Validating Rating Scales for Executive Functioning across Education Levels and Informants

Lucie C. Spreij[1] · Cathy Van Tuijl[2,3] · Paul P. M. Leseman[2]

Abstract

The goal of the study was to deliver and validate a new solution-focused instrument, the Focus Skills Questionnaire (FSQ), to assess the degree of executive functioning skills in the school context, for three different education levels (elementary, secondary, and tertiary education) and informant groups (students, teachers and parents) on a sample of 1109 students from Dutch and Belgian schools. The factor structure was evaluated by confirmative factor analysis (CFA) and the study examined how students' self-reports of executive functioning skills related to outcomes of neuro-psychological tests of executive functions (EF). The CFA results showed a parsimonious model with a four-factor structure of the FSQ that was equivalent for all education levels and informant groups, but that does not correspond with the generally assumed executive functioning factors. There are differences in the perception of executive functioning skills by different informant groups and also differences per education level. Student perceptions of executive functioning skills do not correspond with EF test outcomes and in some subgroups clearly diverge from teacher or parent perceptions of the students' executive functioning skills. Although the new instrument does not converge with laboratory assessments of EF's, the new instrument could be useful in everyday school practice.

Keywords Confirmative factor analysis · Education levels · Executive functioning · Parents · Solution-focused rating scales · Teachers

Introduction

Executive functioning (EF), also called executive or cognitive control (Diamond, 2013), can be regarded as higher level processes that enable individuals to regulate their thoughts and actions during goal-directed behavior in new or complex situations of daily life (Davidson et al., 2006; Friedman, & Miyake, 2017; Miyake et al., 2012). EFs develop from birth during school years to adulthood in relation to the development of neural systems and what is asked from the EFs (Alloway & Alloway, 2013; Best & Miller, 2010; Crone et al., 2017; Davidson et al., 2006; Huizinga et al., 2006). EF processes are taking place in complex neural networks in the pre-frontal cortex interacting with other areas in the brain. These processes create new neurological networks and develop EF skills such as planning and initiating skills to complete chores and homework. These proactive processes of self-regulation, so called self-regulated learning (SRL), enable students to acquire academic skills to reach their learning and achievement goals (Boekaerts & Corno, 2005; Zimmerman, 2008). Therefore, for education it is important to assess to what extent EFs of students are developed to support their self-regulation (Etkin, 2018; Korinek & DeFur, 2016).

EF has been defined by many scientists and there is discussion about the nature and number of executive functions. Based on laboratory research with neuropsychological tests, most researchers assume a system with three basic (but related) components: inhibition, working memory (updating) and cognitive flexibility or shifting (Best et al., 2011; Friedman & Miyake, 2017). Regardless of the number of EF components found in laboratory research, the question is whether EF measured in an

✉ Lucie C. Spreij
  focusluciespreij@gmail.com

1 University of Applied Sciences, Hogeschool Utrecht, Utrecht, The Netherlands

2 Social Sciences, Utrecht University, Utrecht, The Netherlands

3 Saxion University of Applied Sciences, Enschede, The Netherlands

isolated laboratory context is predictive in the context of education, in classes with distraction and group dynamics. Adding teachers and parents reports to students reporting on their EF skills could deliver more valid data in educational practice, especially young students may be unaware of gaps in their EF skills. This study aims to examine how students' self-reports of EF skills relate to teacher and parent reports and to outcomes of neuro-psychological tests of EF.

Assessment of EF

To assess EF for youth and adults, neuropsychological tests such as the Tower of London Test (TLT; Kovács, 2013) are available as well as standardized EF assessment scales such as the Behavior Rating Inventory of Executive Functions (BRIEF; Gioia et al., 2000; Huizinga & Smidts, 2013) for 5- to 18-year-old, and self-assessment scales for 12- to 18-year-old. Both neuropsychological tests and questionnaires are developed to assess EF problems, caused by traumatic brain injuries or impaired development of EF whether or not in relation to psychiatric disorders. To assess the degree of EF skills of students by teachers, neuropsychological tests as well as questionnaires such as the BRIEF are not suitable for several reasons. First, the ecological validity of neuropsychological EF tests and EF questionnaires is questionable: to what extent do EF skills of the student during the test or the items of the questionnaire correspond with behavior in the school context? In addition, the predictive value of EF tests for school outcomes appears to be limited. According to Poon (2018), only inhibition measured by Stroop Color and Word test (Stroop, 1935) and cognitive flexibility contribute to the prediction of school results. Huizinga and Smidts (2011) argue that in research on EF of students in elementary and secondary schools no significant or only low correlations were reported among the BRIEF (representing behavior in the school context) and EF tests like Tower of London, Test of Variables of Attention and Rey Complex Figure. In addition, Becker and Langberg (2014) reported no consistent relation between Sluggish Cognitive Tempo (SCT) and EF in adolescents with Attention-Deficit/Hyperactivity Disorder (ADHD) when laboratory-based neuropsychological tasks of EF were used. These studies hypothesize that this disparity is largely due to the low ecological validity of EF tests. Finally, the existing EF tests have been developed for assessment by psychologists and not by teachers. EF tests administered by teachers are less reliable than those administered by psychologists, because teachers are not trained to administer and interpret EF tests. In addition, EF tests measures individuals potential in a laboratory

setting, without distractions which are part and parcel of the school context, making the results not generalizable to school context.

Dawson and Guare (2010) developed the Executive Skills Questionnaire (ESQ) not only for psychologists but also for teachers in elementary and secondary education. The authors aimed to offer students themselves (10- to 18-year-old), teachers and parents of students in elementary and secondary education, more insight into the weak and strong EF skills. An advantage of the ESQ is the possibility to work from an ecological perspective (Bronfenbrenner, 1975; Bronfenbrenner & Evans, 2000). A disadvantage is that ESQ items are formulated negatively as problem-oriented statements about student behaviors, whereas schools are more interested in solutions than in problems. Although the ESQ is developed for the school context, there is a need for more solution-focused questionnaire to assess the degree of EF skills with student versions for self-assessment for elementary, secondary and tertiary education.

