Relations between Measures of Executive Functions and Self-regulation in Preschoolers

Wiebke F. Evers, Laura M. Walk, Sonja Quante, Katrin Hille

Abstract
Performance on different measures of executive functions (EF) and self-regulation (SR) does not always correspond to the behaviour children show in real-life situations. The present study assesses the relationships between performances on different EF and SR measures and teacher ratings of children's self-control and thoughtfulness.

In total, 217 children between 34 and 72 months (54% boys) were assessed. Four tests measuring cognitive EF (Digit Span backward, Block Recall, Day-Night Stroop, Hearts & Flowers) and two tests measuring behavioural EF (Tower) and SR (Head-Toes-Knees-Shoulders [HTKS]) were administered. Additionally, teachers rated the dimension 'self-control and thoughtfulness' of the German observation scale 'Social-emotional well-being and resilience of children in early childhood settings' (PERiK).
It was found that all measures differentiated with regard to age in the range of three to six years. Correlations between cognitive EF measures with the HTKS were almost twice as high as correlations with the Tower. This indicates that the HTKS taps similar processes as the cognitive EF measures. Teacher ratings did not show higher correlations with behavioural EF and SR than with cognitive EF measures. Also, behavioural EF and SR measures did not predict scores obtained on the teacher rating better than cognitive EF measure. This article discusses to what extent distinctions among measures of EF and SR are possible and useful.

Keywords: Executive Functions, Self-regulation, Test Instruments, Cognitive Development, Pre-school

Zusammenhänge zwischen verschiedenen Maßen der exekutiven Funktionen und Selbstregulation bei Kindergartenkindern

Zusammenfassung
Die Leistung bei standardisierten Verfahren zur Erfassung der exekutiven Funktionen (EF) und Selbstregulation (SR) passt nicht immer zu dem Niveau an Selbstkontrolle, dass Kinder im Kindergartenalter in realen Situationen zeigen. Die vorliegende Studie untersucht die Zusammenhänge zwischen den Leistungen bei verschiedenen Tests zu EF und SR und der Fremdeinschätzung der Selbstkontrolle, die die Kinder im Kindergartenalltag zeigen, durch pädagogische Fachkräfte. Insgesamt 217 Kinder zwischen 34 und 72 Monaten (54% Jungen) nahmen an der Studie teil. Vier Tests zu Erfassung der kognitiven EF (Zahlen nachsprechen rückwärts, Block Recall, Day-Night Stroop, Hearts & Flowers) und zwei Tests zur Erfassung der behavioralen EF (Turmbau) und der behavioralen SR (Head-Toes-Knees-Shoulders [HKTS]) wurden durchgeführt. Darüber hinaus bewerteten pädagogische Fachkräfte die Kinder auf der Skala „Selbstkontrolle und Rücksichtnahme" des deutschen Beobachtungsbogens „Positive Entwicklung

und Resilienz im Kindergartenalltag" (PERiK). Zusammenhänge zwischen den Messinstrumenten zur kognitiven EF mit den HKTS waren fast doppelt so hoch wie ihre Korrelationen mit dem Turmbau. Dieses Ergebnis weist darauf hin, dass der HTKS, ausgewiesen als Instrument zur Erfassung der behavioralen SR, ähnliche Konstrukte wie die Tests zur Erfassung der kognitiven EF misst. Die Bewertung des Kindesverhaltens durch die pädagogischen Fachkräfte zeigte keine höheren Korrelationen mit den behavioralen als mit den kognitiven Messinstrumenten. Auch die Einschätzung durch die pädagogischen Fachkräfte wird durch die behavioralen Messungen nicht besser vorhergesagt als durch die kognitiven Messungen, wie es erwartet wurde. Es wird diskutiert, inwieweit eine Unterscheidung in die beiden Kategorien von Messinstrumenten möglich und sinnvoll ist.

Schlagwörter: Exekutive Funktionen, Selbstregulation, Testinstrumente, kognitive Entwicklung, Kindergarten

1 Introduction and aim

Self-regulation (SR) can be broadly defined as goal-directed behaviour, typically within at least a minimal temporal perspective (*Hofmann/Schmeichel/Baddeley* 2012). The term refers to the ability to manage emotions, directing thoughts and regulate and adapt behavior (*Blair/Razza* 2007; *Smith-Donald* u.a. 2007). In contrast, the term self-control (SC) is often used to describe mainly the ability of response inhibition. A common example is to override an automatic response to activate an alternative more promising response in order to achieve a certain goal (*Diamond* 2013; *Hofmann/Schmeichel/Baddeley* 2012). SR as well as SC are closely connected to executive functions (EF), an umbrella term that refers to cognitive processes of the prefrontal cortex (*Duncan* 1986; *Luria* 1976) that are necessary for goal-directed behaviour. The term EF encompasses a heterogeneous set of cognitive skills such as inhibitory control, working memory, cognitive flexibility, attention, planning, reflection and error detection (*Anderson* 2010; *Welsh/Pennington/Groisser* 1991; *Zelazo* u.a. 1997). They are essential for processes such as effortful control and selective attention as well as adaptively responding to novel or challenging situations when automatic, overlearned responses are inadequate or not existent (e.g. *Miller/Cohen* 2001; *Zelazo* u.a. 2003).

Three of the skills assigned to the term EF are argued to be central for most others (e.g. *Hughes* 1998; *Welsh/Pennington/Groisser* 1991). Following the framework of *Miyake* u.a. (2000) known as the unity and diversity construct of EF (*Miyake/Friedman* 2012; *Miyake* u.a. 2000), those skills are (updating of) working memory, inhibitory control and cognitive flexibility (shifting). Working memory describes the ability to hold, update and monitor information mentally. Therefore, it is crucial for planning and problem solving (*Baddeley* 1986). Inhibitory control is the ability to resist a first impulse and to reflect before acting in order to achieve a desired goal. Inhibitory control is also needed to block out interferences in order to stay focused (*Rothbart/Posner* 1985). Cognitive flexibility (also called shifting) is involved when going back and forth or changing between tasks or mental sets and when adjusting to novel situations. It is also important to take on a new perspective and discover different ways to approach a problem (*Diamond* 2007; *Diamond* u.a. 2007).

SR, SC and EF are considered to have a high influence on the development of children's academic and social-emotional skills. A vast amount of evidence showed that EF are of great importance not only for school readiness (*Bierman* u.a. 2008; *Blair* 2002),

(the emergence of) literacy and numeracy skills (*McClelland/Cameron/Connor* u.a. 2007; *Welsh* u.a. 2010) and academic achievement (e.g. *Duckworth/Seligman* 2005; *Willoughby* u.a. 2012) but also for social competences and socially appropriate behaviour (e.g. *Becker* u.a. 2014; *Kochanska/Murray/Harlan* 2000; *McClelland/Cameron/Wanless* u.a. 2007; *Morgan/Lilienfeld* 2000; *Rhoades/Greenberg/Domitrovich* 2009) as well as for successful emotion regulation (e.g. *Carlson/Wang* 2007; *Hughes* u.a. 2015).

