

Assessing effortful control in typical and atypical development: are questionnaires and neuropsychological measures interchangeable? A latent-variable analysis

Vicky Samyn^{a, *}, Ph.D., Herbert Roeyers^{a, b}, Ph.D., Patricia Bijttebier^c, Ph.D., Yves Rosseel^{a, d}, Ph.D.,
& Jan R. Wiersema^{a, e}, Ph.D.

^aGhent University, Henri Dunantlaan 2, 9000 Ghent, Belgium

^b *E-mail address:* Herbert.Roeyers@UGent.be

^c University of Leuven, Tiensestraat 102 – bus 3720, 3000 Leuven, Belgium. *E-mail address:*

Patricia.Bijttebier@psy.kuleuven.be

^d *E-mail address:* Yves.Rosseel@UGent.be

^e *E-mail address:* Roeljan.Wiersema@UGent.be

* **Correspondence to** Vicky Samyn, Ghent University, Department of Experimental Clinical and Health Psychology, Henri Dunantlaan 2, B-9000 Ghent, Belgium. Telephone: +32(0)9-2649415. Fax: +32(0)92646489. Email: vicky_samyn@hotmail.com

Abstract

Objective: Effortful control (EC), the self-regulation component of temperament, is traditionally measured using questionnaires. Through the years, several neuropsychological measures originating from the cognitive psychology and the executive function (EF) literature have been introduced in the domain of temperament research to tap EC. Although this is not particularly surprising, given the conceptual overlap between EC and EF, it remains unclear whether EC questionnaires and neuropsychological EF tasks can really be used interchangeably when measuring EC. The current study addressed two important aspects in evaluating the interchangeability of both types of measures, that is: (a) do they measure the same construct?, and (b) do they give the same results when comparing clinical populations? **Method:** Three EC questionnaires, two inhibitory control tasks, and two attentional control tasks were administered in 148 typically developing children, 30 children with attention-deficit/hyperactivity disorder (ADHD), and 31 children with autism spectrum disorder (ASD). All children were between 10 and 15 years of age and had a full scale IQ of 80 or higher. **Results:** Confirmatory Factor Analyses revealed that the questionnaires and EF tasks do not capture the same underlying latent variable(s). Groups could not be differentiated from each other based on their performance on EF tasks, whereas significant group differences were found for all EC-reports. **Conclusions:** Overall, our findings show more differences than commonalities between the EC questionnaires and EF tasks and, consequently, suggest that both types of measures should not be used interchangeably. **Keywords:** Effortful Control, temperament questionnaires, neuropsychological measures, ADHD, ASD.

1. Introduction

The construct of effortful control (EC) has received a substantial amount of attention in child development research in the course of the past decade (Bijttebier and Roeyers 2009). EC refers to the self-regulation component of temperament and is defined as “the ability to inhibit a dominant response in order to perform a subdominant response” (Rothbart and Bates 1998, p. 137). EC involves both an attentional aspect (i.e., the ability to focus or shift attention when needed) and a behavioural aspect (i.e., the ability to inhibit or activate behaviour in accordance with situational demands) (Rothbart 1989). EC levels have been linked not only to the positive emotional, social, and cognitive development in children (e.g., Eisenberg et al. 2004), but also to the onset and/or maintenance of both internalizing and externalizing problems (e.g., Muris and Ollendick 2005), as well as to developmental disorders such as attention-deficit/hyperactivity disorder (ADHD; e.g., Martel and Nigg 2006; Roeyers and Bijttebier 2011) and autism spectrum disorders (ASD; e.g., Konstantareas and Stewart 2006; Samyn et al. 2011). In general, within the domain of temperament research, EC is considered a key component in development and it is assumed that a vulnerability to develop psychopathology is largely associated with a temperament characterized by, among other things, low levels of EC (e.g., Lonigan and Phillips 2001).

EC is traditionally measured using questionnaires (e.g., Derryberry and Reed 2002; Ellis and Rothbart 2001; Lonigan and Phillips 2001) or, in young children, by means of Kochanska’s multitask battery (e.g., Walk a Line, Turtle’s House, Telephone Poles, Circle, Star, and Lowering Voice; Kochanska and Knaack 2003). However, through the years, several researchers have introduced neuropsychological measures (e.g., Go/No-Go, Stroop) originating from the domain of cognitive psychology and the executive function (EF) literature to tap EC (e.g., Lengua et al. 2007; for a review, see Zhou et al. 2012). This is not particularly surprising, given the conceptual overlap between EC and EF. Both constructs show considerable similarities in terms of definition and core components (e.g., a focus on inhibition and identifying executive attention as an underlying process; for an extensive review, see Zhou et al. 2012). However, up till now, it remains unclear whether EC questionnaires and neuropsychological EF tasks can really be used interchangeably when measuring

EC. To date, most studies have focused on either EC or EF and studies focusing on both within a single sample are limited. The few studies that did include questionnaires as well as EF tasks at best show small to moderate correlations between both (e.g., Blair and Razza 2007; Verstraeten et al. 2010). Overall, findings are inconclusive and call for additional research. Also, the practice of interchanging these measures hampers the interpretation of and comparison between results of different EC studies (e.g., when comparing different populations on their ability to effortfully control their attention and/or behavior). Therefore, the main aim of our study was to investigate the extent to which EC questionnaires and a selection of neuropsychological EF measures considered to assess inhibitory control, attention focusing, and attention shifting are interchangeable when measuring EC. In the present study we addressed two important aspects in evaluating the interchangeability of measures, that is: (a) do they measure the same construct?, and (b) do they give the same results, for example when comparing clinical populations? First, a necessary (although not sufficient) condition for measures to be interchangeable is that they tap the same (or very similar) construct(s) (Fine 1992). In order to evaluate this, we used latent-variable analyses. Applying a latent-variable approach (as opposed to a correlational approach) has several advantages. In a nutshell, this technique statistically extracts the common variance among multiple measures chosen to tap the same underlying construct while excluding the variance attributable to idiosyncratic task requirements and measurement error (Friedman and Miyake 2004). For the purpose of the current study, two models were analyzed. First, a model was investigated in which the different EC total scores and the relevant variables of the neuropsychological EF tasks all load on the same underlying latent factor and thus are presumed to be best characterized as a unitary factor (see Figure 1 for the hypothesized model). Although this model is in accordance with the practice that both types of measures are used interchangeably to investigate EC, there are some reasons to assume that a single factor approach may not be the best solution. For example, the relatively small correlations between both types of measures (e.g., for a review, see Zhou et al. 2012) and the lack of correlation between scores on EC questionnaires and executive attention (Samyn et al. 2013). One might argue that it is unlikely that scales and neuropsychological tasks designed to tap behavioural control will load on the exact same underlying factor as scales and tasks

considered to tap attentional control. Therefore, we also tested an alternative model in which the different measures were assumed to be best represented by two latent variables, an ‘Attentional Control’ factor and an ‘Inhibitory Control’ factor (see Figure 2).

Secondly, if different measures are interchangeable, they are expected to yield the same results (e.g., Powell et al. 2007). We therefore investigated whether EC questionnaires and neuropsychological tasks give the same results when comparing typically (TD) and atypically developing children on self-regulation abilities. For the purpose of the current study we compared EC scores on questionnaires and EF performance between TD children, children with ADHD, and children with ASD. Both clinical groups have been included in EC as well as EF studies and are known to show difficulties in EC (e.g., Martel and Nigg 2006; Konstantareas and Stewart 2006; Samyn et al. 2011) and EF in comparison with TD peers. For a detailed overview of previous findings regarding group differences in EF, we refer the reader to recent reviews and meta-analyses on this topic (e.g., De La Fuente et al. 2013; Gargaro et al. 2011; Hill 2004; Mullane et al. 2009; O’Hearn et al. 2008; Sergeant et al. 2002). If both types of measures are interchangeable, we expect them to lead to the same (or at least very similar) results. Specifically, we expect that if children with ADHD and/or ASD show difficulties in regulating attention and/or behaviour in comparison with TD peers based on questionnaires, this should also be reflected in the results of the performance based measures and vice versa.

