University of Groningen

Effectiveness of Social Skills Training for children with autism spectrum disorder

Dekker, Vera

DOI:
[10.33612/diss.193918122](https://doi.org/10.33612/diss.193918122)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Dekker, V. (2021). *Effectiveness of Social Skills Training for children with autism spectrum disorder*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.
<https://doi.org/10.33612/diss.193918122>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Effectiveness of Social Skills Training for children with autism spectrum disorder

Vera Dekker

Design/lay-out

Promotie In Zicht | www.promotie-inzicht.nl

Print

Ipskamp Printing

© V. Dekker, 2021

All rights are reserved. No part of this book may be reproduced, distributed, stored in a retrieval system, or transmitted in any form or by any means, without prior written permission of the author.

university of
groningen

Effectiveness of Social Skills Training for children with autism spectrum disorder

Proefschrift

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Groningen
op gezag van de rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

vrijdag 17 december 2021 om 11.00 uur

door

Vera Dekker

geboren op 28 februari 1986
te Assen

Promotores

Prof. dr. P.J. Hoekstra

Prof. dr. M.H. Nauta

Copromotor

Dr. A.A. de Bildt

Beoordelingscommissie

Prof. dr. A. Lichtwarck-Aschoff

Prof. dr. S.M. Begeer

Prof. dr. P.E.H.M. Muris

Table of Contents

Chapter 1	General Introduction	7
Chapter 2	A randomized controlled study of a social skills training for preadolescent children with autism spectrum disorders: generalization of skills by training parents and teachers?	23
Chapter 3	Social skills group training in children with autism spectrum disorder: a randomized controlled trail	47
Chapter 4	Application of latent class analysis to identify subgroups of children with autism spectrum disorders who benefit from social skills training	73
Chapter 5	A fresh pair of eyes: A blind observation method for evaluating social skills of children with ASD in a naturalistic peer situation in school	101
Chapter 6	Differences in parental stress and criticism between fathers and mothers and the association with internalizing and externalizing problems in children with ASD	129
Chapter 7	Summary and General Discussion	151
*	Nederlandse samenvatting	171
	Dankwoord	183
	About the author	191
	Publications	195

1 |

General Introduction

Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that is characterized by impairments in social communication and interaction and by the presence of restricted and repetitive patterns of behaviors, interests, or activities (American Psychiatric Association, 2013). The impairments in social communication and interaction result in limited socio-emotional reciprocity, ranging from not being able to initiate contact to initiating or responding to social contact in a way that is deviant and that may be inappropriate. Other impairments concern non-verbal communication (for example initiating and maintaining eye contact) and developing, maintaining, and understanding relationships (for example engaging in reciprocal conversations, emotion recognition, empathy, joint attention, and recognition of social cues; American Psychiatric Association, 2013).

ASD is a heterogeneous disorder that manifests in many different ways. The expression of impairments differs between children and over time (Lord et al., 2020). Besides ASD, other disorders often co-occur in children with ASD, such as Attention-deficit/hyperactivity Disorder (ADHD), anxiety, depression, and epilepsy. The current mean prevalence of ASD in the United States is approximately 1.5-2% in children aged 8 years (Maenner et al., 2020) and in Europe between 0.4% and 3.1% (Chiarotti & Venerosi, 2020).

Limitations in social-communicative skills in children with ASD

In a considerable part of the children with ASD, the limited social communication is not due to a lack of interest, but to a lack of skills and an insufficient ability to understand and determine when to use which skill (White et al., 2007). Social interaction with others is complex, because it varies depending on the context of the situation and is sometimes unpredictable (Will et al., 2018). Children with ASD tend to have more difficulties in learning social-communicative skills, e.g. they learn these less automatically and slower than typically developing children (Laugeson et al., 2009; Frankel et al., 2010).

The difficulties in applying the appropriate social-communicative skills at the right moment affect the social-communicative development of children with ASD (Rao et al., 2008). Children with ASD have smaller social networks, the friendships they identify are less often perceived as friendship by the other child, and the quality of their friendships is poorer compared to the quality of friendships in typical classmates (Kasari et al., 2011). They are at increased risk of being bullied, of peer rejection and of social isolation, compared to their typically developing peers (Schroeder et al., 2014). Additionally, social difficulties have been associated with mood and anxiety problems, especially when children grow older, as the social environment becomes more complex, with more social demands for children that are harder to fulfill (White et al., 2007). This may lead to more awareness of the child of his

or her social limitations, which in turn can result in increased distress and impairment (White et al., 2007).

Training social-communicative skills

Support to improve social communicative skills of children with ASD is a frequently heard desire of their parents. Parents observe the difficulties of their child and the consequences of their behaviors, and they oversee that the future may demand even more complex social-communicative skills from their child, in order to stay in contact with peers (White et al., 2007). Parents of children with ASD experience insufficient confidence that they can adequately help their child on their own (Karst & van Hecke, 2012), therefore they ask support from professionals in mental health care or school.

As a response to this desire, social skills training (SST) is frequently recommended and applied for children with ASD (McMahon et al., 2013a; Kaat & Lecavalier, 2014; Gates et al., 2017; Wolstencroft et al., 2018). Such training aims to improve the social functioning of the child, e.g. to increase social-communicative skills and quality of life, and to reduce behavior problems and negative consequences of social-communicative impairments (Will et al., 2018). An important assumption when applying SST is that children are able to improve their social skills and that the intervention (possibly with the help of parents, teachers, or peers) leads to such improvement. Another important assumption is that children generalize the learned social skills to situations outside the training, and that they will be able to use learned skills in daily life situations, such as at school and at home (McMahon et al., 2013a; Carruthers et al., 2020). Appropriate social skills in these situations are important because these are the situations in which children interact with other children and in which they can make friends. The third assumption underlying SST is that learned skills will retain after completion of the training and can be the basis for further development of social-communicative skills (McMahon et al., 2013a). Thus, SST aims to improve daily life social-communicative skills in children with ASD, for a longer term.

SSTs are often performed in groups (usually two to six individuals with ASD led by one to three therapists) and are mostly based on behavioral therapeutic principles and social learning theory (White et al., 2007; Reichow et al., 2013). Various methods are used, such as behavior modelling of a specific social skill, practicing the learned social skills through roleplay and receiving feedback on the performance (McMahon et al., 2013a; Wolstencroft et al., 2018). Form, intensity, and duration of SSTs vary (Gates et al., 2017; Wolstencroft et al., 2018). For example, peers or parents may or may not be involved during the training (Koenig et al., 2010; Deckers et al., 2016). Many SSTs have weekly sessions of 60-90 minutes (Laugeson et al., 2009; Soorya et al., 2015; Deckers et al., 2016; Freitag et al., 2016), while other SSTs are intensive summer camps (Lopata et al., 2010).

Effectiveness of Social Skills Training

Although SST is an important part of interventions for children with ASD, until now, the evidence for the effectiveness of an SST for children with ASD is inconclusive (Reichow et al., 2013; Kaat & Lecavalier, 2014; Gates et al., 2017; Wolstencroft et al., 2018). Moderate to large effect sizes have been found between treatment and control groups (e.g. care as usual, waitlist or no treatment) over time (Gates et al., 2017; Wolstencroft et al., 2018), however, measuring the effect of SST in ASD is impeded by several factors.

One of the factors that hinders measuring the effect of social skills training is the complex construct that social skills comprise, which is difficult to operationalize (McMahon et al., 2013a; White et al., 2007). As a result, studies into the effect of SST differ in the social-communicative constructs they investigated, such as social competence (Reichow et al., 2013; Gates et al., 2017) or social knowledge (Gates et al., 2017; McMahon et al., 2013a; Wolstencroft et al., 2018). As a consequence, many different outcome measures have been used. Sometimes studies used more specific intervention related instruments and sometimes more general social skills instruments (McMahon et al., 2013a; Kaat & Lecavalier, 2014) were used. This variation in outcome measures may have contributed to variation in effect of SST over studies. A clear description of the constructs under study and the instruments used to measure them is needed. This would include a description of how the specific social skills learned during the SST are measured, how more general social skills can be measured and how broader social functioning in daily life can be measured.

Another aspect that influences the level of the effect size observed in SST trials is the informant. The behavior of the child with ASD is assessed differently by different informants, such as the child itself, parents, teacher, therapist, or an observer or informant that is blind for treatment allocation (McMahon et al., 2013a; Wolstencroft et al., 2018). The effect of informant may be explained by the various expressions of social skills and social behavior children show among different situations, and by the variation in perceived impact of the intervention across situations and therefore informants (Gates et al., 2017). For example, the generally positive effects reported by parents (Gates et al., 2017) may possibly be related to expectancy effects, based on their investment in and allegiance to the intervention (McMahon et al., 2013a; Gates et al., 2017; Wolstencroft et al., 2018). These aspects probably play a stronger role when raters are not blind to the treatment condition (McMahon et al., 2013a). Gates and colleagues (2017) found large effects based on self-reported measures of social competence in their review. In line with parent report, also child-reported results may have been influenced by expectancy effects (McMahon et al., 2013a). Additionally, there are indications that children and adolescents with ASD may overestimate their social skills compared to their parents (Lerner et al., 2012). Further analysis has

shown that child-reported improvements are especially based on social knowledge, while the performance of social skills in real-life situations has not improved (Gates et al., 2017). Given these findings, it is important to include more objective assessments besides questionnaires for children, parents, and teachers. Coding such assessments by independent observers may reduce biases due to expectancy and allegiance effects (Wolstencroft et al., 2018). The main reason that few studies included such direct observation seems to be a lack of instruments. Deriving reliable information through direct observations is difficult due to the intensity, its limited feasibility, and the difficulty (White et al., 2007; McMahon et al., 2013b).

Last, in addition to insight into the overall effectiveness of an SST, it is also important to know more specifically for which children the training is effective and for which children the training is less or not effective (Gates et al., 2017).

Generalization

Apart from the effects of SST in children with ASD on the specific social skills as learned during the training, a few other aspects should be considered when studying the effect of SST. One of these is the generalization of learned social-communicative skills in order to appropriately interact in daily life situations.

Although generalization of learned social skills to situations beyond the training is an important assumption of SST, only few studies have investigated this generalization to different daily life situations (Owens et al., 2008; Herbrecht et al., 2009; Laugeson et al., 2009; White et al., 2010; Wood et al., 2014). Some studies used information of raters in other situations, such as teacher report (e.g. Herbrecht et al., 2009; Laugeson et al., 2009; White et al., 2010) and other studies used an observation in a naturalistic situations at school recess (Owens et al., 2008; Wood et al., 2014).

Clinicians intuitively expect that involving parents in an SST, i.e. teaching them how to manage their child's social-communicative behavior and discussing the themes that the child is working on at each point, might be valuable to promote generalization of the learned social skills (Karst & van Hecke, 2012). Involving parents in an SST may make them feel more prepared to support their child with ASD in practicing social-communicative skills at home, as learned during the training. This would address the feelings of incapability many parents have when trying to teach their child these skills themselves (Karst & van Hecke, 2012). Additionally, parents may be able to recognize and create situations to practice social skills with their child. In addition to parents, involving teachers may support children in practicing social skills at school and contribute to the generalization of the child's skills in a similar way. Despite the expectations of clinicians, scientific evidence for the added value of the involvement of parents and teachers in SST is insufficient.

The review of Wolstencroft and colleagues (2018) is promising in this respect, reporting that children in an SST with parental involvement improved with large effect

sizes compared to moderate effect sizes for children in an SST without parental involvement. However, since comparison between the various studies is difficult given major differences in study designs, such as participant characteristics and measurement instruments, studies that directly compare training with and without parental involvement are needed.

Long-term effects

Another issue that has not received full attention yet, is that more information is needed on the long-term effects of SST (Karst & van Hecke, 2012). An SST should be a basis for further development of social-communicative skills, so that children continue to use the learned skills or even develop more skills based on the ones they have learned during SST (McMahon et al., 2013a). However, only few studies have investigated the long-term effects of SST by following up on treated and non-treated groups. For example, the review of Gates and colleagues (2017) did not include long-term comparisons between treatment and control groups because most studies did not include follow-up measurements, or only in the treated group. In order to investigate the long term effect of SST compared to care as usual, a control group for the whole follow-up period is needed (Gates et al., 2017).

Additionally, the limited number of studies that have been conducted into the long-term effects report contradictory outcomes. Therefore conclusions about the longer term effects of the SST for children with ASD can hardly be drawn (Gates et al., 2017; Wolstencroft et al., 2018). Two studies reported that the improvement still existed after three months follow-up measurement (Freitag et al., 2016; Deckers et al., 2016). However, since both studies only reported the change from pretreatment to three months follow-up, it is unclear for how long children proceeded their learning curve after the training had ended. Another study reported group improvement in social behavior between pre- and posttreatment, but this improvement no longer existed three months postintervention (Soorya et al., 2015). It is important to note that only about half of the participants were included in the follow-up measurements in this study. In the study of Choque Olsson and colleagues (2017) adolescents showed significant improvement in the SST conditions compared to standard care alone between pre- and posttreatment as well as between posttreatment and three months follow-up. For children, no group differences were observed at all. More research on longer-term effects is needed, both on whether the effects identified between pre- and posttreatment remain after several months and on whether social-communicative behaviors continues to improve further after posttreatment.

Probably the involvement of parents and teachers in SST could also enhance the long-term effects. After the training has ended, parents can continue to encourage their child to use the learned social skills, so that the child can continue to practice social skills after the training (Karst & van Hecke, 2012). This has not been addressed in former studies.

What works for whom?

As is the case for all healthcare, for the treatment of children with ASD it is important to know what approach works for whom. Little is known on whether specific subgroups of participants benefit more from (specific characteristics of) SST than others and further research has been recommended (Gates et al., 2017; Wolstencroft et al., 2018). Most studies investigating the effectiveness of SST in children with ASD in relation to participant and intervention characteristics were based on group level effects.

SST characteristics that have been investigated include parental involvement in SST (Wolstencroft et al., 2018), involving peer tutors (Gates et al., 2017), intensity and duration (Wolstencroft et al., 2018; Jonsson et al., 2019). More intensive SSTs (e.g. summer camp format) and SSTs longer than 40 hours of contact time had large effect sizes compared to the control group post-treatment. Weekly SSTs and SSTs of 40 hours or shorter, had moderate effect sizes compared to the control group post-treatment.

Of participant characteristics, sex and cognitive ability have shown consistent effects (McMahon et al., 2013a; Choque Olsson et al., 2017), with stronger effects for females than males (McMahon et al., 2013a; Choque Olsson et al., 2017) and stronger effects for children with higher IQ's than children with lower IQ's (Herbrecht et al., 2009). However, these findings should be interpreted with caution, since few females were included in the studies and most children were high-functioning. For age, comorbidity, and medication inconsistent effects were reported (Gates et al., 2017).

These findings may increase our insight into factors that may determine to what extent children benefit from SST when treated and control groups are compared. However, possible specific response patterns of individual children or specific subgroups to SST and the relation of these response patterns to multiple participant characteristics are overseen. Identifying subgroups of children with ASD with similar response patterns to SST, and investigating the factors that characterize these subgroups will help us understand which specific subgroups benefit from SST and which do not benefit or do so to a lesser extent. It is important to investigate the factors that are associated with high and low improvements after intervention, in order to be able to adjust interventions for those who may not respond, and to identify factors that may help to predict who will respond to a specific intervention (Paynter et al., 2018).

Outline of this dissertation

This dissertation mainly focuses on the Efficacy of Social Skills Training In Autism (ESTIA) study, a study into the effectiveness of a manualized SST with and without parental and teacher involvement for high-functioning children with ASD in the last two and a half years of primary education, compared to each other and to care-as-usual (CAU).

Chapter 2 provides an extensive description of the ESTIA study. First, a theoretical substantiation of the importance of investigating the effectiveness of social skills training for children with ASD and the limitations in previous studies are described. Additionally, this chapter describes the design and procedures of the study, including the conditions, the randomization, the participants, the social skills training under study (with and without parental involvement), the instruments and measurements.

The immediate and long-term effects of the group-based SST for high-functioning children with ASD compared to care-as-usual are described in **chapter 3**. Additionally, the SST with parental and teacher involvement is compared to SST only. We hypothesized that children in an SST improve more on measures of social skills compared to children without an SST. Additionally, we hypothesized that children with an SST with parent and teacher involvement (SST-PTI) would improve more than children without this support and that children with additional support would better maintain these skills after training, compared to children without such support.

Next, as described in **chapter 4**, we aimed to identify subgroups of participants based on (1) their parent reported social communicative skills before SST and (2) their patterns of response to SST over time (measured immediately and six months after SST). Additionally, we investigated whether and how the response patterns in the subgroups could be related to multiple participant and intervention characteristics. We examined participant and intervention characteristics, such as age, sex, verbal IQ, symptoms of ADHD, symptoms of total and social anxiety, severity of ASD symptomatology, and the presence or absence of involvement of parents and teachers in SST. In case we would find different subgroups, we hypothesized that girls, children with higher verbal IQ, lower ASD symptoms, and involvement of parents and teachers were associated with a more 'successful' response pattern. Due to inconsistent previous findings, we had no hypothesis regarding age, ADHD, and social anxiety.

Chapter 5 reports on the development of an observation instrument for investigating the generalization of learned social skills to daily life situations, i.e. free play at school. With this observation method, we aimed to contribute to the development of a naturalistic, reliable, valid, and feasible measure of social skills in a familiar, unstructured situation with peers, eventually in order to gain insight into the

clinical significance and generalization of learned social skills during SST in peer interactions in daily life. In chapter 5, the evaluation of the psychometric properties is reported, which is needed before the instrument could be used for the evaluation of the effectiveness of an SST for children with ASD.

Chapter 6 focuses on the well-being of parents of children with ASD and stands apart from the chapters on SST. Not only children with ASD themselves experience difficulties and challenges. For parents, raising a child with ASD is complicated and presents several challenges associated with parental stress (Rodriguez et al., 2019) and with the emotional quality of the relationship with their child (Romero-Gonzales et al., 2018).

Parents of children with ASD have reported high levels of stress (Hayes & Watson, 2013; Giovagnoli et al., 2015; Gong et al., 2015) that are reportedly higher than those of parents of typically developing children or children with other disabilities (Estes et al., 2013; Hickey et al., 2019a). Parental stress has been associated with various aspects of being a parent of a child with ASD, such as managing the ASD and comorbid symptoms of the child (i.e. internalizing and externalizing behavior problems; Maskey et al., 2013; Valicenti-McDermott et al., 2015; Karst & van Hecke, 2012; Hastings et al., 2005; Hastings et al., 2006; Lounds et al., 2007; Yorke et al., 2018), finding a way in (special) education (Whitaker, 2002), financial challenges (Sharpe & Baker, 2007), and finding the right support and health care (Nesbitt, 2000). Since most children with ASD need extra support compared to typically developing children, this causes a lifelong burden for parents and also for siblings, which may influence parents' perception of their parenting skills and decrease their optimism about their own future and the future of their child with ASD (Karst & van Hecke, 2012).

The challenges mentioned above are also associated with the emotional quality of the parent-child relationship (Bader et al., 2015), often resulting in high levels of parental criticism. Some studies indicated that parents of children with ASD more often show more criticism towards their child with ASD than towards their other children (Griffith et al., 2015). In turn, high levels of criticism in parents of children with ASD were associated with externalizing behavioral problems of the child (Romero-Gonzales et al., 2018). The association was still present after controlling for parental stress and ASD symptom severity (Bader et al., 2015). The association between parental criticism and internalizing behaviors in children with ASD is less consistent (Romero-Gonzales et al., 2018).

Important to note is that most studies on parental stress and criticism focused on mothers and that fathers have been involved to a limited extent (Romero-Gonzales et al., 2018). The studies that included both mothers and fathers suggested that mothers express higher levels of parental stress (Hickey et al., 2019a; Rodriguez et al., 2019) and criticism (Hickey et al., 2019b; Hickey et al., 2020). More research is needed to

gain insight into the differences between fathers and mothers in parental stress and criticism and in how these are related to child characteristics.

Chapter 6 describes the association between parental stress and parental criticism towards the child in and between mothers and fathers, and between these parental factors and ASD symptoms and externalizing and internalizing behaviors of the child. We hypothesized that the parental stress in our sample would be high and that mothers would experience higher levels of stress than fathers and express more criticism. Second, we hypothesized a positive association between parental stress and criticism. Furthermore, we also hypothesized a positive association between both parental factors and ASD symptoms and externalizing and internalizing behaviors of the child.

Finally, in **chapter 7**, a summary of the findings is provided and general conclusions and implications for clinical practice and research are discussed.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. (5th ed.). Washington DC: American Psychiatry Publishing.
- Bader, S.H., Barry, T.D., & Hann, J.A.H. (2015). The relation between parental expressed emotion and externalizing behaviors in children and adolescents with an autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities*, 30(1), 23–34.
- Carruthers, S., Pickles, A., Slonims, V., Howlin, P., & Charman, T. (2020). Beyond intervention into daily life: A systematic review of generalization following social communication interventions for young children with autism. *Autism Research*, 13(4), 506–522.
- Chiarotti, F., & Venerosi, A. (2020). Epidemiology of Autism Spectrum Disorders: A Review of Worldwide Prevalence Estimates Since 2014. *Brain Sciences*, 10(5), 274.
- Choque Olsson, N., Flygare, O., Coco, C., Görling, A., Råde, A., Chen, Q., Lindstedt, K., Berggren, S., Serlachius, E., Tammimies, K., Kjellin, L., & Bölte, S. (2017). Social Skills Training for Children and Adolescents With Autism Spectrum Disorder: A Randomized Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(7), 585–592.
- Deckers, A., Muris, P., Roelofs, J., & Arntz, A. (2016). A Group-Administered social Skills Training for 8- to 12-Year-Old, high-Functioning Children With Autism Spectrum Disorders: An Evaluation of its Effectiveness in a Naturalistic Outpatient Treatment Setting. *Journal of Autism and Developmental Disorders*, 46(11), 3493–3504.
- Estes, A., Olson, E., Sullivan, K., Greenson, J., Winter, J., Dawson, G., & Munson, J. (2013). Parenting-related stress and psychological distress in mothers of toddlers with autism spectrum disorders. *Brain & Development*, 35(2), 133–138.
- Frankel, F., Myatt, R., Sugar, C., Whitham, C., Gorospe, C.M., & Laugeson, E. (2010). A randomized controlled study of parent-assisted Children's Friendship Training with children having autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 40(7), 827–842.
- Freitag, C.M., Jensen, K., Elsuni, L., Sachse, M., Herpertz-Dahlmann, B., Schulte-Rüther, M., Hänig, S., Von Gontard, A., Poustka, L., Schad-Hansjosten, T., Wenzl, C., Sinzig, J., Taurines, R., Geißler, J., Kieser, M., & Cholemkery, H (2016). Group-based cognitive behavioral psychotherapy for children and adolescents with ASD: the randomized, multicentre, controlled SOSTA-net trial. *Journal of Child Psychology and Psychiatry*, 57(5), 596–605.
- Gates, J.A., Kang, E., & Lerner, M.D. (2017). Efficacy of group social skills interventions for youth with autism spectrum disorder: A systematic review and meta-analysis. *Clinical Psychology Review*, 52, 164–181.
- Giovagnoli, G., Postorino, V., Fatta, L.M., Sanges, V., De Peppo, L., Vassena, L., De Rose, P., Vicari, S., & Mazzone, L. (2015). Behavioral and emotional profile and parental stress in preschool children with autism spectrum disorder. *Research in Developmental Disabilities*, 45, 411–421.
- Gong, Y., Du, Y., Li, H., Zhang, X., An, Y., & Wu, B.L. (2015). Parenting stress and affective symptoms in parents of autistic children. *Science China. Life Sciences*, 58(10), 1036–1043.
- Griffith, G.M., Hastings, R.P., Petalas, M.A., & Lloyd, T.J. (2015). Mothers' expressed emotion towards children with autism spectrum disorder and their siblings. *Journal of Intellectual Disability Research*, 59(6), 580–587.
- Hastings, R.P., Kovsho, H., Ward, N.J., Degli Espinosa, F., Brown, T., & Remington, B. (2005). Systems analysis of stress and positive perceptions in mothers and fathers of pre-school children with autism. *Journal of Autism and Developmental Disorders*, 35(5), 635–644.
- Hastings, R.P., Daley, D., Burns, C., & Beck, A. (2006). Maternal distress and expressed emotion: Cross-sectional and longitudinal relationships with behavior problems of children with intellectual disabilities. *American Journal of Mental Retardation*, 111(1), 48–61.
- Hayes, S.A., & Watson, S.L. (2013). The impact of parenting stress: A meta-analysis of studies comparing the experience of parenting stress in parents of children with and without autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(3), 629–642.

- Herbrecht, E., Poustka, F., Birnkammer, S., Duketis, E., Schlitt, S., Schmötzer, G., & Bölte, S. (2009). Pilot evaluation of the Frankfurt social skills training for children and adolescents with autism spectrum disorder. *European Child & Adolescent Psychiatry*, 18(6), 327–335.
- Hickey, E.J., Hartley, S.L., & Papp, L. (2019a). Psychological well-being and parent-child relationship quality in relation to child autism: an actor-partner modeling approach. *Family Process*, 59(2), 636–650.
- Hickey, E.J., Nix, R.L., & Hartley, S.L. (2019b). Family emotional climate and children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 49(8), 3244–3256.
- Hickey, E.J., Bolt, B., Rodriguez, G., & Hartley, S.L. (2020). Bidirectional relations between parent warmth and criticism and the symptoms and behavior problems of children with autism. *Journal of Abnormal Child Psychology*, 48(6), 865–879.
- Jonsson, U., Olsson, N.C., Coco, C., Göring, A., Råde, A., Chen, Q., Berggren, S., Tammimies, K., & Bölte, S. (2019). Long-term social skills group training for children and adolescents with autism spectrum disorder: a randomized controlled trial. *European Child & Adolescent Psychiatry*, 28(2), 189–201.
- Kaat, A.J., & Lecavalier, L. (2014). Group-based social skills treatment: A methodological review. *Research in Autism Spectrum Disorders*, 8(1), 15–24.
- Karst, J.S., & van Hecke, A.V. (2012). Parent and family impact of autism spectrum disorders: A review and proposed model for intervention evaluation. *Clinical Child and Family Psychology review*, 15(3), 247–277.
- Kasari, C., Locke, J., Gulsrud, A., & Rotheram-Fuller, E. (2011). Social networks and friendships at school: Comparing children with and without ASD. *Journal of Autism and Developmental Disorders*, 41(5), 533–544.
- Koenig, K., White, S.W., Pachler, M., Lau, M., Lewis, M., Klin, A., & Scahill, L. (2010). Promoting social skills development in children with pervasive developmental disorders: a feasibility and efficacy study. *Journal of Autism and Developmental Disorders*, 40(10), 1209–1218.
- Laugeson, E.A., Frankel, F., Mogil, C., & Dillon, A.R. (2009). Parent-assisted social skills training to improve friendships in teens with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39(4), 596–606.
- Lerner, M.D., Calhoun, C.D., Yee Mikami, A., & De Los Reyes, A. (2012). Understanding parent-child social informant discrepancy in youth with high functioning autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 42(12), 2680–2692.
- Lopata, C., Thomeer, M.L., Volker, M.A., Toomey, J.A., Nida, R.E., Lee, G.K., Smerbeck, A.M., & Rodgers, J.D. (2010). RCT of a manualized social treatment for high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 40(11), 1297–1310.
- Lord, C., Brugha, T.S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones, E.J.H., Jones, R.M., Pickles, A., State, M.W., Taylor, J.L., & Veenstra-vanderWeele, J. (2020). Autism spectrum disorder. *Nature reviews. Disease primers*, 6(1), 5.
- Lounds, J., Seltzer, M.M., Greenberg, J.S., & Shattuck, P.T. (2007). Transition and change in adolescents and young adults with autism: Longitudinal effects on maternal well-being. *American Journal of Mental Retardation*, 112(6), 401–417.
- Maenner, M.J., Shaw, K.A., Baio, J., Washington, A., Patrick, M., DiRienzo, M., Christensen, D.L., Wiggins, L.D., Pettygrove, S., Andrews, J.G., Lopez, M., Hudson, A., Baroud, T., Schwenk, Y., White, T., Rosenberg, C.R., Lee, L., Harrington, R.A., Huston, M., ... Dietz, P.M. (2020). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016. *Morbidity and mortality weekly report. Surveillance summaries (Washington, D.C. : 2002)*, 69(4), 1–12.
- Maskey, M., Warnell, F., Parr, J. R., Le Couteur, A., & McConachie, H. (2013). Emotional and behavioral problems in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(4), 851–859.
- McMahon, C.M., Lerner, M.D., & Britton, N. (2013a). Group-based social skills interventions for adolescents with higher-functioning autism spectrum disorder: a review and looking to the future. *Adolescent Health, Medicine and Therapeutics*, 4, 23–38.
- McMahon, C.M., Vismara, L.A., & Solomon, M. (2013b). Measuring changes in social behavior during a social skills intervention for higher-functioning children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(8), 1843–1856.

- Nesbitt, S. (2000). An evaluation of multi-agency service provision for children with autistic spectrum disorders. *The British Journal of Development Disabilities*, 46(90), 43–50.
- Owens, G., Granader, Y., Humphrey, A., & Baron-Cohen, S. (2008). LEGO therapy and the social use of language programme: an evaluation of two social skills interventions for children with high functioning autism and Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 38(10), 1944–1957.
- Paynter, J., Trembath, D., & Lane, A. (2018). Differential outcome subgroups in children with autism spectrum disorder attending early intervention. *Journal of Intellectual Disability Research*, 62(7), 650–659.
- Rao, P.A., Beidel, D.C., & Murray, M.J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: a review and recommendations. *Journal of Autism and Developmental Disorders*, 38(2), 353–361.
- Reichow, B., Steiner, A.M., & Volkmar, F. (2013). Cochrane review: social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD). *Evidence-Based Child Health*, 8(2), 266–315.
- Rodriguez, G., Hartley, S.L., & Bolt, D. (2019). Transactional relations between parenting stress and child autism symptoms and behavior problems. *Journal of Autism and Developmental Disorders*, 49(5), 1887–1898.
- Romero-Gonzalez, M.R., Chandler, S., & Simonoff, E. (2018). The relationship of parental expressed emotion to co-occurring psychopathology in individuals with autism spectrum disorder: a systematic review. *Research in Developmental Disabilities*, 72, 152–165.
- Schroeder, J.H., Cappadocia, M.C., Bebko, J.M., Pepler, D.J., & Weiss, J.A. (2014). Shedding light on a pervasive problem: A review of research on bullying experiences among children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 44(7), 1520–1534.
- Sharpe, D. L., & Baker, D. L. (2007). Financial issues associated with having a child with autism. *Journal of Family and Economic Issues*, 28(2), 247–264.
- Soorya, L.V., Siper, P.M., Beck, T., Soffes, S., Halpern, D., Gorenstein, M., Kolevzon, A., Buxbaum, J., & Wang, A.T. (2015). Randomized comparative trial of a social cognitive social skills group for children with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 54(3), 208–216.
- Valicenti-McDermott, M., Lawson, K., Hottinger, K., Seijo, R., Schechtman, M., Shulman, L. & Shinnar, S. (2015). Parental stress in families of children with autism and other developmental disabilities. *Journal of Child Neurology*, 30(13), 1728–1735.
- Whitaker, P. (2002). Supporting families of preschool children with autism: What parents want and what helps. *Autism*, 6(4), 411–426.
- White, S.W., Koenig, K., & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: a review of the intervention research. *Journal of Autism and Developmental Disorders*, 37(10), 1858–1868.
- White, S.W., Koenig, K., & Scahill, L. (2010). Group social skills instruction for adolescents with high-functioning autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 25(4), 209–219.
- Will, M.N., Currans, K., Smith, J., Weber, S., Duncan, A., Burton, J., Kroeger-Geoppinger, K., Miller, V., Stone, M., Mays, L., Luebrecht, A., Heeman, A., Erickson, C., & Anixt, J. (2018). Evidenced-Based Interventions for Children With Autism Spectrum Disorder. *Current Problems in Pediatric and Adolescent Health Care*, 48(10), 234–249.
- Wolstencroft, J., Robinson, L., Srinivasan, R., Kerry, E., Mandy, W., & Skuse, D. (2018). A Systematic Review of Group Social Skills Interventions, and Meta-analysis of Outcomes, for Children with High Functioning ASD. *Journal of Autism and Developmental Disorders*, 48(7), 2293–2307.
- Wood, J.J., Fujii, C., Renno, P., & van Dyke, M. (2014). Impact of cognitive behavioral therapy on observed autism symptom severity during school recess: a preliminary randomized, controlled trial. *Journal of Autism and Developmental Disorders*, 44(9), 2264–2276.
- Yorke, I., White, P., Weston, A., Rafla, M., Charman, T., & Simonoff, E. (2018). The association between emotional and behavioral problems in children with autism spectrum disorders and psychological distress in their parents: a systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, 48(10), 3393–3415.

2 |

A randomized controlled study of a social skills training for preadolescent children with autism spectrum disorders: generalization of skills by training parents and teachers?

Vera Dekker

Maaïke H. Nauta

Erik J. Mulder

Marieke E. Timmerman

Annelies de Bildt

Abstract

Background

Social skills training (SST) is a common intervention for children with autism spectrum disorders (ASDs) to improve their social and communication skills. Despite the fact that SSTs are often applied in clinical practice, the evidence for the effectiveness of these trainings for children with ASD is inconclusive. Moreover, long term outcome and generalization of learned skills are little evaluated. Additionally, there is no research on the influence of involvement of parents and teachers on effectiveness of SST and on the generalization of learned social skills to daily life. We expect parent and teacher involvement in SST to enhance treatment efficacy and to facilitate generalization of learned skills to daily life.

Method/design

In a randomized controlled trial (RCT) with three conditions, 120 participants with ASD at the end of primary school (10–12 years of calendar age) have been randomized to SST, SST-PTI (SST with Parent & Teacher Involvement), or care-as-usual. The SST consists of 18 group sessions of 1.5 hours for the children. In the SST-PTI condition, parents additionally participate in 8 parent sessions and parents and teachers are actively involved in homework assignments. Assessment takes place at three moments: before and immediately after the intervention period and at 6 months follow-up. Primary outcome is socialization, as an aspect of adaptive functioning. Secondary outcomes focus on specific social skills children learn during SST and on more general social skills pertaining to home and community settings from a multi-informant perspective. Additionally, possible predictors of treatment outcome will be assessed.

Discussion

The current study is an RCT study evaluating SST in a large sample of Dutch children with ASD in a specific age range (10–12 years). Strengths of the study are the use of one manualized protocol, application of standardized and internationally used rating instruments, use of multiple raters, investigation of generalization of learned skills to daily life, and the evaluation of efficacy in the longer term by follow-up measures at 6 months after the end of training.

Trial registration

NTR2405

Introduction

Autism spectrum disorders (ASDs) are defined in the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) as a neurodevelopmental disorder and are characterized by severe and pervasive impairment in two developmental areas: limitations in social communication and the presence of restricted, repetitive behaviors (American Psychiatric Association, 2013). The Social Communication domain as defined in the DSM-5 contains symptoms from the Social and Communication domain which were separate domains in the DSM-IV-TR. In the DSM-IV-TR, classifications within the autism spectrum were separated as well (i.e. autistic disorder [AD], Rett's disorder, childhood disintegrative disorder [CDD], Asperger's disorder and pervasive developmental disorder- not otherwise specified [PDD-NOS; (American Psychiatric Association, 2000), whereas the DSM-5 no longer contains specific ASD classifications. The symptoms are now seen as a continuum, ranging from mild to severe expression (Lauritsen, 2013).