2 Assessing executive functions, self-regulation and self-control

Choosing the right measures to assess EF, SR and SC is not an easy undertaking given the vast variety of the existing instruments (*Carlson* 2005). Any good measure has to fulfil certain criteria that can help identifying useful assessment methods. Firstly, the measure should be appropriate for the age group of interest. Some measures lack relevance to children, which makes it hard for them to stay on task, affects their performance and might keep them from reaching their full potential (*Anderson* u.a. 2002; *Carlson* 2005). Other measures are time consuming or complicated to administer (*Ponitz* u.a. 2008), making them unsuitable for a number of age groups or larger samples. Secondly, the measures should tap as exclusively as possible the construct of interest and keep the influence of other competencies such as language skills, attention, or memory capacity that could moderate performance to a minimum (*Anderson* u.a. 2002; *Espy* u.a. 2008). Thirdly, impairments in one domain of EF do not necessarily imply that other executive domains are also impaired. Hence, it is of interest to include multiple measures especially when measuring EF to tap the different domains across different modalities (*Anderson/Reidy* 2012). It is recommended to use several domain-general and domain-specific measures as well as different levels of analysis (e.g. neurological, physiological, rating measures and questionnaires) (*McClelland* u.a. 2015). And last but not least, the measure used should be reliable and valid, e.g. internally consistent and temporarily stable in their results. One aspect that is often neglected when it comes to choosing appropriate measure is its ecological validity. To ensure ecological validity, *Gioia* and *Isquith* (2004) recommend the use of both assessments of individual components with standardized tests as well as observation of behavioural application of EF in a real-world context.

3 Distinctions between measures of executive functions and self-regulation

3.1 Performance-based measures

Many studies assessing EF use performance-based tests administered by an examiner that measure EF in highly standardized conditions and usually assess accuracy and/or response time (*Pennington/Ozonoff* 1996; *Toplak/West/Stanovich* 2013). Tests that have been used widely are for example the Dots task (*Diamond* u.a. 2007; *Shing* u.a. 2010), the Wisconsin-Card Sorting Test (*Heaton* 1993) and the Stroop test (*Jensen/Rohwer* 1966; *Stroop* 1935). However, performance on such psychometric cognitive tests might not be a

sufficient indicator of behaviour shown in everyday life (*Anderson* u.a. 2002). The administration takes place in a quiet, one-on-one, structured setting with minimal distractions with the test administrator providing support for the child to stay on task and finish each test (*Stuss/Alexander* 2000). This highly standardized laboratory-like testing situation is hardly comparable to situations at home or in classrooms (*Anderson* 1998; *Falk/Heckman* 2009; *Rimm-Kaufmann* u.a. 2009). This raises some doubts on whether the performance shown on tests administered in such controlled environments reflects behaviour shown in real-life situations.

The lack of translation into every-day behaviour makes the usage of more applied behavioural measures, or teacher or parent ratings very important. There are some performance-based measures that assess children's performance in tests that resemble real-life situations. However, those measures rarely assess EF but rather children's application of EF or in other words SR (*Toplak/West/Stanovich* 2013). Whereas EF measures often aim at assessing the three components in isolation, SR measures often rely on the application of EF as a whole. This might heighten the ecological validity but gives less insight on the development of the individual components.

One example of a more realistic EF measure is the Tower (*Kochanska* u.a. 1996) assessing inhibitory control. The task is embedded into a standard play situation in which the child is asked to take turns with the test administrator. Hence, although all assessed in standardized, one-on-one situations, performance-based measures of EF can differ in their level of resemblance to real-life situations and putatively in their ecological validity.

Compared to EF measures, there are more performance-based SR measures that resemble realistic situations. One example of an assessment which aim is to tap the translation of cognitive functioning into behaviour is the Head-Toe-Shoulders-Knees (HTKS; *Ponitz* u.a. 2008; *Ponitz* u.a. 2009). The extended version of the HTKS-task, which is identified as a behavioural measure of SR by the authors, clearly specifies the three individual components of EF: inhibitory control, working memory and cognitive flexibility (*Ponitz* u.a. 2008). In the task, the child has to refrain from following commands given by the test administrator and do something else. After some time, the rules of the game are changed requiring the child to apply the new rules in its behaviour.

So far, there is little agreement on how to label measures of EF and SR that assess the two constructs in more realistic ways in order to distinguish them from the measures that assess EF and SR with tasks further away from real-life situations. The two examples given above, the Tower as a measure of inhibitory control and the HTKS as a measure of behavioural SR, make it clear that the distinction into measures of EF and SR is not always the distinction into more and less realistic measures.

3.2 Hot and cool measures

One way measures are distinguished in is into 'hot' and 'cool' measures. The distinction is based on the assumption that some measures are more motivationally and emotionally relevant to the test taker than others and therefore tap hot EF (*Zelazo/Carlson* 2012; *Zelazo* u.a. 2005). Hot EF measures also often involve a social component although a social aspect alone does not mean that the task is necessarily assessing hot EF (*Zelazo* u.a. 2005). Also, hot measures of EF can also be quite artificial in nature and do not necessarily resemble real-life situations. Measures that are personally not of great relevance, often

presenting abstract and emotionally neutral tasks aim at assessing cool EF (*Zelazo* u.a. 2005). Therefore, measures such as the Tower might fall into the category of hot measures of EF as it contains quite a strong motivational and emotional component with the child's wish to finish the tower. The Tower also has the social component of taking turns with another person. That also makes it emotionally demanding for the child to not get frustrated with the slow progress and to wait for the other person to make his/her turn. The categorization of the HTKS into 'hot' or 'cool' however, is not quite as straight forward. One could argue that the HTKS has some motivational relevance to it as it resembles a game and the child wants to do well at it. However, as the task does not involve any feedback, the child does not necessarily know whether it is doing well or not. Also, the task does not hold a social component as the child is playing alone and carrying out the commands given by the test administrator. In this regard, it actually resembles more tasks such as the Wisconsin-Card Sorting Test (*Heaton* 1993) or the Stroop test (*Jensen/Rohwer* 1966; *Stroop* 1935).

What sets the HTKS apart from most measure, however, is the fact that the child has to act out a behavioural response rather than just pointing to a certain box or calling something out. Hence, it is called a behavioural measure by the authors. In contrast, most other EF measures named before fall into the category of cognitive EF measures (*Bierman* u.a. 2008; *Pennington/Ozonoff* 1996; *Ritter* u.a. 2014).

3.3 Rating scales

The term behavioural measure, however, is almost exclusively used when referring to rating scales. Rating scales, similar to measures of EF and SR that assess the constructs in more realistic ways, are sought to measure the extent to which certain behaviours or competences are shown in complex, everyday situations (*Roth/Isquith/Gioia* 2005). Some researchers argue that the use of rating scales is vital to gain insight into a child's executive and self-regulatory functioning (*Isquith* u.a. 2005). Ratings usually involve an informant reporting on the level of performance with carrying out everyday tasks related to EF (*Miranda* u.a. 2015). One of the most commonly used rating scales in the domain of EF has been the Behavior Rating Inventory of Executive Function (BRIEF; *Gioia* u.a. 2000). Therefore, the ecological validity of ratings such as the BRIEF may be higher than the ecological validity of neuro-psychological measures assessed in standardized test sessions (*Anderson* u.a. 2002). Several studies have shown that performance on cognitive EF tests does not always correspond to performance levels on behavioral measures and ratings (*Anderson* u.a. 2002; *Ponitz* u.a. 2009; *Vriezen/Pigott* 2002). A review by *Toplak/West/Stanovich* (2013) for example based on 20 studies showed that out of the 286 correlations between performance-based and rating measures of EF, only 68 (24%) were statistically significant with a low overall median correlation of only r = .19.

To summarize, different methods to operationalize executive and self-regulatory competences in children exist. They can be distinguished into performance-based measures assessed at child level and teacher or parent ratings. Regarding the construct targeted, performance-based measures and ratings can aim at measuring the subcomponents of EF individually (i.e. working memory, inhibitory control, cognitive flexibility) or in an integrative way. Performance-based measures can differ in their resemblance of real-life situa-

tions and thereby putatively in their ecological validity. They also show differences in their emotional, motivational and social valence as well as in the behavioural level of the participant`s response.

