UvA-DARE (Digital Academic Repository)

Exploring possible predictors and moderators of an executive function training for children with an autism spectrum disorder

de Vries, M.; Verdam, M.G.E.; Prins, P.J.M.; Schmand, B.A.; Geurts, H.M.

[Link to publication](#)

Exploring possible predictors and moderators of an executive function training for children with an autism spectrum disorder

**Marieke de Vries[1,2,3], Mathilde GE Verdam[1,2,4],
Pier JM Prins[1], Ben A Schmand[1,4] and Hilde M Geurts[1,2,3,5]**

Abstract

Previously, a total of 121 children with an autism spectrum disorder (ASD) performed an adaptive working memory (WM)-training, an adaptive flexibility-training, or a non-adaptive control (mock)-training. Despite overall improvement, there were minor differences between the adaptive and mock-training conditions. Moreover, dropout was relatively high (26%). In the current study we explored potential predicting and moderating factors to clarify these findings. The effects of intelligence, autism traits, WM, flexibility, reward sensitivity and Theory of Mind on dropout, improvement during training, and improvement in everyday executive functioning (EF), ASD-like behavior, and Quality of Life (QoL) were studied. None of the predictors influenced dropout or training improvement. However, 1) more pre-training autism traits related to less improvement in EF and QoL, and 2) higher reward sensitivity was related to more improvement in QoL and ASD-like behavior. These findings suggest that these EF-training procedures may be beneficial for children with fewer autism traits and higher reward sensitivity. However, the exploratory nature of the analyses warrant further research before applying the findings clinically.

The autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental disorder with large individual differences in several developmental domains (Geurts et al., 2014; Happé et al., 2006). These individual differences likely cause the often-observed lack of group effects of intervention studies in ASD, suggesting that ASD interventions should be customized. In this study, we explored several individual differences as potential predictors or moderators of the outcome of executive function (EF) training (Prins et al., 2013) for children with ASD.

We compared the effectiveness of adaptive working memory (WM) and flexibility training with an active, nonadaptive control (mock) training in a large randomized controlled trial in children with ASD ($N = 121$, 8–12 years; de Vries et al., 2014). Children in all three conditions improved on most outcome measures. The adaptive conditions (WM and flexibility) did not induce more improvement than the mock training, although there was a trend toward more improvement in WM and attention deficit hyperactivity disorder (ADHD) behavior after the WM training and a trend toward improvement

in flexibility after the flexibility training. The dropout rate was high (26%). These small effects and high dropout indicated that the studied training is not suitable for children with ASD. However, given the heterogeneity of ASD, the training might be promising for some children, while ineffective for others. Individual differences in EF (Diamond, 2012), and also in other areas such as intelligence quotient (IQ; Charman et al., 2011), theory of mind (ToM; Pellicano, 2010), or reward sensitivity (Dawson et al., 2001), may influence compliance and intervention effects.

[1]University of Amsterdam, The Netherlands
[2]Research Priority Area Yield, The Netherlands
[3]Dutch Autism and ADHD Research Center, The Netherlands
[4]Academic Medical Center, The Netherlands
[5]Dr. Leo Kannerhuis, The Netherlands

Corresponding author:
Hilde M Geurts, University of Amsterdam, Amsterdam, Postbus 15915, 1001 NK Amsterdam, The Netherlands.
Email: h.m.geurts@uva.nl

Compliance is considered to contribute to intervention success. Dropout rates in ASD interventions are not well established (McMahon et al., 2013), although low dropout has been reported in studies on EF interventions in ASD (0% in Fisher and Happé, 2005; 5%–9% in Kenworthy et al., 2014) and WM training studies (10% in Melby-Lervåg and Hulme, 2013). To evaluate the usefulness of a treatment for specific subgroups, it is important to unravel factors that influence dropout.

EF training in ASD is a relatively new research topic, and little is known about predictors of dropout and outcome. Therefore, we explored candidate predictors/moderators of compliance, improvement on training tasks, and training effects on everyday EF, ASD-like behavior, and quality of life (QoL). We included a limited amount of predictors/moderators to maintain sufficient statistical power. Given the small amount of research on EF training for children with ASD, we included predictors/modifiers based on findings of other (intervention) studies in children with ASD. Therefore, we studied IQ (Smith et al., 2015), autism symptoms (Eapen et al., 2016), and EF (Pellicano, 2010) as possible predictors/modifiers. Moreover, we included reward sensitivity (Dawson et al., 2001) and ToM (Pellicano, 2010) as possible predictors/modifiers, as children with ASD seem to have difficulties in these areas.

Intelligence might positively influence compliance and training effects. IQ correlates with treatment gains after early behavioral treatment for ASD (Smith et al., 2015; Vivanti et al., 2013) and predicts IQ, language, academic achievement, and adaptive behavior after treatment (Gabriels et al., 2001). We expected that IQ would positively influence effects on all outcome measures.

Less severe autism symptoms (Eapen et al., 2016) and better social engagement (Smith et al., 2015) might predict a better early behavioral treatment outcome, while social anxiety is associated with poorer outcome of an academic intervention (Pellecchia et al., 2016). Besides, dropout rates are higher for children with more complex and severe diagnoses (Barrett et al., 2008), partly due to secondary (e.g. behavioral) problems. We expected that children with less autism traits would drop out less often and improve more in ASD-like behavior.

Children with more severe EF deficits might benefit more from an EF intervention (Diamond, 2012; Karbach and Kray, 2009; Kray et al., 2011; Zinke et al., 2012) but have more difficulty to complete the training. Hence, children with more severe WM or flexibility deficits may not only benefit more from a WM or flexibility training but also drop out more often.