Present Study

The main purpose of the present study is to validate the FSQ for three different education levels (elementary, secondary and tertiary education) and informant groups (students, teachers and parents) and to evaluate the psychometric properties. First, confirmatory factor analysis will be used to investigate to what extent the factor structure of EF from the FSQ is equivalent for all education levels (elementary, secondary or tertiary education) and informant groups (students, teachers, parents) and whether this factor structure delivers a more parsimonious model than the ten-factor model as proposed by Dawson and Guare (2010). We expect to find a factor structure with a limited (e.g., 3–5) number of factors. A second aim of this study is to investigate whether there are significant differences in EF skills among education levels (elementary, secondary and tertiary education) and informant groups respectively (student, teacher, parent). We expect significant differences in EF skills among education levels as well as among informant groups, because of their different knowledge and views of behavior in the school context across different education levels. We expect higher means of EF skills in all informant groups for tertiary than for secondary education, in relation to the development of neural systems and demands from EFs in secondary and tertiary education. We expect higher means of EF skills for students than for teachers because students, especially in elementary education due to their limited knowledge of EFs, may overestimate their EF skills. We also expect higher means of EF scales rated by parents than rated by teachers, due to the different view of behavior at school

and at home. The third aim is to investigate to what extent the FSQ scales of students correspond with concurrently administered external EF tests. We expect weak to moderate associations between the FSQ factors and EF test outcomes because of the different settings to which these instruments apply.

Method

Participants and Procedure

In this study, 29 schools in the Netherlands and Flanders, the Dutch-speaking part of Belgium participated (Tables 1 and 2). Student participants were diverse in age, gender, educational level of the parents, and ethnic background and they visited schools in both urban and rural areas. In the Netherlands, the schools were recruited through the students of the Master Educational Needs program of the Dutch University of Applied Science in Utrecht. Schools for elementary, secondary or tertiary education were asked to participate in this study with whole classes (students, teacher and parents). In Flanders, visitors of the Conference "Action-oriented collaboration on qualitative education" of September 2014 in Antwerp (Belgium) were also asked to participate with their schools. For students of tertiary education, universities and universities of applied science in the Netherlands and in Belgium were contacted and asked to participate. All schools were informed about the study and received an e-mail containing relevant information for the students and the teachers of the classes. After the schools and the teachers gave written permission, the parents of the classes received a letter with information about the study and were asked to give written permission for participation of their children. The students of universities (tertiary education) were asked for permission themselves. In the Netherlands and in Flanders children visit elementary school to the age of 12 and then start secondary school in different tracks: prevocational school for four years to academic education for 5 to 6 years. Tertiary education start to the age of 17 or 18.

All participating schools received an e-mail, with links to the digital questionnaires for students, teachers and parents. All schools were visited by the primary researcher and trained research assistants to assist the students in completing the digital Focus Skills Questionnaire (FSQ) for EF (developed by Spreij and Klapwijk in 2014 as Solution-focused tuning in education at Hogeschool Utrecht) on available devices (computer, laptop or tablet). The primary researcher and trained research assistants tested the students individually in a separated

Table 1 Distribution student sample across education levels in The Netherlands and in Belgium

Education level	The Netherlands				Belgium				Total
	Dutch regular schools	Dutch students regular schools	Dutch special schools	Dutch students special schools	Belgian regular schools	Belgian students regular schools	Belgian special schools	Belgian students special schools	Dutch and Belgian students
Elementary education (8 to 12 years)	6	207	1	15	1	63	1	38	270
Secondary education (12 to 18 years)	7	530	2	56	2	146	1	18	676
Tertiary education (18 to 24 years)	7	145			1	18			163
Total	20	882	3	71	4	227	2	56	1109

Table 2 Distribution boys and girls across education levels and age groups

Education level	Age group	Boys	Girls	Dutch and Belgian students
Elementary education	8 to 12 years	150	120	270
Secondary education	12 to 18 years	351	325	676
Tertiary education	18 to 24 years	59	104	163
Total		560	549	1109

room using the Tower of London Test (TLT) (Kovács, 2013) and the Test of Sustained Selective Attention (TOSSA) (Kovács, 2010). In the same month teachers and parents filled out the digital FSQ questionnaires at home on their own devices.

A total number of 1163 student questionnaires were received, 301 teacher questionnaires and 362 parent questionnaires. Due to internet disruptions, some questionnaires were filled in twice. Only the results of the first and fully completed FSQ questionnaires are included in this study: 1109 students, 235 teachers and 285 parents. TLT and TOSSA results are available for about one third of the students.

Measures

Focus Skills Questionnaire (FSQ)

To construct an EF instrument that reliably reports the degree of EF skills by students, teachers and parents, a modified version of the Dutch ESQ (Dawson et al., 2010), the Focus Skills Questionnaire (FSQ) was developed (Spreij & Klapwijk, personal communication 2014). To construct the FSQ the first nine scales of the ESQ were used: Inhibition (IN), Working memory (WM), Emotion regulation (ER), Sustained attention (SA), Task initiation (TI), Planning (PL), Organizing (OR), Time management (TM) and Flexibility (FB). Two ESQ scales, Metacognition (MC), and Goal oriented perseverance (GP), have been merged to a new scale Monitoring (MO) with three items. The thirty items were formulated positively and as a skill with three items per scale for elementary as well as for secondary education. In addition, because EFs develop into adulthood (Alloway & Alloway, 2013) the same questionnaire has also been constructed for tertiary education. The ESQ problem-oriented 5-point Likert scale (1 = big problem; 2 = moderate problem; 3 = mild problem; 4 = slight problem; 5 = no problem) was changed in a solution-focused 5-point Likert scale (1 = behavior not at all present: 2 = a little present; 3 = more or less present; 4 = more than sufficient present; 5 = fully present). The positively formulated items of the EF questionnaires

for parents (Dawson et al., 2009) served as inspiration to formulate the items in a solution-focused way (Kim & Franklin, 2009). After reformulation of the items, school and educational psychologists and teachers from universities, special schools and school counseling centers in the Netherlands and in Belgium gave feedback on the phrasing of the items of the FSQ and the terminology of the five-point Likert scale. After processing the feedback, all questionnaires were digitized.

To fill in the FSQ students are asked for every item to choose the best description for the way how they work and act (1 = behavior not at all present; 5 = fully present). Teachers and parents are asked to indicate for the same items how the student or their child works and acts. Teachers rates all the students from his or her class. The thirty items are categorized into ten scales with three items and the following description (Table 3).

Tower of London Test

Academic performance take place in the school context and involves a complex of factors, so predicting academic performance requires a test with a complex task. The Tower of London Test (TLT) (Kovács, 2013) was chosen, because it is an often used test to assess planning ability, a complex EF, for use in clinical practice (Michalec et al., 2017). More over the TLT can be administered using a laptop which enhances the reliability of the results. Since most other EF tests measure less complex tasks as attention or inhibition, we chose to use the Tower of London Test.