4   Research Questions

The present study assesses the relations between different measures of EF, SR and SC. By that, it attempts to answer three research questions: (1) To what extend are teacher ratings of SC associated with cognitive measures assessing EF in isolation and with integrative rather behavioural EF and SR measures? It was hypothesized that all tasks would be positively and significantly related to each other. However, it was assumed that integrative behavioural EF and SR and cognitive EF measures show higher correlations among each other. Integrated EF and SR tasks are also expected to correlate higher with teacher ratings than distinct cognitive EF tasks. (2) Is the distinction into integrative behavioural EF and SR and cognitive EF measures supported by a principal components analysis (PCA)? It was hypothesized that the distinction into the two dimensions based on the theory behind the measures involved will be supported by the PCA, meaning that two factors will be found. (3) Do behavioural measures or cognitive measures of EF predict teacher ratings of self-control and thoughtfulness better? It was hypothesized that a multiple regression would show that behavioural measures explain more variance of teacher ratings than cognitive EF measures.

5   Methods

5.1   Participants

The data used for the following analyses were gathered as the pre-assessment evaluating the intervention study 'EMIL', a program to improve self-regulation in pre-school children. In total, 217 children between 34 and 72 months (54% boys, M_{age}= 53 months, SD_{age} = 10.63) participated.

Children were nested in eight pre-schools (range: 15-47 children in each school) located on the outskirts of a middle-sized German city in the state of Baden-Württemberg. All eight pre-schools involved in the study worked according to an open concept called 'infans' that does not involve fixed class rooms but different learning areas (*Andres/Laewen* 2011). Each child is assigned to one teacher as the main care taker, who is responsible for the child's adaption when entering pre-school at age three, keeping track of the child's development and communicating with the parents.

5.2   Instruments

The measures included a battery of cognitive and behavioural tasks as well as teacher ratings of constructs related to self-regulation. The distinction into cognitive EF and behavioural EF and SR measures is based on the theory behind the tasks provided by the authors.

Assessment of cognitive executive functions
Four widely applied tests were administered to measure EF on a cognitive level: Digit Span backward (*Petermann/Petermann* 2008), Block Recall (*Gathercole* u.a. 2004), Day-Night Stroop (*Diamond* u.a. 2007) and Hearts & Flowers (*Shing* u.a. 2010).

Digit Span backward (*Petermann/Petermann* 2008). The Digit Span backward test assesses phonological working memory. The test requires the child to repeat a sequence of digits in backward serial order. Lists of the digits one to nine were read out aloud by the test administrator at the rate of one digit per second. Following a short practice session, the test administrator read out a maximum of four lists of each length starting with two digits. List length was increased by one digit when the child recalled three lists of the same length correctly. Testing continued until the child recalled two lists of one length incorrectly. The number of lists correctly recalled is scored (max. six points).

Block Recall (*Gathercole* u.a. 2004). The Block Recall test assesses visual-spatial memory. It makes use of a plate with nine little blocks. The test administrator taps the blocks with a thin stick in a certain order. The child's task is to remember the sequence to tap the blocks in the same order. The test administrator taps a maximum of three blocks of each length starting with one block. The tapping sequence is increased by one if the child recalled two sequences of the same length correctly. Children could obtain a maximum of 21 points.

Day-Night Stroop (*Diamond* u.a. 2007; *Gerstadt/Hong/Diamond* 1994). This measure assesses inhibitory control. In this version, adapted from *Diamond* u.a. (2007), the test administrator showed pictures displayed in a computer screen that showed either a yellow sun on a white background or a yellow moon on a dark blue background. The children were first asked to react verbally to the pictures by saying 'sun' when the picture of the sun was shown or 'moon' when the picture of the moon was shown. Note that the response words differed from the original tests where the children react to the picture of the moon by saying 'day' and to the picture of the sun by saying 'night'. The decision to use different responses than in the original task was made to heighten the level of difficulty as children older than five years often show ceiling effects. After one test trial, the rules were changed. Now the children were asked to react by saying 'sun' when the picture of the moon was shown and by saying 'moon' when the picture of the sun was shown. Hence, they had to suppress their tendency to name what was displayed on the computer screen and instead name something else. As soon as the child responded, the test administrator pressed a button so the next picture was displayed, making sure the children have to respond only verbally and no fine or gross movements were necessary. 16 pictures were displayed in a fixed order. The children received two points for a correct response, one point for a self-corrected response and zero points for an incorrect response. Therefore, a maximum of 32 points could be obtained.

Hearts & Flowers (*Diamond* u.a. 2007; *Shing* u.a. 2010). The Hearts & Flowers test assesses all three EF components, inhibitory control, working memory and cognitive flexibility. It requires the child to react as fast as possible to a stimulus (red heart or blue flower) presented on a computer screen according to two rules. Depending on the stimulus (red heart or blue flower) and the side of the screen that it appears on (left or right), the child has to press either the left or the right out of two buttons on a small keyboard. In the congruent condition (40 trials), the child is presented with hearts only. The child is in-

structed 'to press the button on the same side as the heart'. The incongruent condition (40 trials) consists of blue flowers only. It requires the children 'to press the button on the side opposite the flower'. In the mixed condition (40 trials), red hearts and blue flowers appear in a random order on the screen needing the children to apply the rules of the congruent and the incongruent condition flexibly. Children were given up to seven seconds to respond. As soon as the child pressed a button, the next stimulus appeared on the screen. If they responded to fast (in less than 0.4 seconds) the response was not taken into account. Children received a point for every correct respond (max. 120 points).

Assessment of behavioural executive functions and self-regulation
Two measures were used to assess behavioural EF and SR, Tower and Head-Toes-Knees-Shoulders (HTKS).

Tower (adapted from *Kochanska* u.a. 1996). For the Tower task, the child was asked to take turns with the test administrator to build a tower out of 15 building blocks. After a brief demonstration of turn-taking, the test administrator began building the tower by lying down the first block. After the child took its turn, the test administrator waited with his/her next turn until the child communicated that it was his/her turn again (e.g. verbally, by handing him/her a block or by waiting). Hence, children had to apply their EF to be successful: They had to resist their urge to place the next block when the response of the test administrator was delayed. After the first tower was erected, the test administrator asked the child to build a second tower with him/her using the blocks of the first. Children received a point for each block that placed correctly either by themselves or by the test administrator. The sum of the blocks placed correctly on both trials was used for analyses.

Head-Toes-Knees-Shoulders (HTKS; *Ponitz* u.a. 2008; *Ponitz* u.a. 2009). The extended version of the HTKS was used to measure behavioural SR. The task requires all three components of EF, inhibitory control, working memory and cognitive flexibility. In the test, children are asked to perform the opposite of a dominant response to four different oral commands. In the first section, when asked to touch their head, they have to touch their toes and vice versa. In the second section, two new rules are introduced in addition to the two of the first section. Now, when asked to touch their shoulders, they have to touch their knees and vice versa. In the third section, four new rules are introduced that replace the ones of the two sections before. Now, when asked to touch their head, they have to touch their knees and vice versa. When asked to touch their toes, they have to touch their shoulders and vice versa. The test was administered following the procedure described in *Ponitz* u.a. (2009) with each section consisting of ten test trials. Children received two points for every correct response, one point for a self-corrected response and zero points for a false response, leading to a maximum score of 60 points.

Teacher ratings of self-control and thoughtfulness
To assess behaviours associated with EF, teachers rated the dimension 'self-control/ thoughtfulness' of the observation scale 'Social-emotional well-being and resilience of children in early childhood settings' (PERiK) (*Mayr/Ulich* 2006, 2009). Two examples of items are 'The child waits until it is his/her turn, for example in group discussions, when handing out materials or food' or 'The child can respect the limits set by other children'. The teacher rating is also regarded as a behavioural assessment form as it is based on the observed behaviour shown by the child in the pre-school setting. Although the testing sessions took part at pre-school, teachers were not present during the assessment of the per-

formance-based measures of the children. Therefore, teacher ratings were not influenced by their performance.