Children with ASD appear to show an altered reward sensitivity (Dawson et al., 2001), but findings are diverse. Some authors report less sensitivity to social rewards but no altered sensitivity to monetary rewards (Delmonte et al., 2012), whereas others suggest a general reward
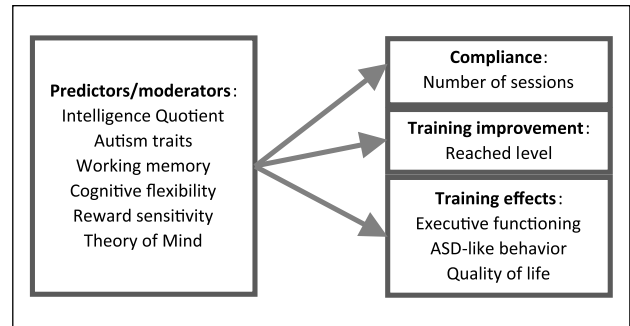

Figure 1. Model of predictors/moderators and outcome measures.

sensitivity deficit in ASD (Kohls et al., 2013). Our EF training includes rewarding elements, and children who are more sensitive to rewards might be less likely to drop out and might show more training-induced improvement.

Finally, an important theory on ASD concerns ToM deficits (Baron-Cohen et al., 1985). EF and ToM appear to be related (Pellicano, 2010), but the influence of ToM capacities on EF training effects is unknown. Functions related to ToM seem to have predictive value (Gillespie-Lynch et al., 2012); for example, goal understanding and imitation predict positive treatment outcome in children with ASD (Vivanti et al., 2013). In our EF training, one needs to help other characters in the game world, and more empathy with the game characters may lead to more willingness to help. A well-developed ToM is expected to positively influence compliance and outcome.

In short, we studied whether IQ, autism traits, EFs (WM and flexibility), reward sensitivity, and ToM influenced compliance, training-induced improvement, and training effects (EF, ASD-like behavior, and QoL; see Figure 1). We expected that (1) children with a higher IQ, less autism traits, a better developed ToM, and greater reward sensitivity would perform better within the training and show more training-induced improvement (EF, ASD-like behavior, and QoL); and (2) children with more WM or flexibility problems would not only more often drop out but also benefit more from the WM training and the flexibility training, respectively, indicating that pre-training WM and flexibility problems are moderators.

Methods

Participants

The training study was reported elsewhere (de Vries et al., 2014). In short, 121 children (8–12 years) were assigned to one of the three versions of Braingame Brian (Prins et al., 2013); 41 WM, 40 flexibility, and 40 mock training. Assignment was double blind. These children all had a clinical ASD diagnosis according to the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric

Association (APA), 2000, 2013), confirmed with the Autism Diagnostic Interview Schedule–Revised (ADI-R; De Jonge and de Bildt, 2007; Lord et al., 1994). All children had an IQ above 80 and no seizure disorder. A total of seven children had incomplete data due to missing values on the questionnaires (two WM, three flexibility, and two mock training). A total of 40 children used psychotropic medication (17 abstained during appointments and 23 continued).

Intervention

Each child performed a WM, flexibility, or mock training (see Prins et al., 2013, for a detailed description of the training). The training consists of 25 sessions, each including two blocks of three training tasks (WM, flexibility, and inhibition).

The WM training includes five adaptive visual–spatial WM training tasks with increasing difficulty (repeating sequences of blocks in a grid), and the flexibility and inhibition task remained at a low, nonadaptive level. The flexibility training includes an adaptive switch task and the WM and inhibition task remained at a low, nonadaptive level. In the mock training, all tasks remain at a low, nonadaptive level; hence, children do not differ in their reached level. In each intervention condition, all tasks are performed, although only the task being trained is adaptive.

Rewarding elements consisted of expanding the game world and extra gadgets in the game world (such as a unicycle) after correctly completing the tasks. Rewards were similar in WM, flexibility, and mock training. Within-training improvement was the sum of the highest levels reached on the five WM tasks (sequence length) and the highest level reached on the flexibility training task.

Predictors/moderators

IQ. The Wechsler Intelligence Scale for Children (WISC-III) subtests Vocabulary and Block Design were used to estimate IQ (Kort et al., 2002). These reliable subtests correlate highly with full-scale IQ (Vocabulary $r = 0.73$, Block Design $r = 0.71$; Legerstee et al., 2004).

Autism traits. The Social Responsiveness Scale (SRS; Constantino et al., 2003; Dutch version: Roeyers et al., 2011) was administered to the parents as indication of the children's autism traits. Higher scores indicate more autism traits.

WM. A computerized version of the Corsi (1972) block tapping task (BTT) was administered to the children to measure visual–spatial WM. A sequence of three of nine blocks light up and has to be repeated. The sequence increases with one block every four trials, up to eight blocks. The task ends when four sequences of the same length are repeated incorrectly. The longest reached sequence length was used as a WM predictor measure.

Flexibility. The gender-emotion switch task, an adaptation of the classical switch task (Rogers and Monsell, 1995), was administered to the children to measure flexibility (for details, see de Vries and Geurts, 2012). The switch cost in error rates was used as a flexibility predictor measure.

Reward sensitivity. Parents filled out the Sensitivity to Punishment and Sensitivity to Reward Questionnaire for Children (SPSRQ-C; Colder and O'Connor, 2004; Luman et al., 2012) to measure reward and punishment sensitivity in the children (33 items, 5-point Likert scale). The scale consists of one sensitivity-to-punishment scale and three sensitivity-to-reward scales (Drive, Reward Responsivity, and Impulsivity/Fun-Seeking). The used outcome measure was the total score of the three sensitivity-to-reward scales. Higher scores indicate higher reward sensitivity.