2. Method

2.1. Participants

209 children aged 10-15 years with an estimated full scale IQ (FSIQ) of 80 or higher participated in our study. 148 children were typically developing controls (TD; 64% boys; age: $M = 12.73$, $SD = 1.48$; estimated FSIQ: $M = 107.21$, $SD = 11.68$), 31 children had a formal diagnosis of ASD (all boys; age: $M = 12.83$, $SD = 1.41$; estimated FSIQ: $M = 101.16$, $SD = 12.48$), and 30 children had a formal diagnosis of ADHD (all boys; age: $M = 13.16$, $SD = 1.61$; estimated FSIQ: $M = 108.20$, $SD = 12.63$). All children with ASD or ADHD were previously diagnosed by a multidisciplinary team using established criteria, as specified in DSM-IV-TR (APA, 2000). Diagnosis of ASD was

confirmed by the Dutch translation of the Social Responsiveness Scale (SRS; Constantino and Gruber 2005; Roeyers et al. 2011). 23 percent of the boys in the ASD group had an SRS Total T-score between 60 and 75, indicating the presence of mild ASD or high functioning autism. 77 percent of the boys had a T-score of over 75, indicating the presence of severe autism. All children with ASD were free of medication. Diagnosis of ADHD was verified using the disruptive behaviour module of the Diagnostic Interview Schedule for Children for DSM-IV (DISC-IV; Shaffer et al. 2000). The DISC-IV was also used for establishing the presence of comorbid Oppositional-Defiant Disorder (ODD) and/or Conduct Disorder (CD). The ADHD group included 13 children with primarily Inattentive type, one with primarily Hyperactive/Impulsive type and 16 with the Combined subtype. Seven boys also met criteria for ODD, and one boy met criteria for both ODD and CD. 24 boys took medication for ADHD symptoms on a regular basis, which was discontinued at least 24 hours prior to the testing. Groups did not differ in age ($F(2, 206) = 1.05, p = .352$), but there was a group difference in estimated full scale IQ (FSIQ; $F(2, 206) = 3.67, p = .027$), with children with ASD scoring lower than their TD peers ($p = .033$).

2.2. Procedure

Once parents were informed about the aims of the study and written consents were obtained, we first asked parents and children to complete a set of questionnaires. Next, parents and children visited the laboratory where the neuropsychological tasks were administered and IQ of the children was estimated based on four subtests (Vocabulary, Similarities, Picture Arrangement and Block Design) of the Wechsler Intelligence Scale for Children III (WISC-III; Kort et al. 2002). The estimated FSIQ correlates strongly with FSIQ (Grégoire 2005).

2.3. Instruments

2.3.1. EC questionnaires

We administered three questionnaires frequently used to tap EC in older children, namely the Effortful Control Scale (ECS), the Attentional Control Scale (ACS) and the self- and parent-report of the Early Adolescent Temperament Questionnaire-Revised (EATQ-R-s and EATQ-R-p, respectively). The ECS (Lonigan and Phillips 2001) measures behavioural and attentional aspects of EC and consists

of 24 self-report items to be rated on a 5-point Likert scale. It yields a total score ($\alpha = .87$) and two subscale scores, namely Persistence/Low Distractibility (12 items; e.g., “I have a hard time concentrating on my work because I’m always thinking about other things” and “I have difficulty completing assignments on time”; $\alpha = .83$) and Impulsivity (12 items; e.g., “I can easily stop an activity when told to do so”; $\alpha = .73$). Lower scores on the Impulsivity subscale indicate higher levels of impulsivity. The ECS shows acceptable internal consistency, one-year test-retest reliability, and construct validity (e.g., Verstraeten et al. 2010).

The ACS (Derryberry and Reed 2002) measures the ability to focus and shift attention by means of 20 self-report items to be rated on a 4-point Likert scale. It yields a total score ($\alpha = .80$) and two subscale scores, namely Attention Focusing (nine items; e.g., “My concentration is good even if there is music in the room around me”; $\alpha = .68$) and Attention Shifting (11 items; e.g., “I can quickly switch from one task to another”; $\alpha = .70$). The ACS shows acceptable internal consistency, one-year test-retest reliability, and construct validity (e.g., Verstraeten et al. 2010).

The EATQ-R (Ellis and Rothbart 2001) self-report consists of 65 items, the parent-report version consists of 62 items. Items are grouped into 12 clusters and four higher-order scales (Positive Reactivity, Negative Affectivity, Affiliativeness and Effortful Control) and have to be rated on a 5-point Likert scale. For the purpose of this study, only the EC scale ($\alpha = .90$ for the EATQ-R-p, $\alpha = .73$ for the EATQ-R-s), consisting of the item clusters Inhibitory Control (e.g., “When someone tells me to stop doing something, it is easy for me to stop”), Attentional Control (e.g., “I pay close attention when someone tells me how to do something”) and Activation Control (e.g., “I put off working on projects until right before they’re due”), was included. The EATQ-R shows acceptable validity and eight-week test-retest reliability (e.g., Muris and Meesters 2009).

2.3.2. Neuropsychological Measures

To tap inhibitory control, we chose a Go/No-Go task and the Animal Stroop (Wright et al. 2003). Both types of tasks have been previously used in EC as well as EF research and tap deliberate, controlled suppression of prepotent responses (e.g., Miyake et al. 2000). To tap attentional control (i.e., attention focusing and attention shifting) we used two tasks of the Amsterdam

Neuropsychological Tasks Program (ANTP; De Sonneville 1999), namely the Focused Attention Task and the Shifting Attention Task-Auditory. Both tasks have been used extensively in TD as well as patient populations and show acceptable to good psychometric properties (e.g., De Sonneville 2005). All tasks were computerized, participants were seated in front of a monitor, approximately 60 cm from the screen.

2.3.2.1. Go/No-Go Task

On each trial, either a white 'X' or 'O' (2 by 2 cm) was randomly presented for 250 ms at the center of a black screen. Children were instructed to press a response button with their dominant hand when the letter X (Go stimulus) appeared on the screen (75% of the trials), but to inhibit their response when the letter O (No-Go stimulus) appeared (25% of the trials). Participants were told to do this as quickly and accurately as possible. The task consisted of 10 practice trials and one block of 200 test trials. The Inter-Stimulus-Interval varied between 1150 ms and 1350 ms, task duration was about 6 minutes. The dependent measure for the Go/No-Go Task was the percentage of commission errors (i.e., pressing the response button after the letter O appeared on the screen; %EOC). Analyses were repeated using mean RT on go-trials as an alternative dependent measure, all results remained the same.

2.3.2.2. Animal Stroop Task

The Animal Stroop (AS) is a computerized, Stroop-like measure of inhibitory control. It was developed by Wright and colleagues (Wright et al. 2003) in order to circumvent problems associated with the standard colour-word Stroop task (e.g., effect of limited/insufficiently 'automatic' reading abilities on performance) and seems to provide a robust measure of inhibitory function in children from 3 to 16 years old (Wright et al. 2003). Stimuli of the AS are based on four exemplar images of animals (a cow, a pig, a sheep, and a duck), all stimuli can be oriented to the right or to the left. The task comprises of three conditions. In the matching condition the animal's body is combined with the appropriate, matching head (i.e., the body of the cow with the head of the cow). In the incongruent, Stroop-like, condition each animal's head was replaced with a head of the other three animal prototypes (i.e., the body of the cow with the head of the duck), thus creating 12 animal-Stroop

stimuli. A control condition was created by substituting each animal's head with a caricature of a face. The control condition is intended to act as a semantic control in that it contains a similar semantic content as a face, but produces less activation of animal representations. Therefore, it is believed to be the most appropriate comparison with the incongruent condition (Wright et al. 2003). The underlying idea of the AS is that facial information is preferentially processed (Johnson 1993) and utilized preferentially in semantic categorization (Quinn and Eimas 1996). Consequently, Stroop-like interference can be elicited as the child is required to name the body of the animal and to inhibit a preferred response based on identification of the animal's head (Wright et al. 2003). The task consisted of six practice trials and three blocks of twenty four test trials each. Block one and block three consisted of incongruent and control stimuli, block two consisted of matching stimuli. Within each block, stimuli were presented randomly. Each stimulus was presented for 3 s and was preceded by a central fixation point with a duration of 1 s. RTs were recorded by a 'voice key', triggered by the participants' vocal responses. Errors were manually recorded. The dependent measure for the Animal Stroop was the mean RT difference between the Stroop condition and the control condition (i.e., $RT_{\text{stroop condition}} - RT_{\text{control condition}}$; interference control), with higher scores indicating less efficient interference control.