The goal of this test is to rearrange three colored cubes (red, yellow and blue) from their initial position on three upright pegs to a new set of predetermined positions on one or more of the pegs, in as few moves as possible. There are 16 test items with a gradually increasing level of complexity. The norms are based on a group of neurological patients (range 25 to 81 year) and a group of healthy controls (range 14 to 91 year). Although the norming sample of the TLT did not include students of elementary education, the test was used exploratively for this age group, so the scores of the different educational levels could be compared with each other. The following indices are used in this study, because these two indices represent the main aspects of planning skills and 'looking ahead' (Kovács, 2013).

1. Total Score (TS): the TS are the total points of all items. The first attempt of rearranging the colored cubes is rewarded much higher than the second attempt and the more difficult items (9 till 16) generate more points than the simpler items.
2. Decision Time (DT): the time between presenting the item and touching the first cube. It is the time to think ahead, to plan the movement(s).

Table 3 Description FSQ scales executive functioning

	FSQ Scales EF	Description FSQ scales	Item example
1	Inhibition (IN)	First thinking before doing something, postponing the reaction, creating the possibility to assess a situation before it is acted upon	I think first before I speak
2	Working memory (WM)	The ability to hold information in memory while performing complex tasks	I remember tasks that I postpone until later
3	Emotion regulation (ER)	The ability to regulate emotions in a way that goals can be realized, tasks can be completed and behavior can be monitored	I understand the behavior of classmates and I adjust my behavior accordingly
4	Sustained attention (SA)	The attention can continue to focus on the situation or on a task, even though there is distraction in the environment, though fatigue increases or even the task is uninteresting	I can concentrate well and keep my focus on what I am doing
5	Task initiation (TI)	Can start directly on a task, on time and in an adequate way	I can easily postpone fun activities and put my homework first
6	Planning (PL)	The ability to create a plan to achieve a goal or complete a task. It is about being able to make decisions about what is important and what is not important	I can manage large assignments well; I know what I to do first and what next
7	Organization (OR)	The ability to organize, arrange or arrange things according to a certain classification or system	I keep a planning overview for all assignments and tests for my school work
8	Time management (TM)	The ability to estimate how much time there is, how the available time can best be divided and what is needed to meet a deadline. The sense of time is also important here	I can make a good estimate of how long something takes, for example doing homework
9	Flexibility (FB)	The ability to change plans based on new information, when there are obstacles or setbacks. The central issue is the ability to adapt the behavior to changed circumstances	I can manage well if the rules or activities change at school
10	Monitoring (MO)	The ability to take a short distance to reflect to the situation and finding out how the problem has been or can be addressed. Important concepts are self-control and self-evaluation	I check my work, especially for a test or exam and correct the mistakes

Test of Sustained Selective Attention (TOSSA)

The Test of Sustained Selective Attention (TOSSA) (Kovács, 2010) is a neuropsychological auditory computer test to assess sustained, selective attention. During a relatively long time period of eight minutes a person has to listen to 240 groups of two, three or four beeps. When a group of three beeps sounds, the spacebar of the computer has to be pushed as quickly as possible. The target is three beeps, the distractors are two and four beeps. During the test the speed of the stimulus presentation varies. The norms are based on groups of several neurological patients (range 12 to 82 year) and a group of healthy controls (range 15 to 93 year). Although the norming sample of the TOSSA did not include students of elementary education, the test was used exploratively for this age group, so the scores of the different educational levels could be compared with each other. The following four of the thirteen indices of the TOSSA are used in this study.

1. Concentration Strength (CS): the CS is the most sensitive TOSSA indices because it represents the two following indices of concentration: the Detection Strength (DS) and the Response Inhibition Strength (RIS).
2. Detection Strength (DS): the number of correct reactions of the three beeps.
3. Response Inhibition Strength (RIS): the number of incorrect reactions on the distractors two or four beeps and on the three beeps.
4. The influence of the Length of presentation on the DS-index (LADS):

The CS indices is used because it represents the main aspects of concentration, DS and RIS, the two other used indices. The DS and RIS are used separately, because the number of correct and incorrect reactions represent more specific the quality of the select attention. LADS is used because this indices represents the quality of the sustained attention (Kovács, 2010).

Statistical Analyses

To examine the factor structure of the FSQ, confirmatory factor analyses (CFA) were used as the main technique using version 7 of the Mplus (Muthén & Muthén, 1998–2015). Starting with 30 items the ten-factor model of EF was tested in the total student sample ($N = 1109$). Based on the outcomes of this first CFA, specifically the fit indices and the standardized factor loadings and covariances among latent variables, more parsimonious models were tested. Then the final student model was tested in the teacher sample ($N = 235$) and the parent sample ($N = 285$).

The fit of the model to the data was evaluated by means of following fit indices: the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), and the Standardized Root Mean Square Residual (SRMR). In general, a model with RMSEA between 0.06 and 0.08 can be considered as fair fit and RMSEA ≤ 0.06 as good fit and RMSEA > 0.10 as poor fit (Hu & Bentler, 1999). In addition, the fit of the model is acceptable when CFI and TLI ≥ 0.90 and good when CFI and TLI ≥ 0.95. Generally, SRMR should not exceed 0.08. To evaluate the model fit of the student samples, cutoff values close to 0.95 were used for TLI and CFI in combination with cutoff values close to 0.09 for SRMR (Hu & Bentler, 1999). The cutoff criteria for the model fit for the teacher and parent sample were used less strictly on the assumption that the informants differ in their image of EF from the student. Furthermore, the fit of the model was acceptable at a factor loading of ≥ 0.5. The standardized residuals for covariance of the latent variables (z-scores ≥—4 or ≥ 4) were used to determine the misfit of the model. The decision to remove or to delete items depended on their content and the degree of overlap between items.

As a final step, measurement invariance of the constructs across educational level and informant groups was examined by means of testing the factorial invariance. The alignment method was used on the assumption that there was a pattern of approximate measurement invariance in the data. In addition, the method has the advantage that not exact measurement invariance is assumed (Asparouhov & Muthén, 2014). Invariance of factor loadings and measurement intercepts as well are required. The maximum likelihood estimation with the fixed identification option was used, because of the little factor loading non-invariance and the small number of groups (Muthén & Muthén, 1998–2015). Based on the

outcomes of the former analyses, new scales were created and their reliability (Cronbach's alpha) was checked.

The education level (elementary, secondary, and tertiary education) and informant group differences (student, teacher and parent) were investigated by a multivariate variance analysis (MANOVA). Because of violation of the assumption of homogeneity of variance–covariance the Mann–Whitney U Test was used, to reveal significant differences in means between the educational levels and between the informant groups.