ToM. The Strange Stories Test (Spek et al., 2010), adapted from Happé (1994), was administered to measure ToM. A total of eight vignettes are read aloud, and two questions are asked: one to verify whether the story was understood and one to inquire about the motivation of the person in the vignette. Answers were scored by two raters independently, and interrater agreement was sufficient ($K = 0.61$, $p < 0.001$). The predictor measure was the number of correct answers including mention of a mental representation.

Outcome measures

Daily EF. The Behavior Rating Inventory of Executive Function (BRIEF; Gioia et al., 2000; Dutch version: Smidts and Huizinga, 2009) was administered to the parents to measure the children's daily EF (75 items, 3-point Likert scale). The total scale was used as the outcome measure. Higher scores indicate more EF problems.

ASD-like behavior. The Children's Social Behavior Questionnaire (CSBQ; Dutch version: Hartman et al., 2007) was administered to measure ASD-like behavior (49 items, 3-point Likert scale). The total scale was used as the outcome measure. Higher scores indicate more social difficulties.

QoL. The Pediatric Quality of Life Inventory (PedsQL; Bastiaansen et al., 2004; Varni et al., 2001) was used as a measure of QoL (23 items, 5-point Likert scale). The total raw scale was used as the outcome measure. Higher scores indicate a lower QoL.

Procedure

Children were tested on four occasions (screening, pre-training, post-training, and 6-week follow-up). All measures

were administered pre-training, and the BRIEF, CSBQ, and PedsQL were also administered post-training and at the 6-week follow-up. The training period took 6 weeks. The four appointments lasted about 100 min each. Children received a small gift for participating. Parents received a report with the ADI-R results and reimbursement of travel expenses.

Statistical analysis

First, to test whether IQ, autism traits, WM, flexibility, reward sensitivity, and ToM predicted dropout, a Cox regression analysis was performed with the number of completed sessions as dependent variable ($N=121$). Second, to test whether predictors (IQ, autism traits, WM, flexibility, reward sensitivity, and ToM) influenced training improvement, we checked which predictors correlated with the level reached in each training version (WM: $N=41$ and flexibility: $N=38$). Third, linear latent curve analyses were used to assess changes in EF, ASD-like behavior, and QoL over time. In the linear latent curve model, the changes across measurements at pre-training, post-training, and follow-up are reflected by two common factors: one factor represents the random intercept (the subjects' initial status), and the other factor represents the random slope (the subjects' curve). IQ, autism traits, WM, flexibility, reward sensitivity, ToM, and version of the training (WM, flexibility, or mock training) were added to the models as predictors of subjects' initial status and curve factors. To explore whether the effects of the predictors on changes in EF, ASD-like behavior, and QoL over time were moderated by the version of the training, we also investigated possible interaction effects between the version of the training and the other predictors ($N=117$). As higher scores on the BRIEF, CSBQ, and PedsQL indicate more problems, decreasing scores indicate improvement in behavior. Finally, to explore whether children who improved significantly in behavior differed from children who did not, we computed reliable change indexes (RCIs; Jacobson and Truax, 1991) for the BRIEF (Smidts and Huizinga, 2009), CSBQ (Hartman et al., 2007), and PedsQL (Varni et al., 2003). We compared predictor and within-training measures of children who significantly improved on one of the questionnaires (pre- to post-training) and children who did not improve.

The Cox regression analysis was performed using SPSS version 22 (IBM Corp., 2013). Linear latent curve analyses were performed using the computer program Lavaan (Rosseel, 2012). The models were fitted to the mean values, variances, and covariances of the scores of EF, ASD-like behavior, and QoL at pre-training, post-training, and follow-up. Analyses were done for each outcome measure separately.[1] Parameter estimates of the final models are reported and interpreted.

Results

The training groups did not differ in demographics and predictors (see Table 1).

Dropout

Dropout[2] did not differ between the three training groups (WM, flexibility, and mock training; see Figure 2) and was not influenced by any of the predictor variables ($ps > 0.05$).

Training improvement

None of the predictor variables were significantly related to the sum of the reached WM training levels or the level reached in the flexibility training ($ps > 0.05$).

Changes in EF, ASD-like behavior, and QoL over time

Overall, subjects' EF, ASD-like behavior, and QoL significantly improved over time (see Table 2). The mean change (standard deviation (SD)), range, and RCI (de Vries et al., 2014; Jacobson and Truax, 1991) from pre- to post-training were $M=5.5$ (11.6), range $=-18$ to 41, and RCI $= 12.4\%$ for the BRIEF; $M=3.8$ (9.3), range $=-13$ to 44, and RCI $= 14.6\%$ for the CSBQ; and $M=3.0$ (8.9), range -21 to 25, and RCI $= 28.1\%$ for the PedsQL. Autism traits and reward sensitivity were significantly related to subjects' initial status of all outcome variables, indicating that more autism traits and a higher reward sensitivity go together with more EF problems, more ASD-like behavior, and lower QoL. In addition, autism traits were a significant predictor of general changes in EF and QoL scores at post-training assessment, indicating that more autism traits lead to less improvement in EF and QoL across time, specifically at post-training assessment. Reward sensitivity was a significant predictor of both QoL and ASD-like behavior at post-training assessment, where higher reward sensitivity leads to relatively more improvement of QoL and ASD-like behavior, in particular immediately after training. There were no other significant main effects of the predictors. There was an effect of training version on the changes in QoL: the WM training led to significantly more improvement across time as compared to the mock training. There was a significant effect of pre-training flexibility on the change in QoL for the subjects in the flexibility training group as compared to the mock training group. Children in the flexibility training group with worse pre-training flexibility improved less in QoL. No other significant interaction effects were found (see Table 3).