2.3.2.3. Focused Attention Task

The Focused Attention Task (FA) employs a four-letter display. Only two diagonal locations are relevant. Children were instructed to attend the relevant diagonal only. A target stimulus was defined as a stimulus that contained the target letter (i.e., the letter 'c') on the relevant diagonal. Participants had to press the 'yes'-key if a target stimulus was presented (see Figure 3, part a) and the 'no'-key if the target was presented at the irrelevant diagonal (a "foil"; e.g., Figure 3, part b) or if the target letter was absent (e.g., Figure 3, part c). The task consisted of 10 practice trials and 80 test trials. The underlying idea of this focused attention task is that foils may break through the focused attending to the relevant diagonal. An attentional shift towards the irrelevant diagonal would then result in an increase of processing time, or when there is a lack of inhibitory control, may lead to error responses to foils. A focused attention deficit would manifest itself in longer RTs to foils than to

relevant targets (e.g., Althaus et al. 1996). Consequently, we used the mean RT difference score between the correct rejection of an irrelevant target (i.e., pressing the “no”-key when a target appears on the irrelevant diagonal; RT_{crit}) and the correct evaluation of a relevant target (i.e., pressing the “yes”-key when a target is presented at the relevant diagonal; RT_{hits}) as a measure of focused attention, with larger scores indicating more difficulties in focusing attention.

2.3.2.4. Shifting Attention Task-Auditory

The auditory version of the Attention Shifting Task of the ANTP (De Sonneville 1999) was used. Stimuli consisted of tones that were either single or double and could be either low-pitched (200 Hz) or high-pitched (400 Hz). Depending on the pitch of the tone, participants had to copy the tone (i.e., press once (twice) when a single (double) tone was presented), or ‘mirror’ the tone (i.e., press once (twice) when a double (single) tone was presented). The task consisted of three parts. In the first part, single or double low-pitch tones are presented and participants have to copy the tones. In the second part, single or double high-pitch tones are presented and participants are required to mirror the tones. In the third part, the pitch of the tones could randomly change from trial to trial, which made it necessary for the participant to adjust his response behaviour and, therefore, required attentional flexibility. Parts one and two consisted of 10 practice trials and 40 test trials. Part three consisted of 16 practice trials and 80 test trials. It is expected that the mirroring of responses will be executed slower than the copying of responses, and that RTs in the third part of the task will be higher than those in part one and two because shifting attentional set will have a negative effect on processing speed (De Sonneville et al. 2002). Consequently, a measure of attentional flexibility can be obtained by calculating the mean RT difference between compatible trials in the third part and compatible trials in the first part of the task, with higher values indicating more difficulties with shifting attention. Analyses were repeated using the difference between the percentage errors on compatible trials in the third part and compatible trials in the first part of the task as an alternative dependent measure, all results remained the same.

2.4. Analytic strategy

2.4.1. Data Trimming and Outlier Analysis

For the RT-based measures, all RTs from errors (voice key or other errors) and all RTs shorter than 150 ms were eliminated¹. To prevent extreme RTs from influencing the means for each participant, we applied a within-subject trimming procedure that is robust to non-normality (Wilcox and Keselman 2003; Friedman et al. 2008): for each participant, observations that deviated from the median by more than 3.32 times the median absolute deviation in each condition were excluded. For each variable used in the analyses, observations farther than 3 SDs from the group means were replaced with values that were 3 SDs from the group mean. This final trimming stage affected no more than 1.8% of the observations for any measure².

2.4.2. Data Analysis

All statistical analyses were conducted with SPSS 21 and R 2.15.2 (R Development Core Team, 2012), the R package lavaan 0.5-11 (Rosseel 2012). Confirmatory factor analyses (CFA) were carried out to test the two hypothesized models. Because the data file contained some missing values (0.95% of all data), full information maximum likelihood estimation (FIML) was used (Enders 2010). Data were not normally distributed, therefore we used robust standard errors and a Yuan-Bentler (YB) scaled chi-square test statistic for non-normality (Yuan and Bentler 2000). In line with theoretical recommendations (Hu and Bentler 1999), we evaluated the fit of each model with multiple indices, namely (a) the scaled chi-square (χ^2) test statistic, (b) the root-mean-square error of approximation (RMSEA), (c) the standardized root-mean-square residual (SRMR), and (d) the comparative fit index (CFI). The most common fit index is the χ^2 statistic, which measures the degree to which the covariances predicted by the specified model differ from the observed covariances. A non-significant difference between model and data suggest a satisfactory fit. However, it is to be noted that the χ^2 statistic is very sensitive to sample size and, with increasing sample size, has an increasing chance of being significant. For the RMSEA, values of .06 or lower are suggested for a good fit (Hu and Bentler 1999), while values up to .08 represent a reasonable model fit (Schreiber et al. 2006). For the SRMR, lower values indicate a closer fit, with values below .08 indicating a fair fit to the data, and values

¹ A RT of less than 150 ms is taken to indicate that the subject's response was anticipatory and not an authentic response per se. Such anticipatory RTs are therefore discarded (Jensen 2006, p. 63).

² Analyses were repeated using the non-trimmed data, all results remained the same. Analyses were repeated excluding observations farther than 3 SDs from the group means, all results remained the same.

below .05 indicating a good fit (Hu and Bentler 1998). Finally, for the CFI, higher values indicate a better fit, with values greater than .95 considered indicative of a good fit and values greater than .90 indicating an acceptable fit (Bentler 1990; Hu and Bentler 1998).

3. Results

3.1. Correlations between EC scales and EF tasks

As a first investigation of the relationship between EC scales³ and EF tasks, bivariate Pearson correlations were computed (Table 1). We found no significant relationship between EC scales and Mean RT on the Go/No-Go Task, interference control as measured by the Animal Stroop, and attention focusing. We did find a significant relationship between parent-reported EC and the %EOC on the Go/No-Go Task and flexibility in time and errors on the Attention Shifting Task. Specifically, higher levels of parent-reported EC were associated with fewer EOC and less problems with shifting attention. We also found a significant relationship between the %EOC on the Go/No-Go Task and self-reported attention focusing (ACS), with higher levels of attention focusing being associated with fewer EOC. Although some correlations between scores on EC questionnaires and performance based measures were significant, it is important to notice that all correlation coefficients were small (r 's ranging from -.14 to -.26).

3.2. CFA of the hypothesized models

3.2.1. CFA of the single factor model

The hypothesized single factor model in which all EC total scales and dependent measures of the different neuropsychological EF tasks are best characterized as a unitary factor, showed a moderate model fit, $YB \chi^2(20) = 47.78$, $p < .001$, $RMSEA = .082$, with 90% CI [.053, .110], $SRMR = .07$, and $CFI = .94$. However, when looking at the factor loadings (see Table 2), only the variables based on the questionnaires loaded significantly on the underlying EC factor. Although the %EOC on the Go/No-Go Task showed a trend towards a significant loading on the EC factor, none of the

³ We found a significant relationship between self- and parent-reported EC. In specific, all scales from the self-report questionnaires were significantly related to all scales from the parent-report questionnaire (with r s ranging from .21 to .61), indicating that higher levels of self-reported EC are associated with higher levels of parent-reported EC.

variables based on the neuropsychological EF tasks showed a significant factor loading. Apparently, these variables do not seem to capture the same latent variable as the questionnaires.

3.2.2. CFA of the two-factor model

The alternative two factor model in which all attention related measures load on an Attentional Control factor and all inhibition related measures load on an Inhibitory Control factor showed a bad model fit, $\chi^2(26) = 99.55$, $p < .001$, RMSEA = .12, with 90% CI [.093, .141], SRMR = .07, and CFI = .83. Estimates, standard errors, p-values, and standardized factor loadings are shown in Table 2.

3.2.3. Additional analyses: alternative models and possible confounding variables

In an attempt to improve the fit of both hypothesized models and to investigate the role of possible confounding variables, additional analyses were performed. First, we investigated if some small alterations to both models, based on the inspection of modification indices, could improve model fit. Allowing Attention Shifting to be correlated to Attention Focusing and interference control, improved the overall model fit for the single factor model (i.e., all fit indices could be evaluated as ‘good’). However, further inspection of the factor loadings showed that still none of the variables based on the EF tasks, loaded significantly on the underlying EC factor. For the two-factor model we tested the impact on model fit if we allowed the error variances of both self-reports of attention to be correlated and if we allowed the error variances of parent-reports on attention and inhibitory control to be related. Allowing for the error variances of both self-reports to be correlated did not improve the overall fit of the two-factor model, whereas allowing for the error variances of both parent-report indicators to be correlated, did (i.e., all fit indices could be evaluated as ‘acceptable’). Further inspection of factor loadings showed that %EOC on the Go/No-Go Task, and Flexibility Time on the Shifting Attention Task loaded significantly on their respective underlying latent variables. However, factor loadings were very low (-.18 and -.17, respectively) and did not even reach the threshold for a weak factor (i.e., .20 to .30; Briggs and MacCallum 2003, p.53), suggesting that these variables are not a good indicator for the underlying latent variables. Additionally, we tested whether the data are better captured by a model in which all questionnaire data load on one underlying factor (e.g., Reports) and all variables based on EF tasks load on another underlying factor (e.g., Performance-based measures).