To examine to what extent the FSQ scales of students correspond with concurrently administered external EF tests, bivariate Pearson correlations of the FSQ scales of students were analyzed with the TLT (Kovács, 2013) and the TOSSA (Kovács, 2010).

Results

Confirmatory Factor Analysis (CFA)

Ten-Factor Model in all Education Levels, Student Sample

First, the ten-factor model (IN, WM, ER, SA, TI, PL, OR, TM, FB, MO) was tested, based on 30 items, in the (total) student sample ($N = 1109$). Although the model fit was satisfactory, RMSEA = 0.051, CFI = 0.905, TLI = 0.885 and SRMR = 0.043, the standardized factor loadings showed low values for five items (varying from 0.44 to 0.49) with correspondingly high standardized residuals (varying from 0.76 to 0.81). In addition, the latent variable covariance matrix (PSI) showed a correlation greater to one (1.038) between the latent variable ER and FL. Furthermore, the covariance for five variables showed z-scores higher than 4 or lower than -4. Because of these findings five items were removed: item 1 from IN, item 8 and 9 from ER, item 18 from PL and item 19 from OR. Of the ten-factor model, six complete factors with three items (WM, SA, TI, TM, FL and MO) and four incomplete factors with only two items (IN, PL and OR) or one item (ER) remained. To construct a more parsimonious model, these 25 items were assigned to five factors and provided with an appropriate name, according to the concept of the PASS model for assessment of cognitive processes: Planning, Attention, Simultaneous, and Successive cognitive processes (Naglieri & Das, 1988) and according to the socio-emotional content of the factors IN and ER. PASS can be considered

Table 4 Model fit indices four-factor model for student, teacher, and parent

	student (N = 1109)	teacher (N = 235)	parent (N = 285)
RMSEA	0.050	0.107	0.082
CFI	0.955	0.916	0.942
TLI	0.944	0.895	0.927
SRMR	0.036	0.057	0.044

21 remaining items disclosed overlap in content and similar phrasing for six items. These six items were removed also.

Four-Factor Model in all Education Levels, Student Sample

Third was tested, whether the remaining 15 items better fitted in a four-factor model: AB, PO (including TI), CCT and RR. The model fit of the four-factor model was satisfactory (see student sample in Table 4). The standardized factor

Table 5 Model fit indices four-factor measurement model per education level

	Elementary education			Secondary education			Tertiary education		
	student (n = 270)	teacher (n = 108)	parent (n = 108)	student (n = 676)	teacher (n = 94)	parent (n = 157)	student (n = 163)	teacher (n = 33)	parent (n = 20)
RMSEA	0.051	0.128	0.103	0.059	0.094	0.091	0.073	0.118	0.191
CFI	0.956	0.896	0.921	0.936	0.930	0.930	0.883	0.900	0.624
TLI	0.945	0.870	0.901	0.921	0.913	0.912	0.853	0.875	0.530
SRMR	0.040	0.062	0.056	0.042	0.072	0.050	0.072	0.087	0.109

parsimonious because all cognitive EF factors involved in academic performance are reflected in the PASS model and EF cognitive tasks are analyzed with the successive steps as a process. The new appropriate names of the five factor-model based on the PASS model were: Adjusting behavior (AB) means being able to adjust social and learning behavior in the school context, in relation to what interaction with teachers and fellow students requires; Planning and organizing (PO) means being able to plan assignments and learning goals and organize the associated learning materials and learning activities to finish learning tasks and achieve learning goals in time; Task initiation (TI) refers to the ability to start directly on a task, on time and in an adequate way; Concentrating and completing tasks (CCT) means maintaining concentration completing tasks, even though there is distraction in the environment, though fatigue increases or even the task is uninteresting; Remembering and revising (RR) refers to the ability to hold information in memory, to reflect and revise while performing complex tasks.

Five-Factor Model in all Education Levels, Student Sample

Second, whether the remaining 25 items were confirmed in the five-factor model was tested. The model fit became less satisfactory, RMSEA = 0.058, CFI = 0.891, TLI = 0.876, and SRMR = 0.043 and the standardized factor loadings of four items showed values lower than 0.49. These four items were removed. Closer inspection of the content of the

loadings varied from 0.51 to 0.80 in this student sample (Appendix Table 12).

Four-Factor Model in all Education Levels, Teacher Sample, and Parent Sample

Two other CFA were performed to validate the fit of the four-factor model in the teacher sample (N = 235) and the parent sample (N = 285). Table 4 shows that the model fit indices were less satisfactory than the student fit indices, especially for the teachers but nonetheless in the right direction and plausible given the fact that especially teachers are not able to compare behaviors and attitudes of students outside the classroom and the school. The standardized factor loadings varied from 0.67 to 0.90 for the teacher sample and from 0.62 to 0.89 for the parent sample (Appendix Table 12).

The Four-Factor Model across Informants and across Three Levels of Education

Then the four-factor model for three levels of education was tested for elementary, secondary and tertiary education and for all three informant samples (student, teacher, parent). Table 5 shows that the model fit indices were satisfactory for all educational levels and informant samples except for the small parent sample of tertiary education (n = 20), whose CFI and TLI values were too low.

Table 6 Means, standard deviations, and Cronbach's alpha coefficients of four executive functioning scales for student, teacher and parent per education level

Education level	informant	n	Adjusting behavior (AB)		Planning and organizing (PO)		Concentrating and completing tasks (CCT)		Remembering and revising (RR)	
			(4 items)		(4 items)		(4 items)		(3 items)	
			M (SD)	α	M (SD)	α	M (SD)	α	M (SD)	α
Elementary education	student	270	3,54 (,80)	,69	3,56 (,88)	,76	3,55 (,85)	,74	3,50 (,87)	,66
	teacher	108	3,26 (,86)	,84	3,30 (,95)	,90	3,28 (,97)	,90	3,09 (,97)	,83
	parent	108	3,57 (,77)	,81	3,31 (,98)	,88	3,46 (1,04)	,90	3,15 (,95)	,82
Secondary education	student	676	3,40 (,71)	,70	3,00 (,78)	,66	3,21 (,86)	,80	3,29 (,80)	,64
	teacher	94	3,45 (,80)	,87	2,99 (,82)	,87	3,39 (,87)	,90	3,16 (,77)	,79
	parent	157	3,50 (,74)	,78	2,96 (,96)	,84	3,48 (,89)	,90	3,10 (,94)	,78
Tertiary education	student	163	3,55 (,57)	,57	3,11 (,74)	,69	3,30 (,73)	,74	3,50 (,69)	,62
	teacher	33	3,99 (,53)	,62	3,53 (,86)	,90	3,67 (,86)	,92	3,64 (,82)	,80
	parent	20	4,18 (,42)	,67	3,98 (,67)	,79	4,15 (,55)	,77	4,33 (,54)	,70

Alignment Method

To compare factor means and variances of the four-factor model across education levels and informant groups, the alignment method was used to test the invariance of intercepts and factor loadings. This requires invariance of both factor loadings and measurement intercepts (Asparouhov & Muthén, 2014). The results of the alignment analyses showed that overall the intercepts of most items were invariant except for three items of elementary education, one item of secondary education and one item of tertiary education. All factor loadings were invariant except for one item of elementary education (Appendix Table 12). These findings mean that the four-factor model is similar across three education levels.