Worse pre-training EF and ASD-like behavior scores were associated with less improvement over time ($r=0.51$ and $r=0.43$), whereas worse pre-training QoL scores were associated with more improvement in QoL over time

Table 1. Mean (*M*) and standard deviation (SD) of demographics, predictors, and training measures.

Measure	Group						Group comparison		
	WM (*N*=41)		Flexibility (*N*=40)		Mock (*N*=40)		χ^2	*p*	
Gender, M/F	37/4		36/4		35/5		(2)0.1	0.91	
Medication use[a]	29/6/6		27/6/7		25/5/10		(4)1.5	0.82	
	M (SD)	Range	*M* (SD)	Range	*M* (SD)	Range	$F_{(2, 118)}$	*p*	η_p^2
Age (years)	10.5 (1.3)	8.1–12.7	10.5 (1.3)	7.9–12.9	10.6 (1.4)	8.1–12.9	0.1	0.94	0.00
Predictor variables									
Intelligence quotient	113.3 (21.3)	81–170	111.0 (20.2)	81–154	108.0 (20.4)	81–154	0.7	0.51	0.01
Autism traits[b]	99.7 (24.2)	58–152	98.0 (24.3)	58–147	93.7 (21.0)	57–149	0.7	0.49	0.01
Working memory[c]	5.2 (0.8)	3–7	5.6 (1.0)	4–8	5.3 (1.2)	3–8	1.3	0.29	0.02
Cognitive flexibility[d]	4.7 (7.6)	−10.8–23.6	7.1 (6.5)	−7.6–23.5	5.5 (7.0)	−4.9–33.3	1.3	0.29	0.02
Reward sensitivity[e]	52.8 (10.3)	30–74	54.5 (11.7)	26–80	54.2 (12.3)	28–83	0.2	0.79	0.00
Theory of mind[f]	5.7 (1.5)	2–8	5.0 (1.7)	1–8	5.0 (2.1)	0–8	2.1	0.13	0.03
Training measures									
Number of sessions	21.3 (7.3)	1–25	19.5 (8.7)	1–25	21.4 (7.7)	1–25	0.7	0.48	0.01
Level[g]	19.4 (6.5)	3.2–27.2	9.2 (4.5)	1–14					

WM: working memory.

[a]No medication use/abstained during test sessions/medicated during test sessions.
[b]Raw score on the Social Responsiveness Scale.
[c]Longest reached sequence length on the Corsi block tapping task.
[d]Error rate switch costs gender-emotion task.
[e]Total score of the three sensitivity-to-reward scales on the Sensitivity to Punishment and Sensitivity to Reward Questionnaire for Children (WM: *N*=40).
[f]Number of correct answers including mention of a mental representation on the Short Stories Test.
[g]Highest levels reached on the five WM tasks (sequence length), and the highest level reached on the flexibility training task. Flexibility: *N*=38, as two children did not finish the first session.
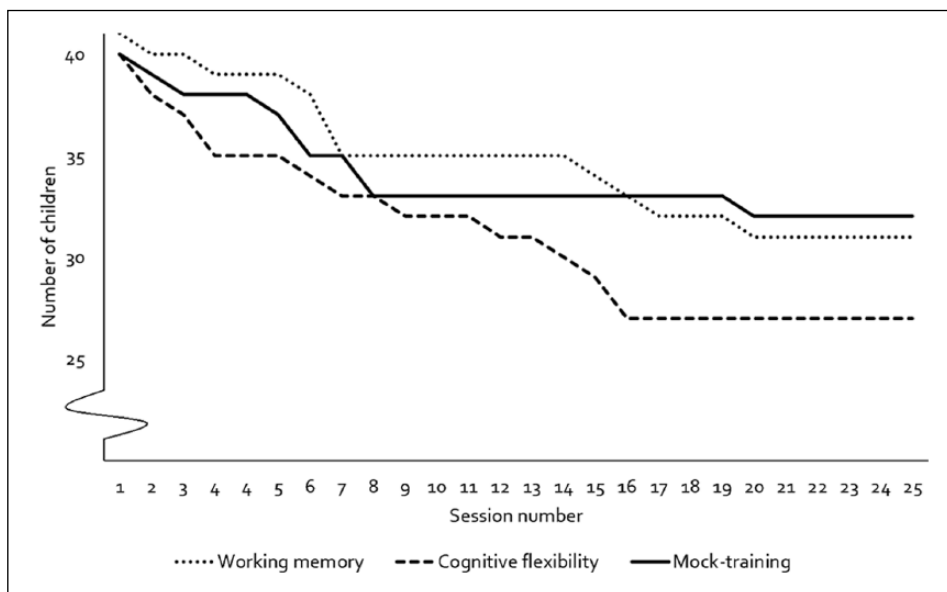

Figure 2. Number of children who completed each session.

(*r*=−0.49). The linear latent curve models were able to explain 71%–91% of the variances of the pre-training, post-training, and follow-up assessments of EF, ASD-like behavior, and QoL. The predictors were able to explain 39%, 28%, and 13% of the variance in the curve factors of EF, ASD-like behavior, and QoL, respectively.

The only outlier was on the CSBQ (change in ASD-like behavior). It appeared to be a child whose mother filled out

Table 2. Goodness of overall fit of the linear latent curve models for change in executive functioning, ASD-like behavior, and quality of life.

Model	df	CHISQ	p	CFI	RMSEA (95% CI)
Executive functioning[a]	9	11.78	0.23	0.99	0.05 (0.00–0.13)
ASD-like behavior[b]	8	8.51	0.39	1.00	0.02 (0.00–0.12)
Quality of life[c]	7	9.75	0.20	0.99	0.06 (0.00–0.15)

ASD: autism spectrum disorder; CHISQ: chi square; RMSEA: root mean square error of approximation; CI: confidence interval.
[a]Behavior Rating Inventory of Executive Function.
[b]Children Social Behavior Questionnaire.
[c]Pediatric Quality of Life Inventory.