Background variables
Background variables of all participants were administered using a caregiver questionnaire on parental education level, migration history and family income.

6 Procedure

After receiving consent of their care givers, the children were seen on two separate occasions in the autumn of the pre-school year. They were administered individually in two one-on-one sessions in a quiet room at their pre-school by research assistants trained in psychology or educational science. The two testing sessions took part within two weeks of each other. Each session lasted about 25 minutes. The tests were administered in a standard order as there was no reason to expect order effects. During Session A, children completed the Tower, Block Recall and the HTKS as well as an interview assessing their social relations to their classmates, which was not analysed for the present study. During Session B, children performed the Day-Night Stroop, Hearts & Flowers, and Digit Span backward. After their assessment, each child received a sticker or a colouring picture. The children's pre-school teachers completed the observation scale within two weeks of the testing session. Parents received the questionnaire on paper via their pre-school within two weeks of their child's testing and were given two weeks to send the questionnaire back in an addressed and stamped envelope.

6.1 Missing data

A number of variables used in the current analyses had missing data (Table 1). For the cognitive EF measures, Digit Span backward data were missing for ten children, Block Recall data were missing for five children, Day-Night Stroop data were missing for eight children, and Hearts & Flowers data were missing for 21 children. Of the integrative behavioural EF and SR measures, Tower data were missing for six children and HTKS data for ten children. The teacher ratings of SC and thoughtfulness were missing for 14 children.

Table 1: Descriptive statistics of cognitive and behavioural measures of executive functions, self-regulation and self-control

Cognitive EF measure	N	Mean	SD	min.	max.
Block Recall (0-21)	212	6.74	3.62	0	15
Day-Night Stroop (0-32)	209	22.75	8.79	0	32
Hearts & Flowers (0-120)	196	77.84	26.41	32	119
Digit Span backward (0-6)	207	.97	1.33	0	5
Behavioural EF &SR measure	**N**	**Mean**	**SD**	**min.**	**max.**
Tower (0-28)	211	23.93	6.71	2	28
HTKS (0-60)	207	20.98	17.08	0	56
Teacher rating	**N**	**Mean**	**SD**	**min.**	**max.**
TR self-control/thoughtfulness (6-30)	203	23.76	4.65	8	30

Note: Scoring range depicted in brackets. Due to missing data, N = 196 to 212 depending on task. TR = teacher rating.

Primary reason for missing data was child refusal to participate in the particular assessment. The high rate of refusals on the Hearts & Flowers was believed to be due to fatigue or boredom on the part of the child, as this was the longest and most monotonous task in the session. The rate of return for the questionnaire on socio-demographic variables was 86.6%. Data were missing for 29 of the 217 children. Data on maternal education was missing for 32 children, on paternal education for 36 children. Data on first language of the child was missing for 29 children. Highest proportion of data was missing for family income (65 children), probably due to the sensitivity of the question.

Data were assumed to be missing completely at random (MCAR). *Little*'s (1988) Missing Completely at Random (MCAR) test performed with IBM SPSS Statistics 22 (IBM Corp., 2013) failed to reach significance, suggesting that the data are indeed missing at random (χ^2 (1, N = 217) = 91.09). Missing data were not imputed.

6.2 Analytic plan

IBM SPSS Statistics 22 (IBM Corp., 2013) was used to obtain descriptive statistics, analyse missing data, and perform data analyses. To address the first research question, Pearson's bi-variate as well as partial correlations controlling for age were run among all variables in the study to investigate the relationship between background variables, cognitive and behavioural measures. For the second research questions, a principal component analysis (PCA) was performed to find support for the distinction between behavioural and cognitive measures. To answer the third research question, whether behavioural measures of EF predict teacher ratings of self-control and thoughtfulness better than cognitive measures of EF, stepwise multiple regressions were carried out.

Maternal and paternal education, family income, child gender, child age and school were included as covariates, as these factors have shown to relate significantly to performance on behavioural and cognitive EF tasks (*Becker* u.a. 2014; *Evans/Rosenbaum* 2008; *Matthews* u.a. 2009; *Wanless* u.a. 2011).

7 Results

7.1 Preliminary analyses

Analyses were based on data from 217 children. Descriptive statistics of socio-economic variables are presented in Table 2. About 53.9% of the participants were male. Although about 31% of the children had a history of migration, only 9.6% of the children did not speak German as their first language.

Table 2: Descriptive statistics: participants and socio-economic variables

		%	N = 217
Child gender			
	male	53.9	117
	female	46.1	100
Age groups		%	N = 217
	34 to 48 months	37.8	82
	49 to 60 months	35.0	76
	61 to 72 months	27.1	59
Migration history		%	N = 188
	no	69.1	130
	yes	30.9	58
First language German		%	N = 188
	yes	90.4	170
	no	9.6	18
Maternal education		%	N = 185
	9 years or less	13.5	25
	10 years	35.1	65
	12 years or more	51.4	95
Paternal education		%	N = 181
	9 years or less	16.6	30
	10 years	27.1	49
	12 years or more	56.4	102
Family income		%	N = 152
	less than 1000	3.3	5
	between 1000 and 2000	19.7	30
	between 2000 and 3000	25.7	39
	between 3000 and 4000	21.7	33
	4000 and more	29.6	45

Descriptive statistics of all measures are presented in Table 1. All measures showed a good range in performance with the exception of the Digit Span backward, where floor effects were found.

There were no age differences between boys and girls (girls: M_{age} = 52.90 months, SD_{age}: 10.74, boys: M_{age} = 52.59, SD_{age} = 10.58; p = .83). No significant performance differences were found between boys and girls across the measures, with one exception: Girls obtained higher scores on Tower than boys, t(209) = 3.56, p < .01.

7.2 Correlations between measures of executive functions, self-regulation and self-control

Pearson's bi-variate correlations were run among all variables in the study to investigate the relationships between background variables, socio-economic status and cognitive and behavioural measures (see Table 3). Child age was highly correlated to all cognitive EF measures (rs ranging from .51 to .64, ps < .001). Correlations between child age and behavioural SR were in general somewhat lower, rs ranging from .22 to .64 (ps < .01). Maternal education was positively related to paternal education (r = .42, p < .001). Both maternal and paternal education were positively correlated to family income (r = .51 and r = .52, respectively, both ps < .001). Background variables also showed significant correlations with several EF measures and teacher ratings. Maternal education showed low but significant positive correlations with performance on Hearts & Flowers (r = .17, p < .05), Digit Span backward (r = .22, p < .01) and teacher ratings of self-control (r = .16, p < .05). Paternal education only showed significant positive correlation with Digit Span backward (r = .17, p < .05). Note that family income was not related to EF measures.

Correlations between EF, SR and SC measures are also presented in Table 3.

Table 3: Bivariate pairwise correlations between variables

Variable	Correlations									
	1	2	3	4	5	6	7	8	9	10
1. Child age	---									
2. Maternal education	.11	---								
3. Paternal education	.00	.42***	---							
4. Family income	.01	.51***	.52***	---						
5. CEF Block Recall	.64***	.10	.09	.08	---					
6. CEF Day-Night Stroop	.51***	.09	.06	.11	.47***	---				
7. CEF Hearts & Flowers	.62***	.17*	.14	.11	.60***	.53***	---			
8. CEF Digit Span backward	.58***	.22**	.17*	.09	.61***	.35***	.64***	---		
9. BEF Tower	.37***	.16	.05	.05	.36***	.34***	.41***	.30***	---	
10. BSR HTKS	.64***	.13	.09	.11	.62***	.51***	.67***	.64***	.37***	---
11. TR Self-control/ thoughtfulness	.22**	.16*	-.02	.11	.32***	.17*	.27***	.26***	.27***	.21**

Note: CEF = cognitive executive functions measure, BEF = behavioural executive functions measure, BSR = behavioural self-regulation measure, TR = teacher rating
*p < .05, **p < .01, ***p < .001

However, due to the high correlations with age, partial correlations controlling for child's age were run for all measures (see Table 4). In the following, partial correlations between measures controlling for age will be presented. When controlling for age, most correlations decrease in height. Between some measures, significant correlations vanished, i.e. between Day-Night Stroop and Digit Span backward as well as between Tower and Digit Span backward. For all but for one measure (Day-Night Stroop and Digit Span backward) significant correlations were found. Significant correlations ranged between .18 and .42. Measures tapping the same EF skill showed higher correlations with one another than with measures tapping different skills. For example, both measures assessing working memory, Block Recall and Digit Span backward, although the former assessing visual-spatial sketchpad and the latter the phonological lope, were moderately correlated (r = .39, p = .000). The Hearts & Flowers, tapping all three components of EF, showed moderate correlations with all measures (rs ranging from .30 to .42, ps < .01).