This model showed a bad model fit (YB $\chi^2(26) = 88.96$, $p < .001$, RMSEA = .11, with 90% CI [.083, .133], SRMR = .07, and CFI = .85).

Second, both hypothesized models were tested using alternative variables for the EF tasks (e.g., mean RT on go-trials instead of %EOC and Flexibility Error instead of Flexibility Time). All results remained the same.

Third, one might argue that the EC questionnaires reflect a more aggregated level of behavioral- and attentional control (across specific self-regulation abilities and across contexts), whereas EF tasks may be more specific and highly susceptible to context effects. To evaluate whether an aggregate measure of the different EF tasks captures a level of measurement that is more similar to the questionnaires, a single factor model was tested in which the different EC total scores and an aggregate score of the (scaled) variables of the EF tasks all load on the same underlying factor. The standardized factor loading of the aggregate EF variable definitely improved (-.18), but still did not reach the threshold for a weak factor, hence suggesting that the aggregate EF score was not a good indicator of the underlying latent variable.

Fourth, we investigated whether optimizing the questionnaire data could improve model fit. In specific, we performed exploratory factor analyses (EFA) on the combined item pool of all EC questionnaires in an attempt to extract 'ideal' questionnaire-based factors for our participant group. The EFA yielded a five-factor model (i.e., Activation Control, Persistence, Attention Focusing, Attention Shifting, and Inhibitory Control). A follow-up CFA of this five-factor model yielded a good fit (YB $\chi^2(109) = 136.29$, $p = .039$, RMSEA = .04, with 90% CI [.012, .052], SRMR = .05, and CFI = .97). The two hypothesized models were tested using the new latent variables instead of the original EC subscales (i.e., in the two-factor model, the latent variables Attention Focusing and Attention Shifting were supposed to load on the underlying latent factor Attentional Control, and the latent variables Activation Control, Persistence, and Inhibitory Control were supposed to load on the underlying latent factor Behavioural Control). The fit of the single factor model did not improve (i.e., fit indices remained 'acceptable' and, crucially, none of the variables based on the EF tasks, loaded significantly on the underlying EC factor). The overall fit for the two-factor model improved

somewhat (i.e., ‘acceptable’ model fit). However, with the exception of %EOC on the Go/No-Go Task, none of the performance based variables loaded significantly on the underlying factors. Also, the factor loading of %EOC was so small (-.18) that this variable does not seem to be a good indicator of the underlying latent variable.

Finally, CFA analyses were repeated including only TD children in order to test whether or not including clinical groups had a negative influence on the model fit. Results remained the same in that the two-factor model showed a bad fit (YB $\chi^2(26) = 71.79$, $p < .001$, RMSEA = .11, with 90% CI [.080, .139], SRMR = .08, and CFI = .82). The overall fit of the single factor model even became worse after excluding the ADHD and ASD groups (i.e., all fit indices changed for the worse and the SRMR no longer fell within the acceptable range; YB $\chi^2(20) = 49.47$, $p < .001$, RMSEA = .10, with 90% CI [.067, .133], SRMR = .08, and CFI = .90).

3.3. Group differences

3.3.1. EC scales

We compared groups on the different EC scales by means of ANCOVAs (i.e., given that some of the EC scales were significantly related to FSIQ, we performed the analyses controlling for the effect of FSIQ) and Hochberg’s GT2 post hoc analyses, using an overall alpha level of .05. Means, standard deviations, and *F* values are shown in Table 3. Significant group differences were found for all scales. Both clinical groups scored significantly lower than the TD group on parent-reported EC (all scales of the EATQ-R-p) as well as on self-reported attentional control (all scales of the ACS) and persistence (ECS). For parent-reported attentional control and for persistence, the deficit was significantly larger in boys with ADHD as compared to boys with ASD. Boys with ADHD scored significantly lower than both other groups on activation control (EATQ-R-s), impulsivity (ECS), and attentional control (EATQ-R-p and -s).

Given that (1) some previous studies found gender differences in terms of EC (e.g., Kochanska et al. 2000), and (2) girls were included in our TD group, but not in our clinical groups, we performed additional analyses to investigate the effects of gender. All results remained the same with the exception that group differences based on self-reported inhibitory control became non-significant

and the differences between children with ADHD and children with ASD on parent-reported attentional control and EC became non-significant.

3.3.2. *EF tasks*

We compared groups on their performance on the different neuropsychological EF task by means of ANOVAs (i.e., given that performance on EF tasks and FSIQ were not significantly related, requirements for analyses of covariance were not met) and Hochberg's GT2 post hoc analyses, using an overall alpha level of .05. Means, standard deviations, and *F* values are shown in Table 4. No significant group differences were found. All groups performed at a very similar level on the Go/No-Go task, the Animal Stroop, the Attention Focusing Task and the Attention Shifting Task.

Additional analyses were performed to investigate potential gender effects on these findings however, all results remained the same when excluding girls.

4. Discussion

The main aim of our study was to investigate the degree to which EC questionnaires and a selection of neuropsychological EF tasks considered to tap inhibitory and attentional control are interchangeable when measuring EC. In order to evaluate the interchangeability of both types of measures, we addressed two important questions, namely: (a) do the questionnaires and EF tasks measure the same construct?, and (b) do they give the same results when comparing different (clinical) groups?

CFAs were used to determine whether or not scores on EC questionnaires and variables based on EF tasks loaded on the same (or very similar) underlying construct(s). Overall, our findings show that this was not the case. Although the single factor model yielded a moderate fit, further inspection of the factor loadings revealed that none of the performance-based variables significantly loaded on the underlying EC factor and, consequently, do not seem to capture the same latent variable as the EC questionnaires. An alternative, two-factor, model that differentiated between more attention related measures and more inhibition related measures showed a bad model fit, suggesting that even when looking at a subcomponent level of EC, the questionnaires and performance based measures did not tap the same constructs. Optimizing model specifications (i.e., based on inspection of modification

indices), performance based data (i.e., data-trimming; alternative variables for the EF tasks), and questionnaire data (i.e., extracting new, better, questionnaire-based latent factors by means of EFA on the combined item-pool of all scales) improved the overall model fit of both hypothesized models somewhat. Nevertheless, evaluation of factor loadings still led to the same conclusion. In the best case scenario, two out of four performance-based variables (i.e., %EOC and Flexibility Time) loaded significantly on the hypothesized underlying factor (i.e., inhibitory control and attentional control, respectively). Nonetheless, factor loadings were so small that they did not even fulfill the conditions for a 'weak' factor and, hence, cannot be considered as good indicators of the underlying latent variables. Given the poor fit of both hypothesized models, one might argue that perhaps the data would be better captured by a model in which all questionnaire data load on one underlying factor (e.g., Reports) and all variables based on EF tasks load on another underlying factor (e.g., Performance-based measures). However, additional CFAs of this alternative model yielded a bad model fit, suggesting that this is not the case. In all, our findings are clear in that they do not support the idea that the EC questionnaires and the EF tasks used in the current study, measure the same underlying construct(s).