Table 3. Influence of intelligence quotient, autism traits, working memory (WM), flexibility, reward sensitivity, theory of mind, and version of the training on change in post-training assessment of quality of life (QoL), executive functioning, and ASD-like behavior.

Predictor variables	Outcome measures					
	QoL[a] (N = 101)		Executive functioning[b] (N = 108)		ASD-like behavior[c] (N = 108)	
	Initial status (R² = 0.363)	Curve (R² = 0.128)	Initial status (R² = 0.473)	Curve (R² = 0.391)	Initial status (R² = 0.683)	Curve (R² = 0.275)
	β	β	β	β	β	β
Intelligence quotient	−0.16	0.17	−0.02	0.37	0.15	−0.22
Autism traits[d]	0.52***	−0.08	0.47***	0.41*	0.76***	−0.06
WM[e]	0.11	0.08	−0.06	−0.04	−0.11	0.06
Flexibility[f]	−0.05	0.00	−0.07	−0.08	−0.09	−0.18
Reward sensitivity[g]	0.24**	−0.13	0.51***	−0.17	0.34***	−0.25
Theory of mind[h]	−0.04	−0.02	−0.10	0.10	−0.11	0.33
WM training versus flexibility training	−0.16	0.17	−0.09	0.08	−0.16	0.24
WM training versus mock training	−0.12	0.31*	−0.17	0.45	−0.04	0.28
Flexibility training versus mock training	−0.04	−0.15	0.08	−0.37	−0.12	−0.04

Additional parameters	Mean	Explained variance (%)	Mean	Explained variance (%)	Mean	Explained variance (%)
Pre-training assessment	32.96	90.6	164.30	84.9	47.51	79.6
Post-training assessment	31.10	71.3	159.97	84.5	45.39	79.3
Follow-up assessment	29.23	77.9	155.63	76.3	43.28	77.9

Additional effects		β
Autism traits on QoL scores at post-training assessment		0.14**
Reward sensitivity on QoL scores at post-training assessment		−0.14***
Reward sensitivity on ASD-like behavior scores at post-training assessment		−0.03*

ASD: autism spectrum disorder; QoL: quality of life; WM: working memory.
[a]Pediatric Quality of Life Inventory.
[b]Behavior Rating Inventory of Executive Function.
[c]Children Social Behavior Questionnaire.
[d]Social Responsiveness Scale.
[e]The longest reached sequence length on the Corsi block tapping task.
[f]The switch cost in error rates on the gender-emotion switch task.
[g]Sensitivity to Punishment and Sensitivity to Reward Questionnaire for Children.
[h]Number of correct answers on the Strange Stories Test.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

the questionnaire pre-training and the father post-training. We reanalyzed the data without children with this same rater pattern (WM training: $N=3$, flexibility training: $N=4^3$) to investigate the robustness of the results. Most results did not change. However, the effect of training version on the changes in QoL (i.e. that the WM training led

to significantly more improvement across time as compared to the mock training) was no longer significant. Besides, we now found a significant effect of ToM on the changes in ASD-like behavior: a better ToM predicted that ASD-like behavior improved less over time.

RCI analyses

Children who improved significantly (based on RCI) on one of the questionnaires (N = 46) performed better on the ToM task than children who did not improve (F(1, 114) = 5.8, p = 0.02). There were no differences observed for other measures (IQ, autism traits, WM, flexibility, reward sensitivity, number of sessions completed, reached WM, or flexibility training level). Children who declined in PedsQL scores (N = 9) showed higher switch costs on the gender-emotion switch task (F(1, 114) = 8.2, p = 0.01).

Discussion

A total of 121 children performed one of the three versions of an EF training: WM, flexibility, or mock training (de Vries et al., 2014). Because of high dropout and minor treatment effects, possible predictors and moderators of compliance and outcomes were explored. We expected that IQ, ToM, and reward sensitivity were positive predictors, and autism traits was a negative predictor of compliance, training performance, and outcome. WM deficits and flexibility deficits were expected to increase dropout, but to positively influence training effects, as these children had more room to improve. We found that none of the predictor variables influenced dropout or within-training improvement. However, more autism traits led to less improvement in EF and QoL after training, while higher reward sensitivity led to more improvement of QoL and ASD-like behavior. Moreover, the WM training led to more improvement in QoL than the mock training, but the latter effect did not appear to be very robust. Moreover, we found some evidence that children with better pre-training flexibility who did the flexibility training improved more in QoL than children in the mock training group.

The large number of analyses, the relatively small effects, and the low statistical power do not allow us to draw firm conclusions. However, the findings suggest that there may be interesting predictive patterns that warrant future research. First, the training may be more appropriate for children who are sensitive to rewards. Although under debate (e.g. Reed et al., 2012), children with ASD are thought to be less sensitive to rewards (Kohls et al., 2013). We currently found that children with higher reward sensitivity also had more EF problems and ASD-like behavior and a lower QoL. Although this subgroup of children appears to be more vulnerable, they also improved more after training. Perhaps, the rewarding elements in the training were more profitable for them. Second, the training

may be more useful for children with less autism traits. Possibly, the training was too difficult or frustrating for children with more autism traits.