Table 4:   Partial correlations between all EF, SR and SC measures controlling for age

Variable	Correlations					
	1	2	3	4	5	6
1. CEF Block Recall						
2. CEF Day-Night Stroop	.18*					
3. CEF Hearts & Flowers	.34***	.30***				
4. CEF Digit Span backward	.39***	.04	.42***			
5. BEF Tower	.19*	.28**	.37***	.12		
6. BSR HTKS	.18*	.21*	.50***	.41***	.23*	
7. TR Self-control/ thoughtfulness	.21**	.07	.15*	.16*	.20**	.07

Note: CEF = cognitive executive functions measure, BEF = behavioural executive functions measure, BSR = behavioural self-regulation measure, TR = teacher rating
*p < .05, **p < .01, ***p < .001

Correlation coefficients between the two behavioural measures, Tower and HTKS, reached .23 (p = .01). Therefore, relations among direct assessments of EF were on average higher than between the two behavioural measures. Significant relations were found between most cognitive and the two behavioural measures. No significant correlation was found between Digit Span backward and Tower (r = .12, p = .19). Correlations between cognitive EF measures and HTKS were on average substantially higher (rs from .21 to .50, all ps < .000) than correlations with Tower (rs from .19 to .37, all ps < .05).

For teacher rating, significant correlations were found with three out of four cognitive EF measures: Block Recall (r = .21, p = .01), Digit Span backward (r = .16, p = .03) and Hearts & Flowers (r = .15, p = .04). A low but significant correlation was found between the Tower and teacher rating (r = .20, p = .01). The correlation between the other behavioural measure, HTKS, and the teacher rating did not reach significance (r = .07, p = .38).

7.3 Distinction between cognitive and behavioural measures

In order to examine the underlying components of the cognitive and behavioural measures of EF related constructs, a principal components analysis (PCA) was carried out. All described measures were included in the analysis. The first unrotated principal component (FUPC) accounts for the maximum amount of variance of the measured variables. Item loadings reflect the correlation between a particular measure and the overall component. A relatively high Kaiser-Meyer-Olkin measure of sampling adequacy (KMO =.69) confirmed the validity of using a factor analysis for structure detection. Loadings of .30 or above are typically considered acceptable (*Tabachnick/Fidell* 1983). All measures had loadings above this threshold on the FUPC, as shown in Table 5.

Principal components analysis (PCA) using direct oblimin rotation extracted two major components. Together they accounted for 61.8% of the variance. All factors of an eigenvalue higher than 1 were selected. Scree plots were examined to confirm factor selections. Measures with factor coefficients higher than .5 were considered to load on a certain factor. The first principal component identified accounted for 47.91% of the variance and included the performance on Block Recall, Day-Night Stroop, Hearts & Flowers, Digits Span and HTKS. The second principal component accounted for 13.86% of the variance and included Tower and teacher rating of 'self-control/thoughtfulness'.

Table 5: Results of Principal Components Analyses (PCA) of executive function and self-regulation measures

Measure	cognitive measures	behavioural measures
CEF Block Recall	.78	---
CEF Day-Night Stroop	.60	---
CEF Hearts & Flowers	.84	---
CEF Digit Span backward	.78	---
BEF Tower	---	.57
BSR HTKS	.83	---
TR Self-control	---	.75

Note: Component loadings >.5 are presented. CEF = cognitive executive functions measure, BEF = behavioural executive functions measure, BSR = behavioural self-regulation measure, TR = teacher rating

7.4 Predicting teacher rating of self-control and thoughtfulness

Results of the correlational analyses indicated that all cognitive and behavioural tasks show significant but low correlations with teacher ratings of self-control and thoughtfulness. To test their predictive value of teacher rating, a Hierarchical Multiple Regression was carried out. We controlled for the effects of the covariates child age, parental education and migration history by entering them into the first model. The second model consisted of all measures of EF and SR. Parameter estimates are presented in Table 6.

Only Block Recall (β = .24, B = .29, p = .032) was significantly associated with teacher rating after adjusting for the covariates child age, child gender, parental education and migration history. Additionally, performance on Tower was marginally associated with teacher rating (β = .15, B = .11, p = .063). The covariates explained 6.4% of the variance in teacher rating of self-control and thoughtfulness. Measures of EF and SR explained an additional 9.3% of the variance.

Table 6: Cognitive and Behavioural Measures Predicting Teacher Rating of Self-control and Thoughtfulness

Measure	B	SE	β
Child age	-.02	.05	-.04
Maternal education	-.39	.56	-.06
Paternal education	.22	.51	.04
Migration history	-1.41†	.76	-.15
CEF Block Recall	.29*	.13	.24
CEF Day-Night Stroop	.01	.05	.01
CEF Hearts & Flowers	.02	.02	.10
CEF Digit Span backward	.32	.35	.10
BSR Tower	.11†	.06	.15
BSR HTKS	-.02	.03	-.09

Note: B = Unstandardized Estimate, SE = Standard Error, β = Standardized Estimate
* $p < .05$, † $p < .10$.

8 Discussion

Choosing appropriate, meaningful and valid measures of EF and SR for young children that reflect their behaviour shown in everyday life can be a difficult undertaking (e.g. *Anderson* 2002; *Carlson* 2005). Researchers are advised to assess both constructs in a comprehensive manner through the inclusion of multiple measures and different levels of analyses (e.g. neurological, physiological, rating measures and questionnaires) (*Anderson/Reidy* 2012; *McClelland* u.a. 2015).

For a better distinction between the numerous existent measures, measures are sorted into various categories such as performance-based measures and behaviour ratings. Within the category of performance–based measures, finer distinctions exist, e.g. into cognitive or behavioural (*McClelland/Cameron/Connor* u.a. 2007; *McClelland/Cameron/Wanless* u.a. 2007; *Ponitz* u.a. 2009) and into hot or cool measures (e.g. *Hongwanishkul* u.a. 2005; *Zelazo/Carlson* 2012).

The aim of the study was to explore the validity of the putative distinction into cognitive and behavioural measures of EF, SR and SC. It was investigated whether measures of the same category really tap similar constructs. Also, we wanted to find out what measures are related closest to behaviour of preschoolers observed by their teachers.

Six cognitive and behavioural EF and SR measures as well as teacher ratings of children's SC and thoughtfulness were used. Performance-based measures were assessed in two testing sessions. It has to be noted that both behavioural measures (Tower and HTKS) were administered within the same testing session. Hence, it could be argued that tasks administered within the same session could correlate higher due to similar levels of motivation or concentration.

All measures were suitable for the age range in question except for the Digit Span backward, which showed to be quite challenging especially for younger children. Gender differences became evident in one measure, the Tower task. It has been found before that girls outperform boys on EF and SR tasks, especially in this age group (e.g. *Carlson/Moses* 2001; *Matthews/Ponitz/Morrison* 2009; *Moffitt* u.a. 2011). The set-up of the

Tower also consists of a strong social component: taking turns and keeping the partner in mind. Studies have shown that girls do not only outperform boys on EF and SR tasks but regarding their social-emotional competences (e.g. *Denham* u.a. 2003; *Walker* 2005).