This revised perspective on ASDs reflects empirical research into the underlying factors for ASDs (Lauritsen, 2013) and is in line with clinical presentation in daily practice. ASDs typically emerge during young childhood and they persist throughout the lifespan. The etiology of ASDs remains unclear. There is a strong involvement of genetic factors in ASD (Devlin & Scherer, 2012). The prevalence of ASDs is estimated at 1 per 150 children and adolescents (Baird et al., 2006), with the ratio of boys to girls being 4:1.

Children with ASDs lack the behavioral repertoire that is necessary to interact with others. Deficits vary over children with ASDs and may include lack of orientation towards social stimuli, inadequate use of eye contact, no or inadequate initiation of social interaction, difficulties in interpreting verbal and nonverbal social cues, inappropriate emotional responses, and lack of empathy for others' distress or emotions (Rao et al., 2008). Besides variation between individuals with ASDs, the social communication within children with ASDs varies over time as well. Although limitations in the communication domain are the core symptoms of ASDs, this does not mean that there can be no development in this area at all (White et al., 2007). However, learning and implementation of the behaviors necessary for social interaction may always remain harder, go slower, and be less automatic for children with ASDs than for other children (Laugeson et al., 2009; Frankel et al., 2010).

At least part of the children with ASDs has a desire for more peer social interaction (White et al., 2007), but they do not act in the right way to generate fluent social interaction with peers. Children with ASDs often show substantial social relational problems when compared to typically developing children, because they have difficulty sharing affective experiences or understanding the perspective of others, which is important for social reciprocity and the development of friendships (Rao et

al., 2008). Additionally, parents and teachers have reported larger deficiencies in cooperation, assertion, and self-control than a matched group of their typically developing peers (Rao et al., 2008). This may result in an increased risk for peer rejection and social isolation.

The social problems encountered already early in life may have serious consequences, often leading to major deviances in development as compared to typical development. From several studies it is known that social problems may also affect the achievement of normal developmental milestones and the establishment of satisfying peer and familial relationships (Rao et al., 2008; Volkmar et al., 2005). Further on, social deficits continue to negatively impact social and occupational functioning into adulthood. Adults with ASDs or high-functioning autism (HFA) are much more likely to be un(der) employed than the general population and are much less likely to have satisfying social relationships (Rao et al., 2008). Apart from the consequences for development in general, social skills deficits predict specific problems later in development such as mood and anxiety problems (White et al., 2007).

There are different types of interventions for individuals with ASDs. Most psychological interventions are developed from a behavioral perspective (Lord et al., 2005). Because of the importance of the social communication skills in ASDs, treatment has focused on improving these skills. Over the last decades, several methods have been developed with that aim, e.g. social stories (Karkhanavich et al., 2010), peer-mediated interventions (Koenig et al., 2010), parents-assisted interventions (Frankel et al., 2010), computer-based interventions (Beaumont & Sofronoff, 2008), and social skills groups (DeRosier et al., 2011).

Specific group-based social skills training (SST) has become an important part of treatment for children with ASDs in clinical practice and in schools (Solomon et al., 2004; Freitag et al., 2013). An SST is a child specific intervention based on behavioral and social learning, during which children are taught specific skills, as for instance making eye contact, initiating a conversation, and cooperation (White et al., 2007). SST is supposed to help children develop those skills that are most severely affected and at the same time most crucial for developing relationships with others. For children with social phobia or specific learning disabilities, SST has been shown to be an effective treatment (Spence et al., 2000; Forness & Kavake, 1996). However, the evidence for the effectiveness for SST for children with ASDs seems inconclusive (Rao et al., 2008; White et al., 2007; Reichow & Volkmar, 2010; Lopata et al., 2010).

The studies into SSTs so far show design limitations with respect to lack of a control group and randomized assignment, lack of adequate measurement of social skills and deficits associated with ASD, limited sample size, not using multiple informants, not including follow-up measurements, and lack of assessment of generalization of learned skills in daily life. Moreover, most of the studies have been performed in the US (Rao et al., 2008; White et al., 2007), except Yoo and colleagues (2014).

In a recent meta-analysis (Reichow et al., 2013), found only five studies, including a total of 196 participants, that were based on a randomized controlled trial design (RCT). All five studies included a control group that did not receive an intervention for social skills. They included studies until December 2011 in their analysis, all performed in the US. To the best of our knowledge, only one RCT study into group SST for children with ASDs has been published since, that included a randomized control group that did not receive an intervention for social skills. Additionally, this was the first study outside the US (Yoo et al., 2014).

A summary of study characteristics and outcome measures of these studies is shown in Table 1. As can be seen in Table 1, the age ranges were broad (ages of included children 6–17 years). Setting, frequency, and duration of the intervention also varied between the studies. One study investigated an intensive five-week summer intervention with five 70-minute treatment cycles every day (Lopata et al., 2010) and the other studies had weekly training sessions of 60–90 minutes (Laugeson et al., 2009; Frankel et al., 2010; Koenig et al., 2010; Solomon et al., 2004; Yoo et al., 2014). The protocol underlying the SSTs differed between the studies and involvement of parents or peers varied as well. Involvement of teachers was scanty, if at all. In two studies (Frankel et al., 2010; Koenig et al., 2010) peers were involved in the sessions of the SST and in four studies the parents were involved in the training (Laugeson et al., 2009; Frankel et al., 2010; Solomon et al., 2004; Lopata et al., 2010; Yoo et al., 2014). Based on their meta-analysis, Reichow and colleagues (2013), concluded that no firm conclusions can be drawn on the efficacy of SSTs for ASDs with respect to improving social competence, social communication, emotion recognition, and quality of life for children with ASDs, due to all the differences between the studies and the outcomes. In more detail: In three studies the children in the SST condition improved significantly (small-medium effect size) compared to the control condition. In the two other studies, no difference was found between the treatment and the control groups. More information on the conclusions of the separate studies is provided in the last column of Table 1.

With respect to limitations of the studies, Reichow and colleagues (2013) reported that the studies in their meta-analysis had the same limitations as mentioned before. Additionally, Gillies and colleagues (2013) commented on their meta-analysis (Reichow et al., 2013) with additional potential sources of bias, regarding families that could not be blinded for their condition, the exclusion of participants with intellectual disability and the inclusion of only group-based skills training.

As reported in Table 1, the studies used different outcome measurements. Some used specific intervention-related instruments, others used more general social skills instruments. The difference between the concepts behind each instrument and thus the exact outcome of each of the studies, complicates comparison between them. The use of standardized, internationally used and well validated instruments would facilitate measuring efficacy of SSTs and comparison between studies.

Table 1 Characteristics of published RCTs with control condition on SSTs in children with ASDs.

Auteur	N (INT/CON)	In- / exclusion criteria	Intervention
Frankel et al. (2010)	76 randomized, 68 completed training (35/33)	ADOS/ADI-R ASD 2 nd -5 th grade regular classroom VIQ > 60 Knowing rules of board and school yard games Able to switch topics in conversation Excl: psychotropic medication Thought disorder Clinical seizure disorder, gross neurologic disease or other medical disorder	Parent-assisted Children's Friendship Training (CFT). 12 weekly sessions of 60 minutes with concurrent sessions for child and parents. Class size was usually 10, with no more than 4 children with ASD.
Laugeson et al. (2009)	33 (17/16)	Clin dx ASD Age 13-17 IQ > 70 English fluency (child and parent) Parent want to participate Social problems Excl: history of major mental illness Hearing, visual or physical impairments	Program for the Education and Enrichment of Relational Skills (PEERS). 12 weekly sessions of 90 minutes with concurrent sessions for child and parents.
Koenig et al. (2010)	44 (25/19)	Clin dx ASD ADOS/SCQ/PDD-BI score ASD Age 8-11 IQ > 70 Excl: need for different treatment ABC irritability > 18 CSI clinically	16 weekly sessions of 75 minutes; 4 to 5 children/2 peer/ 2 licensed clinicians.
Lopata et al. (2010)	36 (18/18)	Clin dx HFASD Age 7-12 IQ > 70 VCI/PRI > 80 Expressive language score > 80	Summer training program. 5 weeks intervention with Five daily 70-minute treatment cycles every day; 3 therapists/6 children. Parents participated in weekly parent training groups on 90-minute sessions.
Solomon et al. (2004)	18 (9/9)	Clin dx ASD ADOS ASD, ADI-R AD Also met DSM-IV criteria of ASD based on a clinical interview Age 8-12 IQ > 75 Able to pass first theory of mind task Excl: serious conduct problems	The Social Adjustment Enhancement Curriculum. 20 weekly sessions of 90 minutes with concurrent sessions for child and parents; 3 therapists/5 children.

	Follow-up	Outcome measures	Conclusion (pre/post treatment)
	Only intervention group after 3 months	<ul style="list-style-type: none"> - Parent <ul style="list-style-type: none"> • QPQ (Ho, Gu, Con, Eng, Dis) • SSRS social skills (As, SC) • SSRS problem behavior (In, Ex) - Child <ul style="list-style-type: none"> • PHS • The Loneliness Scale - Teacher <ul style="list-style-type: none"> o PEI (Wi, Ag) 	<p>$p < .0001$ (Ho, Dis)</p> <p>$p < .05$ (SC)</p> <p>ns</p> <p>$p < .025$</p> <p>$p < .025$</p> <p>ns</p>
	No	<ul style="list-style-type: none"> - Parent <ul style="list-style-type: none"> • SSRS social skills • SSRS problem behavior - vQPQ (Ho, Gu, Con) - Teens <ul style="list-style-type: none"> • QPQ (Ho, Gu, Con) • TASSK • FQS - Teacher (n=13) <ul style="list-style-type: none"> • SSRS social skills • SSRS problem behavior 	<p>$p < .05$</p> <p>ns</p> <p>ns</p> <p>$p < .025$ (Ho)</p> <p>$p < .0001$</p> <p>$p < .05$</p> <p>ns</p> <p>ns</p>
	No	<ul style="list-style-type: none"> - Parent <ul style="list-style-type: none"> • CGI – improvement • SCI (PSI, SI) 	<p>$p = .001$</p> <p>ns</p>
	No	<ul style="list-style-type: none"> - Parent <ul style="list-style-type: none"> • ASC • SRS • BASC-2-PRS – withdrawal • BASC-2-PRS – social skills - Child <ul style="list-style-type: none"> • SKA • DANVA-2 (CF) • CASL 	<p>$p = .006$ $d = .584$</p> <p>$p = .003$ $d = .625$</p> <p>$p < .001$ $d = 1.055$</p> <p>ns $d = .365$</p> <p>$p < .001$ $d = 1.272$</p> <p>ns $d = .532$</p> <p>$p < .001$ $d = .390$</p>
	No	<ul style="list-style-type: none"> - Child <ul style="list-style-type: none"> • DANVA-2 (AF, CF) • Strange Stories Task • Faux Pas Stories Task • TOPS-ER 	<p>$p < .05$ (AF, CF)</p> <p>ns</p> <p>ns</p> <p>$p < .05$</p>

Table 1 Continued.

Auteur	N (INT/CON)	In- / exclusion criteria	Intervention
Yoo et al. (2014)	55 (28/27)	Clin dx ASD Age 12-18 School 6 th grade elementary school to 3 rd grade high school Social difficulties VIQ $\geq$ 65 Substantial treatment motivation No history of major mental illness No current problems with aggressive behavior or severe oppositional tendency No hearing, visual, or physical disabilities preventing outdoor sport activities No clinically significant physical or neurological illnesses inhibiting treatment.	The PEERS Treatment Manual. 14 weekly sessions of 90 minutes with concurrent sessions for child and parents.

ABC, Aberrant Behavior Checklist; AD, Autism Disorder; ADI-R, Autism Diagnostic Interview-Revised; ADOS, Autism Diagnostic Observation Schedule (Ic-a, language and communication algorithm; Ic-t, language and communication total; rsi-a, reciprocal social interaction algorithm; rsi-t, reciprocal social interaction total); AHWA-VABS, Korean Version of the Vineland Adaptive Behavior Scale); ASC, Adapted Skillstreaming Checklist; ASD, Autism Spectrum Disorder; ASDS, Asperger Syndrome Diagnostic Scale (L, Language; SI, Social Interaction; BP, Behavioral Problems; CA, Cognitive Ability; SP, Sensorimotor Problems); BASC-2-PRS, Behavior Assessment System for Children, Second Edition-Parent Rating Scales; BDI, Beck Depression Inventory (F, Father; M, Mother); CASL, Comprehensive Assessment of Spoken Language; CDI, Child Depression Inventory; CGI, Clinical Global Impressions Scale; CON, Control condition; CSI, Children's Symptom Inventory; DANVA-2, Diagnostic Analysis of Nonverbal Accuracy-2 (AF, Adult Facial Expression; CF, Child Facial Expression); FQS, Friendship Qualities Scale; HFASD, High-Functioning Autism Spectrum Disorder; INT, Intervention; K-CBCL, Korean Version of the Child Behavior Checklist (An/dep, anxiety/depression; In, Internalizing problems);

To clarify the effectiveness of SSTs for children with ASD, follow-up measurement is also important (Gillies et al., 2013). None of the studies reported in table 1 had follow-up measurements for the treatment as well as for the control groups beyond post-treatment, therefore no information is available on the longer term effects of the SST compared to not receiving SST.

	Follow-up	Outcome measures	Conclusion (pre/post treatment)		
			I	II	III
Only intervention group after 3 months		- Child			
		• TASSK-R	p < .01	p < .01	p = .01
		• QPQ (Gu, Con)	ns	ns	ns
		• QPQ (Ho)	ns	ns	p = .04
		• K-SSRS (As,Co,Em,SC,T)	ns	ns	ns
		• CDI	p = .04	p = .03	ns
		• STAIC-T	ns	ns	ns
		• STAIC-S	ns	ns	ns
		• ADOS (Ic-a, Ic-t)	p = .01	p < .01	p < .01
		• ADOS (rsi-a, rsi-t)	p < .01	p < .01	p < .01
		- Parents			
		• SCQ	ns	ns	ns
		• SRS	ns	ns	ns
		• QPQ (Gu, Con)	ns	ns	ns
		• QPQ (Ho)	ns	p = .03	p = .03
		• ASDS (L, SI, BP, CA, SP)	ns	ns	ns
		• K-CBCL An/dep	p = .03	ns	p = .02
		• K-CBCL In	p = .02	ns	p = .03
		• AHWA-VABS socialization	p < .01	p < .01	p < .01
		• BDI (F, M)	ns	ns	ns
		• STAI-T (F, M)	ns	ns	ns
		• STAI-S (M)	p < .01	p = .01	p = .04
		• STAI-S (F)	ns	ns	ns

K-SSRS, Korean Version of the Social Skills Rating System (As, Assertion; Co, Cooperation; Em, Empathy; SC, Self-Control; T, Total score); PDD-BI, Pervasive Developmental Disorders: behavior Inventory; PEI, The Pupil Evaluation Inventory (Wi, Withdrawal; Ag, Aggression); PHS, Piers-Harris Self-Concept Scale; PRI, Perceptual Reasoning Index; PSS, Parent Satisfaction Survey; QPQ, Quality of Play Questionnaire (Con, Conflict; Dis, Disengage; Eng, Engage; Gu, Guest; Ho, Host); SCI, Social Competence Inventory (PSI, Pro-Social Index; SI, Social Initiation Index); SCQ, Social Communication Questionnaire; SKA, Skillstreaming Knowledge Assessment; SRS, Social Responsiveness Scale; SSRS, Social Skills Rating System (As, Assertion; SC, Self-Control; In, Internalizing behavior; Ex, Externalizing behavior); STAIC, State and Trait Anxiety Inventory for Children (T, Trait; S, State; M, Mother; F, Father); TASSK, Test of Adolescent Social Skills Knowledge; TASSK-R, Test of Adolescent Social Skills Knowledge-Revised; TOPS-ER, Test of Problem Solving-Elementary Revised; VCI, Verbal Comprehension Index; VIQ, Verbal IQ.

Another important issue, which was not addressed in the evaluated studies, is to find out whether the learned skills generalize to situations in the child's daily life (Gillies et al., 2013). To assess this aspect, one should measure social skills from everyday life of the child instead of very specific and discrete behaviors learned at the investigated intervention (Koenig et al., 2010). The use of multiple informants (e.g.,

parents and teachers) increases the insight into the extent to which the child uses learned social skills in daily life.

Another aspect understudied so far is the benefit of parental involvement in the SST (Gillies et al., 2013). The clinical impression is that parental involvement in interventions increases generalization of the learned skills. Parents may remind children to practice and to apply their learned skills in various situations in daily life. This may be especially beneficial to children with ASDs who generally have difficulty to learn and change behavioral patterns. A part of the studies investigated an SST with parental involvement (Laugeson et al., 2009; Frankel et al., 2010; Solomon et al., 2004; Lopata et al., 2010; Yoo et al., 2014) but only one study compared an SST with and without parental involvement (DeRosier et al., 2011). However, it was unclear whether the significant differences between the two groups, with respect to skills awareness and motivation were due to parental involvement or to the specificity of the training for ASD, since it compared an SST with parental involvement specifically developed for children with HFA to an SST without parent participation not specifically developed for children with ASD.

The best age when to provide an SST is subject to discussion. In clinical practice, SSTs are provided to younger and older children, adolescents, and adults. In early adolescence and adulthood, lack of social skills can result in peer ridicule and rejection (Rao et al., 2008). Therefore, it seems important to offer an SST in preadolescence, to prevent children from experiencing more impairment, distress, and internalizing problems (White et al., 2007; DeRosier et al., 2011). The younger a child masters social and communicational skills and concepts, the earlier he or she can apply these in daily life. If effective, the child's development may then deviate less than without training. Herbrecht and colleagues (2009) have found in their study that the children's group (8 to 13 years old) benefited more from the training than the adolescents' group (13 to 19 years old). Although the interventions differed in frequency and duration, they interpreted this as a possibility that children are more receptive to intervention than adolescents, because of a higher ongoing natural maturation effect in children or because the psychopathology in children is probably less chronic yet than in adolescents. However, a child can be too young for SSTs as well, because the attention span must be long enough (Herbrecht et al., 2009), the didactic presentations must not be overwhelming to allow benefit (Frankel et al., 2010), and the children must be able to read and write their homework.

The current study investigates the efficacy of a comprehensive manualized SST, specifically developed for children with ASDs, in the last two years of primary school. In the Dutch school system, children are typically 10–12 years old in these final years of primary school. The study is designed as an RCT with three conditions: SST consisting of 18 group sessions; SST (same content, frequency and duration) plus parent and teacher involvement (so called SST-PTI); and care-as-usual. The study

has aimed for a large sample size ($n = 120$) and pays particular attention to generalization of learned skills in daily life. Possible improvement is measured with standardized, internationally used instruments for specific behaviors and broader functioning in daily life from a multi-informant perspective. Follow-up measures will add to the knowledge of the efficacy of SSTs on the longer term.

Research aims and hypothesis

The main aim of the current study is to investigate the efficacy of SST as compared to care-as-usual in 10–12 year old children with ASDs. The efficacy is investigated at several levels: on the level of specific behaviors instructed during the intervention and on the level of general social skills at home and at school. Data are collected from multiple informants: children, parents, teachers, and independent observers. The hypothesis is that children participating in SST show greater improvement on all levels than children in the care as usual condition.

The secondary aim is to investigate the efficacy of an enhanced SST (i.e. SST-PTI, involving parents and teachers) as compared to non-enhanced SST, specifically on the generalization of the learned skills. We expect that generalization increases when parents and teachers learn behavioral therapeutic principles to support the child at home or at school in practicing and implementing social and communicative behaviors.

Additionally, the study aims to investigate factors that possibly influence the effectiveness of SST, searching for information on whether and if so which specific groups of children with ASDs improve more or less with SST.

Methods/design

The efficacy of the SST is investigated in an RCT, including two intervention conditions (SST and SST-PTI) and a care-as-usual condition (Efficacy of Social skills Training In Autism [ESTIA]).

Two Dutch child mental health centers participate in the study: Accare University Center for Child and Adolescent Psychiatry (with locations in the cities Groningen and Drachten) and Lentis Jonx Autism Team North-Netherlands (ATN; also with locations in the cities Groningen and Drachten). All four locations have a regional function. Participants have been recruited through the four locations and all four provide training groups.

Participants

Participants in the study are children with a best estimate clinical DSM-IV-TR diagnosis of ASD, including Autistic disorder, Asperger's disorder, or PDD-NOS, assigned in a multidisciplinary team including a child psychiatrist and psychologist. The diagnostic procedure consisted of an interview with parents on the current behavior and on developmental history of the child and observation of the child in a standardized, playful situation (most often the ADOS). For inclusion in the study, the Autism Diagnostic Interview-Revised (ADI-R, Rutter et al., 2003; Dutch version, de Jonge & de Bildt, 2007) and the Autism Diagnostic Observation Schedule (ADOS, Lord et al., 1999; Dutch version, de Bildt & de Jonge, 2008) were administered. Participants had been referred to one of the four participating child mental health care centers. Their clinician advised them to participate in a social skills training. Motivation of the child and parents for training is established. Their IQ is above 80, they are in the last two and a half years of primary education and 10 to 12 years old.

Children with a physical condition that hampers participation or who cannot arrange visiting the child mental health center for the training have been excluded from participation. However, medication or co-morbid disorders were no reason for exclusion.

Informed consent

When an SST was indicated, the therapist introduced this form of treatment to the parents and children and informed them about the study. Sometimes, parents or children were the ones who expressed interest in participating in a social skills training. In that case, the therapist discussed the appropriateness with parents and child and then introduced the study. If parents and children were interested in participating, they agreed that the research team would call them to give more detailed information and would send written information. When the parents and children understood the information from the researchers and were willing to participate, an appointment was scheduled with a therapist. The therapist assessed whether SST was indeed appropriate for the child and the child and parent had the chance to receive more information on the training. If parents, child, and therapist agreed on the appropriateness of treatment and participation in the RCT, the informed consent form was signed by the parents and the child, if aged 12. The researcher informed the teacher and asked for participation after parents had agreed on doing so and the teacher signed informed consent as well. All participants had the right to withdraw from the study at any time, without explanation. Refraining from the study would not affect regular treatment at the child mental health centers. Figure 1 is a flowchart with the phases of the research procedure.

Figure 1 Flow chart of the phases of the RCT.

Therapists

All therapists participating in this study were psychologists who worked at one of the child mental health centers. Two therapists led one group of four to six children. At least one of the (child) psychologists had finished an additional postmaster clinical training. The other had at least a master in Psychology. Both had experience in the broader treatment of children with ASDs and had received training in the theoretical backgrounds and application of cognitive behavioral therapy. Before applying the SST, all therapists participated in a training in the specific protocol. This training lasted six hours for groups of trainers or two hours individually. Trainers received the protocol with all sessions and preparation guidelines. The training focused on the theoretical background, the manualized protocol, the content of the sessions, the physical context of the sessions, the importance of adherence to the protocol, and to the study requirements in general. During the intervention period, therapists had brief meetings or phone calls to discuss the progress of the children or the difficulties in the ongoing groups. These sessions had also been set up to enhance treatment adherence and to prevent therapists to drift from the protocol. In the beginning of the intervention, these meetings were planned weekly, later they were planned once every two weeks. One of the senior therapists of this specific SST supervised these sessions and carefully watched over treatment adherence.

Randomization

The RCT contains three conditions: SST, SST-PTI, and care-as-usual. As soon as 4–6 children in a treatment center had finished the pretreatment assessment, that group was randomized to one of these three conditions. We used a randomized block design with blocks of 5 with a 2-2-1 ratio (2 SST, 2 SST-PTI, 1 CAU). This balanced randomization procedure was used, striving for unbiased comparison groups, but also for comparison groups of about the same size throughout the trial. The training groups were embedded in three treatment locations that have specific characteristics (Accare Groningen; ATN Groningen; Drachten [including ATN and Accare]). Therefore, we stratified the randomization by the three treatment locations. Research assistants logged in to a web-based program, that allocated groups to treatment condition (randomized on group-level in blocks of 5, stratified per treatment location).

During the inclusion period, the recruiting team was blind to the exact details of the randomization process. To conceal the predictability of the randomization process, research assistants not only entered the treatment location, but also information on the sex of the participants and the number of children in each group (4–6). Sex and group size were in fact not weighed in the randomization.

Randomization took place after all measures of the first assessment were completed. Thus, pre-treatment assessments are independent of the participant's knowledge of treatment condition.

Intervention

Social skills training

The manualized SST consists of 15 weekly sessions (except for school vacations), followed by three booster sessions starting two months after the 15th session. An SST group consists of four to six children and two therapists. Each session lasts 90 minutes and is video-taped. The SST is based on behavioral therapeutic principles and the social learning theory. The purpose of the SST for children is to learn to interact with other children and to experience that interacting with other children can be fun. Children are taught through instruction, directed positive feedback, observation, and role-play. The therapists analyze the behavior of the children, define individual positive target behaviors, and elicit positive behavior. Negative behavior is ignored when possible, while differentially reinforcing alternative or incompatible positive behavior.

As shown in table 2, in sessions 1 to 4, the first phase of the SST, the most important aim is to create a safe environment for the children to practice. In sessions 5 to 15, the second phase, nine skills are discussed, practiced, and rehearsed, for instance 'asking something to someone', or 'apologizing to someone'. Sessions 16 to 18 are the so called booster sessions, in which the most relevant skills for each individual child are rehearsed.

Before the start of the training, one of the therapists meets with the child and his/her parents in order to get to know each other and to give practical information. Additionally, this meeting serves as the moment for determining five individual goals of the participant. The parents, child, and therapist together choose five specific behaviors that the child and therapist will focus on during the training.

In order to evaluate the training, the child, parents, and therapists meet and discuss the training and learned skills after the fifteenth session. After the booster sessions, the therapists call the parents for a final evaluation.

For children in this condition, only medication is allowed as an additional treatment.

Social skills training – parent & teacher involvement

In the SST-PTI condition, each child follows the exact same procedure as described above for the SST condition. Additionally, his/her parent(s) participate in eight parent sessions. Preferably, both parents participated, however, in many cases only one parent was able or willing to participate. Another addition is that the teacher is coached to support the child in school. An overview of the parent sessions and teacher involvement is presented in table 2.

The parent sessions are directly linked to the SST sessions and focus on how to support and assist the child in performing the social skills as learned in the training. The sessions consist of instruction on and explanation of behavioral therapeutic principles, behavioral exercises, role-play, and homework. Three parent sessions take place before the children's SST sessions start. In these sessions, antecedent

Table 2 Topics of the social skills training

Children sessions (in SST and SST-PTI)		
1	Phase 1 (weekly):	Introduction
2	Create a safe environment	Saying nice things about yourself and to others
3		Feelings (showing how you feel and see how another feels)
4		Personal presentation (posture, eye contact and use of voice)
5	Phase 2 (weekly):	Asking something to someone
6	Practice skills	Conversation
7		Asking for a play date
8		Asking to participate
9		Discussing with someone
10		Playing a social game
11		Saying no
12		Indicating annoyance
13		Apologizing to someone
14		Responding to bullying
15		Final session, children chose a social activity
16	Phase 3 (2 weekly-monthly):	Repeating the above mentioned skills, focusing on individual goals
17	Booster sessions	Repeating the above mentioned skills, focusing on individual goals
18		Repeating the above mentioned skills, focusing on individual goals
Parents sessions (in SST-PTI)		
1	Phase 1 (weekly):	Psycho-education
2	Before child sessions	Antecedent interventions
3		Consequent interventions
4 (SST 2)	Phase 2 (2 weekly):	Discrimination training
5 (SST 4)	During child sessions	Eliciting desired behaviors and creating opportunities
6 (SST 6)		Energizing desired social behavior
7 (SST 8)		Responding to and redirecting socially awkward behaviour
8 (SST 11)	Phase 3	Continuation and persevere
Teacher (in SST-PTI)		
One meeting with the therapist before the start of the SST for the children en five telephone contacts during the SST (after session 2, 4, 7, 10 and 13).		

and consequent interventions are discussed. The parent(s) learn to distinguish desirable social behavior from undesirable social behavior and to elicit the desirable behavior. The other five sessions take place at determined moments during the SST and these are focused on how to support their child in learning and practicing social skills, based on the parental skills from the first three sessions.

Before the children start with their SST sessions, the teacher has one meeting with the therapists concerning (further) education on ASDs, explanation of the SST and specific behavioral instructions. The teachers receive a file with all child sessions and the weekly homework. Additionally, one of the therapists is in contact with the teacher five times throughout the 15 weeks that the child participates in the SST sessions, in order to discuss the possibilities for practicing the weekly homework skills in school and the individual goals.

Also for children in this condition, only medication is allowed as an additional treatment.

Care-as-usual

Children in the care-as-usual condition receive no SST. Medication, parent training, and other treatments are allowed and content, frequency, and duration are registered by the researchers. However, parents or child do not receive psychological interventions primarily focusing on improving social skills. Participants in the care-as-usual condition can take part in an SST after the follow-up assessment, one year after start of the SST in the treatment conditions.

Assessments

Assessments were planned at three moments: before randomization, immediately after the intervention period, and at 6 months follow-up after the end of intervention. The first assessment (T1) was planned after inclusion of the child and the decision of parents and child to participate. This assessment comprises questionnaires for children, parents, and teachers, interviews with parents, and observations of children. Six months later (after SST session 15) the second assessment (T2) was planned, including questionnaires for parents, child, and teacher, an interview with parents, and observations of the child. Another six months later the follow-up (T3) was planned, with all instruments included at T2 for parents and child. No teacher information was going to be collected at this point, since children often have a different teacher on T3 than on T1 and T2 (due to change of class in the end of the school year). In the care as usual condition, T2 was planned six months after T1 and T3 twelve months after T1. Currently, T1 and T2 have been completed. The last group will complete T3 in the summer of 2014.

Primary outcome

The primary outcome is adaptive functioning, measured with the Dutch translation of the Vineland Adaptive Behavior Scales - Survey Version (VABS (Sparrow et al., 1984); Dutch Version (de Bildt & Kraijer, 2003); measured at T1, T2, and T3), in order to measure the efficacy of the SST in daily life. Adaptive functioning is defined as the performance of daily living activities that are necessary for personal and social functioning of a person (Sparrow et al., 1984) and reflects how well children are able to function in daily life. The Vineland measures this competence with three subscales, in the manual called domains: 'Communication', 'Daily Living Skills', and 'Socialization'. It is an open-ended interview with one or both of the parents. Because the SST does not focus on the domains 'Daily Living Skills' and 'Communication' of the Vineland, only the domain 'Socialization' will be used in the analyses.

Secondary outcomes parent measures

The Social Skills Rating Scale (SSRS, Gresham & Elliott, 1990); measured at T1, T2, and T3) is a 38-item standardized parent questionnaire that measures social skills pertaining to home and community settings for children in primary education, completed by the primary caregiver. The subscales included in the analyses are 'Cooperation', 'Assertion', 'Responsibility' and 'Self-Control'.

The ESTIA - training specific questionnaire (ESTIA-TS; Unpublished manuscript, 2010; measured at T1, T2, and T3) is a 30-item parent-report questionnaire about the specific social skills children learn during the SST. The questionnaire was developed with the aim to investigate change in the specific behaviors as taught in the SST, such as eye contact, recognizing feelings, and apologizing. Parents report on a scale from 0 to 5 how difficult each skill is for their child. They also report on the frequency of each of the skills.

Secondary outcome teacher measure

The Teacher version of the Social Skills Rating Scale (SSRS, Sparrow et al., 1984); measured at T1 and T2) has 30-items in three subscales: 'Cooperation', 'Assertion', and 'Self-Control'. All three will be included in the analyses.

Other measures

Additionally, possible mediators and moderators of treatment outcome will be assessed, e.g. treatment attendance, genetic factors, severity of ASD, intelligence, symptoms of depression, anxiety, attention deficit, hyperactivity, impulsivity, oppositional behaviors, and parental stress.

Drop-outs

Children in the intervention groups who dropped out after the start of the SST were encouraged to still participate in the assessments. Post-treatment assessment was organized as soon as the child dropped out of treatment, if the drop out was before session 8. For the follow-up assessment the regular schedule has been followed. Children who dropped out after session 8 were invited for the regular post-treatment assessments.

Sample size calculation

We computed the minimally required sample sizes for two repeated-measures (RM-) ANOVAs on the primary outcome measure VABS, with one between-subject factor (group; 2 levels) and one within-subject factor (time; 2 levels: T3 versus T1); in one test, the SST and care-as-usual groups will be compared and in the other the SST and SST-PTI groups. For each test, we required a power of .99, a significance level of .01, and assumed a correlation between repeated measures of the VABS of .5. Calculations were performed with the program G*Power 3.1.7 (Faul et al., 2007).

The expected difference between the SST and the care-as-usual group is based on the results of the study into SST of Owens and colleagues (2008), who found a difference between the treated and untreated group of 0.90 SD of the mean VABS outcome. Assuming no effect in the untreated group, the associated effect size $f = .4$, yielding a minimal sample size of 21 per group. We expect that comparing the two treatment conditions yields lower values of difference. No literature on the comparison of two SST training conditions was available, therefore we defined the difference to detect based on clinical relevance of a difference found. This resulted in a difference between SST and SST-PTI of 0.60 SD of the mean VABS outcome, with associated effect size $f = .3$, yielding a minimal sample size of 36 per group. Based on our experiences in past research, we presumed a drop-out rate of 10%, implying a minimal sample size of $21 / .9 = 24$ children in the care-as-usual condition and $36 / .9 = 40$ in each of the treatment conditions. Because we could include more children in the treatment groups, yielding an even larger power, we aimed 48 in each of the treatment conditions and 24 children in the care-as-usual group (total $n = 120$).

Statistical analyses

First, the three groups will be compared with respect to background, age, IQ, and other data in order to assess the comparability between them.

Second, intervention efficacy will be established with hierarchical linear modeling. All data will be analyzed using the intent-to-treat principle. The comparative efficacies of SST, SST-PTI, and care as usual will be investigated with hierarchical linear modeling. The first will test the primary outcome, the socialization domain of the VABS. The other multilevel analyses will be done on each secondary outcome,

including 1) parent reported training-specific social skills and 2) more general social skills pertaining to home and community settings, reported by parents and 3) more general social skills pertaining to home and community settings, reported by teacher.

Third, in the completers, the same analyses will be applied, while taking into account possible drop-out and treatment non-adherence. The possible mediators and moderators will be included as explanatory variables into the hierarchical linear models.

Discussion

The current study investigates the efficacy of SST for children with ASDs in the age of 10–12 years. Focusing on the limitations in knowledge so far, the study also investigates generalization of learned skills, long term efficacy, and the influence of parent and teacher involvement in an RCT in a non-US group of 120 children.

Strengths and limitations

The random assignment of children to one of the treatment conditions or the care-as-usual condition is an important strength of the study. With this design, the two treatment conditions can be compared to the care-as-usual condition and to each other. A detected improvement of efficacy can thus be controlled for development over time. A second important strength of the study is that this is the first study on SST to include a follow-up assessment one year after start of the training in the treated as well as in the control condition. With this follow-up, the long term effect of the SST can be examined. Another strength is the fact that many measures are incorporated in the study, amongst which standardized, international instruments. This may give insight into efficacy at various levels: from specific behaviours and skills to general social functioning in school. Furthermore, if one or both of the interventions show to be effective, the measures may help evaluate why some children improve and others improve less or not at all. Additionally, as recommended in previous studies (Rao et al., 2008; White et al., 2007; Reichow et al., 2013) the current study has a large sample size (120 participants), in a homogeneous age range, uses a manualized protocol, takes place outside the United States, and involves multiple informants (parents, teacher and child).

A limitation of the study is that only the first assessment could be blinded. During the interventions (or care-as-usual) and at the second and third assessment all therapists, most interviewers, some observers, all children, all parents, and all teachers knew the treatment condition. Future research should aim at blinded interviewers and observers, however, this is very difficult in the setting of the current study as it will be in many other settings.