The variables we examined did not predict dropout. Possibly, other factors we did not control for, such as parental support, and children's motivation and attitude toward the training tasks and game world, influenced dropout. Moreover, we currently looked specifically into child characteristics as predictors. As parents might influence the children's compliance, studying parent's characteristics (such as motivation, IQ, or education) as predictors of compliance might be (more) useful. Unfortunately, we did not collect this information in this study. Visual inspection suggests that most children who stopped did so within the first 8 sessions, and little dropout occurred after session 16. We assumed that completing (a large part of) the training was a prerequisite for improvement (Jaeggi et al., 2011) and that completing more sessions would lead to more improvement (dosage effect), although this was not unequivocally found. Some did (Virués-Ortega, 2010) and others did not (Bellini et al., 2007) find a dosage effect of treatment in ASD. Moreover, dosage did not seem to moderate WM training effects (Melby-Lervåg and Hulme, 2013). In line, the number of completed sessions was not related to outcome in this study. How many sessions need to be performed as a threshold for success needs to be unraveled in future research.

We did not find that children with lower initial WM or flexibility improved more in these functions after training, unlike previous research (Diamond, 2013; Kray et al., 2011). In contrast, we found that children with initial poorer flexibility improved *less* in QoL after the flexibility training. Previous research showing that people with less well-developed flexibility improved more after a flexibility training was based on a cross-over design; people who first performed an adaptive training and then a mock training improved more in flexibility than people who trained in the reverse order (Kray et al., 2011). The effect of *initial* flexibility deficits on other outcome variables is unclear. The claim that children with poorer initial EFs benefit more from interventions that target EFs (Diamond, 2013) is based on other treatments than computerized EF training alone. EF might influence reaction to other treatments differently than reaction to a specific EF training and might influence other outcome variables differently than EF itself. Hence, the indications that lower pre-training EF is more susceptible to improvement are not very well substantiated, and baseline EF may not be a very good predictor of general training improvement but might predict different outcome variables in a different way.

The lack of an IQ effect on the outcome measures might be due to the inclusion of only children with an IQ above 80. Although children with an IQ above 70 tend to have a better QoL than children with a lower IQ, IQ did not predict QoL in children with an IQ above 70 (Howlin et al.,

2004). Similarly, the effect of IQ on treatment outcome may disappear for children with an average or higher IQ.

ToM also did not influence training and outcome clearly. However, when reanalyzing the data without children for whom the pre-training questionnaire was filled out by the mother and post-training by the father, a better ToM was related to less improvement in ASD-like behavior. Possibly, children with a better developed ToM had less room for improvement in ASD-like behavior. However, the Short Stories Test (Happé, 1994; Spek et al., 2010) that was used might lack sensitivity in children, as the performance of a control group of typically developing children did not differ from children with ASD ($p > 0.05$, unpublished data obtainable from the first author), and the Short Stories Test might not have captured ToM sufficiently. Nevertheless, in line with previous research that found no influence of a ToM training on EF (Fisher and Happé, 2005), the current results showed no strong association between ToM and training improvement.

Overall, children with ASD improved in EF, ASD-like behavior, and QoL, but unfortunately, for some children, QoL decreased. These children showed more flexibility problems but did not differ on other pre-training measures. Factors that might predict who will decline are contraindications for EF training and important to study in future research.

There are some caveats in this study. First, the relatively low statistical power implies that the findings are too premature to translate into specific clinical advice, although they may shape important future research avenues. Moreover, we cannot be sure that the lack of certain expected effects is due to absence of these effects or to lack of power. Second, although we considered reward sensitivity as a predictor, we did not quantify whether children considered the game elements of the training rewarding. This might be a prerequisite to observe a training effect. Third, we did not include language ability and age as predictors. Language development, particularly poor language abilities (e.g. Schreibman et al., 2009; Sherer and Schreibman, 2005), seems predictive of prospects in ASD (Gillespie-Lynch et al., 2012). The current sample of children had quite well-developed language; hence, variability seemed too small for this variable to be a suitable predictor. Early treatment and increasing treatment length seem most effective in children up to 7 years (Granpeesheh et al., 2009). Moreover, as recent research showed that neither intervention intensity nor age influenced early behavioral intervention outcome in children with ASD (Vivanti et al., 2013), we did not expect an effect of age in the currently included age range. Fourth, fathers and mothers differ in the evaluation of their child. Although we requested that for each test assessment the same rater would be present, this was not always possible. This did not affect the findings but remains important for future research. Finally, other predictors, such as medication use and comorbid ADHD, might have been of influence. To keep the statistical power as high as possible, we included a limited number of predictors. The large variety in medication use (type of medication, dosage, and use during the study) made these data noisy, and analyzing the data of the intervention study without children who used medication did not change the main findings (de Vries et al., 2014). We checked whether ADHD characteristics accounted for the relation between autism traits, reward sensitivity, and EF (post hoc hierarchical regression). Although ADHD characteristics predicted EF, autism traits and reward sensitivity predicted EF above and beyond ADHD characteristics ($ps < 0.01$; unpublished data obtainable from the first author). Therefore, we assume that medication use and ADHD characteristics did not largely influence the current findings.

In sum, we did not find predictors of dropout. Unfortunately, we did not systematically collect data of parent characteristics, and parents might have played a crucial role in the children's compliance. Keeping the aforementioned caveats in mind, our findings suggest that EF training may be more fruitful for children with high reward sensitivity and relatively few autism traits. The lack of power and small effects urge replication of these findings in order to determine their clinical importance. Given that we did not detect clear indicators to predict the high dropout, and as the currently used EF training appeared to have a limited effect, it can be questioned whether rerunning the training in children with high reward sensitivity and few autism traits is a promising future route to take. However, we argue that it would be informative to systematically measure the parents' influence on compliance and include reward sensitivity and autism traits as possible predictors of effectiveness in future studies of (to be developed) EF training programs.

Notes

1. Model fit was evaluated using the chi square (CHISQ) test of exact fit, where a significant CHISQ value indicates that there is a significant deviation between model and data.