Parental education showed significant relations with just one out of seven EF tasks and family income with none. Several studies have found that family income is also related to children's performance on EF task (*Evans/Rosenbaum* 2008). This finding was not supported by the present study.

Most measures were significantly correlated to each other. Some cognitive EF measures did not show significant correlations, e.g. Day-Night Stroop and Digit Span backward. However, since the two measures tap different components of EF (e.g. inhibitory control and working memory) that finding is not really astonishing. It was hypothesized that correlations among the behavioural EF and SR measures and teacher ratings are higher than their correlations with the cognitive EF measures. Conversely, profound differences were found in the strengths of correlations between the different measures. The behavioural SR measure HTKS correlated highly with cognitive EF measures, much higher than with the second behavioural measure, Tower, and the teacher rating of SC and thoughtfulness. Therefore, the first hypothesis was not supported by the data. The unexpected result might be due to the fact, that the HTKS addresses all three EF components (*Ponitz* u.a. 2008; *Ponitz* u.a. 2009) whereas the Tower mainly addresses inhibitory control (*Kochanska* u.a. 1996). Although both tasks require the child to react by carrying out responses behaviourally, the Tower may resemble a real-life situation even better than the HTKS and therefore may be higher in its ecological validity than the HTKS. Its significant correlation with the teacher rating which was not evident for the HTKS supports this assumption.

The principal component analysis supported the results of the correlational analysis. The PCA showed no evidence for the distinction of the measures into cognitive and behavioural stated by the authors. Two distinct components, cognitive measures and behavioural measures, were identified. The first component consisted of all cognitive measures and the HTKS. The second component consisted of Tower and the teacher rating. It was hypothesized that the HTKS as a behavioural SR measure would fall into the second category. Based on the results it may be assumed that the HTKS is more closely related to cognitive than to other behavioural measures.

The question that arises is: What makes a measure behavioural in nature? The authors of the HTKS claim that it measures behavioural self-regulation, although performance seems to rely strongly on the developmental level of the three components of EF that are mentioned explicitly in the task-description (*Ponitz* u.a. 2008; *Ponitz* u.a. 2009). The Tower on the other hand assesses the ability to suppress and initiate activity to signal (*Kochanska* u.a. 1996). In comparison to the HTKS, which taps all three EF components, the component of inhibitory control could be argued to be most central within the Tower as the child has to inhibit itself to lay a block when it's the turn of the test administrator. The child is also required to hold the rule in mind to take turns with placing a block. However, the rule is never changed as it is the case in the HTKS. Therefore, the task also relies only for a small part on working memory capacity.

Looking at the testing situations of both tasks, it becomes evident that the Tower resembles the typical play situation that children experience frequently in pre-school and at home very well. It becomes evident that the Tower has a much stronger social component than the HTKS as the child has to be considered of the test administrator in order to per-

form well. The teacher rating scale that was used has a strong social component too, which could be the common factor that could explain the outcomes of the PCA. In comparison, the HTKS, although it can be regarded as a rather playful measure, is less of a mutual play situation due to the fact that the child is required to act upon the commands given by the test administrator. In fact, this is the only factor that distinguishes the HTKS from all the cognitive EF tasks: The child is required to act out its response by reaching towards a certain body part. Apart from that, the task is very similar to the Hearts & Flowers for example, were the child has to respond as fast and as correct as possible, just not with gross movements but with a fine movement (pressing a button). The HTKS again requires the child to be perseverant and to stay on task to perform well.

Therefore, high correlations between the HTKS and cognitive EF measures might be due to the fact that all of them have a common denominator: perseverance and focus on task. Both factors are thought to be important for academic success but not necessarily for social behaviour (*Rhoades/Greenberg/Domitrovich* 2009). A study by *Ponitz* u.a. (2009) showed that HTKS is a valid predictor of academic but not of social skills of children in primary school. Performance on the HTKS could predict literacy and mathematical skills half a year later, but not their interpersonal competences. On the other hand, inhibitory control, which plays an important role in the Tower task, has shown to be very important for social-emotional competence (*Kochanska/Murray/Harlan* 2000; *Rhoades/Greenberg/Domitrovich* 2009).

The third aim of the study was the predictive value of EF and SR measures when it comes to everyday behaviour. A teacher rating of children's self-control and thoughtfulness was included. One cognitive EF measure of visual-spatial working memory explained a significant proportion of the variance in children's SC and thoughtfulness. One measure of behavioural EF, Tower, was marginally significant. The significant contribution of working memory was quite unexpected. Based on the literature, it was hypothesized that behavioural measures will explain more variance than cognitive measures. Also, out of the three EF components, inhibitory control would be the one closest associated with the rating scale assessing SC and thoughtfulness. Therefore, a cognitive EF measure assessing inhibitory control would have been most likely. The marginal significance of the Tower was therefore much more expected.

Is a distinction between cognitive EF and behavioural SR measures useful? Based on the findings of this study, the distinction of measures into different subcategories (i.e. behavioural and cognitive) claimed by the authors has to be handled with care. Especially when chosen with the aim of ensuring ecological validity of the assessment. More refined categories of measures than merely the distinction between performance-based measure or ratings, EF and SR task or cognitive and behavioural measure could be of good use when aiming at assessing EF related constructs in a comprehensive manner. However, each category has to be defined more clearly so the measures are distinguishable. Regarding the category of behavioural measures for example, it needs to be discussed whether a task in which the response has to be acted out rather than be submitted by an utterance or the press of a button already qualifies as behavioural as it is argued by the authors (*Ponitz* u.a. 2008; *Ponitz* u.a. 2009). Moreover, it should also be considered to distinguish between measures that require social interaction and that do not. Interventions for example that aim at improving the application of EF or self-regulatory competences in school setting would benefit of a more distinctive labelling of measures that assess EF related behaviour in academic as well as in social contexts.

References

Anderson, P.J. (2002): Assessment and development of executive function (EF) during childhood. Child Neuropsychology, 8, 2, pp. 71-82, doi:10.1076/chin.8.2.71.8724.

Anderson, P.J. (2010): Towards a developmental model of executive function. In: *Anderson, V.A./Jacobs, R./Anderson, P.J.* (eds.): Executive functions and the frontal lobes: A lifespan perspective. – New York, London, pp. 3-21. doi:10.4324/9780203837863.

Anderson, P.J./Reidy, N. (2012): Assessing executive function in preschoolers. Neuropsychology Review, 22, 4, pp. 345-360, doi:10.1007/s11065-012-9220-3.

Anderson, V.A. (1998): Assessing executive functions in children: Biological, psychological, and developmental considerations. Neuropsychological rehabilitation, 8, 3, pp. 319-349, doi:10.1080/713755568.

Anderson, V.A./Anderson, P.J./Northam, E./Jacobs, R./Mikiewicz, O. (2002): Relationships between cognitive and behavioral measures of executive function in children with brain disease. Child Neuropsychology, 8, 4, pp. 231-240, doi:10.1076/chin.8.4.231.13509.

Andres, B./Laewen, H.-J. (2011): Das infans-Konzept der Frühpädagogik. Bildung und Erziehung in Kindertagesstätten. – Berlin, Weimar.

Becker, D.R./Miao, A./Duncan, R.J./McClelland, M.M. (2014): Behavioral self-regulation and executive function both predict visuomotor skills and early academic achievement. Early Childhood Research Quarterly, 29, 4, pp. 411-424, doi:10.1016/j.ecresq.2014.04.014.

Baddeley, A.D. (1986): Working memory. – London.