The CFA and alignment method indicate that a four-factor model seems to fit well for the student as a whole as well as for other informants (teachers and parents) and moreover this model also seems satisfactory for different educational levels. In short, the four-factor model is the starting point for further analyses.

Internal Consistency Four-Factor Model

Based on the four-factor model, new scales were created of which the reliability (Cronbach's alpha) was assessed for students, teachers and parents at different educational levels. Table 6 shows that the reliability is satisfactory for the student scales of all three educational levels with the exception of the AB scale for tertiary education, and sufficient to good for almost all teacher and parent scales.

Analyses of Variance across Education Levels and Informant Groups (Student, Teacher, Parent)

The Mann–Whitney U test was used to test the significance of differences, by ranking the factor means of student and teacher groups and student and parent groups per educational level.

With the exception of the AB scale, for all informant groups (student, teacher and parent together) of elementary and secondary education, significant differences ($p \leq 0.05$) were found between elementary and secondary education, with higher means for elementary than for secondary education. With the exception of the CCT scale for all informant groups (student, teacher and parent together) of secondary and tertiary education, significant differences ($p \leq 0.05$) were found between secondary and tertiary education, with higher means for tertiary than for secondary education. Differences in means between the scales for student and teacher were not significant, except for the RR scale, with higher means for students than for teachers. The differences between the scales for student and parent were significant for the AB and the CCT scale, with higher means for parents than for students. No significant differences were found between the scales for teachers and parents.

Concurrent Validity of Correlations among EF Tests and FSQ Scales

Correlations among FSQ Scales Student and Tower of London Test (TLT) and Test of Sustained Selective Attention (TOSSA) per Education level

To examine the convergence of FSQ scales and EF test results, bivariate Pearson correlations were analyzed among the student FSQ scales and two tests for EF: the Tower of London test (TLT) (Kovács, 2013) and the Test of Sustained Selective Attention (TOSSA) (Kovács, 2010). Only one significant correlation ($p < 0.05$) was found between the student FSQ scales for elementary and secondary education and TLT Total Score (TS), and no significant correlations were found between the FSQ scales for elementary and secondary education and TLT Decision Time (DT) (Appendix Table 10). Considering the small sample of students of tertiary education grade 1, 2 ($n = 15$) and grade 3, 4, 5, 6 ($n = 7$) who participated in the TLT, no reliable conclusions can be made for tertiary education.

Most correlations among the student FSQ scales and TOSSA indicators were non-significant (Appendix Table 11). Considering the small sample of students of tertiary education grade 1, 2 ($n = 14$) and grade 3, 4, 5, 6 ($n = 7$) who participated in the TOSSA, no reliable conclusions can be made for tertiary education.

Discussion

Despite the overwhelming research on EF tests and their underlying structure, the question remains how meaningful this information is for EF skills of students and their self-regulation in a school context. Therefore, the main purpose of the present study was to validate a solution-focused instrument to assess the degree of EF skills in the school context for students, teachers and parents in three different education level groups (primary, secondary and tertiary education). The first aim was to investigate to what extent CFA could deliver a more parsimonious model of the ten-factor structure of the FSQ proposed by Dawson and Guare (2010). A second aim was to investigate whether there were significant differences among education level and informant groups respectively in the newly constructed factor structure. The last aim was to investigate to what extent the FSQ scales were corresponding with concurrently recorded external EF tests.

Factor Structure

The CFA results showed a factor structure of the FSQ that was equivalent for all education levels and informant groups. A limited number of factors in the FSQ data was expected to be found using the CFA, although whether the structure would resemble the structure described by Friedman and Miyake (2017) was unclear. Indeed, the number of factors was limited to four but in contrast to their unity/diversity framework, no specific and distinguishable updating and shifting scales were found. This difference was to be expected because the FSQ is not based on EFs measured by laboratory neuropsychological tests but focuses on the behavioral level in the school context. Moreover, the distinction in EFs is artificial because reality always requires a mixture of functions. The content of each of the four factors was indeed mixed, containing items of the three original EF components: inhibition, working memory (updating) and cognitive flexibility or shifting (Best et al., 2011; Miyake et al., 2012). The four new scales are: the Adjusting Behavior (AB) scale, the Planning and Organizing (PO) scale, the Concentrating and Completing Tasks (CCT) scale and the Remembering and Revising (RR) scale. At the behavioral level, these scales are tools for mapping task-related behavior in the school context. With this four-factor model the first aim to investigate whether CFA of the FSQ could deliver a parsimonious model, has been achieved. Future research on EF skills needs to replicate these findings to further strengthen the evidence for this structure of four factors at the behavioral level.

Differences in Means among Educational Level and Informant Groups

A second aim of this study is to investigate whether there are significant differences among education levels (elementary, secondary and tertiary education) and informant groups (respectively student, teacher, parent). Although the FSQ factor structure was equivalent for informant groups, there were significant differences in means among students, teachers and parents. Comparison of the means of informants groups showed no significant differences between the scales for teachers and parents, nor between the scales for student and teacher, except for the RR scale, with higher means for students than for teachers. The lower means of the teachers for these metacognitive remembering and revising skills could mean that teachers perceive the skills differently

than their students do and that students need more insight and training to master these skills. The differences between the scales for student and parent were significant for the AB and the CCT scale, with higher means for parents than for students. These higher estimates of the parents for these behavior and concentrating skills can possibly be explained by the fact that parents assess these EF skills from the home situation in a context with less participants and maybe less distraction than in the school context, whereas these EF skills are the most demanded at school.

The findings of this study justify the conclusion that in the newly constructed four-factor model there are significant differences in means across education levels and to a more limited extent significant differences across informant groups.