In addition, the root mean square error of approximation (RMSEA; Steiger, 1990; Steiger and Lind, 1980) and comparative fit index (CFI; Bentler, 1990) were used to evaluate approximate fit. RMSEA values of below 0.05 indicate "close" approximate fit, and values below 0.08 indicate "reasonable" approximate fit (Browne and Cudeck, 1992). CFI values above 0.95 indicate relatively "good" model fit (Hu and Bentler, 1999). When the model fit was not good, modifications to the model were made based on inspection of correlation residuals ($>$0.10). Each modification to the model was evaluated using the CHISQ test of difference in exact fit, where a significant chi square difference (CHISQdiff) indicates a significant improvement in model fit.

2. The number of completed training sessions did not influence the main outcome variables (de Vries et al., 2014).

3. In all other cases, the same rater filled out the questionnaire pre- and post-training, although in two cases, the mother filled out the questionnaire pre-training and both parents filled out the questionnaire together post-training.

References

American Psychiatric Association (APA) (2000) *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: APA.

American Psychiatric Association (APA) (2013) *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Arlington, VA: APA.

Baron-Cohen S, Leslie AM and Frith U (1985) Does the autistic child have a "theory of mind." *Cognition* 21(1): 37–46.

Barrett MS, Chua W, Crits-Christoph P, et al. (2008) Early withdrawal from mental health treatment: implications for psychotherapy practice. *Psychotherapy: Theory, Research, Practice, Training* 45(2): 247.

Bastiaansen D, Koot HM, Bongers IL, et al. (2004) Measuring quality of life in children referred for psychiatric problems: psychometric properties of the PedsQLTM 4.0 generic core scales. *Quality of Life Research* 13(2): 489–495.

Bellini S, Peters JK, Benner L, et al. (2007) A meta-analysis of school-based social skills interventions for children with autism spectrum disorders. *Remedial and Special Education* 28(3): 153–162.

Bentler PM (1990) Comparative fit indexes in structural models. *Psychological Bulletin* 170(2): 238–246.

Browne MW and Cudeck R (1992) Alternative ways of assessing model fit. *Sociological Methods Research* 21: 230–258.

Charman T, Pickles A, Simonoff E, et al. (2011) IQ in children with autism spectrum disorders: data from the special needs and autism project (SNAP). *Psychological Medicine* 41(3): 619.

Colder CR and O'Connor RM (2004) Gray's reinforcement sensitivity model and child psychopathology: laboratory and questionnaire assessment of the BAS and BIS. *Journal of Abnormal Child Psychology* 32(4): 435–451.

Constantino JN, Davis SA, Todd RD, et al. (2003) Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders* 33(4): 427–433.

Corsi PM (1972) *Human memory and the medial temporal region of the brain*. PhD Thesis, McGill University, Montreal, QC (vol. 34, p. 819B).

Dawson G, Osterling J, Rinaldi J, et al. (2001) Brief report: recognition memory and stimulus-reward associations: indirect support for the role of ventromedial prefrontal dysfunction in autism. *Journal of Autism and Developmental Disorders* 31(3): 337–341.

De Jonge M and De Bildt A (2007) *Nederlandse bewerking van de ADI-R*. Amsterdam: Hogrefe Uitgevers B.V.

De Vries M and Geurts HM (2012) Cognitive flexibility in ASD; task switching with emotional faces. *Journal of Autism and Developmental Disorders* 42(12): 2558–2568.

De Vries M, Prins PJM, Schmand BA, et al. (2014) Working memory and cognitive flexibility-training for children with an autism spectrum disorder: a randomized controlled trial. *Journal of Child Psychology and Psychiatry* 56(5): 566–576.

Delmonte S, Balsters JH, McGrath J, et al. (2012) Social and monetary reward processing in autism spectrum disorders. *Molecular Autism* 3(1): 7.

Diamond A (2012) Activities and programs that improve children's executive functions. *Current Directions in Psychological Science* 21(5): 335–341.

Diamond A (2013) Executive functions. *Annual Review of Psychology* 64: 135–168.

Eapen V, Crncec R and Walter A (2016) There are gains, but can we tell for whom and why? Predictors of treatment response following group early start Denver Model intervention in preschool-aged children with autism spectrum disorder. *Autism Open Access* 6: 168.

Fisher N and Happé F (2005) A training study of theory of mind and executive function in children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders* 35(6): 757–771.

Gabriels RL, Hill DE, Pierce RA, et al. (2001) Predictors of treatment outcome in young children with autism: a retrospective study. *Autism: The International Journal of Research and Practice* 5(4): 407–429.

Geurts HM, Sinzig J, Booth R, et al. (2014) Neuropsychological heterogeneity in executive functioning in autism spectrum disorders. *Autism Research* 60(3): 155–162.

Gillespie-Lynch K, Sepeta L, Wang Y, et al. (2012) Early childhood predictors of the social competence of adults with autism. *Journal of Autism and Developmental Disorders* 42(2): 161–174.

Gioia GA, Isquith PK, Guy SC, et al. (2000) Test review: Behavior rating inventory of executive function. *Child Neuropsychology* 6(3): 235–238.

Granpeesheh D, Dixon DR, Tarbox J, et al. (2009) The effects of age and treatment intensity on behavioral intervention outcomes for children with autism spectrum disorders. *Research in Autism Spectrum Disorders* 3(4): 1014–1022.

Happé F (1994) An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders* 24(2): 129–154.

Happé F, Ronald A and Plomin R (2006) Time to give up on a single explanation for autism. *Nature Neuroscience* 9(10): 1218–1220.

Hartman CA, Luteijn E, Moorlag A, et al. (eds) (2007) *Manual for the CSBQ [handleiding voor de VISK]*. Amsterdam: Harcourt.

Howlin P, Goode S, Hutton J, et al. (2004) Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry* 45(2): 212–229.

Hu L and Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* 6(1): 1–55.