The second question we addressed, is whether or not both types of measures lead to the same results when comparing TD children, children with ADHD, and children with ASD. Our findings suggest that this is not the case. The three groups could not be differentiated from each other based on their performance on the EF tasks, whereas significant group differences were found for all EC scales. If we solely focus on the questionnaire data, we would have to conclude that children with ADHD and children with ASD showed significantly more difficulties in regulating their own behaviour (parent-reports and persistence subscale of the ECS) and attention (parent-reports and the ACS) in comparison with TD peers and that the problems in terms of attentional control (parent- and self-reports) and behavioural regulation (i.e., persistence, impulsivity, and activation control) were more pronounced in children with ADHD than in children with ASD. These findings are in large part consistent with previous research using questionnaires to investigate EC in children with ADHD or ASD (e.g., Konstantareas and Stewart 2006; Martel and Nigg 2006; Samyn et al. 2011; Wiersema and Roeyers

2009). However, if we were to take into consideration only the performance-based measures, we would conclude that children with ADHD and children with ASD perform at an equal level as TD children in terms of attentional and behavioural regulation. Inspection of the large body of literature on EF in ADHD and ASD teaches us that the lack of group differences based on EF tasks, is not that surprising. First, findings based on performance-based measures are often inconclusive and are known to strongly vary depending on, among other things, sampling variation (e.g., ADHD subtypes, comorbidities; e.g., Nigg 2001), confounding variables (e.g., age, intelligence; e.g., Nigg 2001), task characteristics (e.g., presentation rate of the stimuli; working memory load; e.g., Nigg 2001; Van de Voorde et al. 2011; Wiersema et al. 2006;), and even the (calculation of the) dependent measures chosen to reflect inhibitory or attentional control (e.g., Schwartz and Verhaeghen 2008;). Additionally, performance is known to be strongly context dependent, especially in ADHD (e.g., Sonuga-Barke et al. 2010). Second, executive dysfunctions are not universally present in children with ADHD (e.g., Willcutt et al. 2005) or ASD (e.g., Pellicano 2007) and *if* these are present, there is often a large variation in the degree of executive dysfunctioning (e.g., Nigg et al. 2005).

Taking into consideration that our results provide support for neither of the two necessary conditions for the interchangeability of measures (i.e., measuring the same underlying construct, and leading to the same results when comparing groups), we have to conclude that the specific questionnaires and neuropsychological measures used in the current study are not interchangeable when measuring attentional and inhibitory control. Both types of measures seem to tap different kinds of information and both have strengths as well as limitations. Neuropsychological measures give an indication of a child's capacities at a given moment and under certain specific circumstances, hence providing more state-like information. However, they are never 'pure' measures that only tap one aspect of a person's functioning. Although EF tasks are often designed with the intention to tap a single cognitive process such as 'inhibitory control', a child's performance on the task will not only be influenced by the cognitive process intended, but also by other aspects, such as lower-level cognitive skills (e.g., language, memory, and attention) or even contextual factors such as motivation or fatigue (e.g., Anderson et al. 2002; Nigg 2001). The EF tasks administered in the current study seem to tap

processes that are not universally dysfunctional in children with ADHD and/or ASD, at least not in optimal conditions (e.g., a highly structured environment). The EC questionnaires seem to be more ecologically valid, inquiring about a child's specific, daily functioning over a larger time span hence providing more trait-like information. However, reports can be biased and are limited in capturing core processes (e.g., Nigg 2001). In general, the EC questionnaires that were used in the current study seem to be more global measures of self-regulation. Additionally, EC questionnaires seem to focus on self-regulation in a somewhat different context than EF tasks. Recent studies suggest that self-regulation in general (Blair, 2002) and EC in particular (Kim, Nordling, Yoon, Boldt, & Kochanska, 2013) consist of 'hot' (i.e., involving regulation of emotions) as well as 'cool' (i.e., more decontextualized or emotionally neutral) aspects, whereas EF tasks are traditionally 'cool' tasks that do not (at least not intentionally) assess emotional aspects of self-regulation. It is possible that this also had an influence on our findings regarding the relationship between EC reports and performance on EF tasks. Overall, we believe that EC questionnaires and EF tasks should be considered as measures providing information about different aspects of a child's functioning, rather than interchangeable.

Some limitations of the present study need to be acknowledged. First, one might argue that our total sample was rather limited in size. However, the number of participants included in the present study is comparable to the sample size of other studies in this research domain using latent-variable analyses to test models of a similar complexity (e.g., Friedman and Miyake 2004; Miyake et al. 2000). Although it is possible that increasing the sample size would (eventually) lead to more significant factor loadings, it is highly unlikely that this would also considerably improve factor loadings. Therefore, we do not believe that it would lead to major changes in terms of the overall conclusions concerning the models tested. Similarly, it is possible that the sample sizes of our clinical groups were insufficient to detect possibly modest group differences on neuropsychological tasks. Therefore, future research will have to confirm that our findings can be replicated in larger samples. A second limitation may have occurred in the procedure of collapsing the ADHD and ASD groups with the TD group to study the relationship and interchangeability between EC questionnaires and EF

tasks, thus creating a quite heterogeneous sample. It is possible that the inclusion of clinical groups (and, consequently, more extreme scores), may have influenced our findings. However, additional analyses including only TD children yielded the same results, suggesting that this was not the case. Third, for the purpose of the current study we made a selection of EF tasks, proposed to measure attentional and inhibitory control. Despite the fact that previous studies have proven these tasks to be successful in measuring the constructs intended (e.g., De Sonneville 2005; Miyake et al. 2000; Wright et al. 2003), one might argue that using alternative neuropsychological measures (or other manipulations of the tasks administered in the present study) may possibly yield different results. Nonetheless, the fact that our findings are completely in line with studies investigating the relationship between EF tasks and EF questionnaires (for a review, see Toplak et al. 2013), leads us to believe that our general conclusion that EC questionnaires and EF tasks are not interchangeable, is not merely a consequence of our selection of neuropsychological measures.

5. Conclusions

With the above mentioned limitations in mind, our findings do not support the idea that the specific EC questionnaires and neuropsychological measures used in this study are interchangeable when measuring attentional and inhibitory control. Not only did latent-variable analyses reveal that they do not measure the same (or similar) underlying constructs, both types of measures also lead to different results when comparing TD children, children with ADHD and children with ASD. In all, the current study showed more differences than commonalities between both types of measures and, hence, suggests that they should not be considered as being interchangeable.

Disclosure statement

No disclosures.

Funding source

The first author received funding from the research fund of the faculty of Psychology and Educational Sciences of the Ghent University. The funding source had no further involvement.

References

- Althaus, M., De Sonnevile, L. M. J., Minderaa, R. B., Hensen, L. G. N., & Til, R. B. (1996). Information processing and aspects of visual attention in children with the DSM-III-R diagnosis "Pervasive Developmental Disorder Not Otherwise Specified" (PDD-NOS): I. Focused and Divided Attention. *Child Neuropsychology*, 2, 17-29. doi: 10.1080/09297049608401347
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV-TR)*. Washington, DC: Author.
- Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Mikiewicz, O. (2002). Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychology*, 8, 231-240. doi:10.1076/chin.8.4231.13509
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246. doi:10.1037/0033-2909.107.2.238
- Bijttebier, P., & Roeyers, H. (2009). Temperament and vulnerability to psychopathology: Introduction to the special section. *Journal of Abnormal Child Psychology*, 37, 305-308. doi:10.1007/s10802-009-9308-2
- Blair, C. (2002). School readiness – Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, 57, 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 mat hand literacy ability in kindergarten. *Child Development*, 78, 647-663. doi:10.1111/j.1467.8624.2007.01019.x
- Briggs, N. E., & MacCallum, R. C. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. *Multivariate Behavioral Research*, 38, 25-56. doi:10.1207/S15327906MBR3801_2
- Constantino, J. N. & Gruber, C. P. (2005). *Social responsiveness scale. Manual*. Los Angeles, Western Psychological Services.