Implications for practice

Many parents ask for an SST for their child when they have received an ASD diagnosis and have learned about symptoms and forms of treatment. Additionally, many health care institutions provide SSTs because clinical impression indicates that such a treatment is valuable for a child with ASD. Due to the time-consuming character of SSTs for children, parents, and therapists it is very important to evaluate whether these trainings are effective and if so, for whom.

Additionally, in some of the SSTs parents and teachers are involved on top of the child sessions. In that case, they also invest their time, so information is needed on the added value of their participation for the efficacy of an SST and generalization of learned skills. Before the start of the current study, SSTs were already given in the participating child mental health centers, so the results of this study can be implemented directly.

Trial status

Participant recruitment started in May 2010 and was finished in September 2013. Currently, in May 2014, all participants have been randomized and completed assessments at T1 and T2. The measurements in the last groups for T3 will take place in summer 2014.

Ethical approval

The Medical Ethical Committee of the University Medical Center in Groningen has assigned ethical approval for the study (METC nr 2009.320).

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This study has been funded by the Netherlands Organization for Health Research and Development (ZonMw, nr 157003005).

Barbara van den Hoofdakker, Lianne van der Veen, Sjoukje van Warners, and Leonieke Vet have developed the intervention protocols used in the study.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington DC: American Psychiatry Publishing.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: American Psychiatry Publishing.
- Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet*, 368(9531), 210–215.
- Beaumont, R., & Sofronoff, K. (2008). A multi-component social skills intervention for children with Asperger syndrome: the Junior Detective Training Program. *Journal of Child Psychology and Psychiatry*, 49(7), 743–753.
- de Bildt, A., & de Jonge, M.V. (2008). *Autisme diagnostisch observatie schema*. Amsterdam: Hogrefe.
- de Bildt, A.A., & Kraijer D.M. (2003). *Sociale redzaamheid voor kinderen en jeugdigen met een verstandelijke beperking*. Vineland-Z. Leiden: PITS B.V.
- de Jonge, M.V., & de Bildt, A. (2007). *Autisme diagnostisch interview-revised*. Amsterdam: Hogrefe.
- DeRosier, M.E., Swick, D.C., Davis, N.O., McMillen, J.S., & Matthews, R. (2011). The efficacy of a social skills group intervention for improving social behaviors in children with high functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 41(8), 1033–1043.
- Devlin, B., & Scherer, S.W. (2012). Genetic architecture in autism spectrum disorder. *Current Opinion in Genetics & Development*, 22(3), 229–237.
- Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Forness, S.R., & Kavale, K.A. (1996). Treating Social Skill Deficits in Children with Learning Disabilities: A Meta-Analysis of the Research. *Learning Disability Quarterly*, 19(1), 2–13.
- Frankel, F., Myatt, R., Sugar, C., Whitham, C., Gorospe, C.M., & Laugeson, E. (2010). A randomized controlled study of parent-assisted Children's Friendship Training with children having autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 40(7), 827–842.
- Freitag, C.M., Cholemkery, H., Elsuni, L., Kroeger, A.K., Bender, S., Kunz, C.U., & Kieser, M. (2013). The group-based social skills training SOSTA-FRA in children and adolescents with high functioning autism spectrum disorder – study protocol of the randomised, multi-centre controlled SOSTA – net trial. *Trials*, 14, 6.
- Gillies, D., Carroll, L., & Loos, M. (2013). Commentary on 'Social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD)'. *Evidence-Based Child Health*, 8(2), 316–317.
- Gresham, F.M., & Elliott, S.N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Herbrecht, E., Poustka, F., Birnkammer, S., Duketis, E., Schliitt, S., Schmötzer, G., & Bölte, S. (2009). Pilot evaluation of the Frankfurt social skills training for children and adolescents with autism spectrum disorder. *European Child & Adolescent Psychiatry*, 18(6), 327–335.
- Karkhanavah, M., Clark, B., Ospina, M.B., Seida, J.C., Smith, V., & Hartling, L. (2010). Social Stories™ to improve social skills in children with autism spectrum disorder: a systematic review. *Autism*, 14(6), 641–662.
- Koenig, K., White, S.W., Pachler, M., Lau, M., Lewis, M., Klin, A., & Scahill, L. (2010). Promoting social skills development in children with pervasive developmental disorders: a feasibility and efficacy study. *Journal of Autism and Developmental Disorders*, 40(10), 1209–1218.
- Laugeson, E.A., Frankel, F., Mogil, C., & Dillon, A.R. (2009). Parent-assisted social skills training to improve friendships in teens with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39(4), 596–606.
- Lauritsen M.B. (2013). Autism spectrum disorders. *European Child & Adolescent Psychiatry*, 22 Suppl 1, S37–S42.
- Lopata, C., Thomeer, M.L., Volker, M.A., Toomey, J.A., Nida, R.E., Lee, G.K., Smerbeck, A.M., & Rodgers, J.D. (2010). RCT of a manualized social treatment for high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 40(11), 1297–1310.
- Lord, C., Rutter M., DiLavore, P.C., & Risi, S. (1999). *ADOS. Autism Diagnostic Observation Schedule. Manual*. Los Angeles: Western Psychological Services.

- Lord, C., Wagner, A., Rogers, S., Szatmari, P., Aman, M., Charman, T., Dawson, G., Durand, V.M., Grossman, L., Guthrie, D., Harris, S., Kasari, C., Marcus, L., Murphy, S., Odom, S., Pickles, A., Scahill, L., Shaw, E., Siegel, B., Sigman, M., ... Yoder, P. (2005). Challenges in evaluating psychosocial interventions for Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 35(6), 695–708.
- Owens, G., Granader, Y., Humphrey, A., & Baron-Cohen, S. (2008). LEGO therapy and the social use of language programme: an evaluation of two social skills interventions for children with high functioning autism and Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 38(10), 1944–1957.
- Rao, P.A., Beidel, D.C., & Murray, M.J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: a review and recommendations. *Journal of Autism and Developmental Disorders*, 38(2), 353–361.
- Reichow, B., Steiner, A.M., & Volkmar, F. (2013). Cochrane review: social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD). *Evidence-Based Child Health*, 8(2), 266–315.
- Reichow, B., & Volkmar, F.R. (2010). Social skills interventions for individuals with autism: evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, 40(2), 149–166.
- Rutter, M., le Couteur, A., & Lord, C. (2003). *ADI-R Autism Diagnostic Interview Revised Manual*. Los Angeles: Western Psychological Services.
- Solomon, M., Goodlin-Jones, B.L., & Anders, T.F. (2004). A social adjustment enhancement intervention for high functioning autism, Asperger's syndrome, and pervasive developmental disorder NOS. *Journal of Autism and Developmental Disorders*, 34(6), 649–668.
- Sparrow, S.S., Balla, D., & Cicchetti, D.V. (1984). *Vineland adaptive behavior scales*. (Survey ed.). Circle Pines, MN: American Guidance Service.
- Spence, S.H., Donovan, C., & Brechman-Toussaint, M. (2000). The treatment of childhood social phobia: the effectiveness of a social skills training-based, cognitive-behavioural intervention, with and without parental involvement. *Journal of Child Psychology and Psychiatry*, 41(6), 713–726.
- Volkmar, F.R., Paul, R., Kin, A., & Cohen, D. (2005). *Handbook of Autism and Pervasive Developmental Disorders* (3rd ed.) Hoboken, NJ: John Wiley and Sons.
- White, S.W., Koenig, K., & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: a review of the intervention research. *Journal of Autism and Developmental Disorders*, 37(10), 1858–1868.
- Yoo, H.J., Bahn, G., Cho, I.H., Kim, E.K., Kim, J.H., Min, J.W., Lee, W.H., Seo, J.S., Jun, S.S., Bong, G., Cho, S., Shin, M.S., Kim, B.N., Kim, J.W., Park, S., & Laugeson, E.A. (2014). A randomized controlled trial of the Korean version of the PEERS parent-assisted social skills training program for teens with ASD. *Autism Research*, 7(1), 145–161.

3 |

Social skills group training in children with autism spectrum disorder: a randomized controlled trial

Vera Dekker

Maaïke H. Nauta

Marieke E. Timmerman

Erik J. Mulder

Lianne van der Veen-Mulders

Barbara J. van den Hoofdakker

Sjoukje van Warners

Leonieke J.J. Vet

Pieter J. Hoekstra

Annelies de Bildt

Abstract

In 122 high-functioning children with autism spectrum disorder (ASD; 9-13 years; 19 girls), we investigated the effectiveness of a 15-session social skills group training (SST) with and without parent and teacher involvement (PTI) in a randomized controlled trial with three conditions: SST ($n=47$), SST-PTI ($n=51$), and care-as-usual (CAU, $n=24$). Hierarchical linear modeling was used for immediate and 6-month follow-up analyses. Measures were administered before randomization (blind), post-treatment and at follow-up (not blind). Trial registration: Dutch Trial Register; <http://www.trialregister.nl>; NTR2405.

At post-treatment, children in both SSTs had improved significantly more than CAU on the primary outcome, Vineland Socialization (SST: Cohen's $d = 0.39$; 95% CI -2.23-3.11 and SST-PTI: $d = 0.43$; 95% CI -2.19-3.15) and on the secondary outcome parent-SSRS "Cooperation" (SST: $d = 0.43$; 95% CI -0.23-1.15 and SST-PTI: $d = 0.45$; 95% CI -0.21-1.17), with no difference between post-treatment and follow-up. Additionally, children in SST-PTI improved significantly more on the teacher-SSRS than in CAU ["Cooperation" $d = 0.42$ [95% CI -.33-1.13]; "Assertion" $d = 0.34$ [95% CI -.39-1.11]; "Self-Control" $d = 0.61$ [95% CI -.08-1.34]) and in SST ("Cooperation" $d = 0.34$ [95% CI -.37-1.05]; "Self-Control" $d = 0.59$ [95% CI -.13-1.32]).

The current study corroborates earlier findings in smaller samples and wider age ranges, with small but statistically significant effects of SST for higher-functioning pre-adolescent children with ASD. Parental and teacher involvement intensified treatment, yet did not yield an additional effect relative to SST for children only, as reported by parents. Six months after training, no further improvement or decline was found.

Introduction

To improve social-communicative skills in children with autism spectrum disorders (ASD), group-based Social Skills Trainings (SSTs) are widely provided in clinical practice. The short-term effectiveness of SST has been demonstrated in two recent meta-analyses (Gates et al., 2017; Wolstencroft et al., 2018). Effect sizes varied for different sources of treatment evaluation. Parents and external observers generally reported small effects of training, teachers reported no effect, and children and adolescents reported large effects, all compared to no-treatment or waiting-list conditions. The latter effects mainly reflect improvement in social knowledge rather than actual social behavior. Besides informants, the exact instruments used seem to affect the effect, with moderate to large parent reported effects in the meta-analysis (Wolstencroft et al., 2018) on the Social Skills Rating System (SSRS; Gresham & Elliott, 1990), and Social Responsiveness Scale (SRS; Constantino & Gruber, 2012).

One of the ultimate aims of SST in ASD is to improve social skills beyond the duration of the training. Since longer term follow-up data are often lacking, the authors of recent meta-analyses could not draw conclusions regarding the long term effects (Gates et al., 2017; Wolstencroft et al., 2018). For studies with three-months follow-up assessments for treated and non-treated groups, the outcomes varied: the immediate effect of group-based SST reported by Soorya and colleagues (2015) did not sustain at three months' follow-up, whereas Freitag and colleagues (2016) and Deckers and colleagues (2016) reported a significant effect at three months' follow-up. Choque Olsson and colleagues (2017) did not find an effect of training in children, and the effect found for adolescents was shown immediately post treatment only, and not at three months follow-up.

Another important aim of SST is to improve social skills in situations beyond the training situation, i.e. in daily life. Research into SSTs could not draw conclusions about how to reach generalization of skills to real life (Jonsson et al., 2016). One way to improve the generalization of children's social skills may be involving parents and/or teachers in the SST intervention, reasoning that they can directly support children to put their training into practice in daily life. Wolstencroft and colleagues (2018) showed that all children in an SST improved, independent from parental involvement, yet with a large effect size in the group with parental involvement and moderate in the group without. However, due to other differences between the studies (e.g. participant characteristics, measures used), direct comparisons are complicated.

The current study (Efficacy of Social skills Training In Autism [ESTIA]) is a randomized controlled trial (RCT) into the effectiveness of a manualized group SST with and without parental and teacher involvement for high-functioning children with ASD in the last two and a half years of primary education, in a large and well-characterized sample of 122 children, with a six-month follow-up in all conditions.

We aimed to investigate two main questions: 1) what is the immediate and long-term effect of group-based SST for high functioning children with ASD compared to no training (care as usual, CAU) based on social skills applied in school and in home-based daily life situations? and 2) what is the additional immediate and long-term effect of parental and teacher involvement on generalization of social skills in daily life of children with ASD compared to group-based SST for the children only?

We hypothesized that 1) children in an SST would improve more on measures of social skills compared to children without an SST; 2) children with SST and additional parent and teacher involvement (SST-PTI) would improve more on measures of social skills, compared to children without this support (SST); 3) children with additional support (SST-PTI) would better maintain these skills after training, compared to children without such support (SST).

Methods

Design

The RCT had three conditions: group SST only, group SST-PTI, and care-as-usual without SST. We collected measures at three time points: pre-treatment (T1), immediately post-treatment or after the same amount of time in CAU (T2), and follow-up six months post-treatment or after the same amount of time in CAU (T3). We could only collect teacher information at T1 and T2, due to change of class (and teacher) with changing school year. The first measurement took place before randomization, at later measurements informants or interviewers were not blind to condition. Before participation, parents, teacher, and child (if aged 12) signed an informed consent. The study followed the CONSORT guidelines for RCTs (Schulz et al., 2010; see Supplementary Table 1 for CONSORT checklist). The Institutional Review Board of the University Medical Center Groningen had approved the study. The study was registered in the Dutch Trial Register (NTR2405; <http://www.trialregister.nl>). Figure 1 shows the flowchart of study recruitment, treatment allocation, and assessment. For a detailed description, we refer to the research protocol (Dekker et al., 2014).

Participants

Participants were 122 preadolescent high-functioning children with ASD from one of four outpatient mental health care clinics in the northern Netherlands (103 boys, 19 girls). The inclusion criteria were 1) clinical DSM-IV-TR ASD diagnosis (Autistic disorder, Asperger's disorder, or Pervasive Developmental Disorder-Not Otherwise Specified [PDD-NOS]), based on thorough diagnostic procedures (developmental history, current problems, child observation, and information from school) in expert teams including at least a child psychologist and a child psychiatrist; 2) the child's

Figure 1 CONSORT diagram of study recruitment, treatment allocation, and assessment.

^a two children participated in less than half of the child sessions; ^b two children participated in less than half of the child sessions, two parents participated in less than half of the parent sessions (one from the same family as one of the children who participated in less than half of the child sessions), 1 teacher did not participate.

clinician indicated SST as first appropriate treatment; 3) parents and child were motivated for SST, as established during a meeting with the clinician, parents and child; 4) preferably IQ $\geq$ 80. Children with IQs slightly below 80 were included when therapists established they could follow an SST; 5) being in the last two and half years of primary education; 6) no physical condition affecting participation; and 7) the child could travel to the child mental health center for training. Note that three of the criteria

Table 1 Baseline participant characteristics ($n = 122$), including observed sample mean, standard deviation, range and sample size, per condition and per outcome measure.

		SST ($n = 47$; 87% male)		SST-PTI ($n = 51$; 78% male)		CAU ($n = 24$; 92% male)		Statistics
		mean (SD)	Range	mean (SD)	Range	mean (SD)	Range	
Age	Years	10.9 (0.7)	9.9-12.6	10.9 (0.8)	9.6-12.7	11.2 (0.9)	9.8-13.0	
ADOS	Social Affect	8.7 (4.5)	2-20	7.9 (3.8)	0-20	8.6 (3.3)	3-15	0.12
	Restricted and Repetitive behavior	1.0 (0.9)	0-30	1.3 (1.1)	0-50	1.1 (1.2)	0-40	0.58
	Calibrated Severity Score	5.6 (2.4)	1-10	5.4 (2.3)	1-10	5.7 (2.1)	2-90	0.38
ADI-R	Social Interaction	15.0 (6.1)	4-27	13.7 (5.6)	3-26	13.7 (4.5)	2-22	0.85
	Communication	12.5 (4.9)	2-23	11.3 (4.4)	3-21	12.2 (4.6)	0-19	0.46
	Restricted and Repetitive behavior	3.0 (2.1)	0-8	3.4 (2.1)	0-10	3.0 (2.2)	0-8	0.44
Vineland	Communication	111.4 (8.0)	95-126	112.1 (8.6)	88-129	115.3 (8.1)	94-128	0.70
	Daily Living Skills	120.2 (14.7)	70-145	121.2 (13.7)	93-149	125.6 (13.1)	98-150	0.19
	Socialization	79.5 (12.9)	53-107	82.4 (16.1)	26-118	86.8 (14.8)	52-115	0.29
ESTIA-TS	Training-specific social skills	73.0 (15.4)	42-106	68.6 (12.8)	47-102	74.4 (12.5)	53-103	0.14
SSRS-P	Cooperation	7.5 (4.1)	1-17	8.2 (3.2)	2-13	7.8 (4.2)	2-16	0.15
	Assertion	9.8 (3.0)	4-19	10.6 (3.1)	3-19	10.3 (2.6)	4-15	0.67
	Self-Control	7.2 (3.5)	0-16	7.7 (3.3)	1-16	7.5 (3.5)	3-13	0.40
SSRS-T	Responsibility	9.7 (3.6)	0-18	10.9 (3.2)	3-17	10.7 (2.4)	7-17	0.70
	Cooperation	13.5 (4.0)	4-20	13.0 (4.7)	2-20	12.4 (4.5)	7-20	0.20
	Assertion	9.2 (4.1)	0-16	8.2 (4.0)	1-17	8.5 (4.2)	0-17	0.59
IQ	Self-Control	10.5 (4.1)	3-20	9.0 (4.3)	2-18	9.0 (3.7)	4-18	0.52
		102.5 (14.8)	72-135	98.7 (16.4)	73-132	105.6 (13.1)	73-126	0.15
								0.17

ADOS, Autism Diagnostic Observation Schedule; CAU, care-as-usual; ESTIA-TS, Efficacy of Social skills Training In Autism - Training Specific; SSRS-P, Social Skills Rating Scale-Parents; SSRS-T, Social Skills Rating Scale-Teacher; SST, Social Skills Training; SST-PTI, Social Skills Training - parent & teacher involvement; Vineland, Vineland Adaptive Behavior Scales

were slightly different from the original design registered in the trial register, to more closely approximate the regular decisions in clinical practice. The three original inclusion criteria were 1) ASD diagnosis was either supported by an Autism classification on the Autism Diagnostic Interview-Revised (ADI-R) or maximally two points below the cut off for Autism but with an ASD classification on the ADOS, 2) IQ $\geq$ 80, and 3) being in the last two years of primary education.

Seventeen participants had a DSM-IV-TR diagnosis of autistic disorder (14%), 25 of Asperger's disorder (20%), and 80 of PDD-NOS (66%). The conditions did not differ on ASD diagnosis (Pearson χ^2 .45; p =.978). Of all children, 32% had one comorbid diagnosis, 4.1% had two comorbid diagnoses (24 Attention Deficit Hyperactivity Disorder, 19.7% of all participants; 8 Tic disorder, 6.5%; 4 Anxiety Disorder, 3.2%; 4 Oppositional Defiant Disorder, 3.2%; 4 other, 3.2%). The conditions did not differ on comorbid secondary diagnoses (Pearson χ^2 .39.62; p =.166) or tertiary diagnoses (Pearson χ^2 12.61; p =.246). Mean age at start was 11 years (SD = 0.75; range 9.5-13.0), mean total IQ was 101.5 (SD = 15.3; range 72-135). Male sex proportion was similar over the conditions (Pearson χ^2 2.63; p =.268). The conditions did not differ on psychotropic medication use between start and post-treatment. Between post-treatment and follow-up, more children in the CAU condition (22.7%) used anti-psychotic medication compared to SST (6.7%) and SST-PTI (4.4%; Pearson χ^2 6.55; p =.038). Most children had at least one parent of Dutch decent (n =121, including 88 with two Dutch parents). Table 1 presents the participant characteristics at baseline.

Interventions

SST was manualized, based on behavioral therapeutic principles and the social learning theory (van Warners, Vet, van der Veen-Mulders & van den Hoofdakker, 2010; internal publication). The training had 15 weekly 90 minutes basic group sessions and three additional 90 minutes booster group sessions, planned between two and six months after the 15th session. Each session followed a structure: conversation, homework review, introducing a new topic, practice and role-play, new homework, and play-time. Children received a workbook with summaries of the trained skills and homework. The goal of the first four sessions was to create a safe environment. Sessions 5 through 15 covered specific topics, e.g. "asking something to someone", "responding to bullying". A full overview of the session topics can be found in Supplementary Table 2. Children received training through instruction, directed positive feedback, observation, role-play, and homework. The therapists analyzed the behavior of the children, defined individual positive target behaviors and elicited positive behavior. Negative behavior was ignored when possible, while differentially reinforcing alternative or incompatible positive behavior. In the booster sessions, children rehearsed their individual target behaviors. The SST groups

consisted of four to six children, led by two therapists, i.e., psychologists with at least a psychology master. The therapists received training in the SST by behavioral therapists before and supervision during SST to increase treatment integrity. We refer to the research protocol (Dekker et al., 2014) for a more comprehensive description of the procedures for therapists.

The SST-PTI condition consisted of the SST, with additional parent and teacher involvement (van Warners & Vet, 2010; internal publication), aiming to enhance generalization of learned social skills. Parents participated in three group sessions before and five during SST. The first three sessions covered psycho-education and interventions for enhancing desired behaviors. The other sessions were related to the SST sessions and focused on how to support the child in practicing the trained social skills. Parent sessions included instruction, behavioral exercises, role-play, and practicing learned skills at home.

Teacher support was provided through teacher-therapist meetings before the SST, to discuss the training and the skills aimed to address. During SST, the teacher had five telephone meetings with the therapist to discuss opportunities to support the child in practicing skills at school.

Participants in the CAU condition did not receive SST, defined as a manualized, child-specific training or program. Parent counseling, not focused on social skills, was allowed, e.g. psycho-education, counseling for family functioning, support in finding the right school or medication control, depending on the need of each participant. The conditions did not differ in the CAU delivered. Sessions with parents, with the child, or with parents and child, telephone contact, medication control, personalized support at home or school, a special program in school, or other help or support were equally present in all conditions, as reported by parents. After follow-up, children from CAU could enroll in SST.

In all conditions, delivery of SST outside the study was monitored.

Treatment fidelity and adherence

After each session, therapists rated whether they had addressed each component of the session. Therapists adhered to the protocol in 97.6% of the time in the SST basic sessions, 97.5% in the parent sessions, 93.3% in the initial teacher session, and 86.5% in the children's booster sessions.

Baseline assessment

The severity of ASD symptomatology was measured with the Autism Diagnostic Interview-Revised (ADI-R; Rutter et al., 2003) and the Autism Diagnostic Observation Schedule (ADOS) module 3 (Lord et al., 1999; Lord et al., 2012), by trained psychologists who met research requirements for reliability. Cognitive ability was assessed with the Wechsler Intelligence Scale for Children – 3d edition (WISC-III; Wechsler, 1999).

Outcome Measures

Our primary outcome, reflecting the main aim of training and evaluating the effectiveness of SST in daily life, was the level of social functioning as measured with the raw total “Socialization” domain scores of the Vineland Adaptive Behavior Scales – Survey version (Vineland; Sparrow et al., 1984). As secondary outcome measure we evaluated the specific social skills trained during SST with the ESTIA training-specific skills (ESTIA-TS; Vet et al., 2010, unpublished questionnaire). This 30-item parent questionnaire evaluated the difficulty for the child of performing the specific skills taught in the training. We also evaluated the frequency of general social skills in home and school situations as reported by parents and teachers with the subscales “Cooperation”, “Assertion”, “Self-control”, and “Responsibility” of the 38-item Social Skills Rating Scale parent-version (SSRS-P) and corresponding subscales “Cooperation”, “Assertion”, and “Self-control” of the 30-item teacher-version (SSRS-T; Gresham & Elliott, 1990).

Randomization

From September 2010 through September 2013, training groups started (September and February). All children finished the training within one schoolyear. When four to six participants were included in the study in a setting, they were randomized into one of the conditions as a group, in a 2:2:1 ratio (SST:SST-PTI:CAU), after the first assessment. Randomization was done in blocks of five groups per stratum, based on setting, using a computer-generated list of treatment allocations, performed by research assistants, unaware of the randomization algorithm and unable to access the computer-generated list to conceal the sequence of allocation.

We aimed to include 48 children in both SSTs and 24 children in CAU, based on a required power of .99 and a significance level of .01 on the primary outcome measure Vineland for the time effect (i.e., from T1 to T3) and the differential time effects between conditions, taking into account a supposed drop-out rate of 10%. The presumed effect sizes across time in the comparison between SST and CAU groups were based on previous research with the Vineland as outcome measure (Owens et al., 2008). For the comparison between the SST and SST-PTI groups, we considered an effect size of 0.60 SD of the mean to be clinically relevant, of .60 SD of the mean. Due to the variation in defined differences between the compared groups, a larger amount of children was needed for comparing SST and SST-PTI than for comparing SST to CAU. For a detailed description of the sample size calculation, we refer to the research protocol (Dekker et al., 2014). Due to differences in group size (four to six participants), the treatment conditions differed in size (SST $n=47$, ten groups groups, SST-PTI $n=51$, ten groups).

Statistical analyses

First, we tested differences in baseline background variables, age, IQ, and the outcome measures using analyses of variance (ANOVAs) with allocated treatment as independent variable. Second, we evaluated the effectiveness of the treatment conditions compared to CAU with hierarchical linear modeling using the intent-to-treat principle, including all available data points. We built separate models for the Vineland and the subscales of the SSRS-P, SSRS-T, and ESTIA-TS. The structure of the model consisted of three levels: the measurement occasions (level 1) were nested in the participants (level 2) which were nested in the treatment group (level 3). To assess the treatment effects across time, we built a fully multivariate model for each subscale of each outcome measure. The model accounted for effects of measurement occasion (T1, T2, and T3), condition (SST, SST-PTI, and CAU), and its interaction, while only keeping fixed effects that we hypothesized to be present, or that proved significant. We hypothesized time effects for the treatment conditions (SST and SST-PTI) and their interaction, and therefore kept those effects. Furthermore, we expected no difference between conditions at T1, because of randomization, and we expected no improvement on the outcome measures for the children in the CAU condition, thus we only preserved those effects when significant. Dummy coding was used to model the effects of measurement occasion and treatment, taking T1 and the CAU condition, respectively, as reference conditions. We tested the statistical significance of the fixed effects using the approximate *t* test. The exception was testing the differences between the two treatment groups; this was done using the deviance test, comparing the models mentioned above with a model including both treatment groups together in a single category, and contrasted to the CAU. Additionally, we calculated pre-post and post-follow-up effect sizes (ES, Cohen's *d*) for the significant condition differences. ES were derived from differences between conditions at time points, based on the estimated fully multivariate model. In all tests the significance level was $\alpha = 0.05$.

Results

At baseline, the three conditions did not differ significantly on any measure. Therefore we excluded those effects from the model (see table 1). In building the models, the level 3 random effects (i.e., referring to the treatment group) explained only part of the variance for the subscales "Cooperation" and "Self-control" of the SSRS-T, not for the other measures. Therefore we only included level 3 in the hierarchical linear model for these subscales.

Table 2 presents the estimated coefficients, significance levels, and standard errors of the multilevel models for the parent measures. Table 3 presents these for the teacher measure. The random effects of the models are available upon request from

Table 2 Parameter estimates of the multilevel models of the parent measurements (Vineland, SSRS-P, ESTIA-TS).

Fixed effects	Vineland SOC	ESTIA-TS	SSRS-P COO	SSRS-P ASS	SSRS-P SCO	SSRS-P RES
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
Intercept (mean score T1 CAU)	82.6 (1.3) ^a	71.4 (1.3) ^a	7.9 (0.3) ^a	10.3 (0.3) ^a	7.5 (0.3) ^a	10.4 (0.3) ^a
Contrast T1-T2 (CAU)	-	-7.9 (2.1) ^a	-	1.6 (0.5) ^b	1.3 (0.6) ^c	1.8 (0.5) ^a
Contrast T2-T3 (CAU)	-	-	-	-	-	-
SST at T1	-	-	-	-	-	-
SST-PTI at T1	-	-	-	-	-	-
Contrast T1-T2 x SST	5.8 (1.7) ^b	1.2 (2.5)	1.7 (0.4) ^a	-0.5 (0.7)	1.2 (0.7)	-0.6 (0.6)
Contrast T1-T2 x SST-PTI	6.4 (1.7) ^a	2.4 (2.5)	1.7 (0.4) ^a	0.3 (0.7)	1.3 (0.7)	0.0 (0.6)
Contrast T2-T3 x SST	0.6 (1.7)	0.3 (1.4)	0.1 (0.3)	-0.4 (0.4)	-0.5 (0.4)	0.6 (0.4)
Contrast T2-T3 x SST-PTI	2.3 (1.7)	-2.5 (1.4)	-0.4 (0.3)	-0.2 (0.4)	0.2 (0.4)	0.5 (0.4)

^a $p < .001$; ^b $p < .01$; ^c $p < .05$;

CAU, care-as-usual; ESTIA-TS, Efficacy of Social Skills Training In Autism - Training Specific; SSRS-P, Social Skills Rating Scale-Parents (COO, Cooperation; ASS, Assertion; SCO, Self-Control; RES, Responsibility); SST, Social Skills Training; SST-PTI, Social Skills Training - parent & teacher involvement; Vineland, Vineland Adaptive Behavior Scales (SOC, Socialization); - means that this effect was not included in the model, because it was not hypothesized and appeared nonsignificant when including the effect

the first author. The actual scores on the instruments in all conditions can be found in the Supplementary Tables S3 and S4.

Table 3 Parameter estimates of the multilevel models of the teacher measurements (SSRS-T).

	SSRS-T COO	SSRS-T ASS	SSRS-T SCO
Fixed effects	Estimate (SE)	Estimate (SE)	Estimate (SE)
Intercept (mean score T1 CAU)	13.1 (0.5) ^a	8.8 (0.4) ^a	9.6 (0.5) ^a
Contrast T1-T2 (CAU)	-	-	-
SST at T1	-	-	-
SST-PTI at T1	-	-	-
Contrast T1-T2 x SST	0.4 (0.5)	0.2 (0.4)	0.4 (0.5)
Contrast T1-T2 x SST-PTI	1.7 (0.5) ^{a, d}	1.4 (0.4) ^b	2.4 (0.5) ^{a, d}

^a $p < .001$; ^b $p < .01$; ^c $p < .05$; ^d also a significant difference with SST condition; CAU, care-as-usual; SSRS-T, Social Skills Rating Scale-Teachers (COO, Cooperation; ASS, Assertion; SCO, Self-Control); SST, Social Skills Training; SST-PTI, Social Skills Training – parent & teacher involvement; - means that this effect was not included in the model, because it was not hypothesized and appeared nonsignificant when including the effect.

Primary Outcome

Vineland socialization

Children in the SST and the SST-PTI conditions improved significantly more on Vineland Socialization than children in CAU (who did not significantly improve) from T1 to T2 (ES SST: Cohen's $d = 0.39$; 95% CI -2.23-3.11 and ES SST-PTI: Cohen's $d = 0.43$; 95% CI -2.19-3.15). SST and SST-PTI did not differ significantly from each other. All conditions showed stable Vineland Socialization scores from T2 to T3. The trajectory of each condition is presented in figure 2.1.

Secondary outcomes

ESTIA-TS

As shown in figure 2.2, parents in CAU reported significantly lower difficulty scores for the social skills covered in the training at T2 compared to T1 (ES: Cohen's $d = -0.56$; 95% CI -3.02-2.16). From T2 to T3, no significant difference existed for CAU. Parents in both treatment conditions reported a similar pattern from T1 to T2 or T2 to T3.

Figure 2 Expected mean per time point and per condition, as based on the fixed part of the multilevel models of the measurements for parents and teachers

Note Parents: Effect sizes in the figures represent significant changes in the specified condition (by symbol), compared to CAU. Significant effect sizes were only present for the comparison between T1 and T2.

Teachers: Effect sizes in the figures represent significant changes in the social skills training with parent and teacher involvement, compared to CAU (diamond) or compared to SST (square).

SSRS-P

Figures 2.3-2.6 show change as measured with the SSRS. Children in both treatment conditions improved significantly more on "Cooperation" than children in CAU (who did not improve) from T1 to T2 (ES SST: Cohen's $d = 0.43$; 95% CI $-0.23-1.15$ and ES SST-PTI: Cohen's $d = 0.45$; 95% CI $-.21-1.17$). They did not differ significantly from each other. From T2 to T3, "Cooperation" scores were stable for all conditions.

On the subscales "Assertion", "Self-control", and "Responsibility", children in CAU improved significantly from T1 to T2. The ES (Cohen's d) were 0.51 (95% CI $-.02-1.12$), 0.40 (95% CI $-.2-1.01$), and 0.56 (95% CI $-.01-1.17$). Children in the CAU condition did not improve significantly from T2 to T3 on these subscales. Both treatment conditions showed similar patterns as CAU.

SSRS-T

Children in CAU did not improve significantly on the SSRS-T subscales from T1 to T2 (figures 2.7-2.9). Children in the SST condition resembled the CAU condition. Children in the SST-PTI condition improved significantly more between T1 and T2 than CAU, with ES (Cohen's d) 0.42 (95% CI $-.33-1.13$) for "Cooperation", 0.34 (95% CI $-.39-1.11$) for "Assertion", and 0.61 (95% CI $-.08-1.34$) for "Self-Control". On the subscale "Cooperation" and "Self-Control" the children in SST-PTI improved also significantly more than the SST condition, with ES (Cohen's d) 0.34 (95% CI $-.37-1.05$) and 0.59 (95% CI $-.13-1.32$), respectively.

Discussion

This study demonstrated that children improved in social functioning in daily life and broad social skills, reported by parents, immediately after group SST. However, no differences existed between the three conditions on the specifically trained social skills and the other SSRS-P subscales. Contrary to our expectations, actively involving parents and teachers in the training did not increase the immediate effect, nor the generalization of social skills to situations outside the training or beyond the duration of the training, observed by parents. Six months after training, social skills had not further improved in any group. Note that adding three booster sessions (Soorya et al., 2015; Koning et al., 2013) did not contribute to further improvement of social skills either.

The small to medium effect sizes in our study correspond to the findings of Gates and colleagues (2017) in their meta-analysis. They also resemble the results from Deckers and colleagues (2016) in a comparable, Dutch, high-functioning population from a regular outpatient clinic ($n=52$; effect size .34), although measured with a different instrument (the Social Skills Observation [SSO]; Barry et al., 2003).

Compared to the meta-analysis of Wolstencroft and colleagues (2018), we found a smaller effect on the SSRS, and only for the subscale “Cooperation”. Perhaps this difference in effect size is due to the significant improvement of children in the CAU condition in the current study on the subscales “Assertion”, “Self-control”, and “Responsibility”. Although this could reflect natural growth in time, other explanations are also possible. First, parents may have changed in their observation, e.g. the assessment may have included behaviors that parents had not noticed before the first measurement, yet actively looked for after completing the SSRS for the first time. Second, all children in this study, including the children in CAU, wanted to improve social skills and were motivated for training by definition of the inclusion criteria. Parents of children in the CAU condition may thus have tried to improve their child’s social skills in other ways (e.g. reading about social skills/ASD, explaining social situations, stimulating their child to make play dates). Third, the many children with PDD-NOS (66%) and the relatively high-functioning character of the sample, may have affected the outcomes.

