor zorg dat anderen zich slecht voelen of boos worden, merk ik dat pas op als het al te laat is.
	23. I talk at the wrong times.	Ik praat op verkeerde momenten.
	37. When people seem upset by me, I don't understand why.	Als mensen van streek lijken door mij snap ik niet waarom.
	50. I say things without thinking	Ik zeg dingen zonder na te denken
	64. People say I do not think before I do something.	Men zegt dat ik niet nadenk voor ik iets doe.
	70. I don't think about the consequences before I do something.	Ik denk niet na over de gevolgen voordat ik iets doe.
Initiate (INI)	6. I need to be reminded to start a task, even if I am ready to perform the task.	Ik moet eraan herinnerd worden om met een taak te beginnen, zelfs als ik bereid ben om die taak uit te voeren.
	14. I struggle to get ready for the day.	Ik heb moeite om me klaar te maken voor de dag.
	20. I just hang out at home.	Ik hang thuis maar wat rond.
	25. I have trouble getting to work independently.	Ik heb moeite om zelfstandig aan de slag te gaan.
	45. I find it hard to be enthusiastic about things.	Ik vind het moeilijk om enthousiast over dingen te zijn.
	49. I have difficulty starting tasks.	Ik heb moeite om aan taken te beginnen.
	53. I only start things (e.g., assignments, chores, tasks) at the last minute.	Ik begin pas op het nippertje aan dingen (bijvoorbeeld opdrachten, karweitjes, taken).
	62. I have difficulty thinking of things to do in my free time.	Ik heb moeite om dingen te bedenken die ik in mijn vrije tijd kan doen.

Table A1. Continued

Subscale	Items	Dutch translation
Working memory (WM)	4. I have difficulty concentrating on tasks (e.g. chores, reading or work).	Ik heb moeite om me te concentreren op taken (bijvoorbeeld karweitjes, lezen of werk).
	11. I have difficulty with tasks requiring more than one step.	Ik heb moeite met opdrachten of taken die meer dan één stap vereisen.
	17. Halfway through an activity, I forget what I was doing.	Halverwege een activiteit vergeet ik wat ik aan het doen was.
	26. I have difficulty staying on the same topic while talking.	Ik heb moeite om tijdens het praten bij hetzelfde onderwerp te blijven.
	35. I can only concentrate for a short time.	Ik kan me maar kort concentreren.
	46. I forget directions quickly.	Ik vergeet aanwijzingen snel.
	56. I have trouble remembering things even for a few minutes (such as directions and phone numbers).	Ik heb moeite om dingen te onthouden, zelfs voor een paar minuten (zoals aanwijzingen, telefoonnummers).
	68. I have difficulty doing more than one thing at a time.	Ik heb moeite om meer dan één ding tegelijk te doen.
Plan (PL)	9. I get overwhelmed by big tasks.	Ik word overweldigd door grote taken.
	15. I find it difficult to prioritise activities.	Ik vind het lastig om prioriteiten te stellen bij activiteiten.
	21. I start tasks (e.g., cooking and projects) without suitable materials.	Ik begin aan taken (bijvoorbeeld koken, projecten) zonder de juiste materialen ter beschikking te hebben.
	34. I don't plan future activities in advance.	Ik plan toekomstige activiteiten niet van tevoren.
	39. I set unrealistic goals.	Ik stel onrealistische doelen.
	47. I have good ideas but can't get them down on paper.	Ik heb goede ideeën, maar krijg ze niet op papier.
	54. I struggle to complete tasks independently.	Ik heb moeite om taken zelfstandig af te maken.
	63. I do not plan tasks.	Ik plan taken niet vooruit.
	66. I have difficulty organizing activities.	Ik heb moeite bij het organiseren van activiteiten.
	71. I have difficulty organising my work.	Ik heb moeite om mijn werk te organiseren.
Task-monitor (TM)	2. I make carelessness mistakes when making tasks.	Ik maak slordigheidsfouten bij het maken van taken.
	18. I don't check my work for errors.	Ik controleer mijn werk niet op fouten.
	24. I misjudge how difficult or easy tasks will be.	Ik schat verkeerd in hoe moeilijk of gemakkelijk taken zullen zijn.
	41. I make sloppy mistakes.	Ik maak slordigheidsfoutjes.
	52. I have difficulty completing tasks (e.g. chores, work).	Ik heb moeite om taken (bijvoorbeeld karweitjes, werk) af te maken.
	75. I have problems finishing my work.	Ik heb problemen om mijn werk af te maken.
Organization of materials (OM)	3. I cannot organize well.	Ik kan niet goed organiseren.
	7. I have a messy closet.	Ik heb een rommelige kast.
	30. People say I can't organise well.	Men zegt dat ik niet goed kan organiseren.
	31. I lose things (e.g. keys, money, wallet, homework, etc.).	Ik raak dingen kwijt (bijvoorbeeld sleutels, geld, portemonnee, huiswerk, etc.).
	40. I leave the bathroom messy.	Ik laat de badkamer rommelig achter.
	60. I leave my room or house messy.	Ik laat mijn kamer of huis rommelig achter.
	65. I have trouble finding things in my room, closet or desk.	Ik heb moeite om dingen te vinden in mijn kamer, kast of bureau.
	74. I don't tidy up my things.	Ik ruim mijn spullen niet op.

Table A2. Appendix 2: Latent variables and included items.

Latent variables	Included items	Cronbach's alpha or Spearman-Brown coefficient	Factor loadings between
Physical turmoil	INH 5, 16	.71	.64 and .84
Accept change	SH 44, 61, 67	.71	.54 and .76
React easily emotionally	EM 12, 19, 28, 33, 42, 51, 69, 72	.90	.51 and .82
Anger outbursts	EM 1, 51	.69	.74 and .75
React impulsive	SM 50, 64, 70	.72	.62 and .77
Concentrate	WM 4, 35	.74	.76 and .77
Remember things	WM 11, 17, 46, 56, 68	.71	.57 and .63
Initiate	INI 25, 49, 53	.75	.66 and .79
Plan ahead	PL 34, 63	.78	.72 and .89
Make sloppy mistakes	TM 2, 41	.70	.71 and .76
Keep environment organized	OM 7, 40, 60, 65, 74	.82	.51 and .85

Table A3. Appendix 3: Original scales and removed items.

Original scales	Removed items
Inhibition	29, 36, 55, 58, 73
Shift	8, 22, 32
Emotional control	
Self-monitor	13, 23, 37
Working memory	26
Initiate	6, 14, 20, 45, 62
Plan	9, 15, 21, 39, 47, 54, 66, 71
Task-monitor	18, 24, 52, 75
Organization of materials	3, 30, 31