- De La Fuente, A., Xia, S. G., Branch, C., & Li, X. B. (2013). A review of attention-deficit/hyperactivity disorder from the perspective of brain networks. *Frontiers in Human Neuroscience*, 7, 192. doi:10.3389/fnhum.2013.00192
- Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology*, 111, 225-236. doi:10.1037//0021-843X.111.2.225
- De Sonneville, L. M. J. (1999). *Amsterdam neuropsychological tasks: a computer-aided assessment program*. In B. P. L. M. den Brinker, P. J. Brand, S. J. Maarse, & L. J. M. Mulder (Eds.), *Cognitive ergonomics, clinical assessment and computer-assisted learning: Computers in psychology* (Vol. 6, pp. 187-203). Lisse: Swets.
- De Sonneville, L. M. J. (2005). Amsterdam neuropsychological tasks: Scientific and clinical applications. *Tijdschrift voor Neuropsychologie*, 0, 27-41.
- De Sonneville, L. M. J., Boringa, J. B., Reuling, I. E. W., Lazeron, R. H. C., Adèr, H. J., & Polman, C. H. (2002). Information processing characteristics in subtypes of multiple sclerosis. *Neuropsychologia*, 40, 1751-1765. doi:10.1016/S0028-3932(02)00041-6
- Eisenberg, N., Smith, C. L., Sadovsky, A., & Spinrad, T. L. (2004). Effortful control: Relations with emotion regulation, adjustment, and socialization in childhood. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of Self-regulation: Research, Theory, and Applications* (pp. 259-282). New York: The Guilford Press.
- Ellis, L. K., & Rothbart, M. K. (2001, April). *Revision of the early adolescent temperament questionnaire*. Poster presented at the 2001 Biennial Meeting of the Society for Research in Child Development, Minneapolis, Minnesota.
- Enders, C. K. (2010). *Applied missing data analysis*. New York: The Guilford Press.
- Fine, M. A. (1992). On the distinction between two tests as measures of the same construct and as interchangeable. *American Psychologist*, 47, 1146-1146. doi:10.1037/0003-066X.47.9.1146.a

- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology-General*, *133*, 101-135. doi:10.1037/0096-3445.133.1.101
- Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*, 201-225. doi:10.1037/0096-3445.137.2.201
- Gargaro, B. A., Rinehart, N. J., Bradshaw, J. L., Tonge, B. J., & Sheppard, D. M. (2011). Autism and ADHD: How far have we come in the comorbidity debate? *Neuroscience and Biobehavioral Reviews*, *35*, 1081-1088. doi:10.1016/j.neubiorev.2010.11.002
- Grégoire, J. (2005). *Clinical evaluation of the intelligence of the child. Theory and practice of WISC-III, third edition*. Sprimont: Mardaga.
- Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review*, *24*, 189-233. doi:10.1016/j.dr.2004.01.001
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*, *3*, 424-453. doi:10.1037/1082-989X.3.4.424
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55. doi:10.1080/10705519909540118
- Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. New York: Elsevier.
- Johnson, M. H. (1993). *Cortical maturation and the development of visual attention in early infancy*. In M. H. Johnson (Ed.), *Brain development and cognition: A reader* (pp. 167-193). Oxford: Blackwell.
- Kim, S., Nordling, J. K., Yoon, J. E., Boldt, L. J., & Kochanska, G. (2013). Effortful control in “hot” and “cool” tasks differentially predicts children’s behavior problems and academic

- performance. *Journal of Abnormal Child Psychology*, *41*, 43-56. doi:10.1007/s10802-012-9661-4
- Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates, and consequences. *Journal of Personality*, *71*, 1087-1113. doi:10.1111/1467-6494.7106008
- 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*, 220-232. doi:10.1037//0012-1649.36.2.220
- Konstantareas, M. M., & Stewart, K. (2006). Affect regulation and temperament in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *36*, 143-153. doi:10.1007/s10803-005-0051-4
- Kort, W., Schittekatte, M., Compain, E. L., Bosmans, M., Bleichrodt, N., Vermeir, G., Resing, W. C. M., & Verhaeghe, P. (2002). *WISC-III-NL: Manual Dutch adaptation*. London, England: The Psychological Corporation.
- Lengua, L. J., Honorado, E., & Bush, N. R. (2007). Contextual risk and parenting as predictors of effortful control and social competence in preschool children. *Journal of Applied Developmental Psychology*, *28*, 40-55. doi:10.1016/j.appdev.2006.10.001
- Lonigan, C. J., & Phillips, B. M. (2001). Temperamental influences on the development of anxiety disorders. In M. W. Vasey & M. R. Dadds (Eds.), *The developmental psychopathology of anxiety* (pp. 60-91). New York: Oxford University Press.
- Martel, M. M., & Nigg, J. T. (2006). Child ADHD and personality/temperament traits of reactive and effortful control, resiliency, and emotionality. *Journal of Child Psychology and Psychiatry*, *47*, 1175-1183. doi:10.1111/j/1469-7610.2006.01629.x
- 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" task: A latent variable analysis. *Cognitive Psychology*, *41*, 49-100. doi:10.1006/cogp.1999.0734

- Mullane, J. C., Corkum, P. V., Klein, R. M., & McLaughlin, E. (2009). Interference control in children with hand without ADHD: A systematic review of flanker and simon task performance. *Child Neuropsychology, 15*, 321-343. doi:10.1080/09297040802348028
- Muris, P., & Meesters, C. (2009). Reactive and regulative temperament in youths: Psychometric evaluation of the early adolescent temperament questionnaire-revised. *Journal of Psychopathology and Behavioral Assessment, 31*, 7-19. doi:10.1007/s10862-008-9089-x
- Muris, P., & Ollendick, T. H. (2005). The role of temperament in the etiology of child psychopathology. *Clinical Child and Family Psychology Review, 8*, 271-289. doi:10.1007/s10567-005-8809-y
- Nigg, J. T. (2001). Is ADHD a disinhibitory disorder? *Psychological Bulletin, 127*, 571-598. doi:10.1037//0033-2909.127.5.571
- Nigg, J. T., Stavro, G., Ettenhofer, M., Hambrick, D. Z., Miller, T., & Henderson, J. M. (2005). Executive functions and ADHD in adults: Evidence for selective effects on ADHD symptom domains. *Journal of Abnormal Psychology, 114*, 706-717. doi:10.1037/0021-843X.114.3.706
- O'Hearn, K., Asato, M., Ordaz, S., & Luna, B. (2008). Neurodevelopment and executive function in autism. *Development and Psychopathology, 20*, 1103-1132. doi:10.1017/S0954579408000527
- Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology, 43*, 974-990. doi:10.1037/0012-1649.43.4.974
- Powell, R., Johnston, M., & Johnston, D. W. (2007). Assessing walking limitations in stroke survivors: Are self-reports and proxy-reports interchangeable? *Rehabilitation Psychology, 52*, 177-183. doi:10.1037/0090-5550.52.2.177
- Quinn, P. C., & Eimas, P. D. (1996). Perceptual cues that permit categorical differentiation of animal species by infants. *Journal of Experimental Child Psychology, 63*, 189-211. doi:10.1006/jecp.1996.0047

- R Development Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Roeyers, H., Thys, M., Druart, C., De Schryver, M., & Schittekatte (2011). *SRS Screeninglist for autism spectrum disorders: Manual*. Amsterdam: Hogrefe.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1-36. Retrieved from <http://www.jstatsoft.org/>
- Rothbart, M. K. (1989). Temperament and development. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds.), *Temperament in childhood* (pp. 187-247). New York: Wiley.
- Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, personality development* (pp. 105-176). New York: Wiley.
- Samyn, V., Roeyers, H., & Bijttebier, P. (2011). Effortful Control in typically developing boys and in boys with ADHD or autism spectrum disorder. *Research in Developmental Disabilities*, 32, 483-490. doi:10.1016/j.ridd.2010.12.038
- Samyn, V., Roeyers, H., Bijttebier, P., & Wiersema, J. R. (2013). Attentional networks in boys with ADHD or autism spectrum disorder and the relationship with effortful control. *Journal of Attention Disorders*, XX, 1-12.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis: A review. *Journal of Educational Research*, 99, 323-338. doi:10.3200/JOER.99.6.323-338
- Schwartz, K., & Verhaeghen, P. (2008). ADHD and Stroop interference from age 9 to age 41 years: a meta-analysis of developmental effects. *Psychological Medicine*, 38, 1607-1616. doi:10.1017/S003329170700267X
- Sergeant, J. A., Geurts, H., & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for Attention-Deficit/Hyperactivity Disorder? *Behavioral Brain Research*, 130, 3-28. doi:10.1016/S0166-4328(01)00430-2

- Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH diagnostic interview schedule for children version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 28-38. doi:10.1097/00004583-200001000-00014
- Sonuga-Barke, E. J. S., Wiersema, J. R., van der Meere, J. J., & Roeyers, H. (2010). Context-dependent dynamic processes in attention-deficit/hyperactivity disorder: Differentiating common and unique effects of state regulation deficits and delay aversion. *Neuropsychology Review, 20*, 86-102. doi:10.1007/s.11065-009-9115-0
- 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*, 131-143. doi:10.1111/jcpp.12001
- Van de Voorde, S., Roeyers, H., Verte, S., & Wiersema, J. R. (2011). The influence of working memory load on response inhibition in children with attention-deficit/hyperactivity disorder or reading disorder. *Journal of Clinical and Experimental Neuropsychology, 33*, 753-764. doi:10.1080/13803395.2011.554385
- Verstraeten, K., Vasey, M. W., Claes, L., & Bijttebier, P. (2010). The assessment of effortful control in childhood: Questionnaires and the test of everyday attention for children compared. *Personality and Individual Differences, 48*, 59-65. doi:10.1016/j.paid.2009.08.016
- Wiersema, J. R., & Roeyers, H. (2009). ERP correlates of effortful control in children with varying levels of ADHD symptoms. *Journal of Abnormal Child Psychology, 37*, 327-336. doi:10.1007/s10802-008-9288-7
- Wiersema, J. R., van der Meere, J., Roeyers, H., Van Coster, R., & Baeyens, D. (2006). Event rate and event-related potentials in ADHD. *Journal of Child Psychology and Psychiatry, 47*, 560-567. doi:10.1111/j.1469-7610.2005.01592.x
- Wilcox, R. R., & Keselman, H. J. (2003). Modern robust data analysis methods: Measures of central tendency. *Psychological Methods, 8*, 254-274. doi:10.1037/1082-989X.8.3.254

- Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, *57*, 1336-1346. doi:10.1016/j.biopsych.2005.02.006
- Wright, I., Waterman, M., Prescott, H., & Murdoch-Eaton, D. (2003). A new Stroop-like measure of inhibitory function development: typical developmental trends. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *44*, 561-575. doi:10.1111/1469-7610.00145
- Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, *30*, 165-200. doi:10.1111/0081-1750.00078
- Zhou, Q., Chen, S. H., & Main, A. (2012). Commonalities and differences in the research on children's effortful control and executive function: A call for an integrated model of self-regulation. *Child Development Perspectives*, *6*, 112-121. doi:10.1111/j.1750-8606.2011.00176.x

Tables

Table 1. Bivariate correlations between EC scales and performance on EF tasks

Table 2. Estimates, Standard Errors, p-values, and Standardized Factor Loadings for the hypothesized models

Table 3. Descriptive information on the EC questionnaires for the total group and for the three subgroups

Table 4. Descriptive information on the neuropsychological measures for the total group and for the three subgroups

Figure captions

Fig. 1. Hypothesized model linking the EC scales and performance on neuropsychological EF measures to the underlying latent variable EC. GNG = Go/No-Go; AS = Animal Stroop; FA = Focused Attention; SA = Shifting Attention

Fig. 2. Alternative model linking (sub)scales of the EC questionnaires and performance on neuropsychological EF measures to the underlying latent variables Attentional Control and Inhibitory Control

Fig. 3. Example trials of the Focused Attention task (target presented in the [relevant diagonal] (a), target presented in the irrelevant diagonal (b), and no target presented (c))

Table 1

Bivariate correlations between EC scales and performance on EF tasks

EC questionnaires		EF tasks					
		GNG %EOC	GNG mean RT	AS IC	AF RT _{crit} -RT _{hits}	SA Flex Time	SA Flex Error
<i>Parent-rated</i>							
EATQ-R	Total	-.22**	.04	-.02	-.07	-.18*	-.19**
	Inhibitory control	-.17*	.02	.02	-.07	-.13	-.21**
	Activation control	-.16*	.04	-.01	-.02	-.14*	-.14*
	Attentional control	-.26***	.04	-.06	-.11	-.19**	-.18*
<i>Child-rated</i>							
ECS	Total	-.08	-.05	-.04	-.09	-.10	-.03
	Pers./low distr.	-.04	-.11	-.09	-.12	-.13	-.08
	Impulsivity	-.10	.03	.02	-.04	-.05	.04

ACS	Total	-.14*	-.03	-.05	-.09	-.09	-.06
	Focusing	-.21**	-.01	-.07	-.10	-.12	-.07
	Shifting	-.05	-.03	-.02	-.06	-.05	-.04
EATQ-R	Total	-.11	-.08	-.09	-.11	-.08	-.09
	Inhibitory control	-.09	-.06	-.13	-.08	-.01	.03
	Activation control	-.06	-.04	.01	-.07	-.04	-.11
	Attentional control	-.12	-.10	-.11	-.11	-.14*	-.12

Note. Pers. = Persistence/low distractibility; GNG = Go/No-Go task; EOC = errors of commission; AS = Animal Stroop task; IC = Interference control = $RT_{\text{incongruent condition}} - RT_{\text{control condition}}$; FA = Focused Attention task; RT_{crit} = RT trials with correct rejection irrelevant target; RT_{hits} = RT trials with correct response to a relevant target; SA = Shifting Attention task; Flex Time = Flexibility Time; Flex Error = Flexibility Error.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 2Estimates, Standard Errors, *p*-values, and Standardized Factor Loadings for the hypothesized models

Model	Latent variable	Indicator	Estimate	SE	<i>p</i> -value	Standardized factor loading
Single factor model	EC	ECS Total score	12.337	0.670	.000	.92
		ACS Total score	5.746	0.495	.000	.70
		EATQ-R-s EC	0.542	0.029	.000	.92
		EATQ-R-p EC	0.477	0.043	.000	.66
		Prepotent resp. inh. (%EOC)	-1.035	0.538	.054	-.12
		Interference control	-5.858	6.144	.340	-.07
		Focused attention	-10.716	7.264	.140	-.11
		Shifting attention (Flex Time)	-25.944	16.414	.114	-.12
Two-factor model	Attentional Control	EATQ-R-p Attentional Control	0.586	0.055	.000	.71
		EATQ-R-s Attentional Control	0.572	0.039	.000	.85

	ACS Total score	6.457	0.490	.000	.76
	Focused attention	-13.699	7.423	.065	-.14
	Shifting attention (Flex Time)	-38.745	17.641	.028	-.18
Inhibitory Control	EATQ-R-p Inhibitory Control	0.397	0.057	.000	.57
	EATQ-R-s Inhibitory Control	0.337	0.048	.000	.52
	Prepotent resp. inh. (%EOC)	-1.611	0.615	.009	-.19
	Interference control	-9.142	6.633	.168	-.11

Note. Prepotent resp. inh. = Prepotent response inhibition; EOC = Errors of commission; Flex Time = Flexibility Time.

Table 3

Descriptive information on the EC questionnaires for the total group and for the three subgroups

EC questionnaires		Total (<i>n</i> = 209)	TD (<i>n</i> = 148)	ASD (<i>n</i> = 31)	ADHD (<i>n</i> = 30)	Group differences
		<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>F</i> (3,205)
<i>Parent-rated</i>						
EATQ-R	Total	3.10(0.72)	3.39(0.58) ^a	2.54(0.56) ^b	2.27(0.45) ^c	68.81***
	Inhibitory control	3.49(0.70)	3.73(0.54) ^a	2.99(0.69) ^b	2.79(0.67) ^b	44.97***
	Activation control	2.80(0.87)	3.11(0.76) ^a	2.20(0.79) ^b	1.92(0.51) ^b	43.76***
	Attentional control	3.14(0.83)	3.45(0.73) ^a	2.56(0.47) ^b	2.24(0.51) ^c	55.51***
<i>Child-rated</i>						
ECS	Total	83.07(13.42)	86.49(11.88) ^a	79.84(10.14) ^b	68.57(14.06) ^c	27.74***
	Pers./low distr.	44.87(7.83)	46.82(6.68) ^a	42.94(6.64) ^b	36.75(9.09) ^c	25.30***
	Impulsivity	38.20(6.92)	39.68(6.48) ^a	36.90(5.58) ^a	31.82(6.81) ^b	18.47***
ACS	Total	53.89(8.48)	55.74(8.02) ^a	49.48(7.39) ^b	48.96(8.43) ^b	13.06***

	Focusing	23.98(4.32)	24.85(4.03) ^a	22.19(4.54) ^b	21.32(3.97) ^b	11.57***
	Shifting	29.91(5.25)	30.89(5.05) ^a	27.29(4.45) ^b	27.64(5.57) ^b	8.72***
EATQ-R	Total	3.42(0.59)	3.55(0.54) ^a	3.28(0.55) ^b	2.91(0.58) ^c	17.16***
	Inhibitory control	3.49(0.65)	3.57(0.57) ^a	3.32(0.81) ^{a, b}	3.28(0.72) ^b	3.90*
	Activation control	3.13(0.88)	3.29(0.82) ^a	3.05(0.78) ^a	2.38(0.93) ^b	14.25***
	Attentional control	3.61(0.67)	3.74(0.65) ^a	3.44(0.56) ^a	3.03(0.59) ^b	16.67***

Note. Pers./low distr. = Persistence/low distractibility. *P*-values are derived from one-way ANOVA. Superscripts reflect subgroup differences derived from post-hoc Hochberg's GT2 Test; different letters indicate differences between particular groups, identical letters indicate that there were no differences between those particular groups.