Our study is one of few that compared long-term outcomes in the treated and CAU conditions (Gates et al., 2017). Two earlier studies reported a significant effect after three months (Freitag et al., 2016; Deckers et al., 2016), no effect was found after three months in two other studies (Soorya et al., 2015; Choque Olsson et al., 2017). We expected children in the SST-PTI condition to continue improving, based on their parents’ training in how to support them learn and apply social skills. However, this support did not seem to affect further development after training. Practicing may have evanesced after training was over, with its accompanying homework. Moreover, social skills appeal to social insight, and the question is whether that can be trained in an (15-session) SST.

Teachers only reported significant improvement after SST-PTI compared to CAU (all subscales) and SST (“Cooperation” and “Self-control” subscales of the SSRS-T). This finding could indicate an effect of parent and teacher involvement in the training on generalization of learned skills to the school situation. Alternatively, teachers more intensively involved may see improvement for other reasons, e.g. they may have learned how to observe or interpret social behaviors, or they may expect change after their effort in the training, conform the expectancy bias in parents (Gates et al., 2017; Wolstencroft et al., 2018). Although in line with the teacher findings of Deckers and colleagues (2016), our finding is in contrast with the meta-analysis of Gates and colleagues (2017), who found no effect of SST as reported by teachers. Since we have no follow-up data of the teachers, we cannot draw conclusions on generalization to the school situation beyond the duration of the training.

Limitations

Unfortunately, parents and teachers were not blind for condition in the post-treatment measures. A blind observation of the child in a naturalistic situation (Rao et al., 2008; McMahon et al., 2013; Dekker et al., 2016), added to parent and teacher report, would have contributed to a more reliable measure of changes in social skills after training, unbiased by either knowledge of treatment condition or actual contributions to training. However, no well-described and valid instruments were available for such observation. Gates and colleagues (2017) found only five studies in their meta-analysis reporting on observer information. Even in this small sample all studies used other instruments, ranging from the ADOS, measuring social-communication in a semi-structured situation with one adult, to 5-minute observations of mother-child interaction in a structured situation (playing a puzzle), and to peer interactions measured with different instruments.

Additionally, we only used general social skills measures, including social functioning in daily life, and did not focus on ASD specific social skills. For future studies, we recommend to add ASD-specific social skills measures. We are aware that the Vineland was not developed as an outcome measure for SST. However, using it as our primary measure was an approach to assessing improvements in social functioning in daily life; the ultimate aim of SST. Earlier SST studies including the Vineland did not report strong effects either, i.e., only trend significant changes (Owens et al., 2008), no differential changes (Koning et al., 2013), or mixed results on various subscales (Yoo et al., 2014).

Implications

The current study corroborates earlier findings in smaller samples and wider age ranges, indicating small but statistically significant effects of SST in daily life for higher-functioning pre-adolescent children with ASD. Parental and teacher involvement intensified the treatment for therapists, parents and teachers, yet did not yield the expected additional effect relative to SST for children only as reported by parents. More research on who benefits from what form of SST is needed, in order to enable clinicians to decide who to provide with (what form of) SST.

Disclosure of potential conflict of interest

Author AdB received the grant that funded this study (the Netherlands Organization for Health Research and Development; ZonMw, nr 157003005). The agency had no role in data analyses, the decision to publish, or the preparation of the manuscript. She also is first author on the Dutch ADOS manual for which Accare receives enumeration. SvW, LV, BvdH and LvdV were authors of the SST-manual, and LV and SvW were authors of the SST-PTI-manual, both used in the current study. The manuals are not commercially available, hence the authors do not have a

financial interest in the outcomes of the current study. The other authors declare that they have no conflict of interest.

Acknowledgements

The Dutch manual is available upon request from the first author. This work was funded by the Netherlands Organization for Health Research and Development (ZonMw, nr 157003005). The authors gratefully acknowledge all children, parents, teachers, and therapists participating in this study.

References

- Barry, T.D., Klinger, L.G., Lee, J.M., Palardy, N., Gilmore, T., & Bodin, S.D. (2003). Examining the effectiveness of an outpatient clinic-based social skills group for high-functioning children with autism. *Journal of Autism and Developmental Disorders*, 33(6), 685–701.
- Choque Olsson, N., Flygare, O., Coco, C., Görling, A., Råde, A., Chen, Q., Lindstedt, K., Berggren, S., Serlachius, E., Tammimies, K., Kjellin, L., & Bölte, S. (2017). Social Skills Training for Children and Adolescents With Autism Spectrum Disorder: A Randomized Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(7), 585–592.
- Constantino J.N., & Gruber C.P. (2012). *Social Responsiveness Scale (SRS-2)*. Torrance, CA: Western Psychological Services.
- Deckers, A., Muris, P., Roelofs, J., & Arntz, A. (2016). A Group-Administered social Skills Training for 8- to 12-Year-Old, high-Functioning Children With Autism Spectrum Disorders: An Evaluation of its Effectiveness in a Naturalistic Outpatient Treatment Setting. *Journal of Autism and Developmental Disorders*, 46(11), 3493–3504.
- Dekker, V., Nauta, M.H., Mulder, E.J., Sytema, S., & de Bildt, A. (2016). A Fresh Pair of Eyes: A Blind Observation Method for Evaluating Social Skills of Children with ASD in a Naturalistic Peer Situation in School. *Journal of Autism and Developmental Disorders*, 46(9), 2890–2904.
- Dekker, V., Nauta, M.H., Mulder, E.J., Timmerman, M.E., & de Bildt, A. (2014). A randomized controlled study of a social skills training for preadolescent children with autism spectrum disorders: generalization of skills by training parents and teachers?. *BMC psychiatry*, 14, 189.
- Freitag, C.M., Jensen, K., Elsuni, L., Sachse, M., Herpertz-Dahlmann, B., Schulte-Rüther, M., Hänig, S., von Gontard, A., Poustka, L., Schad-Hansjosten, T., Wenzl, C., Sinzig, J., Taurines, R., Geißler, J., Kieser, M., & Cholemkery, H. (2016). Group-based cognitive behavioral psychotherapy for children and adolescents with ASD: the randomized, multicentre, controlled SOSTA-net trial. *Journal of Child Psychology and Psychiatry*, 57(5), 596–605.
- Gates, J.A., Kang, E., & Lerner, M.D. (2017). Efficacy of group social skills interventions for youth with autism spectrum disorder: A systematic review and meta-analysis. *Clinical Psychology Review*, 52, 164–181.
- Gresham, F.M., & Elliott, S.N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Jonsson, U., Choque Olsson, N., & Bölte, S. (2016). Can findings from randomized controlled trials of social skills training in autism spectrum disorder be generalized? The neglected dimension of external validity. *Autism*, 20(3), 295–305.
- Koning, C., Magill-Evans, J., Volden, J., & Dick, B. (2013). Efficacy of cognitive behavior therapy-based social skills intervention for school-aged boys with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 7(10), 1282–1290.
- Lord, C., Rutter M., DiLavore, P.C., & Risi, S. (1999). *ADOS. Autism Diagnostic Observation Schedule. Manual*. Los Angeles: Western Psychological Services.
- Lord, C., Rutter, M., DiLavore, P.C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, second edition (ADOS-2) manual (Part I): Modules 1-4*. Torrance, CA: Western Psychological Services.
- McMahon, C.M., Lerner, M.D., & Britton, N. (2013). Group-based social skills interventions for adolescents with higher-functioning autism spectrum disorder: a review and looking to the future. *Adolescent Health, Medicine and Therapeutics*, 4, 23–38.
- Owens, G., Granader, Y., Humphrey, A., & Baron-Cohen, S. (2008). LEGO therapy and the social use of language programme: an evaluation of two social skills interventions for children with high functioning autism and Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 38(10), 1944–1957.
- Rao, P.A., Beidel, D.C., & Murray, M.J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: a review and recommendations. *Journal of Autism and Developmental Disorders*, 38(2), 353–361.

- Rutter, M., le Couteur, A., & Lord, C. (2003). *ADI-R Autism Diagnostic Interview Revised Manual*. Los Angeles: Western Psychological Services.
- Schulz, K.F., Altman, D.G., Moher, D. (2010). CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*, 63(8), 834-840.
- Soorya, L.V., Siper, P.M., Beck, T., Soffes, S., Halpern, D., Gorenstein, M., Kolevzon, A., Buxbaum, J., & Wang, A.T. (2015). Randomized comparative trial of a social cognitive social skills group for children with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 54(3), 208–216.
- Sparrow, S.S., Balla, D., & Cicchetti, D.V. (1984). *Vineland adaptive behavior scales* (Survey ed.). Circle Pines, MN: American Guidance Service.
- Wechsler, D. (1999). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Wolstencroft, J., Robinson, L., Srinivasan, R., Kerry, E., Mandy, W., & Skuse, D. (2018). A Systematic Review of Group Social Skills Interventions, and Meta-analysis of Outcomes, for Children with High Functioning ASD. *Journal of Autism and Developmental Disorders*, 48(7), 2293–2307.
- Yoo, H.J., Bahn, G., Cho, I.H., Kim, E.K., Kim, J.H., Min, J.W., Lee, W.H., Seo, J.S., Jun, S.S., Bong, G., Cho, S., Shin, M.S., Kim, B.N., Kim, J.W., Park, S., & Laugeson, E.A. (2014). A randomized controlled trial of the Korean version of the PEERS parent-assisted social skills training program for teens with ASD. *Autism research*, 7(1), 145–161.

Supplementary Table 1 CONSORT checklist.

CONSORT 2010 checklist of information to include when reporting a randomised trial*			
Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	p.47 (p. 48/55)
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	p.48
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	p.49-50
	2b	Specific objectives or hypotheses	p.49-50
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	p.50,55
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	p.51,53
Participants	4a	Eligibility criteria for participants	p.50,51,53
	4b	Settings and locations where the data were collected	p.50
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	p.53,54
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	p.55
	6b	Any changes to trial outcomes after the trial commenced, with reasons	NA
Sample size	7a	How sample size was determined	p.55
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	p.55
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	p.55

Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	p.55		
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	p.50,55		
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	NA		
	11b	If relevant, description of the similarity of interventions	p.53,54		
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	p.56		
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	p.56		
Results					
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	Figure 1, p.51		
	13b	For each group, losses and exclusions after randomisation, together with reasons	Figure 1, p.51		
Recruitment	14a	Dates defining the periods of recruitment and follow-up	p.50,55		
	14b	Why the trial ended or was stopped	p.55		
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1, p.52		
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Figure 1, p.51		
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Tables 2 and 3; p.57,58 supplementary tables S3 and S4; p.70,71		
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	NA		
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	NA		
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	NA		

Supplementary Table 1 Continued.

CONSORT 2010 checklist of information to include when reporting a randomised trial*			
Section/Topic	Item No	Checklist item	Reported on page No
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	p.62
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	p.62
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	p.60-62
Other information			
Registration	23	Registration number and name of trial registry	p.48,50
Protocol	24	Where the full trial protocol can be accessed, if available	Dutch Trial Register; http://www.trialregister.nl ; NTR2405; Dekker et al., 2014
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	p.62,63

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

Supplementary Table 2 Topics of the social skills training.

Children sessions (in SST and SST-PTI)		
1	Phase 1 (weekly): Create a safe environment	Introduction
2		Saying nice things about yourself and to others
3		Feelings (showing how you feel and see how another feels)
4		Personal presentation (posture, eye contact and use of voice)
5	Phase 2 (weekly): Practice skills	Asking something to someone
6		Conversation
7		Asking for a play date
8		Asking to participate
9		Discussing with someone
10		Playing a social game
11		Saying no
12		Indicating annoyance
13		Apologizing to someone
14		Responding to bullying
15		Final session, children chose a social activity
16	Phase 3 (2 weekly-monthly): Booster sessions	Repeating the above mentioned skills, focusing on individual goals
17		Repeating the above mentioned skills, focusing on individual goals
18		Repeating the above mentioned skills, focusing on individual goals
Parents sessions (in SST-PTI)		
1	Phase 1 (weekly):	Psycho-education
2	Before child sessions	Antecedent interventions
3		Consequent interventions
4 (SST 2)	Phase 2 (2 weekly): During child sessions	Discrimination training
5 (SST 4)		Eliciting desired behaviors and creating opportunities
6 (SST 6)		Energizing desired social behavior
7 (SST 8)		Responding to and redirecting socially awkward behaviour
8 (SST 11)	Phase 3	Continuation and persevere
Teacher (in SST-PTI)		
One meeting with the therapist before the start of the SST for the children en five telephone contacts during the SST (after session 2, 4, 7, 10 and 13).		

Supplementary table 3 Per outcome measure and condition, the observed sample mean, standard deviation, range and sample size at T2

		SST			SST-PTI			CAU		
		mean (SD)	Range	n	mean (SD)	Range	n	mean (SD)	Range	n
Age	Years	11.5 (0.6)	10.5-12.9	47	11.4 (0.8)	10.1-13.2	47	11.8 (0.9)	10.3-13.7	22
ESTIA-TS	Training-specific social skills	65.8 (16.3)	34-98	46	63.3 (13.6)	38-94	45	66.8 (14.2)	44-106	22
SSRS-P	Cooperation	9.3 (4.2)	1-19	47	9.7 (3.6)	4-18	46	8.9 (4.1)	2-16	22
	Assertion	11.1 (3.6)	0-19	47	12.3 (3.1)	6-20	46	11.8 (3.6)	6-19	22
	Self-Control	9.9 (3.5)	2-19	47	10.2 (2.8)	3-16	46	9.4 (4.2)	0-16	22
	Responsibility	11.2 (3.7)	2-19	47	12.5 (2.9)	6-18	46	12 (3.4)	6-17	22
SSRS-T	Cooperation	13.7 (4.1)	4-20	46	14.7 (4.4)	2-20	47	13.3 (4.0)	5-20	23
	Assertion	9.2 (4.1)	2-19	46	9.7 (4.3)	3-20	47	10.0 (4.1)	1-20	23
	Self-Control	10.7 (4.5)	1-19	46	11.4 (4.2)	2-20	47	9.3(4.0)	3-18	23
Vineland	Socialization domain	86.3 (15.8)	47-114	47	88.7 (15.0)	34-122	47	88.6 (12.5)	58-105	22

Supplementary table 4 Per outcome measure and condition, the observed sample mean, standard deviation, range and sample size at T3

		SST			SST-PTI			CAU		
		mean (SD)	Range	n	mean (SD)	Range	n	mean (SD)	Range	n
Age	Years	11.9 (0.7)	10.9-13.6	45	11.9 (0.8)	10.5-13.9	45	12.3 (0.8)	10.8-14.0	22
ESTIA-TS	Training-specific social skills	67.3 (16.8)	36-101	44	61.5 (12.7)	38-85	44	65.9 (15.1)	42-96	20
SSRS-P	Cooperation	9.3 (3.8)	2-18	44	9.3 (3.6)	3-17	44	8.5 (4.0)	2-16	20
	Assertion	10.6 (3.7)	4-18	44	12.3 (3.0)	6-19	44	12.3 (3.9)	5-20	20
	Self-Control	9.5 (3.8)	2-19	44	10.4 (2.9)	5-17	44	8.5 (3.7)	2-13	20
	Responsibility	11.9 (3.2)	7-19	44	13.1 (3.0)	5-18	44	13.0 (3.6)	8-19	20
Vineland	Socialization domain	86.8 (16.5)	49-115	45	91.5 (15.1)	41-120	45	89.8 (16.0)	61-120	22

4 |

Application of Latent Class Analysis to identify subgroups of children with Autism Spectrum Disorders who benefit from Social Skills Training

Vera Dekker

Maaïke H. Nauta

Marieke E. Timmerman

Erik J. Mulder

Pieter J. Hoekstra

Annelies de Bildt

Abstract

With Latent Class Analysis applied on data of 98 children with autism spectrum disorder (ASD) (9-12 years; 17 girls) participating in social skills training (SST) in a randomized controlled trial (Dekker et al., 2019), four subgroups were detected, based on social-communicative skills before, and response patterns to training. Two subgroups improved after SST. Characterizing the subgroups based on participant and intervention characteristics showed that improvement was related to lower parent-reported perceived difficulty of social-communicative skills at start, higher verbal ability, younger age and milder symptoms of ASD and anxiety. The lowest performing non-improving subgroup participated more often in SST without parent/teacher involvement, compared to all other subgroups. Response to SST in ASD seems to vary depending on participant characteristics.

Introduction

Since limitations in social communication and interaction are core characteristics of an Autism Spectrum Disorder (ASD), many of the treatment approaches for ASD focus on social communicative skills, often in the form of Social Skills Training (SST). Many studies have been conducted into the effect of SST for children and adolescents with ASD and overall, on group-level, moderate to large effect sizes have been found (Gates et al., 2017; Wolstencroft et al., 2018). Much less is known on whether specific subgroups of participants exist that benefit more from specific (characteristics of) SST. The current study aimed to contribute to this subject by investigating whether subgroups can be identified of participants who respond differently to SST with or without parent and teacher involvement, and by relating these subgroups to multiple dimensions such as participant and intervention characteristics.

The authors of two recent meta-analyses that investigated the effect of SST on group level (Gates et al., 2017; Wolstencroft et al., 2018) recommended to further investigate the benefit of specific forms of SST for specific individuals. To this end, they investigated the effect of SST in relation to several *intervention characteristics*. Further, Gates and colleagues (2017) also investigated *participant characteristics* in relation to effect. In their meta-analysis they could only analyze the effects on the self report measures, because of too little heterogeneity in the studies with parent report, and a too small number of studies with teacher report included.

With respect to the *intervention characteristics*, Wolstencroft and colleagues (2018) found that *parental involvement* appeared to have a surplus value, with large effect sizes for SST with parental involvement, and moderate ones for SST without parental involvement. Furthermore, *intensity* was related to effect size: an intensive SST, in the form of a summer camp, had a large effect size, compared to a moderate effect size for weekly sessions. Last, *duration* affected the outcome of SST in ASD, with a large effect size for SSTs that lasted over 40 hours, compared to a moderate effect size of SSTs of 40 hours or less. The effect of duration was corroborated in a more recent study (Jonsson et al., 2019). Gates and colleagues (2017) found no effect of intervention length. They found no effect of involving *peer tutors* in SST either.

Several studies investigated *participant characteristics* in relation to the effectiveness of SST in ASD in earlier studies. Consistent effects have been found for gender and cognitive ability, that is, larger effects have been found for females than males (McMahon et al., 2013; Choque Olsson et al., 2017). However, as Gates and colleagues (2017) indicated, many studies included too few female participants to reliably investigate the effect of gender on the outcome of SST. Further, participants with higher IQ's benefited more from SST than participants with lower IQ's (Herbrecht et al., 2009). It is important to note that most studies into SST in ASD were conducted with relatively high-functioning participants. Inconsistent effects have been found for

age, comorbidity and medication. That is, older children and adolescents benefited more from SST in some studies (Mathur et al., 1998; Herbrecht et al., 2009; Choque Olsson et al., 2017), while younger children did in others (Wang et al., 2011; McMahon et al., 2013). Further, comorbid ADHD has been found to decrease the effect of SST in children with ASD (Antshel et al., 2011), or to have no effect (Deckers et al., 2016). Comorbid anxiety has been reported to increase the effect of SST in ASD (Antshel et al., 2011), to decrease the effect (Pellecchia et al., 2016), or to have no effect on outcome (Deckers et al., 2016). Finally, children on medication profited more from SST in one study (Herbrecht et al., 2009), whereas children without medication did better in another study (Frankel et al., 2007). In the meta-analysis of moderation effects of Gates and colleagues (2017), none of the participant characteristics were reported to significantly affect the outcome of SST in ASD when based on self report. The authors emphasized that their findings do not mean that these factors are not related to outcomes. Based on information from other raters (parent, teacher, external observer) these variables may play a role in the outcome.

It is important to note that all relations found in the studies described were based on group level effects in relation to specific characteristics. This increases our insight into factors that may affect how much children benefit from SST, yet it overlooks possible specific response patterns to SST in relation to multiple characteristics. Identifying subgroups of children with ASD with similar response patterns to SST, and investigating the factors that characterize these subgroups will help us understand which specific subgroups benefit from SST and which do not benefit or do so to a lesser extent.

To our knowledge, no studies are available that investigated whether subgroups of children with ASD can be defined that respond differently to SST. Amongst young children with ASD, research has been conducted on identifying subgroups and their development (Stevens et al., 2000; Kim et al., 2016; Paynter et al., 2018). This research has been conducted with different approaches. The first of these studies (Stevens et al., 2000) defined two subgroups of children with ASD at school-age, i.e. a higher functioning group and a lower functioning group, based on social behavior, language and cognitive functioning. As a next step, they investigated which characteristics, as measured during pre-school, were related to group-membership at school-age. They found that pre-school non-verbal cognition, language and the social domain of adaptive functioning were most related to group-membership, with the lowest functioning group at school-age having the lowest scores at pre-school. Kim and colleagues (2016) identified four subgroups of toddlers with ASD, a higher functioning one, a lower functioning one and two groups in between, based on their levels of social communication, rigid repetitive behaviors, nonverbal and verbal skills and adaptive functioning at one point. Their next step was to compare how these groups developed over time. The four groups differed in the stability of ASD symptoms,

adaptive functioning, and verbal cognitive functioning over time. The lowest functioning group gained the least or even decreased in these areas. Paynter and colleagues (2018) used the development of toddlers over time to define their subgroups. They found two subgroups of toddlers with ASD based on their different patterns of response to early intervention over time: a subgroup that responded with change and a subgroup that did not change so much. The subgroups differed in cognitive, verbal and adaptive functioning, with the lowest skills in the low change group. As pointed out by Paynter and colleagues (2018), it is important to investigate the factors that are associated with high and low change after intervention, in order to be able to adjust intervention for those who may not respond, and to identify the factors that may help to predict who will respond to a specific intervention.

With the current study we aimed to increase the insight into patterns of response to SST of children with ASD, and their relation with multiple characteristics. Therefore, we applied a Latent Class Analysis (LCA) to a randomized controlled trial (RCT) dataset on the effectiveness of group SST with and without parental and teacher involvement, compared to each other and to care as usual (CAU) (Efficacy of Social skills Training In Autism [ESTIA]; Dekker et al., 2019). The RCT was conducted in 122 high-functioning pre-adolescents with ASD, and involved pre, post, and six months follow-up measures in all conditions (Dekker et al., 2014; Dekker et al., 2019). The primary analyses focused on the effect of the intervention on group level. Small to moderate parent reported effects were found on some of the outcome measures (Vineland socialization and SSRS-parents; cooperation). On these measures children in both SST conditions improved more than children in the CAU from pre to post treatment (Vineland socialization SST ES = 0.39 and SST-PTIES = 0.43; SSRS-parents SST ES = 0.43 and SST-PTI ES = 0.45). Outcome of the SST conditions did not differ from each other. On the other subscales of the parent SSRS or the specifically trained social skills, no difference between the three conditions was found. Predictors were not included in the primary analyses.

In this next step, we combined the approaches as used by Kim and colleagues (2016) and Paynter and colleagues (2018). We aimed to identify subgroups of participants based on 1) their parent reported social communicative skills before SST and 2) their patterns of response to SST over time (measured immediately and six months after SST). We did so using Latent Class Analysis, resulting in subgroups of individuals with the same patterns of response to SST. We used three measures of social communicative skills, which cover slightly different aspects, thereby allowing for detecting clinically relevant subgroups that are defined by higher-order concepts which only become visible when all characteristics are combined in one analysis. Additionally, we investigated whether and how the response patterns in the subgroups could be related to multiple participant and intervention characteristics. We examined relevant participant characteristics from the literature on SST (age, gender, verbal IQ,

symptoms of ADHD, and symptoms of total and social anxiety), and from studies on change over development (verbal IQ, severity of ASD symptomatology). Assuming we would find subgroups of children with ASD with different response patterns to SST, we hypothesized that being a girl (McMahon et al., 2013; Choque Olsson et al., 2017), having a higher verbal IQ (Herbrecht et al., 2009) and lower ASD symptoms (Kim et al., 2016; Paynter et al., 2018) would be related to a more 'successful' response pattern. Regarding age, ADHD and social anxiety, the results have been too inconsistent to form a hypothesis. As intervention characteristic, we examined the presence or absence of involvement of parents and teachers.

Methods

Design

The study was based on data from an RCT on the effectiveness of group SST with and without parental and teacher involvement, with three conditions: Social Skills Training (SST; $n=47$), Social Skills Training with Parent and Teacher Involvement (SST-PTI; $n=51$) and CAU ($n=24$) (Dekker et al., 2014; Dekker et al., 2019). From September 2010 through September 2013, training groups had started (September and February), data collection therefore started in May 2010 (first pre test) and ended in October 2014 (last follow-up measurement).

The current study only included the children in the treatment conditions ($n=98$; 47 SST and 51 SST-PTI), in order to be able to investigate their progress after SST. Three measurements were conducted: before randomization (T1), immediately after the SST (T2), and six months after the end of SST (T3). Before participation, all parents and children above 12 signed an informed consent. The study followed CONSORT guidelines for RCTs, was approved by the Institutional Review Board of the University Medical Center Groningen, and was registered in the Dutch Trial Register (NTR2405; <http://www.trialregister.nl>). For a detailed description of study recruitment and treatment allocation, we refer to the research protocol (Dekker et al., 2014) and the report on the findings of the primary analyses (Dekker et al., 2019).

Participants

The participants of the ESTIA-study ($n=122$) were recruited in four outpatient mental health care clinics in the northern part of the Netherlands. The current study only included the 98 participants who participated in the intervention ($n=24$ were in the control condition). All were preadolescent high-functioning children with a clinical diagnosis of ASD (81 boys, 17 girls). As reported in Dekker and colleagues (2019) all participants met the following inclusion criteria, which were slightly adjusted from the original design registered in the trial register: 1) pre-existing clinician-based

DSM-IV-TR ASD diagnosis (Autistic disorder, Asperger's disorder, or Pervasive Developmental Disorder-Not Otherwise Specified [PDD-NOS]), based on developmental history, current problems, child observation (ADOS), and information from school in expert teams including at least a child psychologist and a child psychiatrist (original criterion: ASD diagnosis either supported by an Autism classification on the Autism Diagnostic Interview-Revised (ADI-R) or maximally two points below the cut off for Autism on the ADI-R but with an ASD classification on the ADOS); 2) the child's clinician indicated SST as the first appropriate treatment due to social problems in school and other contexts; 3) parents and child were motivated for SST, as established during a meeting with the clinician, the parents, and the child; 4) preferably $IQ \geq 80$ (original criterion: $IQ \geq 80$); children with IQs slightly below 80 were included when therapists established they were able to follow an SST; 5) being in the last two and half years of primary education (original criterion: being in the last two years of primary education); 6) no physical condition affecting participation; and 7) the child could travel to the child mental health center for training. The original criteria were broadened to more closely approximate the regular decisions in clinical practice. As a result of the change in the first criterion, the clinical diagnoses in the sample had not been corroborated with classifications on standardized instruments. This indicates that the included sample functioned at or towards the less autistic end of the spectrum.

Of the 98 participants in the intervention, 65 had a DSM-IV-TR diagnosis of PDD-NOS (66%), 20 of Asperger's disorder (20%), and 13 of autistic disorder (13%). ASD diagnoses did not differ between the two SST conditions (Pearson $\chi^2 .25$; $p = .881$). One comorbid diagnosis was established in 32.7% of the children, 3.1% had two comorbid diagnoses (20 ADHD, 20.4% of all participants; 7 a tic disorder, 7.1%; 3 an Anxiety Disorder, 3.1%; 4 an Oppositional Defiant Disorder, 4.1%; 2 other, 2%). The SST conditions did not differ in comorbid secondary diagnoses (Pearson $\chi^2 15.19$; $p = .296$) or tertiary diagnoses (Pearson $\chi^2 2.94$; $p = .402$). At the start of the SST the mean age was 10.9 years ($SD = 0.7$; range 9.5-12.7). Most children had two Dutch parents ($n = 72$), all had at least one Dutch parent. The characteristics of the participants at baseline are presented in Table 1.

Intervention

Children in both training conditions participated in a manualized SST, based on behavioral therapeutic principles and the social learning theory (van Warners, Vet, van der Veen-Mulders & van den Hoofdakker, 2010; internal publication), in groups of 4-6 children, led by two therapists. The training consisted of 18 sessions of 90 minutes. The first 15 were weekly sessions, followed by three booster sessions after two months, to maintain trained social skills. During the sessions, the children learned how they could perform social skills, based on behavioral therapeutic principles and the social learning theory. The therapists created a safe situation for

Table 1 Baseline participant characteristics ($n = 98$, 83% male)

		mean (SD)	Range
Age	Years ^a	10.9 (0.7)	9.5-12.7
ADOS	Social Affect ^a	8.3 (4.2)	0-20
	Restricted and Repetitive behavior ^a	1.2 (1.0)	0-5
	Calibrated Severity Score ^a	5.5 (2.3)	1-10
ADI-R	Social Interaction ^a	14.3 (5.9)	3-27
	Communication ^a	11.9 (4.7)	2-23
	Restricted and Repetitive behavior ^a	3.2 (2.1)	0-10
	Total score ^a	31.2 (11.0)	8-55
ESTIA-TS	Training-specific social skills ^a	70.7 (14.2)	42-106
SSRS-P	Total ^a	35.9 (10.3)	10-59
Vineland	Socialization ^a	81.0 (14.7)	26-118
IQ	Verbal IQ ^a	102.9 (16.1)	72-145
	Performal IQ ^a	98.0 (16.9)	60-139
	Total IQ ^a	100.5 (15.7)	72-135
RCADS-C	Social Phobia ^a	7.3 (4.0)	0-16
RCADS-C	Total anxiety score ^a	21.6 (11.5)	2-54
SNAP-IV-P	ADHD Inattention ^a	12.6 (6.6)	1-26
	ADHD hyperactivity/impulsivity ^a	9.3 (5.8)	0-27

ADHD, Attention Deficit Hyperactivity Disorder; ADI-R, Autism Diagnostic Interview – Revised; ADOS, Autism Diagnostic Observation Schedule; ESTIA-TS, Efficacy of Social skills Training In Autism - Training Specific; RCADS-C, Revised Child Anxiety and Depression Scale-Children; SNAP-IV-P, Swanson, Nolan and Pelham Questionnaire-Parents; SSRS-P, Social Skills Rating Scale-Parents; Vineland, Vineland Adaptive Behavior Scales

^a no difference between SST and SST-PTI

the children so that they could experience that social interaction can be enjoyable. The sessions had a recurring structure: conversation, homework review, introducing a new topic, practice and role-play, new homework, and play-time. In the first four sessions the focus was on creating a safe environment for the children. In the other sessions specific topics were discussed, e.g. “asking something to someone”, “responding to bullying”.

In the SST with Parental and Teacher Involvement (SST-PTI; van Warners & Vet, 2010; internal publication) parents and teachers of the 51 children were additionally involved in the training, as opposed to SST only. Parents received 8 additional parent sessions, specifically related to the SST. Additionally, therapists met with teachers at the start of SST-PTI, followed by five telephone meetings during SST-PTI. The aim

was to teach parents and teachers how to support children in practicing social skills in daily life. Therapists used instruction, behavioral exercises, and role-play. This involvement was added to advance the generalization of social skills to situations beyond the training. A more comprehensive description of the SST and SST-PTI can be found elsewhere (Dekker et al., 2014; Dekker et al., 2019).

Outcome measures

We included three parent measures as indicators of social communicative skills, pertaining to the application and the perceived difficulty of social skills by their child, at three moments: before randomization (i.e. before SST; T1), immediately after SST (T2), and six months after the end of SST (T3). The measures used in this study rate social performance rather than social competence.

Vineland Adaptive Behavior Scales

The Vineland Adaptive Behavior Scales – Survey version (Vineland; Sparrow et al., 1984) is a semi-structured parent interview. We used the “Socialization” domain (66 items), because this domain corresponded most with the social skills in the SST, in its Dutch version (de Bildt & Kraijer, 2003). Raw scores were used (range 0-132), since no Dutch normed scores are available for children with IQ's above 70. In order to clinically interpret raw scores on the Vineland, estimated age equivalents had been developed (de Bildt & Kraijer, 2003), and these will be used in the current study to interpret change. The psychometric properties of the Vineland are considered good (Sparrow et al., 1984; de Bildt & Kraijer, 2003).

Social Skills Rating System

The Social Skills Rating System (SSRS, Gresham & Elliott, 1990; van der Oord et al., 2005) is a standardized questionnaire about general social skills in home situations. Parents rated the frequency of behavior on a 3-point Likert scale. We used the total raw score of the 38-item parent-version (SSRS-P), since no Dutch normed scores are available. The total raw score may range from 0 to 76. For interpretation of change, we applied the American norm scores for boys in Elementary schools. The psychometric properties of the SSRS are good (Gresham & Elliott, 1990).

ESTIA-Training Specific

The ESTIA-Training Specific (ESTIA-TS; Vet et al., 2010; unpublished questionnaire) is a parent questionnaire about the difficulty of the specific social skills trained during the SST for the child. It consists of 30 training specific social skills that are taught explicitly (e.g., ‘recognizing emotions’, ‘asking something to someone’ and ‘saying no when you don't want something to happen’). The parents reported how difficult each of the social skills was for their child, on a 4-point Likert scale (total range 30 to 120).

With these three instruments we aimed to measure slightly different aspects of social communicative skills. That is, we aimed to measure the specific skills that were explicitly taught during the SST, focusing on the perceived difficulty as reported by parents (ESTIA-TS). Moving further away from the training we aimed to measure the actual application of social skills in daily life, with a slight difference in level of generalization of application between the Vineland and the SSRS.

Participant characteristics

Various participant characteristics were collected that have been shown to impact the effect of SST in earlier research. *Verbal cognitive ability* was assessed with the Dutch Wechsler Intelligence Scale for Children- 3th edition (WISC-III, Wechsler, 1999; Dutch version, Kort et al., 2005). *Severity of ASD symptoms* was measured with the Autism Diagnostic Observation Schedule (ADOS, Lord et al., 1999; Dutch version, de Bildt & de Jonge, 2008) and Autism Diagnostic Interview-Revised (ADI-R, Rutter et al., 2003; Dutch version, de Jonge & de Bildt, 2007). For the ADOS, we applied the calibrated severity scores (CSS; Lord et al., 2012). Symptoms of ADHD were measured with the two subscales *Inattention* and *Hyperactivity/impulsivity* of the 26-item parent-version of the Swanson, Nolan, and Pelham questionnaire (SNAP-IV; Swanson, 1997). *Symptoms of Social Phobia*, and *Total Anxiety* were measured with the Social Phobia subscale and the total score of the five anxiety subscales of the child-version of the Revised Child Anxiety and Depression Scale (RCADS; van Oort et al., 2009). All data were categorized as described in Table 2.

Intervention characteristics

Parent and teacher involvement was based on the condition that participants were randomized to: SST with or without additional parent and teacher sessions.