Concurrent Validity of FSQ Scales

The investigation of the relationship among the FSQ scales of students with concurrently recorded external EF tests and EF skills reported by teachers and parents, revealed limited significant correlations. The hypothesized relationship of the FSQ scales with a concurrently performed external EF test, the TLT, was not found, and relationship with the TOSSA was limited and some significant correlations were in the wrong direction. Student reported EF skills in the school context measured by a questionnaire, are to a very limited degree represented by EF measured by laboratory tests such as the TLT and the TOSSA. The findings show the gap between laboratory test outcomes and students self-reports of EF skills in a real-life context, but at the same time reflect the conceptual differences between indicators of neuro-psychological functioning on a test (e.g., decision time) and broader estimates of goal-directed behavior in a real classroom setting.

Limitations

Some limitations of the present study should be noted. First the FSQ gives insight into how different informants rate EF skills in the school context but EF test outcomes and data from teachers and parents have been retrieved in a much smaller sample than the student data. Therefore, data to examine the concurrent validity were limited. Second the norming sample of the EF standardized tests (TLT and TOSSA) did not include the students of elementary education, therefore the scores of this age group may be more difficult to interpret. Another limitation is that the categorization of educational levels in three levels was broad and could obscure meaningful relationships.

Implications

In education, learning to exhibit goal-directed behaviors and skills is important for school success. Providing teachers with instruments to assess students' skills and to communicate with students and their parents about these skills in a solution-focused way could foster students' development of goal-directed behavior in the school context. The significant differences in means between the rating of informants, make it necessary that after completing the FSQ, the informants share their different views on EF skills of the student with each other and explain them orally. This could be part of the problem analysis and solution in real-life situations. Students from tertiary education can use the FSQ themselves and share their rating on EF skills with other students or their teachers to gain insight into their own EF skills.

Conclusions

This study has indicated first, a more parsimonious model of EF skills, however with a structure that does not correspond with the generally assumed EF factors. Second, this study shows that there are differences in the perception of EF skills by different informant groups and also differences per education level. Third, the student perceptions of EF skills do not correspond with EF test outcomes and in some subgroups clearly diverge from teacher or parent perceptions of the students' EF skills. Therefore, the FSQ has shown insufficient validity but nonetheless could be a useful addition for teachers and students. The FSQ[1] could serve as a starting point for conversations between teachers, students and parents on task-related behavior in the school context. Comparing their different perspectives on engagement can support teachers and students in adjusting their planning and organizing and other task-related behavior, like participation, attention, persistence, self-regulated learning and exerting the necessary effort for comprehension of complex ideas (Fredericks et al., 2016).

---

[1] In 2023 the FSQ will be published by Paragin in The Netherlands (Paragin.nl).

Appendix

Table 7 Scales, standardized factor loadings, alignment intercepts and alignment factor loadings of the Measurement model of all education levels

Scales	Factor Loading Student	Factor Loading Teacher	Factor Loading Parent	Alignment Intercepts students invariant (+) variant (-) elementary(1) secondary(2) tertiary (3) education	Alignment factor loadings students invariant (+) variant (-) elementary(1) secondary(2) tertiary (3) education
Adjusting behavior					
1	0.512	0.685	0.619	+	+
2	0.643	0.841	0.814	+	------ (1)
3	0.655	0.777	0.715	+	+
4	0.601	0.779	0.673	+	+
Planning and organizing					
1	0.621	0.850	0.827	+	+
2	0.609	0.674	0.686	- (1)	+
3	0.607	0.853	0.789	- (1)	+
4	0.626	0.887	0.773	+	+
Concentrating and completing tasks					
1	0.612	0.857	0.805	+	+
2	0.732	0.879	0.881	+	+
3	0.796	0.904	0.888	- (1)—(2) - (3)	+
4	0.617	0.707	0.737	+	+
Remembering and revising					
1	0.620	0.724	0.706	+	+
2	0.569	0.806	0.794	+	+
3	0.659	0.793	0.804	+	+

Table 8 Scales, items, standardized factor loadings, alignment intercepts and alignment factor loadings of the Measurement Model for elementary education

Scales	Factor Loading Student	Factor Loading Teacher	Factor Loading Parent	Alignment Intercepts Invariant (+) Variant (-) Student (s) Teacher (t) Parent (p)	Alignment Factor loadings Invariant (+) Variant (-) Student (s) Teacher (t) Parent (p)
Adjusting behavior					
1. I think before I speak	0.468	0.660	0.674	- (t)	+
2. I understand what my classmates are doing and I adapt if necessary	0.650	0.848	0.845	+	+
3. I adapt easily when a classmate wants something different than me	0.714	0.752	0.766	+	+
4. If there is a problem, I can come up with different solutions and choose a good solution	0.628	0.785	0.628	+	+
Planning and organizing					
1. I can manage large assignments well; I know what to do first and what next	0.697	0.920	0.896	+	+
2. I keep my backpack / school bag and my locker at school neat and tidy	0.588	0.595	0.616	- (s)	+
3. I carry out my tasks on time	0.726	0.871	0.839	- (p)	+
4. I can finish three to four assignments one after the other in the classroom	0.672	0.930	0.874	+	+
Concentrating and completing tasks					
1. I can concentrate well and keep my focus on what I am doing	0.580	0.852	0.840	+	+
2. I enjoy my schoolwork or other tasks from the beginning to the end	0.734	0.874	0.894	+	+
3. I keep at my school work or with other tasks until they are finished	0.783	0.899	0.890	+	+

Table 8 (continued)

Scales	Factor Loading Student	Factor Loading Teacher	Factor Loading Parent	Alignment Intercepts	Alignment Factor loadings
4. I always get my homework done at home	0.533	0.723	0.720	+	+
Remembering and revising					
1. I remember assignments that I have to do after school	0.610	0.672	0.748	- (s)	+
2. I check my work when I am done and correct the mistakes	0.594	0.890	0.804	+	+
3. If my approach to a task does not work, then I look for a different approach or a better solution	0.681	0.851	0.791	+	+

Table 9 Scales, items, standardized factor loadings, alignment intercepts and alignment factor loadings of the Measurement Model for secondary education