**p* < .05.

***p* < .01.

****p* < .001.

Table 4

Descriptive information on the neuropsychological measures for the total group and for the three subgroups

Neuropsychological measures		Total ($n = 208$)	TD ($n = 148$)	ASD ($n = 31$)	ADHD ($n = 30$)	Group differences
		$M(SD)$	$M(SD)$	$M(SD)$	$M(SD)$	$F(2,206)$
GNG	% EOC	28.5(17.15)	27.14(16.76) ^a	35.16(20.25) ^a	28.33(14.26) ^a	2.86
	Mean RT	327.69(55.15)	328.38(56.47) ^a	313.28(48.33) ^a	339.17(53.65) ^a	1.73
AS	Interference control	56.20(86.30)	56.41(85.17) ^a	62.90(104.11) ^a	48.25(72.89) ^a	0.22
FA	RT _{crit} -RT _{hits}	134.37(100.22)	132.71(93.43) ^a	129.35(121.79) ^a	147.75(110.52) ^a	0.33
SA	Flexibility time	351.98(214.80)	339.60(205.88) ^a	346.75(194.03) ^a	418.30(267.33) ^a	1.70
	Flexibility error	0.18(0.50)	0.14(0.35) ^a	0.24(0.80) ^a	0.33(0.70) ^a	2.07

Note. GNG = Go/No-Go task; EOC = errors of commission; AS = Animal Stroop task; Interference control = $RT_{\text{incongruent condition}} - RT_{\text{control condition}}$; FA = Focused Attention task; RT_{crit} = RT trials with correct rejection irrelevant target; RT_{hits} = RT trials with correct response to a relevant target; SA = Shifting Attention task. *P*-values are derived from one-way ANOVA. Superscripts reflect subgroup differences derived from post-hoc Hochberg's GT2 Test; identical letters indicate that there were no differences between those particular groups. Results remained the same after controlling for differences in IQ.

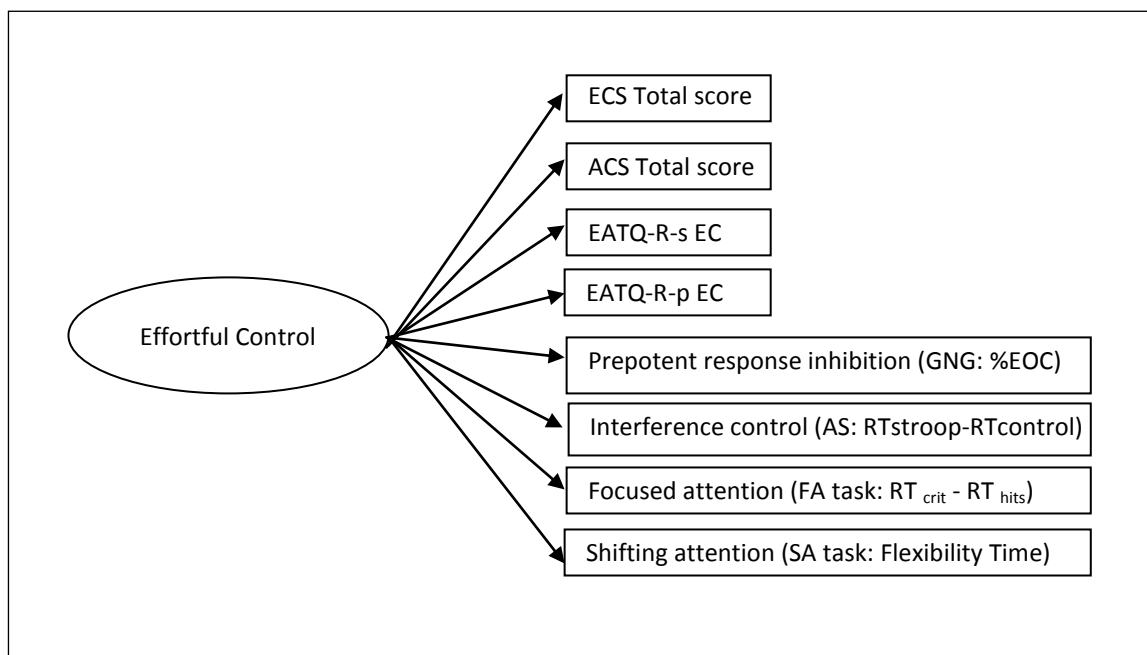


Figure 1. Hypothesized model linking the EC scales and performance on neuropsychological EF measures to the underlying latent variable EC. GNG: Go/No-Go; AS: Animal Stroop; FA: Focused Attention; SA: Shifting Attention

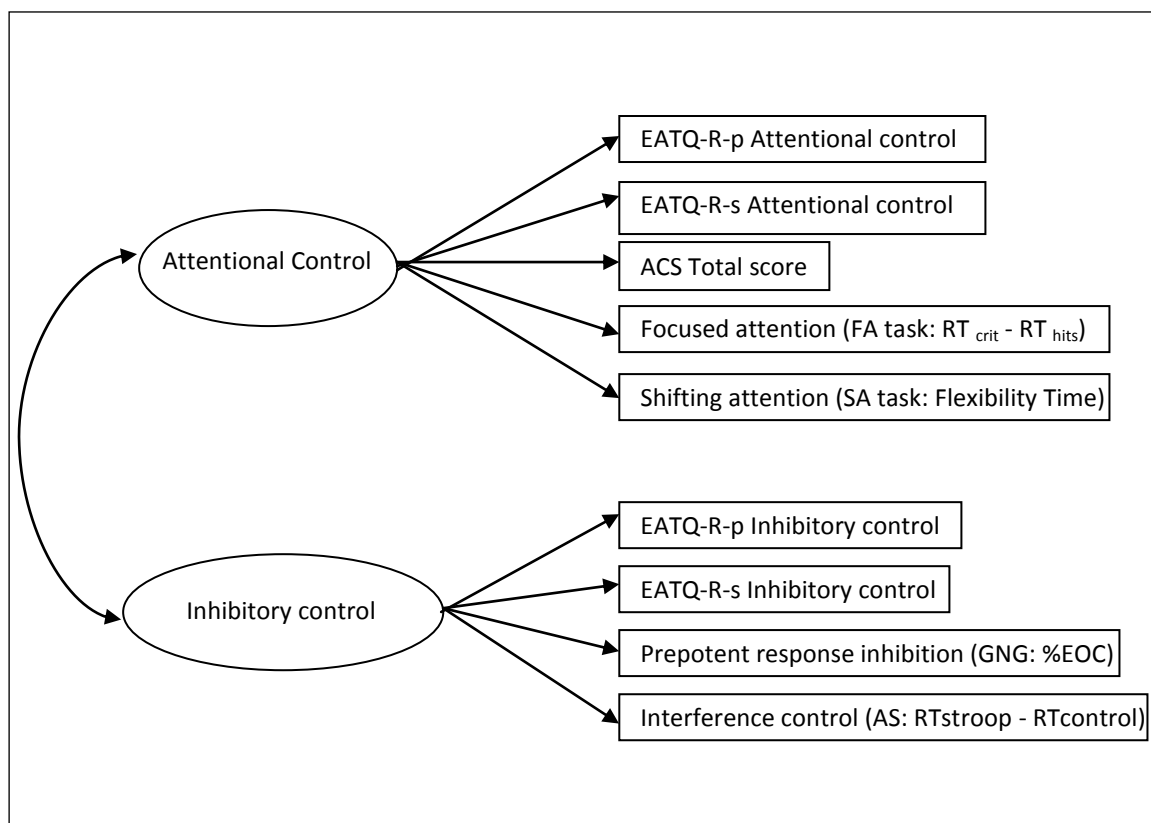


Figure 2. Alternative model linking (sub)scales of the EC questionnaires and performance on neuropsychological EF measures to the underlying latent variables Attentional Control and Inhibitory Control

(a)		(b)		(c)	
b	[m]	c	[z]	a	[h]
[c]	l	[k]	e	[m]	d

Figure 3. Example trials of the Focused Attention task (target presented in the [relevant diagonal] (a), target presented in the irrelevant diagonal (b), and no target presented (c))