Analyses

We used Multilevel Latent Class Growth Analysis (MLCGA; Muthén, 2004; Palardy & Vermunt, 2010) to identify distinct subgroups of participants based on their level and perceived difficulty of social communicative skills as reported by parents at start of SST and their response to SST. In the MLCGA the three outcome measures (Vineland Socialization, SSRS-P, and ESTIA-TS) at the three moments (before (T1), immediately after (T2) and six months after end of SST (T3)) were jointly analyzed, using for each outcome variable a regression equation with normal error, including an intercept and two dummy variables. The latter were coded such that the intercept referred to T1, and the slopes for the dummy variables referred to the difference between T1 and T2, and to the difference between T2 and T3. MLCGA models with one to seven classes were estimated (i.e., response patterns). From these models, we selected the model with a minimal class size of 10% and the lowest value of the Bayesian Information

Table 2 Categorization of the participant characteristics

Participant characteristic	Measurement	Categories
Gender		1) Male 2) Female
Age		1) 9 year 2) 10 year 3) 11 year 4) 12 year
Verbal cognitive ability	WISC-III	1) Below average (< 90) 2) Average (90-110) 3) Above average (> 110)
Severity of ASD symptoms	ADOS	1) Minimal evidence ASD (CSS 1-2) 2) Low level ASD (CSS 3-4) 3) Moderate level ASD (CSS 5-7) 4) High level ASD (CSS 8-10)
	ADI-R	1) No symptoms of ASD/AD 2) Symptoms of ASD as defined by Risi and colleagues (2006) ¹ 3) Symptoms of AD as defined in the ADI-R manual (Rutter et al., 2003) ²
Symptoms of ADHD	SNAP-IV Inattention	1) Below population mean score ³ 2) Above population mean score Male means 9 y 9.09; 10 y 7.02; 11 y 6.48 Female means 9 y 4.5; 10 y 4.68; 11 y 8.37 For children aged 12, no population data were available ⁴
	SNAP-IV Hyperactivity/ Impulsivity	1) Below population mean score ³ 2) Above population mean score Male means 9 y 7.2; 10 y 6.12; 11 y = 6.93 Female means 9 y 3.78; 10 y 3.96; 11 y 6.12 For children aged 12, no population data were available ⁴
Symptoms of Social Phobia	RCADS	1) Below population mean score ⁵ 2) Above population mean score Male and female means 10 y 7.2; 11-12 y 6.84 For children aged 9, no population data were available. ⁶
Symptoms of total anxiety	RCADS	1) Below population mean score ⁵ 2) Above population mean score Male and female means 10 y 22.0; 11 y 20.33; 12 y 18.37 For children aged 9, no population data were available. ⁶

Note ¹ child meets one of the following three criteria a) cut-off social interaction domain ≥ 10 & cut-off communication domain ≥ 6 ; b) cut-off social interaction domain ≥ 8 & cut-off communication domain ≥ 8 ; c) social interaction domain ≥ 9 & communication domain ≥ 7 ; ² cut-off social interaction domain ≥ 10 and cut-off communication domain ≥ 8 ; ³ categorization based on population data from American children with a Caucasian background, separated for age and sex (Bussing et al., 2008); ⁴ excluded for the analyses regarding ADHD symptomatology; ⁵ categories were based on data in the Dutch population for each age (van Oort et al., 2009); ⁶ excluded for the analyses concerning symptoms of Social Phobia and Total Anxiety

Criterion (BIC; Lanza et al., 2007). The BIC indicates a model with an optimal balance between fit and model complexity. Of the selected model, we calculated within group effect sizes (ES) for all slopes and classes (i.e., per class change between T2-T1 and/or T3-T2, as $d = \frac{(\hat{\rho}_{Tj} - \hat{\rho}_{T(j-1)})}{\hat{\sigma}_{T1}}$, with $j=2,3$). ES were based on the estimated model, and were interpreted as ES = 0.2 as small, ES = 0.5 medium, ES = 0.8 large and ES = 1.3 very large (Cohen, 1988), and applied as indicators for the clinical relevance of the findings.

For each participant we established the posterior probability of membership of each class based on the selected model above. Each participant was assigned to the class that matched with his or her highest probability.

In a follow-up analysis, we predicted the probability of class membership with each of the selected participant and intervention characteristics via univariable regression analyses, using the so-called three-step procedure as described in Vermunt and Magidson (2013). The participant and intervention characteristics were categorized as described in Table 2. These analyses provide insight into the distribution of the characteristics in each class, indicating whether differences between classes can be detected, based on other characteristics than outcome. For the characteristics that appeared to be significant, we performed a post-hoc analysis, to examine possible differences among the different classes with paired comparisons. Additionally, the characteristics that appeared to be significantly related to class membership were included in a multivariable regression analysis, to assess their unique contribution to predict class membership. The analyses were performed with Latent Gold 5.0 (Vermunt & Magidson, 2013), using all available data.

Results

Missing data

For the first measurement (T1) no missing data existed. Four participants dropped-out after the first measurement, but before the start of the actual SST, for these participants T2 and T3 data were not available. Additionally, at T2, two parents did not complete the ESTIA-TS and one did not complete the ESTIA-TS and the SSRS. At T3, for four additional participants, no data were available and two parents did not complete the ESTIA-TS and the SSRS.

Multilevel Latent Class Growth Analysis

A four-class model was selected and interpreted, as this model had the lowest BIC (see Table 3) and a minimal class size of 10%, with the smallest of the four classes representing 17.4% of the children.

Table 3 BIC values for the MLCGA models with one up to seven classes

Number of Classes	1	2	3	4	5	6	7
BIC	6791.2	6580.9	6531.3	6528.4	6531.4	6550.5	6572.1

BIC = Bayesian Information Criterion. The smallest BIC value is printed in bold face.

Interpretation of the classes

The four classes represent four underlying subgroups of children with ASD with different levels and perceived difficulty of social skills as reported by parents at start and different patterns of response during and after SST. Table 4 and Figure 1 present the information for interpreting and naming the four classes.

Class 1 is characterized by significantly higher levels of social communicative skills (SSRS percentile 19 and Vineland Socialization age equivalent 6y8m – 6y10m) and significantly lower perceived difficulty of these skills as reported by parents than the mean across all children before the SST, even though the levels of these skills are still low as expressed in the norm scores or age equivalents. Additionally, the children significantly improved in their social communicative skills (SSRS: from percentile 19 to percentile 53 from T1 to T2, ES 1.12; and Vineland Socialization: improvement of 14 months between T1 and T2, to 7y10m/8y, ES 0.77) after SST and their perceived difficulty as reported by parents significantly decreased after SST (ES -0.94). The children showed no further change six months after SST. We named this class *best performers (high starters, perceived difficulty below the mean; improvement)* and it consisted of 28.6% of the children.

Class 2 is characterized by significantly lower scores on Vineland Socialization (age equivalent 5y2m) than the mean across all children before SST. The perceived difficulty of social skills as reported by parents was similar to the mean across all children before SST. Additionally, the children significantly improved in their social communicative skills after SST (SSRS: from percentile 2 to percentile 14 between T1 and T2, ES 1.15; and Vineland Socialization: improvement of 6-8 months between T1 and T2, to 5y8m/5y10m, ES 0.53) and their perceived difficulty as reported by parents significantly decreased after SST (ES -0.71). The children showed no further change six months after SST. We named this class *improvers (low starters, mean perceived difficulty; improvement)*, and it was the biggest class with 29.2% of the children.

Class 3 is characterized by significantly higher levels than the mean across all children before SST on social communicative skills (SSRS percentile 8 at T1; and Vineland Socialization age equivalent 6y6m) and significantly higher perceived difficulty of these skills, as reported by parents, than the mean across all children before SST. No significant change appeared in skills (SSRS from percentile 8 to

Table 4 Perceived difficulty reported by parents and level of social skills at start and during SST for each class

Dependent variable	Class	Intercept	SE Intercept	Slope T1-T2 (ES) ¹	SE Slope T1-T2	Slope T2-T3 (ES) ^c	SE Slope T2-T3
ESTIA-TS Difficulty	Class 1	58.1***a	1.8	-9.0*** (d = -0.94)	2.9	-1.5 (d = -0.16)	2.7
	Class 2	69.2***	1.9	-7.3** (d = -0.72)	2.6	0.1 (d = 0.01)	2.7
	Class 3	76.2***b	2.1	-1.6 (d = -0.15)	2.9	-2.5 (d = -0.24)	2.9
	Class 4	86.3***b	2.5	-5.2 (d = -0.50)	3.3	1.5 (d = 0.15)	3.4
SSRS-Parents Total	Class 1	44.2***a	1.5	8.9*** (d = 1.12)	2.1	0.1 (d = 0.01)	2.2
	Class 2	32.2***	1.6	9.9*** (d = 1.16)	2.2	-1.6 (d = -0.19)	2.2
	Class 3	38.8***a	1.8	4.2 (d = 0.47)	2.3	1.6 (d = 0.18)	2.5
	Class 4	24.4***b	1.9	4.7 (d = 0.60)	2.7	1.1 (d = 0.14)	2.7
Vineland Socialization	Class 1	90.9***a	2.1	8.6** (d = 0.77)	2.9	4.5 (d = 0.40)	3.0
	Class 2	74.3***b	2.3	6.6* (d = 0.54)	3.0	2.2 (d = 0.18)	3.1
	Class 3	88.4***a	2.4	5.4 (d = 0.46)	3.3	-0.5 (d = -0.04)	3.3
	Class 4	65.4***b	2.8	4.9 (d = 0.42)	3.7	-1.0 (d = -0.09)	3.8

* $p < .05$; ** $p < .01$; *** $p < .001$; ^a significantly better score than the mean across all children before SST; ^b significantly weaker score than the mean across all children before SST; ^c within group effect size (ES)

Figure 1 Perceived difficulty as reported by parents and level of social skills at start and during SST for each class

percentile 18 between T1 and T2; Vineland Socialization: improvement of 4 months between T1 and T2, to 7y2m) or perceived difficulty as reported by parents, immediately and six months after training. We named this class *stable class (high starters, perceived difficulty above the mean; no improvement)*, and it consisted of 24.8% of the children.

Class 4 is characterized by significantly lower levels of social communicative skills (SSRS below percentile 2 at T1; and Vineland Socialization age equivalent 4y6m) and significantly higher perceived difficulty of these skills, as reported by parents, than the mean across all children before SST. No significant change appeared in skills (SSRS: did not exceed below percentile 2 from T1 to T2; Vineland Socialization: improvement of 4 months between T1 and T2, to 4y10m) or perceived difficulty as reported by parents, immediately and six months after training. We named this class *poor performers (low starters, perceived difficulty above the mean; no improvement)*, and it represented the smallest proportion with 17.4% of the children.

Of the four classes, the two classes with average or lower perceived difficulty of social skills as reported by parents (i.e., Classes 1 and 2) showed significant improvement of high to very high effect size during the SST. The other two classes, with higher perceived difficulty as reported by parents (i.e., Classes 3 and 4) showed no significant improvement during the SST, independent of the actual level of social communicative skills at start; yet the effect sizes of improvement during SST were small to medium, suggesting that this lack of significance may be due to a too small power. None of the four classes showed a clear change from T2-T3, as indicated by nonsignificant tests and generally very small effect sizes. See also Figure 2.

Figure 2 Distribution classes

Relationship of participant and intervention characteristics with class membership

In Table 5, we present the results of the univariable regression analyses predicting class membership based on each *participant and intervention characteristic*. The mean scores on these characteristics in each of the four classes is presented in supplementary Table 1. We here report the characteristics that showed a significant relationship with class membership. See also Figure 3.

Table 5 Relationship between participant and intervention characteristics and class membership

Class	1 vs 2	1 vs 3	1 vs 4	2 vs 3	2 vs 4	3 vs 4
Gender	-	-	-	-	-	-
Age	**	-	-	-	*	-
Verbal IQ	-	**	**	-	-	-
ASD symptoms: ADOS	***	*	*	*	*	-
ASD symptoms: ADI-R	***	-	***	*	-	***
ADHD Inattention	**	-	**	-	-	-
ADHD Hyperactivity/impulsivity	*	-	**	**	-	***
Social Phobia	-	**	-	**	-	*
Total Anxiety	-	-	*	*	***	-
Parent and teacher involvement	-	-	***	-	***	***

* $p < .05$; ** $p < .01$; *** $p < .001$; - no significant difference between the classes

For age, 'improvers' were the youngest and differed significantly from 'best performers' and 'poor performers'. On verbal ability, 'best performers' scored higher than the 'stable class' and 'poor performers'. On ASD symptoms measured with the ADOS, 'best performers' showed the least severe symptoms and differed significantly from all other three classes. 'Improvers' showed the most severe symptoms, compared to all other classes. On ASD symptoms measured with the ADI-R, 'poor performers' showed the most severe symptoms and differed significantly from 'best performers' and 'stable class'. 'ADHD Inattention and Hyperactivity/impulsivity symptoms were most apparent among the 'improvers' and 'poor performers', compared to both other classes for Hyperactivity/impulsivity, and to best performers on Inattention. Social phobia symptoms were most present in the 'stable class' compared to all other classes. 'Stable class' also showed more total anxiety symptoms than 'improvers', as did 'poor performers' compared to 'improvers' and 'best performers'.

Figure 3 Distribution of participant and intervention characteristics over the classes

Gender was the only participant characteristic that was not significantly related to class membership.

On the intervention characteristic (parent and teacher involvement), the 'poor performers' participated more often in SST without parent and teacher involvement than in SST with parent and teacher involvement, compared to all other classes.

In the multivariable regression we included all participant and intervention characteristics except gender. All but ADHD inattention symptoms, appeared to be significant in the multivariable regression analysis and thus relate uniquely to class membership.

Discussion

In the current study, four subgroups of children with ASD were identified with MLCGA, based on their level and perceived difficulty of social performance before SST as reported by parents, and on their progress over time (measured immediately and six months after SST). Two of these subgroups improved after SST (6 to 14 months on the Vineland Socialization domain in the 6 months between pre and posttest, and 12 to 34 percentiles on the SSRS), the other two showed no significant improvement (4 to 6 months on the Vineland Socialization and 0 to 10 percentiles on the SSRS). None of the subgroups improved or declined between the end of SST and follow-up six months later. These findings indicate an effect of SST for specific subgroups of children with ASD. In our primary analyses only small to moderate parent reported effects were found of the intervention on group level, and only on part of the outcome measures (Vineland socialization and SSRS-parents; cooperation; Dekker et al., 2019). The current analysis indicates that two specific subgroups improved on all three outcome measures (57.8% of the children), whereas two subgroups (42.2%) did not improve.

The improvement in the two improving subgroups seems of clinical relevance and is most clear in the first, with the highest improvement (large to very large effect for ESTIA-TS and SSRS, medium (close to large) for Vineland Socialization). The other improving subgroup improved to a lesser extent (medium effect size for ESTIA-TS, large to very large for SSRS, large for Vineland Socialization). Note that all subgroups had low levels of social behavior at start, which stayed relatively low after SST, even after improvement.

The findings of improving subgroups become even more interesting when relating the response patterns in the subgroups to the multiple participant and intervention characteristics in order to detect differences between them, based on other characteristics than outcome. Improvement did not seem to be related to level of social performance at start, as the two improving subgroups differed in their level

of social performance at baseline (above mean versus low to mean scores). Similarly, the non-improving subgroups differed on social performance compared to each other. In contrast, perceived difficulty of social skills for the child, as reported by parents, showed to be more crucial, as the improving groups had the lowest perceived difficulty (low and mean scores, respectively), and the non-improving subgroups showed high perceived difficulty of these skills. This may be due to the possibility that two children can perform equally well in how and how often they interact (expressed in comparable scores on actual performance as measured with SSRS/Vineland Socialization), even though they can differ enormously in the difficulty they have in reaching that behavior (expressed in significantly different scores on perceived difficulty as reported by parents on the ESTIA-TS). The latter may be of greater importance for benefiting of SST. Perhaps the children with ASD whose parents reported that their children did not perceive difficulty in interacting with others, may already interact more and benefit from the current SST by changing their manner of doing so to a more adequate one, without having to increase the frequency with, or the situations in which they interact.

Additionally, regarding participant characteristics, the current study did not find an effect of *gender* in relation to outcome on SST or basal social performance level. Earlier studies showed larger effects of SST for females than for males (McMahon et al., 2013; Choque Olsson et al., 2017). The absence of such a finding in the current study may be due to limited power to detect gender differences, as we included relatively few females (only 17%). Many earlier studies also included too few female participants to reliably investigate the effect of gender on the outcome of SST (Gates et al., 2017), so more research into the effect of SST in relation to gender should be conducted. Younger *age* was related to improvement, but only in the subgroup with lower levels of social performance at start. That is, younger children with lower levels of social skills improved more from training than older ones with lower levels of social skills. Former research did not result in firm conclusions on age and its influence on outcome in SST for ASD, and our study does not lead to a final conclusion either. We had a rather narrow age range in our study (9-12 years), which may decrease the influence of age on the outcomes of SST and the power to detect such an influence. *Verbal cognitive ability* was highest in the improving subgroup with high levels of social performance at start, illustrating the relatively high level of functioning of these children. Less verbally able children were more often in the non-improving subgroups, either with higher or lower levels of social communicative skills at start. Even in this relatively high functioning sample, this finding seems to corroborate former research concluding that participants with higher developmental levels benefit more from SST (Herbrecht et al., 2009) or from early intervention and development (Stevens et al., 2000; Kim et al., 2016; Paynter et al., 2018) than participants with lower developmental levels. This finding probably also relates to the character of the currently studied

intervention that has a high verbal component in its explanation of skills and the homework assignments. However, larger samples with larger variation in are needed to further investigate the role of verbal cognitive ability.

Contrary to our hypothesis, *ASD symptom severity* was not unequivocally related to improvement. The improving subgroup with high levels of social performance at start showed the least severe ASD symptoms on the ADOS and ADI-R, also compared to the subgroup with high levels of social performance at start without improvement (for the ADOS). This is in line with findings on early intervention and development (Stevens et al., 2000; Kim et al., 2016; Paynter et al., 2018). However, we also identified a subgroup with lower social performance at start that improved, even though this subgroup showed the most severe symptoms on the ADOS within our sample and the most severe but one on the ADI-R within our sample. We would have expected the scores on the ADI-R and ADOS at start to be more directly related to the response pattern to SST, and thus to contribute more clearly to describing the subgroups that benefit more or less from SST. Perhaps the fact that the current sample only consisted of relatively high functioning children with ASD, including many with PDD-NOS (66%) and Asperger's Disorder (20%), and that the mean scores on ADI-R and ADOS were relatively low, has led to a too small range and variation in ASD symptom severity in the current sample leading to limited power to allow drawing strong conclusions.

Regarding comorbidity, the relation found between *symptoms of ADHD* and subgroups seems to indicate that ADHD affects basal levels of social performance in the current sample, yet not outcome. Children with higher scores on inattention and hyperactivity/impulsivity symptoms had lower social performance levels at start than children with lower scores on these domains; however, improvement is not related to the level of these symptoms. This is in line with some earlier research reporting no effect of ADHD on outcome of SST (Deckers et al., 2016), yet it is in contrast with other findings (Antshel et al., 2011) reporting a decreasing effect of ADHD on the effect of SST in children with ASD. In the current sample inattention symptoms were relatively frequent in general. *Anxiety* was also related to subgroups, indicating an effect on outcome of social performance, independent from social performance levels at start. The low starting, no improvement subgroup had higher scores on total anxiety than the subgroups that did improve, including the low starting subgroup that improved. Perhaps the level of anxiety explains why some children with ASD with lower basal social performance benefit from SST (lower level of anxiety), whereas others do not (higher level of anxiety). Earlier research was inconclusive on this issue, and our study seems most in line with the study that reported that anxiety decreased the effect of SST (Pellecchia et al., 2016). Notably, the level of anxiety seemed to vary over the subgroups in the classes in a similar way as the perceived difficulty of social performance as reported by parents. The poor performers subgroup (low starters, no improvement) contained children with higher levels of anxiety and higher perceived

difficulty as reported by parents, the best performers subgroup (high starters, improvement) contained children with lower scores on each. The improvers subgroup (low starters, improvement) consisted of children with mean perceived difficulty as reported by parents compared to the whole group and low levels of anxiety, and the stable subgroup (high starters, no improvement) showed higher perceived difficulty as reported by parents and higher anxiety (compared to the improvers). Possibly anxiety affects how children participate in SST. Anxious children may hesitate to perform crucial parts of training during the sessions (role play, discussing homework, answering questions, etc.) or even more so beyond the session (practice in school or at home), or doing so may take much more energy that they cannot spend on the actual content of the assignments. This could lead to less benefit from training for these children in its current form. Of note, as mentioned for all other participant characteristics, comorbidity should be addressed in larger studies in order to optimize the power to detect its influence.

Regarding the intervention characteristic presence or absence of *parent and teacher involvement*, we found this to be related to the subgroup with low skills at start and no improvement. That is, more children in this subgroup participated in SST without parent and teacher involvement than in SST with such involvement. In the comparison between the three conditions of the RCT that the current paper was based on, we found no difference across all participants between the effect of SSTs with or without parent and teacher involvement compared to CAU reported by parents (Dekker et al., 2019). Yet, taking into account individual differences, the current study indicates that no formal parent and teacher involvement negatively affects outcome on SST for those children with poor social communicative skills and high perceived difficulty of these skills as reported by parents at start.

In order to understand what factors affect improvement after SST, we compared the subgroups with similar basal levels of social performance. In the subgroups with relatively high basal levels, verbal cognitive ability, lower ASD symptoms on the ADOS and ADI-R, and lower social phobia symptoms seem to be related to improvement after SST. Additionally, in the subgroups with low basal levels, improvement seems to be most clearly related to lower perceived difficulty as reported by parents, younger age, lower anxiety and presence of parent and teacher involvement.

Strengths and limitations

The discussion of the findings in our study indicates several strengths that contribute to the existing literature; from the manualized training with and without parent and teacher involvement, to the variety of participant characteristics that we could include in the analyses. However, the current study also had several limitations.

First, for the type of analyses, the sample size of the children included is relatively small, which limits the power to detect the role of all participant and treatment characteristics. This implies that further studies using a similar design are needed to corroborate the current findings. Second, we considered a relatively high functioning sample, both with respect to ASD symptom level and to verbal cognitive level. Although the sample represents children diagnosed with ASD (based on DSM-IV-TR criteria) who were referred for SST in clinical practice, their clinical diagnosis had not been corroborated by standardized instruments, such as the ADOS or ADI-R. Additionally, the majority of children (66%) received a clinical diagnosis of PDD-NOS. This impacts the generalizability of the findings as the results of the current study may perhaps be more applicable to children in clinical practice with autism traits and social skills impairments rather than a strict ASD diagnosis as defined in the DSM-5 criteria. Third, we could only include parent reported measures, completed by parents who were only blind to condition at pretest, which may have led to bias, since they knew in which condition their child had participated after training. We did not include objective measures, completed by independent raters. Adding blinded outcome measures by someone not involved in the training (unlike parents) is currently an important focus of research, in order to increase the highly needed insight into the objective effect of SST or other interventions for children with ASD. However, none was available at the time the current study started. For this aim we had developed a measure of naturalistic behavior in school, observed by an independent rater (Dekker et al., 2016), yet questions have risen on the feasibility during the process of data collection. Other measures are for younger children (e.g. the Brief Observation of Social Communication Change; Grzadzinski et al., 2016). The current study still is informative and adds to the meta-analysis of Gates and colleagues (2017), as they focused on the relation between participant characteristics and self reported improvement. Additionally, we used the ESTIA-TS to measure perceived difficulty of social performance as reported by parents, a simple, yet non-validated instrument. Fourth, we could not investigate medication as a participant characteristic, or duration and intensity as intervention characteristics. We did not have data on the exact medication at baseline of the children. Additionally, all children participated in the same 15 sessions of 90 minutes (22.5 hours of duration), in the same weekly pattern. Our duration and the intensity belong to the less effective end according to the literature (Wolstencroft et al., 2018). Last, the concept measured as the outcome in the current study is closer to social performance than social competence. Social performance is a narrower concept, yet it is in line with the aims of the training, i.e. applying social skills in role play and real life.

The current study provides insight into differences between subgroups of children with ASD in their response to SST, even within a relatively homogeneous sample, thereby specifying the effects found on group level. By doing so the current study contributes to the desired development towards personalizing SST for children with ASD. As improvement was not related to level of social communicative skills at start of the intervention but showed to be related much more to perceived difficulty of social communicative skills as reported by parents at start, the subgroups with higher perceived difficulties and anxiety may benefit more from an SST that focuses explicitly on perceived difficulty and anxiety in interaction. Additionally, subgroups with lower verbal levels probably benefit more from a training that addresses learning of social communicative skills through other means than mainly verbally. Last, explicit parent and teacher involvement in SST only seems to enhance the outcome of training for children with low levels of social communicative skills before training. Adapting training from the perspectives of specific needs of children with ASD may add to the benefit of SST. However, the current study concerned only one training (with and without parent and teacher involvement) and a specific sample. Replication in other samples and with other training approaches is needed, refining our understanding of what works for whom in SST, in order to reach the ultimate goal of personalizing social skills treatment for children with ASD.

Ethical approval

This study was approved by the ethics committee of the Institutional Review Board of the University Medical Center Groningen. Therefore, all procedures performed in this study were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

The study was registered in the Dutch trial register (NTR2405).

Disclosure of potential conflict of interest

Author AdB received the grant that funded this study (the Netherlands Organization for Health Research and Development; ZonMw, nr 157003005). The agency had no role in data analyses, the decision to publish, or the preparation of the manuscript. She also is first author on the Dutch ADOS manual for which Accare receives enumeration. The other authors declare that they have no conflict of interest.

Acknowledgments

The authors gratefully acknowledge all the children, their parents, teachers, and therapists participating in this study. This study has been funded by the Netherlands Organization for Health Research and Development (ZonMw, nu 157003005).

References

- Antshel, K.M., Polacek, C., McMahon, M., Dygert, K., Spenceley, L., Dygert, L., Miller, L., & Faisal, F. (2011). Comorbid ADHD and anxiety affect social skills group intervention treatment efficacy in children with autism spectrum disorders. *Journal of Developmental and Behavioral Pediatrics*, 32(6), 439–446.
- Bussing, R., Fernandez, M., Harwood, M., Hou, W., Garvan, C.W., Swanson, J.M., & Eyberg, S.M. (2008). Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: psychometric properties and normative ratings from an school district sample. *Assessment*, 15(3), 317–328.
- Choque Olsson, N., Flygare, O., Coco, C., Görling, A., Råde, A., Chen, Q., Lindstedt, K., Berggren, S., Serlachius, E., Tammimies, K., Kjellin, L., & Bölte, S. (2017). Social Skills Training for Children and Adolescents With Autism Spectrum Disorder: A Randomized Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(7), 585–592.
- Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic.
- de Bildt, A.A., & Kraijer, D.M. (2003). *Sociale redzaamheid voor kinderen en jeugdigen met een verstandelijke beperking*. Vineland-Z. Leiden: PITS B.V.
- de Bildt, A., & de Jonge, M.V. (2008). *Autisme Diagnostisch Observatie Schema*. Amsterdam: Hogrefe.
- de Jonge, M.V., & de Bildt, A. (2007). *Autisme Diagnostisch Interview-Revised*. Amsterdam: Hogrefe.
- Deckers, A., Muris, P., Roelofs, J., & Arntz, A. (2016). A Group-Administered social Skills Training for 8- to 12-Year-Old, high-Functioning Children With Autism Spectrum Disorders: An Evaluation of its Effectiveness in a Naturalistic Outpatient Treatment Setting. *Journal of Autism and Developmental Disorders*, 46(11), 3493–3504.
- Dekker, V., Nauta, M.H., Mulder, E.J., Sytema, S., & de Bildt, A. (2016). A fresh pair of eyes: A blind observation method for evaluating social skills of children with ASD in a naturalistic peer situation in school. *Journal of Autism and Developmental Disorders*, 46(9), 2890–2904.
- Dekker, V., Nauta, M.H., Mulder, E.J., Timmerman, M.E., & de Bildt, A. (2014). A randomized controlled study of a social skills training for preadolescent children with autism spectrum disorders: generalization of skills by training parents and teachers? *BMC Psychiatry*, 14, 189.
- Dekker, V., Nauta, M.H., Timmerman, M.E., Mulder, E.J., van der Veen-Mulders, L., van den Hoofdakker, B.J., van Warners, S., Vet, J.J., Hoekstra, P.J., & de Bildt, A. (2019). Social skills group training in children with autism spectrum disorder: a randomized controlled trial. *European Child & Adolescent Psychiatry*, 28(3), 415–424.
- Frankel, F., Myatt, R., & Feinberg, D. (2007). Parent-assisted friendship training for children with autism spectrum disorders: effects of psychotropic medication. *Child Psychiatry and Human Development*, 37(4), 337–346.
- Gates, J.A., Kang, E., & Lerner, M.D. (2017). Efficacy of group social skills interventions for youth with autism spectrum disorder: A systematic review and meta-analysis. *Clinical Psychology Review*, 52, 164–181.
- Gresham, F.M., & Elliott, S.N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Grzadzinski, R., Carr, T., Colombi, C., McGuire, K., Dufek, S., Pickles, A., & Lord C. (2016). Measuring changes in social communication behaviors; preliminary development of the Brief Observation of Social Communication Change (BOSCC). *Journal of Autism and Developmental Disorders*, 47(7), 2464–2479.
- Herbrecht, E., Poustka, F., Birnkammer, S., Duketis, E., Schlitt, S., Schmötzer, G., & Bölte, S. (2009). Pilot evaluation of the Frankfurt social skills training for children and adolescents with autism spectrum disorder. *European Child & Adolescent Psychiatry*, 18(6), 327–335.
- Jonsson, U., Olsson, N.C., Coco, C., Görling, A., Råde, A., Chen, Q., Berggren, S., Tammimies, K., & Bölte, S. (2019). Long-term social skills group training for children and adolescents with autism spectrum disorder: a randomized controlled trial. *European Child & Adolescent Psychiatry*, 28(2), 189–201.
- Kim, S.H., Macari, S., Koller, J., & Chawarska, K. (2016). Examining the phenotypic heterogeneity of early autism spectrum disorder: subtypes and short-term outcomes. *Journal of Child Psychology and Psychiatry*, 57(1), 93–102.
- Kort, W., Schittekatte, M., Dekker, P.H., Verhaeghe, P., Compaaan, E.L., Bosmans, M., & Vermeir, G. (2005). *Wechsler Intelligence Scale for Children* (3rd ed. Nederlandse vertaling). Amsterdam: Harcourt.

- Lanza, S.T., Collins, L.M., Lemmon, D., & Schafer, J.L. (2007). PROC LCA: A SAS procedure for latent analysis. *Structural Equation Modeling*, 14(4), 671–694.
- Lord, C., Rutter, M., DiLavore, P.C., & Risi, S. (1999). *ADOS. Autism Diagnostic Observation Schedule. Manual*. Los Angeles: Western Psychological Services..
- Lord, C., Rutter, M., DiLavore, P.C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, second edition (ADOS-2) manual (Part I): Modules 1-4*. Torrance, CA: Western Psychological Services.
- Mathur, S.R., Kavale, K.A., Quinn, M.M., Forness, S.R., & Rutherford, R.B. (1998). Social skills interventions with students with emotional and behavioural problems: A quantitative synthesis of single-subject research. *Behavioural Disorders*, 23(3), 193–201.
- McMahon, C.M., Vismara, L.A., & Solomon, M. (2013). Measuring changes in social behavior during a social skills intervention for higher-functioning children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(8), 1843–1856.
- Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Thousand Oaks, CA: Sage publications.
- Palardy, G.J., & Vermunt, J.K. (2010). Multilevel growth mixture models for classifying groups. *Journal of Educational and Behavioral Statistics*, 35(5), 532–565.
- Paynter, J., Trembath, D., & Lane, A. (2018). Differential outcome subgroups in children with autism spectrum disorder attending early intervention. *Journal of Intellectual Disability Research*, 62(7), 650–659
- Pellecchia, M., Connell, J.E., Kerns, C.M., Xie, M., Marcus, S.C., & Mandell, D.S. (2016). Child characteristics associated with outcome for children with autism in a school-based behavioral intervention. *Autism*, 20(3), 321–329.
- Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., Cook, E.H. Jr, Leventhal B.L., & Pickles, A. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45(9), 1094–1103.
- Rutter, M., le Couteur, A., & Lord, C. (2003). *ADI-R Autism Diagnostic Interview Revised Manual*. Los Angeles: Western Psychological Services.
- Sparrow, S.S., Balla, D., & Cicchetti, D.V. (1984). *Vineland adaptive behavior scales (Survey ed.)*. Circle Pines, MN: American Guidance Service.
- Stevens, M.C., Fein, D.A., Dunn, M., Allen, D., Waterhouse, L.H., Feinstein, C., & Rapin, I (2000). Subgroups of children with autism by cluster analysis: a longitudinal Examination. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(3), 346–352.
- Swanson, J.M (1997). *The SNAP-IV teacher and parent rating scale*. University of California, Irvine, CA.
- van Oort, F.V.A., Greaves-Lord, K., Verhulst, F.C., Ormel, J., & Huizink, A.C. (2009). The developmental course of anxiety symptoms during adolescence; the TRAILS study. *Journal of Child Psychology and Psychiatry*, 50(10), 1209–1217.
- van der Oord, S., van der Meulen, E.M., Prins, P.J.M., Oosterlaan, J., Buitelaar, J.K., & Emmelkamp, P.M.G. (2005). A psychometric evaluation of the social skills rating system in children with attention deficit hyperactivity disorder. *Behaviour Research and Therapy*, 43(6), 733–746.
- Vermunt, J.K., & Magidson, J. (2013). *Latent GOLD 5.0 Upgrade Manual*. Belmont, MA: Statistical Innovations Inc.
- Wang, S., Cui, Y., & Parrila, R. (2011). Examining the effectiveness of peer-mediated and video-modeling social skills interventions for children with autism spectrum disorders: A meta-analysis in single-case research using HLM. *Research in Autism Spectrum Disorders*, 5(1), 562–569.
- Wechsler, D. (1999). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Wolstencroft, J., Robinson, L., Srinivasan, R., Kerry, E., Mandy, W., & Skuse, D. (2018). A Systematic Review of Group Social Skills Interventions, and Meta-analysis of Outcomes, for Children with High Functioning ASD. *Journal of Autism and Developmental Disorders*, 48(7), 2293–2307.

Supplementary table 1 Mean scores (SD) on all participant characteristics in the four classes, based on the modal class membership

		Class 1		Class 2		Class 3		Class 4	
		mean (SD)	Range	mean (SD)	Range	mean (SD)	Range	Mean (SD)	Range
<i>n</i>	Male:Female	22:7		22:6		22:2		15:2	
Treatment condition	SST:SST-PTI (<i>n</i>)	12:17		11:17		10:14		14:3	
Age	Years	10.9 (0.7)	9.6-12.4	10.8 (0.6)	9.7-12	10.9 (0.9)	9.5-12.7	11.1 (0.6)	9.9-12.6
Vineland	Socialization	90.7 (9.6)	68-108	74.1 (8.9)	55-95	88.8 (10.7)	71-118	64.9 (15.1)	26-90
SSRS-P	Total	44.0 (8.1)	25-59	31.9 (7.1)	15-43	39.1 (7.5)	28-53	24.1 (6.6)	10-35
ESTIA-TS	Training-specific social skills	58.1 (9.0)	42-74	69.6 (9.4)	49-86	76.3 (10.6)	53-97	86.0 (13.6)	66-106
ADOS	Social Affect	6.7 (3.1)	0-13	9.4 (4.0)	3-19	8.4 (4.7)	2-20	8.9 (4.7)	2-20
	Restricted and Repetitive behavior	1.0 (0.8)	0-3	1.5 (1.2)	0-5	1.3 (1.0)	0-4	0.9 (0.9)	0-3
	Calibrated Severity Score	4.6 (2.1)	1-8	6.5 (2.0)	3-10	5.5 (2.5)	2-10	5.5 (2.6)	1-10
ADI-R	Social Interaction	10.6 (5.2)	3-21	15.3 (5.6)	4-24	14.1 (4.9)	5-22	19.4 (5.9)	8-27
	Communication	10.1 (4.7)	2-21	13.21 (3.9)	5-20	10.3 (4.6)	3-20	14.9 (3.8)	8-23
	Restricted and Repetitive behavior	3.1 (2.2)	0-10	2.9 (1.7)	0-7	3.25 (2.4)	0-9	3.8 (2.3)	0-7
	Total	25.5 (10.3)	9-55	33.1 (9.8)	11-51	29.5 (10.6)	8-46	40.1 (8.9)	20-55
RCADS-C	Social Phobia	6.9 (3.7)	0-15	7.3 (4.2)	0-15	8.0 (4.4)	1-16	7.0 (3.7)	1-15
	Total anxiety	20.3 (10.3)	2-42	21.8 (13.5)	3-54	22.1 (11.2)	6-44	22.8 (11.2)	6-47
SNAP-IV-P	ADHD Inattention	10.3 (6.7)	1-24	13.2 (6.3)	5-26	13.3 (7.0)	2-26	14.7 (5.6)	6-24
	ADHD hyperactivity/impulsivity	8.7 (5.1)	1-20	9.9 (7.3)	0-27	8.3 (5.1)	0-18	10.9 (4.8)	3-19
IQ	Verbal	107.4 (14.3)	80-132	101.8 (16.7)	73-139	101.8 (18.1)	74-145	98.4 (14.6)	72-121
	Perfomal	99.5 (19.8)	70-133	96.8 (15.3)	72-139	97.7 (17.1)	60-126	98.2 (14.6)	76-135
	Total	104 (16.2)	73-135	99.4 (14.7)	77-132	99.5 (16.9)	72-123	98.1 (15.1)	77-128

5 |

A fresh pair of eyes: a blind observation method for evaluating social skills of children with ASD in a naturalistic peer situation in school

Vera Dekker
Maaïke H. Nauta
Erik J. Mulder
Sjoerd Sytema
Annelies de Bildt

Abstract

The Social skills Observation Measure (SOM) is a direct observation method for social skills used in naturalistic everyday situations in school. This study describes the development of the SOM and investigates its psychometric properties in 86 children with Autism spectrum disorder (ASD), aged 9.8-13.1 years. The interrater reliability was found to be good to excellent. The convergent validity was low in relation to parent and teacher reports of social skills, and also to parent interview on adaptive social functioning. Therefore this direct observation seems to provide additional information on the frequency and quality of social behaviors in daily life situations. As such it contributes to parent and teacher information as a blind measurement to evaluate Social Skills Training (SST).