Scales	Factor Loading Student	Factor Loading Teacher	Factor Loading Parent	Alignment Intercepts Invariant (+) Variant (-) Student (s) Teacher (t) Parent (p)	Alignment Factor loadings Invariant (+) Variant (-) Student (s) Teacher (t) Parent (p)
Adjusting behavior					
1. I think first before I speak	0.547	0.694	0.576	+	+
2. I understand the behavior of classmates and I adjust my behavior accordingly	0.625	0.768	0.826	+	+
3. I adapt easily when a classmate wants something different than me	0.635	0.860	0.733	+	+
4. I can manage well if the rules or activities change at school	0.637	0.901	0.721	+	+
Planning and organizing					
1. I can manage large assignments well; I know what to do first and what next	0.590	0.745	0.787	+	+
2. I keep a planning overview for all assignments and tests for my school work	0.582	0.841	0.691	- (t)	+
3. I can make a good estimate of how long something takes, for example doing homework	0.549	0.839	0.772	+	+
4. I can easily postpone fun activities and put my homework first	0.578	0.825	0.781	+	+
Concentrating and completing tasks					
1. I can concentrate well and focus on what I am doing	0.660	0.874	0.776	+	+
2. I enjoy doing my schoolwork or other tasks from the beginning to the end	0.719	0.879	0.878	- (s)	+
3. I keep at my homework or other tasks until they are finished	0.821	0.911	0.906	+	+
4. I always finish my homework	0.631	0.714	0.770	+	+

Table 9 (continued)

Scales	Factor Loading Student	Factor Loading Teacher	Factor Loading Parent	Alignment Intercepts	Alignment Factor loadings
Remembering and revising					
1. I remember tasks that I postpone until later	0.607	0.726	0.684	- (s)	+
2. I check my work, especially for a test or exam and correct the mistakes	0.579	0.731	0.798	+	+
3. If my approach to a task does not work, then I change my approach to increase the chance of success	0.657	0.769	0.777	+	+

Table 10 Scales, items, standardized factor loadings, alignment intercepts and alignment factor loadings of the Measurement Model for tertiary education

Scales	Factor Loading Student	Factor Loading Teacher	Factor Loading Parent	Alignment Intercepts	Alignment Factor loadings
				Invariant (+) Variant (-) Student (s) Teacher (t) Parent (p)	Invariant (+) Variant (-) Student (s) Teacher (t) Parent (p)
Adjusting behavior					
1. I think first before I speak	0.317	0.643	0.618	+	+
2. I understand the behavior of my fellow students and I adjust my behavior accordingly, where necessary	0.676	0.740	0.467	+	+
3. I adapt easily when a classmate wants something different than me	0.636	0.420	0.793	+	+
4. I can handle manage if the rules or activities change at school	0.476	0.409	0.682	+	+
Planning and organizing					
1. I can manage large assignments well; I know what I to do first and what next	0.694	0.758	0.724	+	+
2. I keep a planning overview for all assignments and tests for my study or schoolwork	0.586	0.878	0.611	+	+
3. I can estimate well how long something takes, for example carrying out school or study assignments	0.568	0.786	0.754	+	+
4. I can easily postpone fun activities and put my work for the study program or study first	0.581	0.900	0.754	+	+
Concentrating and completing tasks					
1. I can concentrate well and keep my focus on what I am doing	0.583	0.868	0.666	+	+

Table 10 (continued)

Scales	Factor Loading Student	Factor Loading Teacher	Factor Loading Parent	Alignment Intercepts	Alignment Factor loadings
2. I enjoy my schoolwork, study assignments or other tasks from the beginning to the end	0.663	0.828	0.512	+	+
3. I keep at my schoolwork, study assignments or other tasks until they are finished	0.668	0.929	0.475	+	+
4. I always finish my work for training or study on time	0.682		0.851	- (t)	+
Remembering and revising					
1. I remember tasks that I postpone until later	0.579	0.827	0.596	+	+
2. I check my work, especially for a test or an exam, and correct the mistakes	0.585	0.767	0.710	+	+
3. If my approach to a task does not work, then I change my approach to increase the chance of success	0.628	0.691	0.744	+	+

Table 11 Bivariate Pearson Correlations FSQ Scales Student and Tower of London Test (TLT) per education level

Education level	Informant scale	TLT_Decision time (DT)	TLT_Total score (TS)
Elementary Education (EE)			
Secondary Education (SE)			
Tertiary Education (TE)			
EE, 3, 4 Student (TLT, n=77)	Student AB	-,18	,02
	Student PO	-,12	,12
	Student CCT	-,13	,12
	Student RR	-,27*	-,01
EE, 5, 6 Student (TLT, n=55)	Student AB	,02	,22
	Student PO	-,16	,02
	Student CCT	,01	-,04
	Student RR	-,16	-,03
SE, 1 Student (TLT, n=35)	Student AB	-,28	,09
	Student PO	-,31	,02
	Student CCT	-,33	-,07
	Student RR	-,21	,19
SE, 2 Student (TLT, n=79)	Student AB	-,11	-,01
	Student PO	-,17	,02
	Student CCT	-,21	,05
	Student RR	-,06	,20
SE, 3, 4, 5, 6 Student (TLT, n=43)	Student AB	,03	-,12
	Student PO	-,19	-,08
	Student CCT	-,09	-,26
	Student RR	-,11	-,10

**. Correlation is significant at the 0.01 level (2-tailed)

*. Correlation is significant at the 0.05 level (2-tailed)

Table 12 Bivariate Pearson correlations FSQ Scales Student and Test of Sustained Selective Attention (TOSSA) per education level.

Education level	Informant scale	TOSSA_Concentration Strength (CS)	TOSSA_Detection Strength(DS)	TOSSA_Response inhibition Strength (RIS)	TOSSA_influence length on DS (LADS)
Elementary Education (EE) Secondary Education (SE) Tertiary Education (TE)					
EE, 3, 4 Student (TOSSA, $n=69$)	Student AB	,08	,13	-,09	,10
	Student PO	-,14	-,10	-,16	,05
	Student CCT	,02	,10	-,16	-,11
	Student RR	-,09	-,10	-,13	,26*
EE, 5, 6 Student (TOSSA, $n=53$)	Student AB	-,19	-,21	-,05	-,15
	Student PO	,04	,05	-,05	-,12
	Student CCT	-,01	-,03	,02	-,06
	Student RR	-,03	-,04	,01	-,22
SE, 1 Student (TOSSA, $n=35$)	Student AB	-,13	-,10	-,07	,03
	Student PO	,21	,24	,09	,34*
	Student CCT	,06	,13	-,07	,29
	Student RR	,08	,17	-,11	,40*
SE, 2 Student (TOSSA, $n=68$)	Student AB	-,02	-,04	,05	-,09
	Student PO	-,20	-,27*	,05	-,25*
	Student CCT	-,15	-,21	,07	-,08
	Student RR	-,02	-,07	,13	-,04
SE, 3, 4, 5, 6 Student (TOSSA, $n=29$)	Student AB	-,20	-,22	-,14	,33
	Student PO	-,41*	-,38*	-,43*	-,04
	Student CCT	-,31	-,32	-,34	-,08
	Student RR	-,32	-,32	-,37	-,22

**. Correlation is significant at the 0.01 level (2-tailed)

*. Correlation is significant at the 0.05 level (2-tailed)

Author Contribution All three authors:
- were involved in the creation of this manuscript;
- have read and approved the manuscript;
- believe that the manuscript represents honest work and the information is not provided in another form.