Bierman, K.L./Nix, R.L./Greenberg, M.T./Blair, C./Domitrovich, C.E. (2008): Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. Development and Psychopathology, 20, pp. 821-843, doi:10.1017/s0954579408000394.

Blair, C. (2002): School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. American Psychologist, 57, 2, pp. 111-127, doi:10.1037//0003-066X.57.2.111.

Blair, C./Razza, R.P. (2007): Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. Child Development, 78, 2, pp. 647-663, doi:10.1111/j.1467-8624.2007.01019.x.

Carlson, S.M. (2005): Developmentally sensitive measures of executive function in preschool children. Developmental Neuropsychology, 28, 2, pp. 595-616, doi:10.1207/s15326942dn2802_3.

Carlson, S.M./Moses, L.J. (2001): Individual Differences in Inhibitory Control and Children's Theory of Mind. Child Development, 72, 4, pp. 1032-1053, doi:10.1111/1467-8624.00333.

Carlson, S.M./Wang, T.S. (2007): Inhibitory control and emotion regulation in preschool children. Cognitive Development, 22, 4, pp. 489-510, doi:10.1111/1467-8624.00333.

Denham, S.A./Blair, K.A./DeMulder, E./Levitas, J./Sawyer, K./Auerbach-Major, S./Queenan, P. (2003): Preschool Emotional Competence: Pathway to Social Competence? Child Development, 74, 1, pp. 238-256, doi:10.1111/1467-8624.00533.

Diamond, A. (2007): Consequences of variations in genes that affect dopamine in prefrontal cortex. Cerebral Cortex, 17 (suppl. 1), pp. 161-170, doi:10.1093/cercor/bhm082.

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

Diamond, A./Barnett, W.S./Thomas, J./Munro, S. (2007): Preschool program improves cognitive control. Science, 318, 5855, pp. 1387-1388, doi:10.1126/science.1151148.

Duckworth, A.L./Seligman, M.E. (2005): Self-discipline outdoes IQ in predicting academic performance of adolescents. Psychological Science, 16, 12, pp. 939-944, doi:10.1111/j.1467-9280.2005.01641.x.

Duncan, J. (1986): Disorganisation of behaviour after frontal lobe damage. Cognitive Neuropsychology, 3, 3, pp. 271-290, doi:10.1080/02643298608253360.

Espy, K.A./Bull, R./Kaiser, H./Martin, J./Banet, M. (2008): Methodological and conceptual issues in understanding the development of executive control in the preschool period. In: *Anderson, V.A./Jacobs, R./Anderson, P.J.* (eds.): Executive functions and the frontal lobes: A lifespan perspective. – New York, London, pp. 105-121, doi:10.4324/9780203837863.

Evans, G.W./Rosenbaum, J. (2008): Self-regulation and the income-achievement gap. Early Childhood Research Quarterly, 23, 4, pp. 504-514, doi:10.1016/j.ecresq.2008.07.002.

Falk, A./Heckman, J.J. (2009): Lab experiments are a major source of knowledge in the social sciences. Science, 326, 5952, pp. 535-538, doi:10.1126/science.1168244.

Gathercole, S.E./Pickering, S.J./Ambridge, B./Wearing, H. (2004): The Structure of Working Memory From 4 to 15 Years of Age. Developmental Psychology, 40, 2, pp. 177-190, doi:10.1037/0012-1649.40.2.177.

Gerstadt, C./Hong, Y./Diamond, A. (1994): The relationship between cognition and action: performance of children 3 1/2-7 years old on a Stroop-like day-night test. Cognition, 53, 2, pp. 129-153, doi:10.1016/0010-0277(94)90068-x.

Gioia, G.A./Isquith, P.K. (2004): Ecological Assessment of Executive Function in Traumatic Brain Injury. Developmental Neuropsychology, 25, 1-2, pp. 135-158, doi: 10.1080/87565641.2004.9651925.

Gioia, G.A./Isquith, P.K/Guy, S.C/Kenworthy, L. (2000): Behavior Rating Inventory of Executive Function: BRIEF. – Odessa.

Heaton, R.K. (1993): Wisconsin card sorting test: Computer version 2. Odessa: Psychological Assessment Resources.

Hofmann, W.J./Schmeichel, B.J./Baddeley, A. (2012): Executive functions and self-regulation. Trends in Cognitive Sciences, 16, 3, pp. 174-180, doi:10.1016/j.tics.2012.01.006.

Hongwanishkul, D./Happaney, K.R./Lee, W.S./Zelazo, P.D. (2005): Assessment of hot and cool executive function in young children: Age-related changes and individual differences. Developmental Neuropsychology, 28, 2, pp. 617-644, doi: 10.1207/s15326942dn2802_4.

Hughes, C. (1998): Executive function in preschoolers: Links with theory of mind and verbal ability. British Journal of Developmental Psychology, 16, 2, pp. 233-253, doi:10.1111/j.2044-835x.1998.tb00921.x.

Hughes, S.O./Power, T.G./O'Connor, T.M./Fisher, J.O. (2015): Executive functioning, emotion regulation, eating self-regulation, and weight status in low-income preschool children: How do they relate? Appetite, 89, pp. 1-9, doi:10.1016/j.appet.2015.01.009.

Isquith, P.K./Crawford, J.S./Espy, K.A./Gioia, G.A. (2005): Assessment of executive function in preschool-aged children. Mental retardation and developmental disabilities research reviews, 11, 3, pp. 209-215, doi:10.1002/mrdd.20075.

Jensen, A.R./Rohwer, W.D. (1966): The Stroop color-word test: A review. Acta psychologica, 25, pp. 36-93.

Kochanska, G./Murray, K.T./Harlan, E.T. (2000): Effortful control in early childhood: continuity and change, antecedents, and implications for social development. Developmental Psychology, 36, 2, pp. 220-232, doi:10.1037/0012-1649.36.2.220.

Kochanska, G./Murray, K.T./Jacques, T.Y./Koenig, A.L./Vandegeest, K.A. (1996): Inhibitory control in young children and its role in emerging internalization. Child Development, 67, 2, pp. 490-507, doi:10.2307/1131828.

Little, R.J. (1988): A test of missing completely at random for multivariate data with missing values. Journal of the American Statistical Association, 83, 404, pp. 1198-1202, doi:10.1080/01621459.1988.10478722.

Luria, A.R. (1976): The working brain: An introduction to neuropsychology. – New York.

Matthews, J.S./Ponitz, C.C./Morrison, F.J. (2009): Early gender differences in self-regulation and academic achievement. Journal of Educational Psychology, 101, 3, pp. 689-704, doi:10.1037/a0014240.

Matthews, J.S./Ponitz, C.C./Morrison, F.J./National Institute of Child, H./Human, D./National Science, F./Institute of Education, S. (2009): Early gender differences in self-regulation and academic achievement. Journal of Educational Psychology, 101, 3, pp. 689-704, doi: 10.1037/a0014240.

Mayr, T./Ulich, M. (2006): Perik: positive Entwicklung und Resilienz im Kindergartenalltag: Beobachtungsbogen.

Mayr, T./Ulich, M. (2009): Social-emotional well-being and resilience of children in early childhood settings – PERIK: An empirically based observation scale for practitioners. Early Years, 29, 1, pp. 45-57, doi:10.1080/09575140802636290.

McClelland, M.M./Cameron, C.E./Connor, C.M./Farris, C.L./Jewkes, A.M./Morrison, F.J. (2007): Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. Developmental Psychology, 43, 4, pp. 947-959, doi:10.1037/0012-1649.43.4.947.

McClelland, M.M./Cameron, C.E./Wanless, S.B./Murray, A. (2007): Executive function, behavioral self-regulation, and social-emotional competence. Contemporary perspectives on social learning in early childhood education, pp. 113-137.