Introduction

In recent years, there has been an increase in research into the effectiveness of social skills training (SST) for children with autism spectrum disorders (ASDs). SST is a common intervention for children with ASD, aiming to improve social skills (e.g. Koenig et al., 2010; Freitag et al., 2013). Although many parents desire such training for their children, and many professionals have the clinical impression that these trainings help improve the child's social communication, the scientific evidence for the effectiveness of such training is still inconclusive (Dekker et al., 2014; Reichow et al., 2013; White et al., 2007; Reichow & Volkmar, 2010; Lopata et al., 2010). One of the main problems is the limited number of studies into the generalization of learned social skills to daily social life (Owens et al., 2008).

In order to measure the effectiveness of SST for children with ASD, researchers and clinicians use a variety of instruments. One reason for this variety is the complex construct that social skills comprise, which is difficult to operationalize (McMahon et al., 2013a; White et al., 2007). In most studies, researchers use parent or teacher questionnaires on social skills to evaluate the intervention. However, these informants are rarely blind to treatment condition and therefore possibly biased (McMahon et al., 2013a). Even in the most recent study on SST in ASD, Freitag and colleagues (2016) conclude that the main limitation of their rather well designed randomized controlled trial (RCT) is that they could not use blinded raters of the primary outcome measures. In other words, objective evaluation of SST requires direct observational methods of children's social skills in daily life situations involving peers, rated by blinded observers. Such observation would provide crucial information on the actual application of skills taught during SST, indicating the clinical significance outside the therapy sessions and the generalizability to daily life (McMahon et al., 2013a). Although the cited studies concerned efficacy instead of effectiveness, blinded observations are needed in studies into effectiveness as well as into efficacy.

Blinded, direct observations in naturalistic settings with peers have not regularly been performed to evaluate SST for children with ASD. The main reasons seem to be the intensity of direct observation, its limited feasibility and the difficulty of deriving reliable information from the observations (White et al., 2007; McMahon et al., 2013b). However, some studies included a direct observation for the evaluation of SST in a naturalistic setting with peers (LeGoff, 2004; Owens et al., 2008; Hopkins et al., 2011; Shih et al., 2016; Wood et al., 2014). The methods used in these studies reveal two main characteristics that differentiate between them: the interval of observation and the focus on frequency and/or quality of behaviors. Three studies (LeGoff, 2004; Owens et al., 2008; Hopkins et al., 2011) used an overall observation period to observe general social behavior, for example the duration of social interaction with peers. Of these three studies, two studies reported a good to excellent reliability of

the observation measures (LeGoff, 2004; Hopkins et al., 2011). With respect to validity, only LeGoff (2004) correlated the observation outcomes (i.e. self-initiated contact and duration of social interactions) with another measure, the Social Interaction subscale of the Gilliam Autism Rating Scale. The correlations found were weak to moderate ($-.46$ for contact and $-.37$ for duration).

Two other studies (Wood et al., 2014; Shih et al., 2016) divided the total observation into smaller intervals and coded presence of general social behavior after each interval. The intervals did not affect the interrater agreement, which was $.84$ – $.87$. No information was available on the validity of the observation. The smaller intervals enable the evaluation of the extent to which a child shows social behavior. Regardless of the observation interval, observing the quality of behaviors seems equally important as the frequency. In SST the focus is on teaching specific behaviors and assertive interactions instead of more sub-assertive or negative and aggressive interactions, i.e. quality of behaviors. Hopkins and colleagues (2011) focused on the quality of behaviors in their observation, differentiating positive social interaction, low-level interaction and negative social interaction. They found that these qualities could be observed reliably, with Cohen's kappa's ranging from $.74$ (low-level interaction) to $.95$ (positive social interaction). Wood and colleagues (2014) also focused on the quality of behaviors in their observation, differentiating positive or appropriate interaction (or response) from negative behaviors. They found an overall interrater agreement of $.84$, however, the interrater agreement for each quality separately was not reported.

From the previous studies into direct observation methods in naturalistic settings with peers, several limitations and recommendations appear repeatedly. One limitation was the short duration of some observation methods. Some lasted 10 minutes or less (Owens et al., 2008; Hopkins et al., 2011) and the authors were concerned about whether such short observations were enough to sufficiently distinguish behaviors between children (Owens et al., 2008; Frankel et al., 2011). Another limitation was the unvaried situation in which the observations took place (Hopkins et al., 2011; Wood et al., 2014): all on the playground at school. Hopkins and colleagues (2011) recommended to observe in other real life situations as well, for example in the classroom, to evaluate the generalization of learned skills in more varied natural situations. Additionally, observers were not always blind to the treatment condition or group status, indicating a possible influence on the observation through the expectations of the observer (LeGoff, 2004; Owens et al., 2008; Humphrey & Symes, 2011). Moreover, children in the study of Humphrey and Symes (2011), were not blind to the fact that they were being observed. Before the observation, they signed an informed consent. In the studies of LeGoff (2004) and Owens and colleagues (2008), the observer was also the therapist of the children, so these children were probably also aware of the fact that they were being observed. The awareness of being observed may cause children to express less naturalistic and

less representative behavior compared to a situation, in which children are not aware that they are being observed. Finally, two recommendations concerned the content of the observations were also made. Hopkins and colleagues (2011) recommended to observe the duration or frequency of social skills, instead of only coding the presence or absence of each skill during the overall observation. Additionally, Frankel and colleagues (2011) recommended to observe the number of peers with whom the participating child interacted. They argued that a positive interaction with one or two other children would not imply a more widespread general acceptance by a broader peer group, and that the observation therefore should indicate the total number of peers interacted with.

The current study describes the development of the Social skills Observation Measure (SOM), as a method to observe social behavior of children with ASD in a naturalistic, unstructured and familiar setting with peers in school. This direct observation method built upon the limitations and recommendations from previous studies with naturalistic observation methods as described above. We observed children for 15 minutes, divided into 15 one-minute intervals, in two unstructured situations in school, extending the situations as recommended. We selected the three most important aspects from the earlier observation methods to focus on, and thus recorded the occurrence of specific behaviors, the quality of behavior, and the number of peers with whom the participating child interacted. The observer was unfamiliar to the participant, and the children were unaware of being the focus of the observation. Additionally, blindness of the observer for treatment allocation was adhered to as much as possible. The method was based on the observation methods of Humphrey and Symes (2011) and Frankel and colleagues (2011) and was further developed to tailor it to our specific study, in which we investigate the effectiveness of SST in children with ASDs ('Efficacy of Social Skills Training In Autism' study (ESTIA study); Dekker et al., 2014).

With the observation method described here, we aimed to contribute to the development of a naturalistic, reliable, valid, and feasible measure of social skills in a familiar, unstructured situation with peers, eventually in order to gain insight into the clinical significance and generalization of learned social skills during SST in peer interactions in daily life. The current study reports on the interrater reliability and the convergent validity of the SOM and does not involve analyses related to SST. These psychometric properties need to be examined before this observation can be used for the evaluation of the effectiveness of an SST for children with ASD.

Method

Participants

The participants in the current study participated in the ESTIA study (Dekker et al., 2014). The main aim of that study was to investigate the effectiveness of SST in children with ASD in an RCT. The study included two intervention conditions: Social Skills Training (SST) and Social Skills Training with Parent and Teacher Involvement (SST-PTI), and a care-as-usual (CAU) condition without SST. The children in the ESTIA study were recruited through four out-patient mental health clinics in the northern part of the Netherlands: two clinics of Accare University Center for Child and Adolescent Psychiatry and two clinics of Lentis Jonx Autism Team North-Netherlands. All were seeking treatment to assist with social skills. The children had a clinical DSM-IV-TR diagnosis of ASD, established by a child psychiatrist, based on multi-disciplinary assessment of the development and current behavior of the child by interviewing the parents, observing the child (most often based on the ADOS) and information from school. The total IQ was between 72 and 135 (mean 101.6), and the children were in the last two and half years of primary education (common age in the Netherlands in these grades is approximately 10 to 12 years). The ESTIA study was approved by the Institutional Review Board of the University Medical Center Groningen, and all parents, teachers, and children (if aged 12) gave written informed consent. This form was signed in the beginning of the study and included approval for the entire study, including the observation. For a detailed description of the procedures and training in the ESTIA study, please refer to Dekker and colleagues (2014; 2019).

To investigate the observation method in a naturalistic situation with peers in daily life we observed two groups of children from the ESTIA study. The first consisted of 86 observations (71 boys and 15 girls) used to determine the prevalence of behaviors in the chosen settings and validity. The second involved 39 observations (29 boys, 10 girls; of which 9 boys and 4 girls had been included in the prevalence/validity study as well) for interrater reliability.

The 86 children in the prevalence/validity study were observed before the initiation of treatment or CAU period (T1), in their own school setting. Of these 86 children, 12 attended special education and the other 74 children went to mainstream schools. All 86 children were Caucasian and most had at least one parent of Dutch decent ($n=85$, including 60 with two Dutch parents). In table 1 the characteristics of these 86 children are presented.

For the interrater reliability study, we included 39 complete observations by two observers in the classroom, playground or both situations (29 boys, 10 girls; see table 1 for characteristics). We collected 13 observations with two independent observers before treatment or CAU (i.e. T1; these were also included in the validity

Table 1 Characteristics of the children (*n*=86).

	Prevalence/validity study			Interrater reliability study*				
	<i>n</i> (%)	Mean	SD	Range	<i>n</i> (%)	Mean	SD	Range
Age	86	11.1	.8	9.8-13.1	39	11.3	.8	10.0-13.5
Male/Female	71 (82.6)/15 (17.4)				29 (74.4)/10 (25.6)			
Total IQ (WISC)	86	101.6	15.4	72-135	39	99.9	17.0	72-132
Verbal IQ (WISC)	86	103.7	16.1	72-145	39	104.2	16.8	74-145
Performal IQ (WISC)	86	99.1	16.7	60-139	39	95.5	18.7	60-135
Special education	12 (14.0)				10 (25.6)			
Maternal education	85 (98.8)				39 (100)			
Low ¹	1 (1.2)				1 (.03)			
Middle ²	53 (61.6)				24 (61.5)			
High ³	31 (36.0)				14 (35.9)			

¹ no or primary education; ² secondary or intermediate vocational education; ³ higher education
* Note: 10 children from the interrater reliability study were also included in the prevalence/validity study

study). Additionally, we collected twenty-six observations on children from the ESTIA sample post-treatment or post-CAU (i.e. T2; for 10 children we used observations from T1 as well as T2 for reliability). Thus, unlike the analyses on prevalence and validity, involving only T1 data ($n=86$), the reliability analyses were based on data from T1 for some ($n=13$ from the 86) and T2 for other observations ($n=26$ additional observations).

Instruments

Social skills Observation Measure (SOM) – initial content

The SOM is based on the observation model used by Frankel and colleagues (2011) and Humphrey and Symes (2011). Both previous methods used 1-minute observation intervals. Frankel and colleagues (2011) observed children on the playground for at least 15 minutes. Their observations focused on seven interactive states (solitary, proximity, onlooker, parallel, parallel aware, joint engagement, and games with rules) and three discrete behaviors (child initiates, positive response to social initiation, and positive shared affect). Humphrey and Symes (2011) observed children during two to four break and lunchtime sessions for 15 to 45 minutes during two days. Their observation focused on 22 observable behavior codes (15 behaviors exhibited by the focal student and 7 related to behavior of their peers). In order to develop an observation method as close to the SST under study as possible, we based our observation codes on the method of Humphrey and Symes (2011). The behaviors in their observation varied from solitary behaviors to interactive behaviors and aggressive behaviors, covering a large range of (non-)social behaviors that were also the focus of our SST. We made a few small changes in their model for an even better connection to our SST. First, we selected eight behaviors from the original 15 behaviors (Humphrey & Symes, 2011; see table 2 for the original items and the final items in our observation). Second, four of the remaining seven behaviors were combined into two items in our observation: instrumental aggression verbal and instrumental aggression physical were merged into aggression verbal/physical and reactive aggression verbal and reactive aggression physical were merged into reactive aggression verbal/physical. We omitted social initiation and acceptance of social initiation, because of the reported low proportion scores of both behaviors added to the difficulty of differentiating between the two when not being close and without recording. We decided to code social interaction as a broader behavior, without differentiating initiation or response. Additionally, we omitted submissive/passive from the *specific behaviors*. However, we added this aspect in the quality of behaviors, with which we aimed to describe the *quality* of the behaviors the children showed, defined as sub-assertive behavior (comparable to submissive/passive), dominant or aggressive behavior, and assertive behavior. In the SST that is evaluated in the ESTIA study, children practice with behaviors that accompany each of these

Table 2 Definitions of the specific behaviors*; quality of the behaviors; number of children.

Specific behavior	Description observed behavior	SOM
Solitary unoccupied	Child is alone and not looking at others, e.g. daydreaming	Included
Solitary engaged	Child is alone but engaged with some kind of activity, e.g. reading a book	Included
Solitary onlooker	Child is watching, but not interacting with another child	Included
Parallel	Child is next to another child, both engaged in an activity, but not interacting	Included
Co-operative interaction / joint engagement	Child is engaged in reciprocal interaction with another child	Included
Co-operative game / game with rules	e.g. football or a card game	Included
Rough / vigorous play	e.g. play fighting or wrestling	Included
Locomotor	e.g. running around – but not as part of a game with rules	Included
Instrumental aggression verbal	Not in reaction to aggression by peers, e.g. telling another child to 'F*** off' without provocation	Combined with Instrumental aggression physical
Instrumental aggression physical	Not in reaction to aggression by peers, e.g. pushing another student without provocation	Combined with Instrumental aggression verbal
Reactive aggression verbal	As above but in reaction to aggression by peer	Combined with reactive aggression physical
Reactive aggression physical	As above but in reaction to aggression by peer	Combined with reactive aggression verbal
Social initiation by focal student	e.g. offers invitation to play chess	Omitted**
Acceptance of social initiation by peer	Child demonstrates positive social behavior, either verbal or physical in reaction to social initiation	Omitted**
Submissive/passive	No retaliation, shows pain, cries, submission, tries to leave	Omitted (is captured in the quality of behavior)

Table 2 Continued.

Quality of behavior	Description of behavior	
Sub-assertive	Child behaves sub-assertive	
Aggressive/dominant	Child behaves dominant or aggressive	
Assertive	Child behaves assertive	
Number of children with whom the child interacted		

* The specific behaviors are based on the study of Humphrey and Symes (2011).
** Omitted due to a combination of low proportion scores in Humphrey and Symes (2011) and the distance to the child in the current observation, which resulted in questionable differentiation between these two specific behaviors

ways of behaving and discuss the effects they have on their own behavior and on the interaction with others. Therefore we aimed for a measure that would be able to identify these important behaviors. Another addition was that we counted the number of children with whom the child interacted, as recommended (Frankel et al., 2011).

Each observation consisted of two separate observation periods of 15 one-minute intervals: once in a naturalistic, unstructured situation in the classroom (e.g. free play-time), and once in a naturalistic, unstructured situation in the playground during a break. The observation was direct and live. The observer was introduced as someone who came to observe how things go in the classroom or on the playground, without specifying the target child or the aim of the observation. With this explanation, we aimed to create a situation as natural as possible. Since children and parents knew that observation was part of the study, and had given informed consent for that at the time they entered the study, this was the closest we could get to a natural observation. Parents were informed about when the observer would visit the school. We did not record the observations nor closely follow the target child around. This means that we may have missed remarks from children on some occasions, i.e. we may have seen them talking to others, at a distance from which we could not hear what they were saying. However, following a specific child with camera and microphone would not have added to a naturalistic setting nor to unawareness of being the focus of observation.

The observers were blinded for the condition of the child as much as possible. In a few cases, the observer knew about the condition of the child, due to logistical reasons, that could not be altered. The observer was not involved in the treatment of the child.

Each one-minute interval consisted of 40 seconds of observation and 20 seconds of coding three aspects, namely the *specific behaviors*, the *quality of behavior* and

the *number of children* the child interacted with. The *specific behavior* that appeared most in the 40 seconds was scored 1, behavior that occurred second most was scored 2, etc. Behavior that did not occur was scored 0. The most prevalent *quality of behavior* in an interval was scored 1, the one that occurred second most was scored 2, etc., with a 0 for qualities that did not occur. Additionally, the *number of children* with whom the target child interacted with during each interval was noted. At the end of the 15 minutes, the total number of children with whom the target child interacted during the whole observation was also noted.

The observers were psychologists or master students in psychology and were trained during two sessions of three hours. During the training they received explanation about all the specific behavior codes and they watched various unstructured situations of SST sessions to practice with the quality ratings. The children in these sessions were not included in the study, because the observers then knew the conditions with possible bias as a result. After the training sessions, they accompanied a trained observer to a school to observe in the classroom and on the playground for practice. Only after all these steps, did observers carry out observations by themselves that were used in the analyses.

Social skills Observation Measure (SOM) – further development

The SOM focused on three aspects of social skills, namely *specific behaviors*, the *quality of behavior*, and the *number of children* with whom the child interacted. Before outcome variables of these three aspects could be defined some further development was needed.

The original scores of the *specific behaviors* and the *quality of behavior* were transformed into 1 for present and 0 for non-present behavior, to create dichotomous codes per interval for each behavior and each quality. In order to gain insight into the frequency with which a *specific behavior* or *quality of behavior* occurred in relation to all the specific behavior/qualities, proportion scores were calculated. These proportion scores were calculated by dividing the sum score of one specific behavior/quality by the sum of all observed specific behaviors/qualities (in line with Humphrey & Symes, 2011).

Of the 86 children in the prevalence/validity study, 61 observations in the classroom and 55 on the playground lasted for the complete 15-minutes. Most observations lasted at least 10-minutes (84 observations in the classroom and 83 observations on the playground). The missing intervals were sometimes in the middle of the observation (e.g. child went inside to the bathroom or to get something), but most missed intervals were at the end (e.g. break situation lasted shorter than 15 minutes). The behaviors in the missing intervals were imputed, with the mean of the same behavior in the other intervals. For the imputation, at least 8 of the 15 intervals needed to be observed.

The 39 observations (22,8% of all observations) in the interrater reliability study contained all 15 intervals in the classroom, in the playground, or in both situations. Of these observations, three scores were available in the end: one of observer 1, one of observer 2, and a consensus score after discussion between both observers. For the interrater reliability the scores of observer 1 and 2 were used. The consensus scores were used for validity purposes and as final scores for the children.

Vineland Adaptive Behavior Scales

The Vineland Adaptive Behaviour Scales (Vineland; Sparrow et al., 1984) is a widely used (valid and reliable) parent interview that measures adaptive behavior. We used the Dutch version (de Bildt & Kraijer, 2003) of this parent interview. Adaptive functioning is represented in three domains in the Vineland: "Communication" (67 items), "Daily living skills" (92 items), and "Socialization" (66 items; Sparrow et al., 1984). The Vineland was administered by trained psychologists, social workers or last-year master psychology students. The "Socialization" domain is the primary outcome measure of the ESTIA study, because this domain is closely related to social skills. Therefore it is also relevant to use this domain and its subdomains ["Interpersonal relationships" (28 items), "Play & leisure time" (20 items), and "Coping skills" (18 items)] to evaluate the validity of the SOM. Although the "Socialization" domain of the Vineland and the SOM aim to measure social skills outside the training situation, the focus of both instruments is different. The Vineland measured different aspects of social skills (e.g. makes friends easily, or apologizes for mistakes) and the SOM is a more general measure about the extent to which the child had contact with other children.

Social Skills Rating Scales (SSRS)

The Social Skills Rating Scale (SSRS; Gresham & Elliott, 1990) is a widely used, valid, reliable and standardized questionnaire that measures social skills pertaining to home and community settings for children at elementary school. It has a 38-item parent-version (SSRS-P), with subscales "Cooperation" (e.g. sharing and helping others), "Assertion" (e.g. initiating behaviors), "Responsibility" (e.g. ability to communicate with adults, and regard for property or work), and "Self-control" (e.g. behaviors that emerge in conflict situations). The 30-item teacher-version resembles the parent version for three subscales, namely: "Cooperation", "Assertion", and "Self-control".

In the current study, the subscale "Assertion" of the parent and teacher version was used to evaluate the validity of the SOM, because the items in this scale fit best with the behaviors in naturalistic situations with peers. However, the items of the subscale "Assertion" of the SSRS as well as the Vineland measured different aspects of social skills, instead of a more general measure about the presence of interactive behavior, with the SOM. Additionally, with the SSRS it is possible to compare consistency between the SOM and the rating of the parents and the rating of the teacher.

ADOS

The Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999; Dutch version, de Bildt & de Jonge, 2008) is a semi-structured observational instrument for children, adolescents, and adults referred for ASD. It focuses on social and communication behavior in structured and unstructured situations (Lord et al., 1999). In the current study, module 3 was administered by ADOS-trained and certified psychologists who fulfilled research requirements of standardized administration and reliability in scoring. The ADOS was administered within 2.5 months before the start of the study, except for 3 children for whom the ADOS was administered within 5 months before the start. We applied the revised algorithms (ADOS-2; Lord et al., 2012): “Social Affect” (SA) and “Restricted and Repetitive behavior” (RRB). For examination of the validity of the SOM, we only used the SA domain, since its content seems most close to that of the SOM.

ADI-R

The Autism Diagnostic Interview-Revised (ADI-R, Rutter et al., 2003; Dutch version, de Jonge & de Bildt, 2007) measures the severity of ASD symptomatology with an extended semi-structured parent-interview and consists of 93 items. All items are coded for current and past behaviour. The algorithm consists of 4 domains: “Social Interaction”, “Communication”, “Restricted, Repetitive, and Stereotypes Patterns of Behavior”, and “Age of First Manifestation”. All ADI-R’s were administered by certified psychologists, who met research requirements of standardized administration and scoring reliability. In order to evaluate the validity of the SOM, we will use the current scores on the domain “Social Interaction” and associated subdomains (“Failure to use nonverbal behaviors to regulate social interaction”, “Failure to develop peer relationships”, “Lack of shared enjoyment”, and “Lack of socioemotional reciprocity”). Therefore, we included only ADI-R assessments that were conducted within five months of the school observation before the treatment or care-as-usual period, discarding data of 9 children who were administered with the ADI-R earlier.

Design and statistics

Figure 1 presents a schedule of the analyses. First, we wanted to limit the analyses to those behaviors that occurred in at least one third of the participants, because behavior needs to be prevalent enough to be observed reliably (in line with Frankel and colleagues, 2011).

Next, the interrater reliability was examined for the *specific behaviors and qualities of behavior* that were prevalent in one third or more of the cases observed by two observers. Additionally, we investigated the interrater reliability on the *total number of children* with whom the child interacted. We calculated the average measures (AM) of the intraclass correlation coefficient (ICC) in a two way random model for absolute

* 'Aggressive or dominant behavior' was included in the analyses (prevalent in 30.8% of the participants)

Figure 1 Analysis schedule for the specific behaviors and the quality of the behaviors of the SOM.

agreement. The ICC interrater reliability was calculated with the sum score of the dichotomous scores. For these analyses, we only included children with all 15 intervals available without imputation, since it is important to use real scores for the interrater reliability ($n=39$). An ICC above .60 is considered good, above .75 it is considered excellent (Cicchetti, 2001).

Additionally, we defined outcome variables of the SOM, depending on the proportion scores of the *specific behaviors* and the *quality of behavior* that were prevalent in more than one third of the participants and coded reliably. Additionally, the total number of children with whom the child interacted was defined as an outcome variable. We used the total number since the numbers per interval gave limited information on whether the children counted in each interval were the same as or different from those in the other intervals. To decide whether the observation in the classroom and the observation on the playground should be analyzed separately or combined, Spearman's rho correlation coefficient was calculated between both situations for the overlapping *specific behaviors*, the overlapping *qualities of behavior*, and the *total number of children* the child interacted with. A Spearman's rho correlation of .40-.59 is considered moderate and between .60-.79 is considered strong. We used a significance level of $\alpha = .05$ for the correlations between the situations.

Finally, to establish the convergent validity of the SOM, we exploratorily calculated correlations between the defined SOM outcome variables and other measures of social

behavior. To reduce the probability of coincidental findings, we used a significance level of $\alpha = .01$.

Results

Prevalence of the specific behaviors and the quality of behavior in the classroom and on the playground

The prevalence of each *specific behavior* and each *quality of behavior* is presented in Table 3. In the playground situation many behaviors occurred in one third of the participants or more, except "Solitary unoccupied", "Aggression verbal/physical", and "Reactive aggression". In the classroom, the specific behaviors "Co-operative game", "Rough / vigorous play", "Locomotor", "Aggression verbal/physical", and "Reactive aggression" were observed in less than one third of the participants. The further analyses were limited to the items occurring in one third or more of the participants, omitting the ones occurring less frequently.

Table 3 Prevalence of behaviors and quality of the behaviors of the SOM.

	Classroom		Playground	
	<i>n</i>	Prevalence	<i>n</i>	Prevalence
Specific behavior				
Solitary unoccupied	85	44.7*	85	21.2
Solitary engaged	85	84.7*	85	49.4*
Solitary onlooker	85	72.9*	85	75.3*
Parallel	85	52.9*	85	65.9*
Co-operative interaction	85	94.1*	85	96.5*
Co-operative game	85	23.5	85	64.7*
Rough / vigorous play	85	9.4	85	38.8*
Locomotor	85	24.7	85	51.8*
Aggression verbal/physical	85	10.6	85	22.4
Reactive aggression	85	7.1	85	15.3
Quality of behavior				
Sub-assertive	78	94.9*	77	93.5*
Aggressive/dominant	78	30.8#	77	57.1*
Assertive	78	96.2*	77	97.4*

* = sufficiently prevalent (33%)

= conceptually important even though infrequently present

For the *quality of behavior* in the classroom, "Aggressive/dominant behavior" occurred in just less than one third (30.8%) of the participants, the other qualities were more prevalent. We decided not to exclude the quality "Aggressive/dominant behavior" from the analyses despite its low prevalence, due to the importance of this quality in the SST that we wanted to evaluate. On the playground, all *qualities of behavior* occurred in more than one third of the cases.

Interrater reliability in the classroom and on the playground

Table 4 presents the ICC between two observers. The interrater reliability is based on real scores, therefore missing intervals were not imputed. Thus we could only include participants with all the 15 intervals available. Additionally, we have only calculated the ICC for the specific behaviors and the qualities of behavior that were prevalent in one third or more of the participants and "Aggressive/dominant behavior".

In the classroom, all *specific behaviors* in the analysis had ICC's ranging from .93-.96 (Solitary engaged .95, Solitary onlooker .94, Parallel .93, Co-operative interaction .96) , except "Solitary unoccupied" (.74). The ICC of all the *qualities of behavior* ranged from .86-.93 (Sub-assertive .86, Aggressive/dominant .93, and Assertive .86). The ICC of the *total number of children* with whom the child interacted was .90.

On the playground, the ICC's of the *specific behaviors* in the analysis ranged from .84-.92 (Solitary onlooker .94, Parallel .84, Co-operative interaction .87, Co-operative game .92, Locomotor .90) . Only "Solitary engaged" (.52) and "Rough and vigorous play" (.72) had lower ICCs. The ICC of the *qualities of behavior* were .66 for "Aggressive/dominant behavior" and .91 for the other two qualities (Sub-assertive and Assertive). The *total number of children* with whom the child interacted on the playground had an ICC of .67.

SOM outcome variables

The results of table 3 and 4 revealed eight *specific behaviors* that were prevalent and reliable. We divided these eight specific behaviors into interactive behaviors and non-interactive behaviors. Table 5 shows which specific behaviors were classified as interactive and non-interactive, and which occurred in which observational situation. Only four of the eight specific behaviors were observed in both situations. The last column of table 5 shows that the Spearman rho correlation of the proportion scores of these specific behaviors in the classroom and in the playground ranged from .14-.22. Because of the low correlation between the two situations, we decided to differentiate between them in the analyses, resulting in four outcome variables of the SOM based on the scores of *specific behaviors*: "Interactive behavior in the classroom", "Non-interactive behavior in the classroom", "Interactive behavior on the playground", and "Non-interactive behavior on the playground".

Table 4 Interrater agreement: Intraclass correlation coefficient in a two way random model.

Specific behavior	Classroom			Playground		
	<i>n</i>	ICC	95% CI	<i>n</i>	ICC	95% CI
Solitary unoccupied	29	.74	.44-.88	-	-	-
Solitary engaged	29	.95	.9-.98	26	.52	-.07-.79
Solitary onlooker	29	.94	.87-.97	26	.94	.88-.98
Parallel	29	.93	.85-.97	26	.84	.64-.93
Co-operative interaction	29	.96	.92-.98	26	.87	.72-.94
Co-operative game	-	-	-	26	.92	.82-.96
Rough / vigorous play	-	-	-	26	.72	.37-.87
Locomotor	-	-	-	26	.90	.78-.96
Aggression verbal/physical	-	-	-	-	-	-
Reactive aggression	-	-	-	-	-	-
Quality of behavior						
Sub-assertive	29	.86	.71-.94	25	.91	.8-.96
Aggressive/dominant	29	.93	.85-.97	25	.66	.24-.85
Assertive	29	.86	.7-.93	25	.91	.79-.96
Total children with whom the child interacted	47	.90	.81-.94	48	.67	.41-.81

ICC, intraclass correlation coefficient (average measure); - behaviors were excluded, because they occurred in less than one third of the participants

Table 5 shows that all *qualities of behavior* were included in the analyses in both situations. The last column of this table shows a Spearman rho correlation of the proportion scores of the quality “Aggressive/dominant behavior” of .44 between the classroom and the playground, .24 for “Sub-assertive behavior”, and .25 for “Assertive behavior”. Based on these weak correlations, we decided to analyze all *qualities of behavior* separately for both situations, resulting in six outcome variables of the SOM based on the *qualities of behavior*: “Sub-assertive behavior in the classroom”, “Aggressive/dominant behavior in the classroom”, “Assertive behavior in the classroom”, “Sub-assertive behavior on the playground”, “Aggressive/dominant behavior on the playground”, and “Assertive behavior on the playground”.

The *total number of children* with whom the child interacted during the observation showed a correlation of .41 between the situations. We added the number of children as two separate subscales, constructing all SOM outcome variables in the same way. This resulted in “number of children in the classroom” and “number of children on the playground”.

Table 5 SOM behaviors observed frequently and reliably.

	Situation			Correlation
	Class	Play	Class & Play	Class & Play
	<i>Interactive behaviors</i>			
Co-operative interaction	X	X	X	.22*
Co-operative game		X		
Rough / vigorous play		X		
	<i>Non-interactive behaviors</i>			
Solitary unoccupied	X			
Solitary engaged	X	X	X	.12
Solitary onlooker	X	X	X	.19*
Parallel	X	X	X	.14
Locomotor		X		
	<i>Quality of behavior</i>			
Sub-assertive	X	X	X	.24*
Aggressive/dominant ¹	X	X	X	.44**
Assertive	X	X	X	.25*
	<i>Total children with whom the child interacted</i>			
	X	X	X	.41**

Class: Classroom; Play: Playground

** Correlation is significant at the .01 level (1-tailed)

* Correlation is significant at the .05 level (1-tailed)

¹ Prevalence in the classroom was 30.8% of the participants and on the playground 57.1% of the participants

Validity

As shown in table 6, the correlations of the SOM outcome variables with the Vineland ranged from .01-.21 and with the ADI-R domains from .00-.27. All were non-significant. SOM "Non-interactive behavior" in the classroom correlated significantly with the Social Affect domain of the ADOS (.35). We found a significant correlation between the subscale "Assertion" of the SSRS-T and SOM "Interactive behavior" (.44), "non-interactive behavior" (-.31), "sub-assertive" (-.34) and "number of children"(.34); all in the classroom. No significant correlations were found between the SSRS-T and the playground observations. On the playground, the SOM outcome variable "number of children the child interacted with" correlated significantly with the subscale "Assertion" (.31) of the SSRS-P.

Table 6 Spearman's rho correlation between subscales SOM and other measures of social behavior.

	ADOS		SSRS-P		SSRS-T	ADI-R current (< 6 months before SOM)					Vineland			
	SA	AS	AS	AS	AS	SOC	NVB	PR	SE	SR	SOC	IR	PLT	CS
Mean (sd)	7.9 (4.0)	10.3 (3.2)	8.6 (4.2)	8.6 (4.2)	7.7 (4.4)	7.7 (4.4)	1.8 (1.3)	1.0 (1.1)	1.7 (1.7)	3.2 (1.8)	80.7 (15.7)	34.7 (7.0)	25.8 (5.5)	20.2 (6.4)
Range	0-20	3-19	0-17	0-17	0-19	0-19	0-4	0-5	0-6	0-7	26-118	8-48	10-39	2-34
Classroom														
Interactive	-.24	-.02	.44**		-.15	-.22	-.22	-.07	.05	-.19	.10	.11	.15	.08
Non-interactive	.35**	-.01	-.31**		.19	.26	.26	-.02	-.06	.27	-.03	-.05	-.02	-.07
Sub-assertive	.01	.06	-.34**		-.06	.10	.10	.01	-.23	-.04	-.01	.07	.03	-.18
Assertive	.00	-.06	.26		.05	-.08	-.08	-.02	.23	.03	.02	-.06	-.01	.16
Aggressive/dominant	-.13	.01	.27		-.02	-.14	-.14	-.08	.03	.00	.06	.04	.12	.07
Number of children	-.07	.07	.34**		.07	-.05	-.05	-.02	.19	.07	-.15	-.14	-.12	-.06
Playground														
Interactive	-.02	.12	.24		-.10	-.18	-.18	-.06	.01	-.11	.05	.14	.01	-.03
Non-interactive	.01	-.23	-.21		.09	.16	.16	.10	.03	.08	-.07	-.14	-.06	.01
Sub-assertive	-.11	-.15	-.16		.01	.17	.17	.11	-.17	-.01	-.01	.01	.02	-.03
Assertive	.15	.07	.17		.05	-.13	-.13	-.02	.21	.02	-.04	-.05	-.11	.01
Aggressive/dominant	-.08	.09	.11		-.15	-.12	-.12	-.20	-.12	-.05	.13	.10	.21	.07
Number of children	-.20	.31**	.21		-.18	-.19	-.19	-.20	-.10	-.11	.13	.13	.12	.11

ADI-R, Autism Diagnostic Interview-Revised (SOC, Total score of Social domain; NVB, Failure to use nonverbal behaviors to regulate social interaction; PR, Failure to develop peer relationships; SE, Lack of shared enjoyment; SR, Lack of socioemotional reciprocity); ADOS, Autism Diagnostic Observation Schedule (SA, Social Affect); sd, standard deviation; SSRS-P, Social Skills Rating Scale-Parents (AS, Assertion), SSRS-T, Social Skills Rating Scale-Teacher (AS, Assertion); Vineland, Vineland Adaptive Behavior Scales (SOC, Total score of Socialization domain; IR, Interpersonal Relationships; PLT, Play & Leisure Time; CS, Coping Skills)
** Correlation is significant at the .01 level (2-tailed)

Discussion

One of the most important limitations in the evaluation of the clinical significance of an SST outside the training for children with ASD is the absence of a blinded primary outcome measure (Freitag et al., 2016). In the current study we built upon various attempts that have been made to develop a direct observation measure in a daily life situation with peers by blind observers, and on the recommendations of previous studies (LeGoff, 2004; Owens et al., 2008; Frankel et al., 2011; Hopkins et al., 2011; Humphrey & Symes, 2011; Shih et al., 2016; Wood et al., 2014). We investigated the psychometric properties of a direct observation method (Social skills Observation Measure; SOM) based on the recommendations from the earlier studies, and found that the SOM seems to be a reliable instrument to observe specific behaviors and the quality of behavior in two naturalistic situations with peers in school for children with ASD. The convergent validity was relatively small. The results indicate added value of the SOM for evaluating social skills in SST effect studies in children with ASD.