IBM Corp (2013) *IBM SPSS Statistics for Windows*. 22nd ed. New York: Armonk.

Jacobson NS and Truax P (1991) Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* 59(1): 12.

Jaeggi SM, Buschkuehl M, Jonides J, et al. (2011) Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the United States of America* 108(25): 10081–10086.

Karbach J and Kray J (2009) How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science* 12(6): 978–990.

Kenworthy L, Anthony LG, Naiman DQ, et al. (2014) Randomized controlled effectiveness trial of executive function intervention for children on the autism spectrum. *Journal of Child Psychology and Psychiatry* 55(4): 274–383.

Kohls G, Schulte-Ruther M, Nehrkorn B, et al. (2013) Reward system dysfunction in autism spectrum disorders. *Social Cognitive and Affective Neuroscience* 8(5): 565–572.

Kort W, Schittekatte M, Compaan EL, et al. (2002) *WISC-III NL. handleiding. nederlandse bewerking*. London: The Psychological Corporation.

Kray J, Karbach J, Haenig S, et al. (2011) Can task-switching training enhance executive control functioning in children with attention deficit/-hyperactivity disorder? *Frontiers in Human Neuroscience* 5: 180.

Legerstee JS, Van der Reijden-Lakeman IA, Lechner-van der Noort MG, et al. (2004) Bruikbaarheid verkorte versie WISC-RN in de kinderpsychiatrie. *Kind En Adolescent* 25(4): 178–182.

Lord C, Rutter M and Couteur A (1994) Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders* 24(5): 659–685.

Luman M, van Meel CS, Oosterlaan J, et al. (2012) Reward and punishment sensitivity in children with ADHD: validating the sensitivity to punishment and sensitivity to reward questionnaire for children (SPSRQ-C). *Journal of Abnormal Child Psychology* 40(1): 145–157.

McMahon CM, Lerner MD and Britton N (2013) Group-based social skills interventions for adolescents with higher-functioning autism spectrum disorder: a review and looking to the future. *Journal of Adolescent Health, Medicine and Therapeutics* 2013(4): 23.

Melby-Lervåg M and Hulme C (2013) Is working memory training effective? A meta-analytic review. *Developmental Psychology* 49(2): 270.

Pellecchia M, Connell JE, Kerns CM, et al. (2016) Child characteristics associated with outcome for children with autism in a school-based behavioral intervention. *Autism: The International Journal of Research and Practice* 20(3): 321–329.

Pellicano E (2010) Individual differences in executive function and central coherence predict developmental changes in theory of mind in autism. *Developmental Psychology* 46(2): 530–544.

Prins PJM, Brink ET, Dovis S, et al. (2013) "Braingame brain": toward an executive function training program with game elements for children with ADHD and cognitive control problems. *Games for Health Journal* 2(1): 44–49.

Reed P, Hawthorn R, Bolger S, et al. (2012) Disrupted stimulus control but not reward sensitivity in individuals with autism spectrum disorders: a matching law analysis. *Journal of Autism and Developmental Disorders* 42(11): 2393–2403.

Roeyers H, Thys M, Druart C, et al. (2011) *Screeninglijst voor autismespectrumstoornissen (SRS) handleiding*. Amsterdam: Hogrefe uitgevers bv.

Rogers RD and Monsell S (1995) Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General* 124(2): 207.

Rosseel Y (2012) Lavaan: an R package for structural equation modeling. *Journal of Statistical Software* 48(2): 1–36.

Schreibman L, Stahmer AC, Barlett VC, et al. (2009) Brief report: toward refinement of a predictive behavioral profile for treatment outcome in children with autism. *Research in Autism Spectrum Disorders* 3(1): 163–172.

Sherer MR and Schreibman L (2005) Individual behavioral profiles and predictors of treatment effectiveness for children with autism. *Journal of Consulting and Clinical Psychology* 73(3): 525.

Smidts D and Huizinga M (eds) (2009) *BRIEF executieve functies gedragsvragenlijst: Handleiding*. Amsterdam: Hogrefe Uitgevers.

Smith T, Klorman R and Mruzek DW (2015) Predicting outcome of community-based early intensive behavioral intervention for children with autism. *Journal of Abnormal Child Psychology* 43(7): 1271–1282.

Spek AA, Scholte EM and Van Berckelaer-Onnes IA (2010) Theory of mind in adults with HFA and Asperger syndrome. *Journal of Autism and Developmental Disorders* 40(3): 280–289.

Steiger JH (1990) Structural model evaluation and modification: An interval estimation approach.

Steiger JH and Lind JC (1980) Statistically based tests for the number of common factors. *Paper presented at the annual meeting of the Psychometric Society*, Iowa City, IA. Multivariate Behavioral Research 25: 173–180.

Varni JW, Burwinkle TM, Seid M, et al. (2003) The PedsQL™* 4.0 as a pediatric population health measure: feasibility, reliability, and validity. *Ambulatory Pediatrics* 3(6): 329–341.

Varni JW, Seid M and Kurtin PS (2001) PedsQL (TM) 4.0: reliability and validity of the pediatric quality of life inventory (TM) version 4.0 generic core scales in healthy and patient populations. *Medical Care* 39(8): 800–812.

Virués-Ortega J (2010) Applied behavior analytic intervention for autism in early childhood: meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical Psychology Review* 30(4): 387–399.

Vivanti G, Dissanayake C, Zierhut C, et al. (2013) Brief report: predictors of outcomes in the early start Denver model delivered in a group setting. *Journal of Autism and Developmental Disorders* 43(7): 1717–1724.

Zinke K, Zeintl M, Eschen A, et al. (2012) Potentials and limits of plasticity induced by working memory training in old-old age. *Gerontology* 58(1): 79–87.