McClelland, M.M./John Geldhof, G./Cameron, C.E./Wanless, S.B. (2015): Development and Self-Regulation. In: *Lerner, R.M./Liben, L.S./Mueller, U.* (eds.): Handbook of child psychology and developmental science. – Hoboken, New Jersey, pp. 1-43, doi:10.1002/9781118963418.childpsy114.

Miller, E.K./Cohen, J.D. (2001): An integrative theory of prefrontal cortex function. Annual Review of Neuroscience, 24, 1, pp. 167-202, doi: 10.1146/annurev.neuro.24.1.167.

Miranda, A./Colomer, C./Mercader, J./Fernández, M.I./Presentación, M.J. (2015): Performance-based tests versus behavioral ratings in the assessment of executive functioning in preschoolers: Associations with ADHD symptoms and reading achievement. Frontiers in psychology, 6, pp. 1-10, doi:10.3389/fpsyg.2015.00545.

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, pp. 8-14, doi:10.1177/0963721411429458.

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, pp. 49-100, doi:10.1006/cogp.1999.0734.

Moffitt, T.E./Arseneault, L./Belsky, D./Dickson, N./Hancox, R.J./Harrington, H./Houts, R./Poulton, R./Roberts, B.W./Ross, S./Sears, M.R./Thomson, W.M./Caspi, A. (2011): A gradient of childhood self-control predicts health, wealth, and public safety. Proceedings of the National Academy of Sciences, 108, 7, pp. 2693-2698, doi:10.1073/pnas.1010076108.

Morgan, A.B./Lilienfeld, S.O. (2000): A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. Clinical Psychology Review, 20, 1, pp. 113-136, doi:10.1016/s0272-7358(98)00096-8.

Pennington, B.F./Ozonoff, S. (1996): Executive functions and developmental psychopathology. Journal of child psychology and psychiatry, 37, 1, pp. 51-87, doi:10.1111/j.1469-7610.1996.tb01380.x.

Petermann, F./Petermann, U. (2008): HAWIK-IV: Hamburg-Wechsler Intelligenztest für Kinder-IV. – New York, doi:10.1026/0942-5403.17.2.71.

Ponitz, C.E.C./McClelland, M.M./Jewkes, A.M./Connor, C.M./Farris, C.L./Morrison, F.J. (2008): Touch your toes! Developing a direct measure of behavioral regulation in early childhood. Early Childhood Research Quarterly, 23, 2, pp. 141-158, doi:10.1016/j.ecresq.2007.01.004.

Ponitz, C.E.C./McClelland, M.M./Matthews, J.S./Morrison, F.J. (2009): A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. Developmental Psychology, 45, 3, pp. 605-619, doi:10.1037/a0015365.

Rhoades, B.L./Greenberg, M.T./Domitrovich, C.E. (2009): The contribution of inhibitory control to preschoolers' social-emotional competence. Journal of Applied Developmental Psychology, 30, 3, pp. 310-320, doi: 10.1016/j.appdev.2008.12.012.

Rimm-Kaufmann, S.E.C./Curby, T.W./Grimm, K.J./Nathanson, L./Brock, L.L. (2009): The Contribution of Children's Self-Regulation and Classroom Quality to Children's Adaptive Behaviors in the Kindergarten Classroom. Developmental Psychology, 45, 4, pp. 958-972, doi:10.1037/a0015861.

Ritter, B.C./Perrig, W./Steinlin, M./Everts, R. (2014): Cognitive and behavioral aspects of executive functions in children born very preterm. Child Neuropsychology, 20, 2, pp. 129-144, doi:10.1080/09297049.2013.773968.

Roth, R.M./Isquith, P.K./Gioia, G.A. (2005): BRIEF-A: Behavior Rating Inventory of Executive Function – Adult Version: Professional Manual. – Odessa.

Rothbart, M.K./Posner, M.I. (1985): Temperament and the development of self-regulation. In: *Hartlage, L.C./Telzrow, C.F.* (eds.): The neuropsychology of individual differences. – Heidelberg, NY, S. 93-123, doi:10.1007/978-1-4899-3484-0_5.

Shing, Y.L./Lindenberger, U./Diamond, A./Li, S.C./Davidson, M.C. (2010): Memory maintenance and inhibitory control differentiate from early childhood to adolescence. Developmental Neuropsychology, 35, 6, pp. 679-697, doi:10.1080/87565641.2010.508546.

Smith-Donald, R./Raver, C.C./Hayes, T./Richardson, B. (2007): Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. Early Childhood Research Quarterly, 22, 2, pp. 173-187, doi:10.1016/j.ecresq.2007.01.002.

Stroop, J.R. (1935): Studies of interference in serial verbal reactions. Journal of experimental psychology, 18, 6, pp. 643.

Stuss, D.T./Alexander, M.P. (2000): Executive functions and the frontal lobes: A conceptual view. Psychological research, 63, 3-4, pp. 289-298, doi:10.1007/s004269900007.

Tabachnick, B.G./Fidell, L.S. (1983): Using multivariate statistics. – New York.

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, pp. 131-143, doi:10.1111/jcpp.12001.

Vriezen, E.R./Pigott, S.E. (2002): The Relationship Between Parental Report on the BRIEF and Performance-Based Measures of Executive Function in Children with Moderate to Severe Traumatic Brain Injury. Child Neuropsychology, 8, 4, pp. 296-303, doi:10.1076/chin.8.4.296.13505.

Walker, S. (2005): Gender differences in the relationship between young children's peer-related social competence and individual differences in theory of mind. The Journal of genetic psychology, 166, 3, pp. 297-312, doi:10.3200/gntp.166.3.297-312.

Wanless, S.B./McClelland, M.M./Tominey, S.L./Acock, A.C. (2011): The Influence of Demographic Risk Factors on Children's Behavioral Regulation in Prekindergarten and Kindergarten. Early Education and Development, 22, 3, pp. 461-488, doi: 10.1080/10409289.2011.536132.

Welsh, J.A./Nix, R.L./Blair, C./Bierman, K.L./Nelson, K.E. (2010): The Development of Cognitive Skills and Gains in Academic School Readiness for Children from Low-Income Families. Journal of Educational Psychology, 102, 1, pp. 43-53, doi:10.1037/a0016738.

Welsh, M.C./Pennington, B.F./Groisser, D.B. (1991): A Normative Developmental-Study of Executive Function – A Window on Prefrontal Function in Children. Developmental Neuropsychology, 7, 2, pp. 131-149, doi:10.1080/87565649109540483.

Willoughby, M.T./Blair, C.B./Wirth, R./Greenberg, M. (2012): The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. Psychological assessment, 24, 1, pp. 226-239, doi:10.1037/a0025361.

Zelazo, P.D./Carlson, S.M. (2012): Hot and Cool Executive Function in Childhood and Adolescence: Development and Plasticity. Child Development Perspectives, 6, 4, pp. 354-360, doi:10.1111/j.1750-8606.2012.00246.x.

Zelazo, P.D./Carter, A./Reznick, J.S./Frye, D. (1997): Early development of executive function: A problem-solving framework. Review of general psychology, 1, 2, pp. 198-226, doi:10.1037/1089-2680.1.2.198.

Zelazo, P.D./Müller, U./Frye, D./Marcovitch, S. (2003): The development of executive function in early childhood. Monographs of the Society for Research in Child Development, 68, 3, pp. 93-119, doi:10.1111/j.0037-976x.2003.00266.x.

Zelazo, P.D./Qu, L./Müller, U./Schneider, W./Schumann-Hengsteler, R./Sodian, B. (2005): Hot and cool aspects of executive function: Relations in early development. In: *Schneider, W./Schumann-Hengsteler, R./Sodian, B.* (eds.): Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind. – New York, London, S. 71-93, doi:10.4324/9781410612007.