The largest coherence we had expected was between the observation and the report of teachers on the child's social skills in school. However, we only found one moderate correlation, between "interactive behavior" of the SOM in the classroom and assertive behavior reported by the teacher. Surprisingly, teacher reports on the SSRS-T did not resemble the observations with the SOM in the playground situation at all. This finding indicates that teachers seem to report mainly on classroom behavior rather than on playground behavior when they are asked to report on social behaviors of their pupils in school. From that perspective, an observation method for the playground situation seems most valuable as a contribution to existing instruments. However, we would recommend to also observe in a classroom situation, since the correlation between SOM and SSRS-T was also weak. Moreover, we would encourage further study into specific classroom behaviors that may improve the classroom observation, thereby increasing the contribution of that situation. Additionally, standardizing the classroom situation to a greater extent than we could, would be helpful.

Compared to another study into direct observation measures of social behavior in children with ASD, our findings are similar. LeGoff (2004) found a moderate correlation between the frequency of self-initiated contact from the observation method and the Social Interaction subscale of the Gilliam Autism Rating Scale (GARS).

Various aspects influence the rating of social behavior of children, also in a broader perspective. The context, the informant, and the focus of measurement can influence the rating (see also De Los Reyes and Kazdin, 2005). The weak coherence found in the current study, between the "socialization" domain of the Vineland and the SSRS-P "Assertion" subscale with the SOM was therefore not unexpected, due to the nature, the difference in focus and the informant of the instruments. The SOM measured the

extent to which the child interacted with other children in a 15 minute situation as observed by an independent observer, whereas the other instruments measure more complex aspects of social skills (e.g. collaboration, bullying) over a longer period of time, reported by parents. We also found a weak coherence between the ADOS and the SOM, even though both instruments involved an observation of social behavior by a trained and independent observer. However, the ADOS is administered in a semi-structured situation with one adult who interacts with the child, whereas the SOM is a naturalistic situation in the classroom or in the playground with peers and a distant, non-involved observer for the SOM. Additionally, the ADOS focuses on many more and other aspects of social behaviors (e.g., eye contact, facial expressions directed to others, quality of social overtures, and amount of reciprocal social communication). The finding that the SOM adds value from a different perspective supports the need to involve multiple informants in the evaluation of social skills training (Murray et al., 2009).

Also important for the value of an instrument is its reliability. All but one aspect of the SOM were reliably rated, i.e. the specific behaviors (except solitary engaged), the quality of behaviors, and the number of children the participating child interacted with. This indicates that the outcome variables of the SOM are consistent and independent from the observer. The interrater reliability found in the current study resembles the reliability reported in previous studies (Hopkins et al., 2011; Cohen's kappa between .74-.95; Frankel et al., 2011, average ICC of .88; Shih et al., 2016, average accuracy of .87; Wood et al., 2014, interrater agreement of .84).

Based on the prevalence of the specific behaviors, we were able to identify two outcome variables: interactive and non-interactive behavior. These resemble other methods of observation in school (Shih et al., 2016; Wood et al., 2014). These two factors may be of value in measuring whether an SST does what it intends: improving socially interactive behaviors and decreasing solitary non-interactive behavior. However, the classroom situation elicited other behaviors than the playground situation, with more behavioral variation in the latter. As a result, the outcome variable 'interactive behavior' in the classroom consists of only one specific behavior (cooperative interaction). This indicates that interactive behavior is less present in a classroom situation, even when it is an unstructured situation. The fact that the behaviors were more varied and prevalent in the playground situation may be due to the fact that the observation was based on Frankel and colleagues (2011) and Humphrey and Symes (2011), who observed in a playground situation only. To increase the range of specific social behaviors to be observed in a classroom situation, specific behaviors could be added to the observation (e.g. number of initiatives of the child, collaboration), replacing the more physical behaviors seen in the playground. These findings indicate that observing children in two different situations has the advantage of being able to see a broader spectrum of behaviors,

corroborating earlier recommendations to observe in different situations (Hopkins et al., 2011).

A remarkable finding was the infrequent occurrence of specific aggressive behaviors, being prevalent in only 7.1-22.4% of the participants. Additionally, in terms of quality of behavior, "Aggressive/dominant behavior" occurred least of all the qualities rated, and more often in the playground (57.1% of the participants) than in the classroom (30.8% of the participants). However, evaluating aggressive behaviors is important when investigating the effect of an SST, since one of the aims of an SST is to replace possible aggressive behaviors with social or assertive behaviors. Therefore, it should be included in an observation method. In the current study, we decided to not include the specific aggressive behaviors in the reliability analyses, in the outcome variables or in the validity analyses. However, we have included the quality "Aggressive/dominant behavior" in all analyses for both situations, despite its relatively low prevalence in the classroom, in order to address aggressive behaviors in school. In order to better measure aggressive behaviors in future studies, the observation could either be prolonged to prevent aggressive behavior from not occurring due to the shortness of the observation; a third, less structured, situation could be chosen (for example cooperating on a specific project without supervision); or the method of observing aggressive behaviors could be altered (e.g. observer closer to the child, prolonging intervals, camera's instead of direct observers). The fact that aggressive behavior did not occur very frequently indicates that merging the aggressive items from the observation of Humphrey and Symes (2011) in our observation did not obscure our findings.

The quality of the behavior codings in the SOM correspond in part with the observation codes used by Hopkins and colleagues (2011) and Wood and colleagues (2014). The SOM was able to reliably measure the assertive and subassertive quality of behaviors. Although aggressive behaviors occurred infrequently, even this quality still seems to be measured reliably with the SOM. Thus the SOM seems informative for measuring the quality of behaviors as taught in an SST.

In the current study, we are unable to draw conclusions on the character of the relationship between the participating child and the other children. Frankel and colleagues (2011) recommended to observe the number of children the child interacted with, to increase the insight in the acceptance of the participating child in a larger group. However, although we have information about how many children the participating child interacted with, we don't have information on whether they were real friends and were together during the whole break, or whether the participating child interacted with many children more superficially. Moreover, children with ASD may approach many children frequently without taking the interaction somewhere. Oblivious to rejection, they just continue interacting. It would be interesting to observe typically developing children of the same age on this variable, and compare children

with ASD to them. Perhaps typically developing children have longer, more reciprocal interactions with fewer children, whereas children with ASD may have many single superficial interactions. Such future research would shed light onto the usefulness of this variable in a direct observation of social behavior in a naturalistic setting with peers. Information on the nature of relationships between the participating children and those they interact with is needed, as well as on the content of the observed interaction. Additionally, investigations into whether the number of children the participating child interacts with is sensitive to change after an SST and whether it is related to social functioning at all are required.

Limitations

Some aspects of the SOM could be improved or need further investigation. The timeframe of 15 minutes we aimed for was not feasible for all children, due to the fact that the observation needed to fit in with the regular schedule. However, we were able to observe relevant behaviors in shorter timeframes as well. Therefore 15 minutes seems a good aim, and we would recommended to strive for the total of 15 minutes if possible.

Due to the feasibility issue of 15 minutes, we have imputed data for the validity study, based on mean value. We did so in 24 observations from the classroom and 30 observations from the playground. We were aware that mean value imputation results in a larger number of participants with identical scores. This obviously affects the outcome of the analyses because of smaller variance and inflation of the apparent precision of the estimates (Newgard & Lewis, 2015). However, this inflation increases with increasing proportions of imputed data. In our study, only 5.85% of the specific behaviors and 6.33% of the quality of the behaviors were imputed. Therefore we decided to use the imputed data for the validity study in order to prevent the alternative bias occurring by losing too many observations.

With respect to the two situations of observation, we found that the degree of freedom during classroom observations was not fully comparable over all schools. Some schools showed difficulties in creating an unstructured situation in the classroom and asked the children to draw something or to play a game. In other schools, children were completely free to choose what they wanted to do and with whom. The unstructured and unstandardized situations may have led to differences between children as coded during the observation, although the outcomes may not have been attributable to differences between the children, but more so to differences between the situations. Although in all classroom situations the children were allowed to freely speak or otherwise interact with each other, we would recommend a more standardized situation in the classroom, that more explicitly stimulates social interaction. Additionally, the situations in the playground may have varied as well, because of available play materials, the form of the playground, the number of

children at the site, the weather, etc. Thus, perhaps a standardized situation would be better at the playground as well. However, the balance between standardized and naturalistic should be carefully kept in mind.

In order to observe with blind, non-involved observers, we trained observers in the observation before they carried out observations by themselves. However, after the training we did not measure interrater reliability. We only sent observers to schools if they had shown understanding of the items and good coding during exercise sessions. Nevertheless, despite the fact that the interrater reliability in current study was good to excellent, we would recommended to measure interrater agreement before performing the real observations in future studies.

Although we aimed to have observers blind for treatment allocation, unfortunately this was only feasible in the majority of cases, for logistical reasons as described in the methods section. Of the observations, we do not have clear information on which observation were blind for treatment allocation and which were not. Future studies need to record this or, even better, strive to blind all observations.

A possible limitation of the SOM is that no video recordings were made, which means that we were unable to reassess the data. However, it was a considered choice, because the camera would then follow one specific child leading to a much less naturalistic situation. Besides, receiving permission to record whole classes would have been unfeasible. Therefore, the observer stood at an appropriate distance from the child, even though this resulted in the fact that not all conversations could be understood.

The SOM was only used in children with ASD. Therefore, nothing is known about the prevalence of the specific behaviors and the quality of the behaviors in typically developing children and in children with other mental health disorders. Future research may examine the measure in various groups, and provide norm scores.

Conclusion

The development and investigation of the SOM in the current study is an additional step in the search for instruments to evaluate social skills of children with ASD in a naturalistic situation. Incorporating recommendations from the literature we have reached an observation method with which we observed children with ASD for fifteen consecutive minutes in two naturalistic situations within their own peer group. The SOM seems to be of added value to existing instruments in measuring social skills. All outcome variables of the SOM were observed reliably and showed similarities with recently published observation measures at school that evaluate an SST (Hopkins et al., 2011; Wood et al., 2014; Shih et al., 2016). Due to their resemblance to recently published methods, we expect the outcome variables to be sensitive to change during an SST, similar to these methods. The SOM had three advantages compared to the existing observation measures, based on the recommendations of earlier

research: 1) the observation took place in two naturalistic situations; 2) the observer was blind to treatment allocation in the majority of cases; and 3) the observer was not involved in the treatment and was unfamiliar to the participating child. Moreover, the SOM complements existing measures of social behaviors reported by parents and teachers, which are usually used as instruments to evaluate social skills in children with ASD.

Ethical approval

This study was approved by the ethics committee of the Institutional Review Board of the University Medical Center Groningen. Therefore, all procedures performed in this study were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was registered in the Dutch trial register (NTR2405).

Disclosure of potential conflict of interest

Author AdB received a grant by the Netherlands Organization for Health Research and Development (ZonMw, nr 157003005) which funded this study. The agency had no role in the data analyses, the decision to publish, or the preparation of the manuscript. EJM is in the board of directors of Accare. However, he had no financial interest in the outcomes of the study. The other auteurs declare that they have no conflict of interest.

Acknowledgments

The authors gratefully acknowledge all the children, their parents, teachers, and therapists participating in this study.

References

- Cicchetti, D.V. (2001). The precision of reliability and validity estimates re-visited: distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clinical and Experimental Neuropsychology*, 23(5), 695–700.
- de Bildt, A., & de Jonge, M.V. (2008). *Autisme diagnostisch observatie schema*. Amsterdam: Hogrefe.
- de Bildt, A.A., & Kraijer DM (2003). *Sociale redzaamheid voor kinderen en jeugdigen met een verstandelijke beperking*. Vineland-Z. Leiden: PITS B.V.
- de Jonge, M.V., & de Bildt, A. (2007). *Autisme diagnostisch interview-revised*. Amsterdam: Hogrefe.
- De Los Reyes, A. & Kazdin, A.E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, 131(4), 483–509.
- Dekker, V., Nauta, M.H., Mulder, E.J., Timmerman, M.E., & de Bildt, A. (2014). A randomized controlled study of a social skills training for preadolescent children with autism spectrum disorders: generalization of skills by training parents and teachers? *BMC Psychiatry*, 14, 189.
- Dekker, V., Nauta, M. H., Timmerman, M. E., Mulder, E. J., van der Veen-Mulders, L., van den Hoofdakker, B. J., van Warners, S., Vet, L., Hoekstra, P. J., & de Bildt, A. (2019). Social skills group training in children with autism spectrum disorder: a randomized controlled trial. *European Child & Adolescent Psychiatry*, 28(3), 415–424.
- Frankel F.D., Gorospe, C.M., Chang, Y., & Sugar, C.A. (2011). Mothers' reports of play dates and observation of school playground behavior of children having high-functioning autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 52(5), 571–579.
- Freitag, C.M., Cholemkery, H., Elsuni, L., Kroeger, A.K., Bender, S., Kunz, C.U., & Kieser, M. (2013). The group-based social skills training SOSTA-FRA in children and adolescents with high functioning autism spectrum disorder – study protocol of the randomised, multi-centre controlled SOSTA – net trail. *Trials*, 14, 6.
- Freitag, C.M., Jensen, K., Elsuni, L., Sachse, M., Herpertz-Dahlmann, B., Schulte-Rüther, M., Hänig, S., Von Gontard, A., Poustka, L., Schad-Hansjosten, T., Wenzl, C., Sinzig, J., Taurines, R., Geißler, J., Kieser, M., & Cholemkery, H (2016). Group-based cognitive behavioral psychotherapy for children and adolescents with ASD: the randomized, multicentre, controlled SOSTA-net trial. *Journal of Child Psychology and Psychiatry*, 57(5), 596–605.
- Gresham, F.M., & Elliott, S.N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Hopkins, I.H., Gower, M.W., Perez, T.A., Smith, D.S., Amthor, F.R., Wimsatt, F.C., & Biasini, F.J. (2011). Avatar Assistant: improving social skills in students with an ASD through a computer-based intervention. *Journal of Autism and Developmental Disorders*, 41(11), 1543–1555.
- Humphrey, N., & Symes, W. (2011). Peer interaction patterns among adolescents with autistic spectrum disorders (ASDs) in mainstream school settings. *Autism*, 15(4), 397–419.
- Koenig, K., White, S.W., Pachler, M., Lau, M., Lewis, M., Klin, A., & Scahill, L. (2010). Promoting social skills development in children with pervasive developmental disorders: a feasibility and efficacy study. *Journal of Autism and Developmental Disorders*, 40(10), 1209–1218.
- LeGoff, D.B. (2004). Use of LEGO as a therapeutic medium for improving social competence. *J Journal of Autism and Developmental Disorders*, 34(5), 557–571.
- Lopata, C., Thomeer, M.L., Volker, M.A., Toomey, J.A., Nida, R.E., Lee, G.K., Smerbeck, A.M., & Rodgers, J.D. (2010). RCT of a manualized social treatment for high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 40(11), 1297–1310.
- Lord, C., Rutter, M., DiLavore, P.C., & Risi, S. (1999). *ADOS. Autism Diagnostic Observation Schedule. Manual*. Los Angeles: Western Psychological Services.
- Lord, C., Rutter, M., DiLavore, P.C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, second edition (ADOS-2) manual (Part I): Modules 1-4*. Torrance, CA: Western Psychological Services.
- McMahon, C.M., Lerner, M.D., & Britton, N. (2013a). Group-based social skills interventions for adolescents with higher-functioning autism spectrum disorder: a review and looking to the future. *Adolescent Health, Medicine and Therapeutics*, 4, 23–38.

- McMahon, C.M., Vismara, L.A., & Solomon, M. (2013b). Measuring changes in social behavior during a social skills intervention for higher-functioning children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(8), 1843–1856.
- Murray, D.S., Ruble, L.A., Willis, H., & Molloy, C.A. (2009). Parent and teacher report of social skills in children with autism spectrum disorders. *Language, speech, and hearing services in schools*, 40(2), 109–115.
- Newgard, C.D. & Lewis, R.J. (2015). Missing data: How to best account for what is not known. *JAMA*, 314(9), 940–941.
- Owens, G., Granader, Y., Humphrey, A., & Baron-Cohen, S. (2008). LEGO therapy and the social use of language programme: an evaluation of two social skills interventions for children with high functioning autism and Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 38(10), 1944–1957.
- Reichow, B., Steiner, A.M., & Volkmar, F. (2013). Cochrane review: social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD). *Evidence-Based Child Health*, 8(2), 266–315.
- Reichow, B., & Volkmar, F.R. (2010). Social skills interventions for individuals with autism: evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, 40(2), 149–166.
- Rutter, M., le Couteur, A., & Lord, C. (2003). *ADI-R Autism Diagnostic Interview Revised Manual*. Los Angeles: Western Psychological Services.
- Shih, W., Patterson, S.Y., & Kasari, C. (2016). Developing an adaptive treatment strategy for peer-related social skills for children with autism spectrum disorders. *Journal of Clinical Child & Adolescent Psychology*, 45(4), 469–479.
- Sparrow, S.S., Balla, D., & Cicchetti, D.V. (1984). *Vineland adaptive behavior scales (Survey ed.)*. Circle Pines, MN: American Guidance Service.
- White, S.W., Koenig, K., & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: a review of the intervention research. *Journal of Autism and Developmental Disorders*, 37(10), 1858–1868.
- Wood, J.J., Fujii, C., Renno, P., & van Dyke, M. (2014). Impact of cognitive behavioural therapy on observed autism symptom severity during school recess: a preliminary randomized, controlled trial. *Journal of Autism and Developmental Disorders*, 44(9), 2264–2276.

6 |

Parental stress and criticism of fathers and mothers of children with autism spectrum disorder in association with co-occurring internalizing and externalizing problems

Vera Dekker

Maaïke H. Nauta

Pieter J. Hoekstra

Annelies de Bildt

Abstract

Parental stress and criticism may influence development and maintenance of comorbid behaviors in autism spectrum disorder (ASD). In relatively high functioning children with ASD ($n=122$, aged 9-13 years), differences in parental stress and criticism between parents were examined, and their association with ASD symptoms and internalizing and externalizing problems.

Father and mother mean stress levels were comparable, and below average to average for a clinical sample. Few parents (9.6%) were critical. Parental stress, but not criticism, was associated with externalizing and internalizing child behaviors. None were related to ASD severity. Mothers and fathers had similar patterns of results. Clinical practice should consider the relation between parental stress and comorbid behaviors, even in relatively high functioning children and mildly stressed parents.

Introduction

Children with autism spectrum disorder (ASD) often experience emotional and/or behavioral problems (Lai et al., 2019; Chandler et al., 2016; Maskey et al., 2013; Simonoff et al., 2008), negatively affecting their development and functioning (Joshi et al., 2010; Gillberg et al., 2016; Helles et al., 2017; Chiang et al., 2016). Parental stress and the emotional quality of the parent-child relationship have been associated with the development or maintenance of such problems in children with ASD (Romero-Gonzalez et al., 2018; Bader & Barry, 2014; Bader et al., 2015; Baker et al., 2011). The current study builds upon these earlier findings by investigating how paternal and maternal stress and the emotional quality of the father-child and mother-child relationship are associated with each other, and with ASD symptoms, externalizing and internalizing behaviors.

Stress in parents of children with ASD has been reported to be high (Hayes & Watson, 2013; Giovagnoli et al., 2015; Gong et al., 2015) and to be higher than in parents of typically developing children or children with other disabilities (Hickey et al., 2019a; Estes et al., 2013). Various challenges that come with raising a child with ASD have been associated with parental stress, as parents need to find their way in managing the child's ASD symptoms and co-occurring behavior problems (Maskey et al., 2013; Valicenti-McDermott et al., 2015; Hastings et al., 2005; Hastings et al., 2006; Lounds et al., 2007), healthcare (Nesbitt, 2000), (special) education (Whitaker, 2002), and financial challenges (Sharpe & Baker, 2007). Parental stress is reported to be higher in mothers of children with ASD than in fathers (Hickey et al., 2019a; Rodriguez et al., 2019). In a meta-analysis of 34 studies on parental stress and behavior problems in children with ASD (Yorke et al., 2018), parental stress has shown to be moderately and bidirectionally associated with internalizing and externalizing behavior problems. ASD severity accounted for the relationship between behavior problems and parental stress in some studies (Bader et al., 2015; Falk et al., 2014; Firth & Dryer, 2013), yet not in others (Giovagnoli et al., 2015; Hastings et al., 2005; Manning et al., 2011; McStay et al., 2014; Weiss et al., 2015; Zaidman-Zait et al., 2014). Regarding the direction of the association between parental stress and child behaviors, in a recent study (Rodriguez et al., 2019), parental stress predicted internalizing child behaviors over time, most prominently stress of mothers. For externalizing behaviors, a change in direction was found over time for paternal stress: starting with paternal stress being predictive of externalizing behaviors, followed by externalizing behaviors being predictive of paternal stress. Maternal stress was steadily predictive of externalizing behaviors and not the other way around (Rodriguez et al., 2019).

The various challenges that parents experience when raising their child with ASD may also affect the emotional quality of the relationship with their child (Bader et al.,

2015). To investigate the association between the emotional quality of the parent child relationship and internalizing and externalizing behaviors of the child, most studies in ASD have focused on parental criticism, as a measure of the emotional valence (Bader and Barry, 2014). Research has indicated that criticism in parents is higher towards children with ASD than towards other children of the same parent (Griffith et al., 2015). High levels of criticism in parents of children with ASD was related to externalizing behavioral problems in a recent meta-analysis (Romero-Gonzalez et al., 2018). This association was also found after controlling for ASD symptom severity, parental distress, parenting practice, and demographic variables related to child externalizing behaviors (Bader et al., 2015). Some studies suggest a directionality from parental criticism to later externalizing behaviors in children with ASD (Baker et al., 2011; Bader & Barry, 2014), while a recent study reported the opposite: ASD symptom severity and externalizing behavioral problems predicted criticism in mothers (and not fathers) at a later time, while both maternal and paternal criticism were not predictive of problem behaviors of the child (Hickey et al., 2020). These differences may be due to age differences between the samples, as the latter study was conducted in a younger sample (age 5-12 years) and the earlier studies reported on a broader age range (6-18 years; Bader & Barry, 2014) or older age range (11-49 years; Baker et al., 2011), pointing to a change in directionality of the association from young childhood to later in life.

Findings on the association between parental criticism and internalizing behaviors in children with ASD are less consistent (Romero-Gonzalez et al., 2018), although some indication exists for such an association in adolescents and adults with ASD (Greenberg et al., 2006). More studies are needed to clarify this association (Bader & Barry, 2014). Most findings on parental criticism were based on mothers' reports (Romero-Gonzalez et al., 2018) and only few studies included both mothers and fathers. These studies suggest that, in general, mothers express higher levels of criticism than fathers (Hickey et al., 2019b; Hickey et al., 2020).

To our current knowledge, there have been only two studies that investigated the association between parental stress and parental criticism in parents of children with ASD (Hickey et al., 2019a; Sher-Censor et al., 2018). One study (Hickey et al., 2019a), in school-aged children (5-12 years) showed in mothers and fathers alike that parenting stress predicted the level of criticism towards the child with ASD, 12 months later. However, that study only controlled for ASD symptoms and behavior problems of the child, but did not investigate the relation between the parental factors and ASD symptoms and externalizing and internalizing behaviors of the child. The other study (Sher-Censor et al., 2018), in toddlers, was cross-sectional and only included mothers. They found that criticism was associated with maternal stress while both were related to externalizing behaviors of the child. Only maternal stress was related to internalizing behaviors. It may well be that the combination of two parents with high stress levels

or high criticism may be stronger associated with externalizing and internalizing behaviors in children with ASD.

The objective of the current study was to investigate the association between parental stress and parental criticism towards the child in and between mothers and fathers, and between these parental factors and ASD symptoms and externalizing and internalizing behaviors of the child. First, we aimed to investigate cross-sectional associations of maternal and paternal stress with maternal and paternal criticism, and how these factors and associations differed between mothers and fathers. Second, we examined the cross-sectional associations between these parental factors and ASD symptoms and externalizing and internalizing behaviors of the child, and tested whether differences existed between maternal and paternal associations. Additionally, we aimed to exploratively examine whether having two critical parents or two parents with a high degree of stress was related to the child's problems.

Based on preceding studies, we hypothesized that reported stress of fathers and mothers would be high (Hickey et al., 2019a), that mothers would report higher levels of stress than fathers (Hickey et al., 2019a; Rodriguez et al., 2019), and express more criticism (Hickey et al., 2019b). For both mothers and fathers, we hypothesized a positive association between stress and criticism (Hickey et al., 2019a), a positive association between parental stress and the child's ASD symptoms (Rodriguez et al., 2019) as well as internalizing and externalizing behaviors (Yorke et al., 2018; Rodriguez et al., 2019; Sher-Censor et al., 2018) and between parental criticism and the child's ASD symptoms (Greenberg et al., 2006), and externalizing behaviors (Romero-Gonzalez et al., 2018; Greenberg et al., 2006; Bader & Barry, 2014; Bader et al., 2015; Baker et al., 2011; Baker et al., 2019; Sher-Censor et al., 2018). We also hypothesized a positive association of parental criticism with the child's internalizing behaviors, yet weaker than with externalizing behaviors (Rea et al., 2020). Previous research on this association is inconsistent (Romero-Gonzalez et al., 2018; Rea et al., 2020) and based on adults with ASD (Greenberg et al., 2006) and children with an intellectual disability and ASD (Hastings et al., 2006).

Methods

Design

The current study used a cross-sectional design, based on data from the first measurement of a randomized controlled trial on the effectiveness of group social skills training with and without parental and teacher involvement, compared to care as usual (Dekker et al., 2014; Dekker et al., 2019). The Institutional Review Board of the University Medical Center Groningen approved the study. Informed consent was signed by all parents and children above 12 before participation. The study was preregistered in the Dutch Trial Register (NTR2405; <http://www.trialregister.nl>).

Participants

In the current study, 122 preadolescent high-functioning children (103 boys, 19 girls) with a clinical ASD diagnosis participated. They were recruited from four clinics for outpatient mental health care in the Northern Netherlands. Children were included in the study if: 1) the child had a clinician-based DSM-IV-TR ASD diagnosis (Autistic disorder, Asperger's disorder, or Pervasive Developmental Disorder-Not Otherwise Specified [PDD-NOS]), established in an expert team including at least a child psychiatrist and a child psychologist, and based on information on developmental history, current problems, child observation (ADOS), and information from school; 2) social skills training was indicated by the child's clinician as the next appropriate treatment in stepped care, based on parent and child reported social problems in school and other situations; 3) motivation of parents and child for social skills training had been established during a meeting between the child, the parents, and the clinician; 4) IQ was preferably 80 or above; although children with IQs at a level that was slightly below 80 participated when their clinician established that the child was able to participate in social skills training; 5) the child was in the last two and half years of primary education; 6) the child had no known physical condition that affected participation; and 7) it was possible for the child to travel to the clinic for training. For more information on the design and outcomes of the original study, please see Dekker and colleagues (2019).

Of the 122 participants, 17 were clinically diagnosed with DSM-IV-TR autistic disorder (14%), 25 with Asperger's disorder (20%), and 80 with PDD-NOS (66%). In 32% of the children one comorbid diagnosis was reported, in 4.1% two comorbid diagnoses were reported. Diagnosed disorders were Attention Deficit Hyperactivity Disorder ($n=24$; 19.7% of all participants), Tic disorder ($n=8$; 6.5%), Anxiety Disorder ($n=4$; 3.2%), Oppositional Defiant Disorder ($n=4$; 3.2%), and another diagnosis ($n=4$; 3.2%). The age of the participants in this study ranged from 9.5-13.0 years, with a mean of 11 years ($SD = .75$).

Eighty-eight of the children had two parents of Dutch descent, 33 had one Dutch parent. Ninety-one children lived with both parents. In 24.6% (30 participants) of the families the parents were divorced, of which 13 children lived with both parents alternatively (co-parenting), and 17 lived entirely with their mother. Of the 17 children who lived entirely with their mother, 10 children had no contact with their father. Most children grew up in a family with siblings (66 (54.1%) with one sibling, 27 (22.1%) with two siblings, 17 (13.9%) with three siblings, 2 (1.6%) with four siblings, 1 grew up with siblings, but the number was unknown).

Instruments

Parenting Stress (Parenting Stress Index short form; PSI-SF)

The Dutch version of the Parenting Stress Index - Short Form (PSI-SF, Abidin, 1995; de Brock et al., 1992) measured the level of parental stress. Both biological parents scored 25 statements about parenting stress on a 6-point Likert scale (ranging from 1 (totally disagree) to 6 (totally agree)). In case a parent had no contact with their child at all, the PSI-SF was not administered. The PSI-SF has two major domains, a parent and a child domain. The parent domain measures stress in parents and the effect on their behavior (for example: 'parenting this child is more difficult than I thought it would be' and 'Often I feel like giving up on it'), the child domain reflects behaviors of the child that may be stress evoking in parents (for example: 'if I forbid my child to do something, he will do it again later' and 'my child's attention wanders more often than I expected'). In accordance with the advice for future research by Yorke and colleagues (2018), in the current study the parent domain (10 items, range 10 to 60) was used to evaluate parental stress, because some items of the child domain overlap with the symptom measures used for measuring child factors. However, for a comparison of the level of stress with that of other parents (normal and clinical population), we used total scores including the child domain, since norms are only available for that total score, which enables a comparison. In the normal population Dutch population, the average score (the middle 30%) for mothers is between 43-61 and for fathers between 39-53. This means that 35% of the parents scores below 43 (mothers) or 39 (fathers) and another 35% scores above 61 or 53. In a clinical Dutch population, the average score (the middle 30%) for mothers is between 73-97 and for fathers between 59-79 (de Brock et al., 1992). Criterion validity and reliability of the PSI-SF are high (de Brock et al., 1992).

Parental criticism (Five Minute Speech Sample; FMSS)

The Five Minute Speech Sample (FMSS) measured parental criticism (Magaña et al., 1986). The FMSS was introduced to parents with the following explanation: *'I'd like to hear your thoughts and feelings about (child's name), in your own words and without me interrupting you with any questions or comments. When I ask you to begin, I'd like you to speak for 5 minutes, telling me what kind of person (child's name) is and how the two of you get along together. After you begin to speak, I prefer not to answer any questions until after 5 minutes. Do you have any questions before we begin?'*

The fragments were recorded and coded afterwards with respect to the first statement about the child, the relationship with the child, the number of critical remarks, the number of positive remarks, attitude, emotions, excessive details about the past and self-sacrifice or overprotective behavior toward the child. The current study focused on 'Criticism', because this aspect is most associated with various child characteristics (Romero-Gonzalez et al., 2018; Hickey et al., 2020; focus of this

paper). Criticism is high when the initial statement is negative (other possibilities for the initial statement are neutral or positive), the relationship is described as negative (other possibilities for the relationship are neutral or positive), or when the parent has made one or more critical comments about the child during the FMSS. In all other cases, there is no criticism. The FMSS was administered to both biological parents separately, as long as they had regular contact with their child. In case a parent had no contact with their child at all, the FMSS was not administered. The fragments were coded by independent raters and randomly divided over the raters, that were different for the mother and father of the same child. The raters were psychologists or master students in psychology. Before coding, the raters received training that included discussion of the manual of the FMSS with the instructions for coding, and practice fragments were coded and discussed with a certified supervisor. If the rater had difficulties or questions about the coding of a fragment in the study, the fragment was coded by other raters and discussed in the presence of the FMSS certified supervisor to reach consensus. We assessed inter-rater reliability of "criticism" by double coding 10% of FMSS samples ($n=24$; $n=13$ of mothers and $n=11$ of fathers) by an independent rater. For mothers the agreement of 'criticism' was 76.9% and for fathers 90.9%. The inter-rater reliability corresponded to those of other studies (Griffith et al., 2015; Benson et al., 2011; Baker et al., 2011; Hickey et al., 2019a; Hickey et al., 2019b).

Externalizing behaviors (Swanson, Nolan, and Pelham questionnaire; SNAP-IV)

The Swanson, Nolan, and Pelham questionnaire (SNAP-IV; Swanson, 1997) is a standardized questionnaire about Attention Deficit Hyperactivity Disorder (ADHD) and Oppositional Defiant Disorder (ODD) behaviors. Parents rated the frequency of behaviors (26-items) based on the past week on a 4-point Likert scale (0 for never to 3 very clear). We used the dimensional scores of the three different subscales: *Inattention* (9 items, range 0 to 27), *Hyperactivity-impulsivity* (9 items, range 0 to 27), and *Oppositional Defiant Disorder (ODD)* (8 items, range 0 to 24) of the parent-version in the regression analyses. To provide an indication of the severity of comorbid problems in the current sample, we used available cut-off scores (95th percentile score), as cut-off for high level of behaviors, namely 16.02 for Inattention behaviors, 12.96 for Hyperactivity-Impulsivity behaviors, and 15.04 for ODD behaviors (Swanson, 1997). The SNAP-IV has acceptable internal consistency and acceptable reliability (Bussing et al., 2008).

Internalizing behaviors (Revised Child Anxiety and Depression Scale; RCADS)

The Revised Child Anxiety and Depression Scale (RCADS; Chorpita et al., 2000; van Oort et al., 2009) is a standardized questionnaire about anxiety and depression behaviors. We used the child version, in which children rated the frequency of 47 behaviors on a 4-point Likert scale (0 for never to 3 always). We used the depression

subscale (10-items, range 0 to 30) and the total score of the four anxiety subscales (discarding the OCD-subscale, in line with DSM-5: Generalized Anxiety Disorder, Social Phobia, Separation Anxiety Disorder, Panic Disorder; 31-items, range 0 to 93). We used the dimensional scores in the regression analyses. To provide an indication of the severity of comorbid problems in the current sample, we used available cut-off scores (80th percentile) of the RCADS subscales anxiety and depression, separately for boys and girls in a large Dutch cohort of young adolescents ($n=2230$, mean age 11.09 years ($SD=.56$), the TRAILS-study; Huisman et al., 2008), as cut-off for low versus high level of behaviors. The 80th percentile for the anxiety behaviors was 23 for boys and 27 for girls and for depression the 80th percentile was 9 for both boys and girls. The child version of the RCADS showed favorable psychometric properties in a clinical sample (Chorpita et al., 2005).