Data Availability The research data are in the possession of the first author and will be archived at DANS (https://dans.knaw.nl) an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Organization for Scientific Research (NWO). This will make the data available for reuse and published research verifiable and repeatable.

Declarations

Ethical Approval There is ethical approval of the research project and of the manuscript.

Consent to Participate Informed consent was obtained from all individual participants included in the study. Written informed consent was obtained from the parents and from the school directors.

Statement about Submissions and Previous Reports There are no submissions or previous reports that might be regarded as redundant publication of the same or very similar work.

Conflict of Interest The authors declare no competing interests.

References

Alloway, T. P., & Alloway, R. G. (2013). Working memory across the lifespan: A cross-sectional approach. *Journal of Cognitive Psychology, 25*(1), 84–93. https://doi.org/10.1080/20445911.2012.748027

Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling, 21*(4), 495–508. https://doi.org/10.1080/10705511.2014.919210

Becker, S. P., & Langberg, J. M. (2014). Attention-deficit/hyperactivity disorder and sluggish cognitive tempo dimensions in relation to executive functioning in adolescents with ADHD. *Child Psychiatry & Human Development, 45*, 1–11.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences, 21*(4), 327–336.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review, 54*(2), 199–231. https://doi.org/10.1111/j.1464-0597.2005.00205.x

Bronfenbrenner, U. (1975). *The ecology of human development in retrospect and prospect.* Retrieved from. https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED128387&scope=site. Accessed in 2017.

Bronfenbrenner, U., & Evans, G. W. (2000). Developmental science in the 21st century: Emerging questions, theoretical models, research designs and empirical findings. *Wiley-Blackwell.* https://doi.org/10.1111/1467-9507.00114

Crone, E. A., Peters, S., & Steinbeis, N. (2017). Executive function development in adolescence.. Chapter in *Executive function, Development Across the Life Span, e*dited by S. A. Wiebe & J. Karbach (pp. 58–72). Routledge.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia, 44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

Dawson, P., & Guare, R. (2010). *Executive skills in children and adolescents: A practical guide to assessment and intervention* (2nd ed.). Guilford.

Dawson, P., Guare, R., & Scheen, W. (2009). *Slim maar...: Help kinderen hun talenten benutten door hun executieve functies te versterken. [Smart but scattered: The revolutionary "executive skills" approach to helping kids reach their potential].* Amsterdam: Hogrefe.

Dawson, P., Guare, R., & Scheen, W. (2010). Executieve functies bij kinderen en adolescenten: Een praktische gids voor diagnostiek en interventie. Amsterdam: Hogrefe. [Translation of *Executive skills in children and adolescents: A practical guide to assessment and intervention.* Dawson & Guare, 2010 (2nd ed.). Guilford.].

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Etkin, J. (2018). Understanding self-regulation in education. *BU Journal of Graduate Studies in Education, 10*(1), 35–39. Retrieved from. https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1230272&scope=site. Accessed in 2019.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction, 43*, 1–4.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychology, 6*, 235–238. https://doi.org/10.1076/chin.6.3.235.3152

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.

Huizinga, M., & Smidts, D. P. (2011). Age-related changes in executive function: A normative study with the dutch version of the behavior rating inventory of executive function (BRIEF). *Child Neuropsychology, 17*(1), 51–66. https://doi.org/10.1080/09297049.2010.509715

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.

Huizinga, M., & Smidts, D. P. (2013). *BRIEF. Vragenlijst executieve functies voor 5- tot 18-jarigen.* Amsterdam: Hogrefe. [Translation of *BRIEF. Executive functions questionnaire for 5 to 18 year olds.* Gioia et al. 2000].

Kim, J. S., & Franklin, C. (2009). Solution-focused brief therapy in schools: A review of the outcome literature. *Children and Youth Services Review, 31*(4), 464–470.

Korinek, L., & DeFur, S. H. (2016). Supporting student self-regulation to access the general education curriculum. *Teaching Exceptional Children, 48*(5), 232–242. https://doi.org/10.1177/0040059915626134

Kovács, F. (2010). TOSSA Test of Sustained Selective Attention. Retrieved from https://pyramidproductions.nl.server41.firstfind.nl/Bijlage/TOSSA_manual.pdf. Accessed in 2017.

Kovács, F. (2013). *TLT Tower of London Test.* Retrieved from. https://pyramidproductions.nl.server41.firstfind.nl/Bijlage/TLT_Manual2013.pdf. Accessed in 2017.

Michalec, J., Bezdicek, O., Nikolai, T., Harsa, P., Jech, R., Silhan, P., & Shallice, T. (2017). A comparative study of Tower of London scoring systems and normative data. *Archives of Clinical Neuropsychology, 32*(3), 328–338.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Muthén, L. K., & Muthén, B. O. (1998–2015). *Mplus User's Guide. Seventh Edition.* Los Angeles: Muthén & Muthén. https://digital.library.temple.edu/digital/api/collection/p245801coll10/id/430538/download

Naglieri, J. A., & Das, J. P. (1988). Planning-arousal-simultaneoussuccessive (PASS): A model for assessment. https://doi.org/10.1016/0022-4405(88)90030-1

Poon, K. (2018). Hot and cool executive functions in adolescence: Development and contributions to important developmental outcomes. *Frontiers in Psychology, 8*, 2311.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662. https://doi.org/10.1037/h0054651

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166–183.

L.C. Spreij has worked for more than 30 years as a teacher, school psychologist and researcher. She developed this instrument because there is a need to stimulate conversations between teachers, students and parents on task-related behavior in the school context.

C. van Tuijl is assistant professor of special education. Her topics are on executive function development, language and literacy and the role of the environment in early childhood education.

P. P. M. Leseman is a full professor of special education. His research concerns executive function development, and language, literacy and math learning in early and middle childhood. He directed large-scale cohort studies on these topics.