ASD severity

The severity of ASD symptomatology was measured with the Autism Diagnostic Observation Schedule (ADOS) module 3 (Lord et al., 1999; Lord et al., 2012; de Bildt & de Jonge, 2008) and the Autism Diagnostic Interview-Revised (ADI-R; Rutter et al., 2003; de Jonge & de Bildt, 2007), by trained psychologists who met research requirements for reliability. The “Calibrated Severity Scores” (CSS; Lord et al., 2012) of the ADOS were used in the analyses, as a measure of symptom severity independent from the parent.

The ADI-R was administered to describe the sample based on the total score as well as the domains “Social Interaction” (SI), “Communication” (Com), and “Restricted, Repetitive, and Stereotyped Patterns of Behavior” (RRB).

Cognitive ability

Cognitive ability was assessed with the Dutch version of the Wechsler Intelligence Scale for Children – 3d edition (WISC-III; Wechsler, 1999; Kort et al., 2005).

Analyses

Difference between fathers and mothers on parental stress and criticism

We used a paired sample t-test to investigate whether mothers and fathers differed in levels of stress. To investigate the differences between fathers and mothers on criticism, we used the chi-square test for dichotomous measures. In order to investigate the association between parental stress and criticism in fathers and in mothers, we used a t-test to examine a difference in stress level of fathers/mothers with criticism compared to fathers/mothers without criticism.

The association between parental stress, criticism, and child behaviors

First, we investigated the associations between parental stress or criticism on the one hand and child behaviors (ASD symptoms, ADHD, ODD, Anxiety, and Depression behaviors) on the other. For parental stress, we did so with Pearson r correlations. For criticism, we used t -tests to examine a difference in child behaviors of mothers with criticism compared to mothers without criticism, and of fathers with and without criticism.

Second, we performed multiple regression analyses to investigate the association between parental stress and criticism and the child behaviors. We performed separate regression analyses for each child behavior (dependent variable), with maternal stress and criticism, paternal stress and criticism, and the interaction between maternal stress and paternal stress, and between maternal criticism and paternal criticism. Before the calculation of the interaction variable we centered both stress variables. Regression analyses were performed with the ENTER method and were only reported if they were significant.

We did not include age, sex, IQ or SES in our analyses due to the homogeneity of our sample on these variables: the sample had a limited age range (9.5-13 years; of which 81.1% were 10 or 11 years old), was predominantly male (84.4%), only 18% had a total IQ below the 85, and only 12% had a low income.

Results

Missing data

The FMSS was collected from 121 mothers and 107 fathers. One maternal recording was lost due to technical problems with recording; 15 paternal recordings were missing due to no regular contact between father and child ($n=10$), technical problems with recording ($n=4$), and decline to participate ($n=1$). The PSI-SF questionnaire was collected from 118 mothers (3 missing due to completion by father instead of mother, 1 due to completion by both parents together), and from 111 fathers (10 fathers had no regular contact with their child and 1 was completed by both parents together).

Descriptives

The mean total IQ of the participants was 101.5 ($SD = 15.3$; range 72-135). ASD severity as measured with the CSS of the ADOS was moderate (Mean = 5.5; $SD = 2.3$), which was corroborated with relatively low scores on the ADI-R total score (Mean = 31.2; $SD = 10.7$), the "Social Interaction" domain (Mean = 14.2; $SD = 5.6$), the "Communication" domain (Mean = 11.9; $SD = 4.6$), and the "Restricted, Repetitive, and Stereotyped Patterns of Behavior" domain (Mean = 3.2; $SD = 2.1$).

Table 1 Differences between parents (fathers and mothers) with and without criticism

		Total sample		Criticism father		No criticism father		t-test fathers criticism (criticism vs no criticism)		Criticism mother		No criticism mother		t-test mothers criticism (criticism vs no criticism)	
		Mean (SD)	Range	Mean (SD)	n=10	Mean (SD)	n=97	t (p-value)	Mean (SD)	n=12	Mean (SD)	n=109	t (p-value)	Mean (SD)	n=109
Stress father (PSI-SF)	(n=111)	24.9 (9.0)	10-48	28.9 (11.4)		24.3 (8.7)		-1.53 (.13)	25.7 (9.7)		24.9 (9.1)		-.25 (.80)		
Stress mother (PSI-SF)	(n=118)	25.0 (10.1)	10-56	24.1 (11.7)		25.6 (10.1)		.45 (.65)	25.0 (10.3)		25.1 (10.1)		.04 (.97)		
Externalizing behavior (SNAP-IV-P)	(n=122)	12.8 (6.5)	1-26	12.8 (7.9)		12.9 (6.5)		.03 (.97)	14.7 (5.1)		12.6 (6.6)		-1.09 (.28)		
		9.7 (5.9)	0-27	7.2 (5.2)		9.8 (5.7)		1.39 (.17)	10.2 (6.1)		9.7 (5.9)		-.278 (.78)		
		7.8 (5.8)	0-23	6.6 (7.4)		7.7 (5.6)		.59 (.56)	8.1 (6.2)		7.8 (5.8)		-.17 (.87)		
Internalizing behavior (RCADS-C)	(n=121)	7.9 (3.5)	0-18	7.8 (4.4)		7.7 (3.5)		-.10 (.92)	7.9 (5.6)		7.9 (3.3)		-.06 (.95)		
		18.4 (11.4)	2-67	17.6 (10.4)		17.7 (11.5)		.02 (.99)	20.7 (18.0)		18.2 (10.6)		-.71 (.48)		
ASD severity (ADOS)	(n=122)	5.5 (2.3)	1-10	6.8 (2.7)		5.3 (2.2)		-1.93 (.06)	6.7 (2.2)		5.4 (2.3)		-1.84 (.07)		

ADOS, Autism Diagnostic Observation Schedule (CASS, Calibrated Severity Score); ASD, Autism Spectrum Disorder; PSI-SF, Parental Stress Index – Short Form; RCADS-C, Revised Child Anxiety and Depression Scale-Children; SNAP-IV-P, Swanson, Nolan and Pelham Questionnaire-Parents (ADHD In, Attention Deficit Hyperactivity Disorder Inattention; ADHD H/I, Attention Deficit Hyperactivity Disorder Hyperactivity and Impulsivity; ODD, Oppositional Defiant Disorder)

Based on the FMSS, 10 fathers (9.3%) and 12 mothers (9.9%) were classified as critical. Only one participant had two critical parents. Compared to a normal population, the level of total parental stress was above average (mother; mean = 71.1; SD = 24.1) to high (fathers; mean = 70.4; SD = 21.5). Compared to a clinical population, stress levels were below average (mothers) to average (fathers).

Regarding the child behaviors, 37 participants (30.3%) scored high on Inattention (i.e., 95th percentile or above; mean = 12.8; SD = 6.5), 38 participants (31.1%) scored high on Hyperactivity-impulsivity (mean = 9.7; SD = 5.9), and 17 participants (13.9%) scored high on ODD (mean = 7.8; SD = 5.8). Regarding anxiety symptoms 31 participants (25.4%) scored high (i.e. 80th percentile or above; mean = 18.4; SD = 11.4) and regarding depression this was the case for 36 participants (29.5%; mean = 7.9; SD = 3.5).

Difference between fathers and mothers on parental stress and criticism

The level of parental stress did not differ between fathers and mothers ($t(107) = .8$, $p = .43$), nor did the level of criticism ($X^2(1, N=107) = .01$; $p = .94$, $df=1$).

Additionally, levels of parental stress did not differ between mothers with and without criticism ($t(115) = .043$, $p = .97$), or fathers with and without criticism ($t(104) = -1.53$, $p = .13$).

The associations between parental stress, criticism and child behaviors

As shown in table 2, maternal and paternal stress did not correlate significantly with severity of the child's ASD symptoms. Maternal stress correlated significantly with all other child behaviors on a weak to moderate level. Paternal stress correlated significantly with hyperactivity-impulsivity, ODD, and anxiety behavior. Additionally, none of the child behaviors differed between mothers with criticism and mothers without criticism, or between fathers with and without criticism. See table 1 for all comparisons between critical and non-critical fathers and mothers.

The data were checked for the assumptions for regression analyses. The assumption of multicollinearity was not completely met, because paternal and maternal stress were not independent. However, the correlation between paternal and maternal stress was acceptable ($r = .44$, $p = .000$) and the Variance Inflation Factor (VIF) was between 1.010 and 1.357. These scores are low enough to perform the regression analyses. For the assumption on linearity, homoscedasticity and normality, no serious violations were found. The residuals were normally distributed and the variance of the residuals was constant.

Multiple regression analyses were applied to investigate the association between parental stress (father, mother and the interaction between both), criticism (father and

Table 2 Associations between parental stress and child behaviors

	Maternal stress	Paternal stress
Anxiety (RCADS-C)	.27**	.21*
Depression (RCADS-C)	.28**	.16
Inattention (SNAP-IV-P)	.36**	.16
Hyperactivity-impulsivity (SNAP-IV-P)	.27**	.21*
ODD (SNAP-IV)	.43**	.30**
ASD symptoms (ADOS Severity score)	.05	.04

ADOS, Autism Diagnostic Observation Schedule; ASD, Autism Spectrum Disorder; RCADS-C, Revised Child Anxiety and Depression Scale-Children; SNAP-IV-P, Swanson, Nolan and Pelham Questionnaire-Parents

mother) and child behaviors. We could not investigate the interaction between criticism of father and mother, since the distribution among the cells was very uneven. Of the available data, only one participant had two critical parents and 20 participants had one critical parent.

The regression analyses showed that only maternal stress was significantly related to child behaviors: children with mothers who experienced more stress, had higher levels of anxiety, inattention, hyperactivity-impulsivity, and ODD behaviors. Table 3 presents all significant multiple regression analyses. For conciseness, only the R square of the total model was reported (and not the regression weights) if the regression was not significant (in case of depression, and ASD symptoms).

Table 3 Predicting child behaviors from parenting stress and criticism: multiple regression analyses

Child behavior	Intercept	Criticism mother	Criticism father	Stress mother	Stress father	Stress moth*fath	R ²
RCADS anxiety (child)	7.52	4.23	-.36	.29*	.07	.01	.124*
RCADS depression (child)	-	-	-	-	-	-	.096
SNAP-IV Inattention	6.91	3.29	.58	.26**	-.03	-.00	.169**
SNAP-IV Hyperactivity-Impulsivity	4.57	2.34	-2.64	.15*	.04	.01	.143*
SNAP-IV ODD	-.05	-.29	-1.11	.23**	.07	.00	.242**
ASD symptoms	-	-	-	-	-	-	.071

- the regression analysis was not significant

Discussion

In the current study we examined for fathers and mothers of children with ASD whether parental stress and criticism were related to each other and to ASD symptoms and internalizing and externalizing behaviors in the child. In this sample, parental stress was below average to average compared to that of parents of children from the clinical population and above average to high compared to the normal population. Only a limited number of parents was critical (9.6%).

The levels of parental stress and parental criticism did not differ between fathers and mothers, and parental stress and parental criticism were not associated with each other. Moreover, no association was found between parental stress and the severity of ASD symptoms. However, parental stress had a small to moderate association with comorbid internalizing and externalizing behaviors, which was slightly more pronounced in mothers than in fathers. Stress reported by mothers was significantly associated with child anxiety, depression, hyperactivity, inattention, and oppositional behaviors, whereas paternal stress was only significantly associated with anxiety, hyperactivity, and oppositional behaviors. Contrary to our expectations, we found no evidence that high stress in fathers combined with high stress in mothers would be especially associated with child problems. There was no association between parental criticism and ASD symptoms or comorbid behaviors of children with ASD.

The levels of total stress and the rates of high criticism were relatively low in our study. Although parents of children with autism have been reported to vary in rates of high criticism (18.1%-45.5%, Romero-Gonzalez et al., 2018), the rate of high criticism in our sample was even lower than the lowest rate reported (9.6%). The low levels of stress and the low levels of high criticism may be explained by the fact that most children in the study were verbal and higher-functioning, with moderate ASD severity levels, and many participated in regular education. Besides, parents and children in this study were referred for a social skills training, implying that a mental health professional saw the potential of learning social skills, and families were motivated to start social skills training. These characteristics indicate that the parents in this study may have found their way in education (Whitaker, 2002) and health care, which may reduce parental stress levels (Nesbitt, 2000). Additionally, a considerable part of the parents had already received psychoeducation and/or parent training, which may have decreased stress by increasing the knowledge and understanding of the child (Patra et al., 2015).

The low rate of high criticism may be explained by the interpretation of the FMSS, that differed across studies. Some studies applied a 'lower cut-off' for criticism, including borderline criticism (a parent made one or more statements of dissatisfaction) in the rate of high criticism (Sher-Censor et al., 2018; Hickey et al., 2019b). Doing so, the rate of high criticism is higher. We applied the cut-off criteria for high criticism as proposed in the manual, including only parents who had a negative initial statement, described their relationship with their child as negative, or made one or more critical comments about the child during the FMSS. This resulted in a stricter criterion for high criticism than in some other studies, which may explain the low percentage of critical parents in our study.

The finding that parental stress and criticism were not related was in contrast with earlier findings (Hickey et al., 2019a; Sher-Censor et al., 2018). Possibly, the low percentage of critical parents explains the low and non-significant association in our study. Another explanation could be the combination of the age of the children and the severity of the behavior problems. One study showed comparable levels of behavior problems as in our study, however, the children were much younger (Sher-Censor et al., 2018). Perhaps parents of these young children may still have been learning how to approach their child, whereas the parents in our study may have found a way at the end of primary school. Another study had children in the same age as our study, but some of the children in that study had intellectual disability and the behavior problems were more severe (Hickey et al., 2019a) than in our sample.

The sample characteristics may also play a role in the fact that we found no association between parental stress and ASD symptoms, in contrast to other studies (Rodriguez et al., 2019; Greenberg et al., 2006). Another possible explanation is that

the level of parental stress in this study was not extremely high in general for a clinical population. Nevertheless, even these relatively modest levels of stress showed to be associated with comorbid behaviors, somewhat stronger with externalizing behaviors than with internalizing behaviors. Additionally, the association was slightly stronger and more consistent for mothers compared to fathers. However, the pattern of the findings may indicate reporter bias. Perhaps, the association between maternal stress and externalizing behaviors as reported by mothers was stronger due to the same informant, which was not the case for paternal stress and internalizing child behaviors, or parental stress and ASD symptoms. This may be a reason why the relationship between parental stress and depression behaviors did not remain significant in the regression analyses, even though a correlation existed between them.

Strengths and limitations

The current study involved mothers and fathers, which can be considered a major strength. Second, we used a commonly used, coded interview, the FMSS, to measure parental criticism, asking about the child and the relationship with the child in a general way, leaving few possibilities for socially desirable answers.

One of the limitations of this study is that the study was cross-sectional. Because of this, no conclusions can be drawn on the direction of the relationship between parental stress and externalizing and internalizing behaviors. Longitudinal research is needed to investigate the direction of the effects and the time point in which parental stress may play a larger or smaller role on child behaviors (Bader & Barry, 2014; Baker et al., 2019). This can provide information for the clinical practice about the level and moment to intervene. Another limitation is the relatively small sample size of the study, which was not sufficient to detect associations in the smaller range. Future research could include a larger group that would allow to study more subtle associations and interactions between variables.

Additionally, in this study we used a global measure of parental stress. For a better understanding of the relation between parental stress and child behaviors, it would be interesting to investigate how stress fluctuates during the day and how these fluctuations relate to actual behaviors of the child, such as social communication, restricted and repetitive behaviors, internalizing and externalizing behaviors. This may provide insight in the situations in which parents experience most stress during the course of the day. This could probably lead to interventions that specifically focus on the moments and situations that parents are in highest need (Khor et al., 2014).

Conclusion

The current study provides insight into the associations between parental stress and high criticism in parents of children with ASD, and between these parental factors and ASD symptoms and externalizing and internalizing behaviors of the child. Parental stress showed no association with the severity of the ASD symptoms of the child, yet was associated with comorbid behaviors of the child. For high parental criticism no associations were found. The pattern of results was similar in mothers and fathers. Although the stress level of parents was not extremely high and the ASD symptoms were relatively mild, there still was an association with comorbid behaviors. This is a strong indication that clinical practice should take this into account, even if children and parents function relatively well. Insight into the relationship between parental stress and externalizing and internalizing behaviors in children with ASD is important for the entire ASD spectrum.

Ethical approval

The ethics committee of the Institutional Review Board of the University Medical Center Groningen approved the study. All procedures performed in this study were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Disclosure of potential conflict of interest

Author AdB received the grant that funded this study (the Netherlands Organization for Health Research and Development; ZonMw, nr 157003005). The agency had no role in data analyses, the decision to publish, or the preparation of the manuscript. She also is author on the Dutch ADOS and ADI-R manuals for which Accare receives enumeration. The other authors declare that they have no conflict of interest.

References

- Abidin, R.R. (1995). *Parenting Stress Index (PSI) manual (3rd ed.)*. Charlottesville, VA: Pediatric Psychology Press.
- Bader, S.H., & Barry, T.D. (2014). A longitudinal examination of the relation between parental expressed emotion and externalizing behaviors in children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44(11), 2820–2831.
- Bader, S.H., Barry, T.D., & Hann, J.A.H. (2015). The relation between parental expressed emotion and externalizing behaviors in children and adolescents with an autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities*, 30(1), 23–34.
- Baker, J.K., Fenning, R.M., Howland, M.A., & Huynh, D. (2019). Parental criticism and behavior problems in children with autism spectrum disorder. *Autism*, 23(5), 1249–1261.
- Baker, J.K., Smith, L.E., Greenberg, J.S., Seltzer, M.M. & Taylor, J.L. (2011). Change in maternal criticism an behavior problems in adolescents and adults with autism across a 7-year period. *Journal of Abnormal Psychology*, 120(2), 465–475.
- Benson, P.R., Daley, D., Karlof, K.L., & Robison, D. (2011). Assessing expressed emotion in mothers of children with autism: the Autism-Specific Five Minute Speech Sample. *Autism*, 15(1), 65–82.
- Bussing, R., Fernandez, M., Harwood, M., Hou, W., Garvan, C.W., Eyberg, S.M., & Swanson, J.M. (2008). Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: psychometric properties and normative ratings from an school district sample. *Assessment*, 15(3), 317–328.
- Chandler, S., Howlin, P., Simonoff, E., O'Sullivan, T., Tseng, E., Kennedy, J., Charman, T., & Baird, G. (2016). Emotional and behavioral problems in young children with autism spectrum disorder. *Developmental medicine and child neurology*, 58(2), 202–208.
- Chiang, H.L., & Gau, S.S.F. (2016). Comorbid psychiatric conditions as mediators to predict later social adjustment in youths with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 57(1), 103–111.
- Chorpita, B.F., Yim, L.M., Moffitt, C.E., Umemoto L.A., & Francis, S.E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A Revised Child Anxiety and Depression Scale. *Behaviour Research and Therapy*, 38(8), 835–855.
- Chorpita, B.F., Moffitt, C.E., & Gray, J. (2005). Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. *Behaviour Research and Therapy*, 43(3), 309–322.
- de Bildt, A., & de Jonge, M. V. (2008). *Autisme diagnostisch observatie schema*. Amsterdam: Hogrefe.
- de Jonge, M.V., & de Bildt, A. (2007). *Autisme diagnostisch interview-revised*. Amsterdam: Hogrefe.
- Dekker, V., Nauta, M.H., Mulder, E.J., Timmerman, M.E., & de Bildt, A. (2014). A randomized controlled study of a social skills training for preadolescent children with autism spectrum disorders: generalization of skills by training parents and teachers? *BMC Psychiatry*, 14, 189.
- Dekker, V., Nauta, M.H., Timmerman, M.E., Mulder, E.J., van der Veen-Mulders, L., van den Hoofdakker, B.J., van Warners, S., Vet, J.J., Hoekstra, P.J., & de Bildt, A. (2019). Social skills group training in children with autism spectrum disorder: a randomized controlled trial. *European Child & Adolescent Psychiatry*, 28(3), 415–424.
- de Brock, A.J.L.L. de, Vermulst, A.A., Gerris, J.R.M., & Abidin, R.R. (1992). *NOSI, Nijmeegse Ouderlijke Stress Index. Experimentele versie. Handleiding*. Lisse: Swets & Zeitlinger.
- Estes, A., Olson, E., Sullivan, K., Greenson, J., Winter, J., Dawson, G., & Munson, J. (2013). Parenting-related stress and psychological distress in mothers of toddlers with autism spectrum disorders. *Brain & Development*, 35(2), 133–138.
- Falk, N.H., Norris, K., & Quinn, M.G. (2014). The factors predicting stress, anxiety and depression in the parents of children with autism. *Journal of Autism and Developmental Disorders*, 44(12), 3185–3203.
- Firth, I., & Dryer, R. (2013). The predictors of distress in parents of children with autism spectrum disorder. *Journal of Intellectual & Developmental Disability*, 38(2), 163–171.
- Gillberg, I.C., Helles, A., Billstedt, E., & Gillberg, C. (2016). Boys with asperger syndrome grow up: psychiatric and neurodevelopmental disorders 20 years after initial diagnosis. *Journal of Autism and Developmental Disorders*, 46(1), 74–82.
- Giovannoli, G., Postorino, V., Fatta, L. M., Sanges, V., De Peppo, L., Vassena, L., de Rose, P., Vicari, S., & Mazzone, L. (2015). Behavioral and emotional profile and parental stress in preschool children with autism spectrum disorder. *Research in Developmental Disabilities*, 45–46, 411–421.

- Gong, Y., Du, Y., Li, H., Zhang, X., An, Y., & Wu, B.L. (2015). Parenting stress and affective symptoms in parents of autistic children. *Science China. Life Sciences*, 58(10), 1036–1043.
- Greenberg, J.S., Seltzer, M.M., Hong, J., & Orsmond, G.I. (2006). Bidirectional effects of expressed emotion and behavioral problems and symptoms in adolescents and adults with autism. *American Journal of Mental Retardation*, 111(4), 229–249.
- Griffith, G.M., Hastings, R.P., Petalas, M.A., & Lloyd, T.J. (2015). Mothers' expressed emotion towards children with autism spectrum disorder and their siblings. *Journal of Intellectual Disability Research*, 59(6), 580–587.
- Hastings, R.P., Daley, D., Burns, C., & Beck, A. (2006). Maternal distress and expressed emotion: Cross-sectional and longitudinal relationships with behavior problems of children with intellectual disabilities. *American Journal of Mental Retardation*, 111(1), 48–61.
- Hastings, R.P., Kovsho, H., Ward, N.J., degli Espinosa, F., Brown, T., & Remington, B. (2005). Systems analysis of stress and positive perceptions in mothers and fathers of pre-school children with autism. *Journal of Autism and Developmental Disorders*, 35(5), 635–644.
- Hayes, S.A., & Watson, S.L. (2013). The impact of parenting stress: A meta-analysis of studies comparing the experience of parenting stress in parents of children with and without autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(3), 629–642.
- Helles, A., Gillberg, I.C., Gillberg, C., & Billstedt, E. (2017). Asperger syndrome in males over two decades: quality of life in relation to diagnostic stability and psychiatric comorbidity. *Autism*, 21(4), 458–469.
- Hickey, E.J., Hartley, S.L., & Papp, L. (2019a). Psychological well-being and parent-child relationship quality in relation to child autism: an actor-partner modeling approach. *Family Process*, 59(2), 636–650.
- Hickey, E.J., Nix, R.L., & Hartley, S.L. (2019b). Family emotional climate and children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 49(8), 3244–3256.
- Hickey, E.J., Bolt, B., Rodriguez, G., & Hartley, S.L. (2020). Bidirectional relations between parent warmth and criticism and the symptoms and behavior problems of children with autism. *Journal of Abnormal Child Psychology*, 48(6), 865–879.
- Huisman, M., Oldehinkel, A. J., de Winter, A., Minderaa, R. B., de Bildt, A., Huizink, A. C., Verhulst, F. C., & Ormel, J. (2008). Cohort profile: the Dutch 'Tracking Adolescents' Individual Lives' Survey; TRAILS. *International Journal of Epidemiology*, 37(6), 1227–1235.
- Joshi, G., Petty, C., Wozniak, J., Henin, A., Fried, R., Galdo, M., Kotarski, M., Walls, S., & Biederman, J. (2010). The heavy burden of psychiatric comorbidity in youth with autism spectrum disorders: a large comparative study of a psychiatrically referred population. *Journal of Autism and Developmental Disorders*, 40(11), 1361–1370.
- Khor, A.S., Gray, K.M., Reid, S.C., & Melvin, G.A. (2014). Feasibility and validity of ecological momentary assessment in adolescents with high-functioning autism and Asperger's disorder. *Journal of Adolescence*, 37(1), 37–46.
- Kort, W., Schittekatte, M., Dekker, P.H., Verhaeghe, P., Compas, E.L., Bosmans, M., & Vermeir, G. (2005). *Wechsler Intelligence Scale for Children* (3rd ed. Nederlandse vertaling). Amsterdam: Harcourt.
- Lai, M., Kassee, C., Bensay, R., Bonato, S., Hull, L., Mandy, W., Szatmari, P., & Ameis, S.H. (2019). Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis. *Lancet Psychiatry*, 6(10), 819–829.
- Lord, C., Rutter, M., DiLavore, P.C., & Risi, S. (1999). *ADOS. Autism Diagnostic Observation Schedule. Manual*. Los Angeles: Western Psychological Services.
- Lord, C., Rutter, M., DiLavore, P.C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, second edition (ADOS-2) manual (Part I): Modules 1–4*. Torrance, CA: Western Psychological Services.
- Lounds, J., Seltzer, M.M., Greenberg, J.S., & Shattuck, P.T. (2007). Transition and change in adolescents and young adults with autism: Longitudinal effects on maternal well-being. *American Journal of Mental Retardation*, 112(6), 401–417.
- Magaña, A.B., Goldstein, M.J., Karno, M., Miklowitz, J.J., Jenkins, J., & Falloon, I.R.H. (1986). A brief method for assessing expressed emotion in relatives of psychiatric patients. *Psychiatry Research*, 17, 203–212.
- Manning, M.M., Wainwright, L., & Bennett, J. (2011). The double ABCX model of adaptation in racially diverse families with a school-age child with autism. *Journal of Autism and Developmental Disorders*, 41(3), 320–331.

- Maskey, M., Warnell, F., Parr, J.R., le Couteur, A., & McConachie, H. (2013). Emotional and behavioural problems in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(4), 851–859.
- McStay, R.L., Dissanayake, C., Scheeren, A., Koot, H.M., & Begeer, S. (2014). Parenting stress and autism: The role of age, autism severity, quality of life and problem behavior of children and adolescents with autism. *Autism*, 18(5), 502–510.
- Nesbitt, S. (2000). An evaluation of multi-agency service provision for children with autistic spectrum disorders. *The British Journal of Development Disabilities*, 46(90), 43–50.
- Patra, S., Arun, P., & Chavan, B.S. (2015). Impact of psychoeducation intervention module on parents of children with autism spectrum disorders: A preliminary study. *Journal of Neurosciences in Rural Practice*, 6(4), 529–535.
- Rea, H.M., Factor, R.S., Kao, W., & Shaffer, A. (2020). A meta-analytic review of the five minute speech sample as a measure of family emotional climate for youth: Relations with internalizing and externalizing symptomatology. *Child Psychiatry & Human Development*, 51, 656–669.
- Rodríguez, G., Hartley, S.L., & Bolt, D. (2019). Transactional relations between parenting stress and child autism symptoms and behavior problems. *Journal of Autism and Developmental Disorders*, 49(5), 1887–1898.
- Romero-Gonzalez, M.R., Chandler, S., & Simonoff, E. (2018). The relationship of parental expressed emotion to co-occurring psychopathology in individuals with autism spectrum disorder: a systematic review. *Research in Developmental Disabilities*, 72, 152–165.
- Rutter, M., Le Couteur, A., & Lord, C. (2003). *ADI-R Autism Diagnostic Interview Revised Manual*. Los Angeles: Western Psychological Services.
- Sharpe, D. L., & Baker, D. L. (2007). Financial issues associated with having a child with autism. *Journal of Family and Economic Issues*, 28(2), 247–264.
- Sher-Censor, E., Shulman, C., & Cohen, E. (2018). Associations among mothers' representations of their relationship with their toddlers, maternal parenting stress, and toddlers' internalizing and externalizing behaviors. *Infant Behavior & Development*, 50, 132–139.
- Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(8), 921–929.
- Swanson, J.M. (1997). *The SNAP-IV teacher and parent rating scale*. University of California, Irvine, CA.
- Valicenti-McDermott, M., Lawson, K., Hottinger, K., Seijo, R., Schechtman, M., Shulman, L., & Shinnar, S. (2015). Parental stress in families of children with autism and other developmental disabilities. *Journal of Child Neurology*, 30(13), 1728–1735.
- van Oort, F.V.A., Greaves-Lord, K., Verhulst, F.C., Ormel, J., & Huizink, A.C. (2009). The developmental course of anxiety symptoms during adolescence; the TRAILS study. *Journal of Child Psychology and Psychiatry*, 50(10), 1209–1217.
- Wechsler, D. (1999). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Weiss, J.A., Cappadocia, M.C., Tint, A., & Pepler, D. (2015). Bullying victimization, parenting stress, and anxiety among adolescents and young adults with autism spectrum disorder. *Autism Research*, 8(6), 727–737.
- Whitaker, P. (2002). Supporting families of preschool children with autism: What parents want and what helps. *Autism*, 6(4), 411–426.
- Yorke, I., White, P., Weston, A., Rafla, M., Charman, T., & Simonoff, E. (2018). The association between emotional and behavioral problems in children with autism spectrum disorders and psychological distress in their parents: a systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, 48(10), 3393–3415.
- Zaidman-Zait, A., Mirenda, P., Duku, E., Szatmari, P., Georgiades, S., Volden, J., Zwaigenbaum, L., Vaillancourt, T., Bryson, S., Smith, I., Fombonne, E., Roberts, W., Waddell, C., & Thompson, A. (2014). Examination of bidirectional relationships between parent stress and two types of problem behavior in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44(8), 1908–1917.

7 |

Summary and General Discussion

Summary and General Discussion

This dissertation aimed to investigate the effectiveness of a Social Skills Training (SST) for children with ASD, applying three steps. First, I set up a randomized controlled trial (RCT) to investigate the effectiveness of SST with child sessions only by comparing them to care-as-usual (CAU). Second, in this RCT, we compared the SST with child sessions only to SST with parent and teacher involvement, to evaluate whether children with such additional support would improve more during and after the training. Third, through latent class analysis, we investigated which specific subgroups benefit from SST and which do not (or to a lesser extent).

Furthermore, related to the overall aim, we developed an observational instrument and investigated its psychometric properties, in order to improve independent observation of social skills in naturalistic daily life school situations. Last, in order to increase the understanding of the burden for parents of children with ASD the last part of the dissertation focused on the well-being of parents of children with ASD in relation to ASD problems and comorbid mental health problems.

In this chapter, I have summarized the findings from each chapter followed by a discussion of the main findings in relation to the literature and clinical practice.

Summary of the findings

Chapter 2 describes the setup and design of the study, with details about the conditions, randomization, participants, the SST under study, the instruments, and measurements.

In Chapter 3, I investigated the effectiveness of SST in a sample of 122 high-functioning children with ASD, aged 9-13 years in an RCT with three conditions: SST ($n=47$), SST with parent and teacher involvement ($n=51$), and care as usual ($n=24$). Parents completed instruments that report on three levels of behavior, i.e. i) directly related to the social skills trained in the SST sessions (ESTIA-TS, perceived difficulty rated by parents; e.g., 'recognizes emotions', 'asking something to someone', and 'saying no when you don't want something to happen'), ii) more general social skills as measured with the parent- and teacher-rated SSRS (subscales 'Cooperation', e.g. sharing and helping others; 'Assertion', e.g. initiating behaviors; and 'Responsibility', e.g. ability to communicate with adults; 'Self-Control', e.g. behaviors that emerge in conflict situation) and iii) broader social functioning, indexed by the parent interview Vineland (socialization domain, e.g. makes friends easily or apologizes for mistakes).

In all conditions, children improved equally well on the parent rated perceived difficulty of specific social skills learned during the SST (ESTIA-TS) and more general social skills (the Assertion, Self-Control, and Responsibility subscale of the SSRS; except the Cooperation subscale). Children in the SST condition improved significantly more from pre- to post treatment than children in the CAU condition on the Vineland Socialization domain (Cohen's $d = .39$) and on the SSRS Cooperation subscale

(Cohen's $d = .43$). Children in both conditions improved on these measures. Also, children in the SST-PTI condition improved significantly more than children in the CAU condition on Vineland Socialization domain (Cohen's $d = .43$) and on the SSRS Cooperation subscale (Cohen's $d = .45$). Unexpectedly, no differences existed between the SST and SST-PTI condition on any of the instruments.

Findings based on teacher outcomes were slightly different. Teachers reported no significant changes from pre- to post treatment in children in the SST and CAU condition. However, based on teacher report, a significant improvement from pre- to post treatment existed in the SST-PTI condition compared to the SST condition, with respect to the SSRS Cooperation subscale (Cohen's $d = .34$) and the Self-Control subscale (Cohen's $d = .59$). Additionally, children in the SST-PTI condition improved significantly more than children in the CAU condition with respect to all SSRS-teacher subscales (Cooperation, Cohen's $d = .42$; Assertion, Cohen's $d = .34$; Self-Control, Cohen's $d = .61$).

For evaluation of a long-term effect, parent data were collected. These showed no significant differences on any of the measurement instruments at follow-up compared to posttreatment between any condition.

Chapter 4 showed that specific subgroups benefited more from SST than others. We found four subgroups with different response patterns to SST, based on the parent reported social-communicative skills and perceived difficulty before SST and on the patterns of response to SST over time (immediately and six months after SST). Two groups changed immediately after SST: they showed an increase of social communicative skills and a decrease of perceived difficulty as reported by parents. One of the improving groups (named '*best performers*', and comprising 28.6% of the sample) started with higher levels of social skills and lower perceived difficulty of these skills compared to the mean across all children before SST, as reported by parents. The other improving group ('*improvers*', 29.2%) started with lower levels of social skills and an average perceived difficulty of social skills as reported by parents. Two other groups showed no change after SST. One of these ('*stable class*', 24.8%) started with higher levels of social skills and higher perceived difficulty of these skills, as reported by parents. The other ('*poor performers*', 17.4%) started with lower levels of social skills and higher perceived difficulty on these skills, as reported by parents. After six months follow-up, none of the subgroups improved or declined as compared to immediately after SST.

Next, the response patterns in the subgroups were related to participant and intervention characteristics. Results showed that the level of social skills at the beginning of the SST was not related to change, since change appeared in subgroups with higher and lower levels of social skills before SST. Comparing the subgroups with similar baseline levels of social skills at the start of SST increased the insight into

what participant characteristics affect improvement immediately after SST. In the subgroups with relatively high baseline levels, improvement after SST was related to higher verbal cognitive ability, lower ASD symptoms on the ADOS and ADI-R, and lower social phobia symptoms. Additionally, in the subgroups with low baseline levels, lower perceived difficulty as reported by parents, younger age, and lower anxiety most clearly affected improvement. Regarding the participant characteristics, we found no effect of sex in relation to effectiveness of SST. *Symptoms of ADHD* were associated with the baseline level of social skills but not with the effectiveness of SST. Regarding intervention characteristics, the study showed that *parent and teacher involvement* was related to improvement in the subgroups with lower baseline levels of social skills before SST, with more involvement being associated with better outcomes for this group.

In order to overcome expectancy and allegiance effects because of parents' and teachers' investment in an SST intervention, an independent assessment of social skills is of added value (McMahon et al., 2013a; Gates et al., 2017; Wolstencroft et al., 2018). Observation in a naturalistic daily life situation gives insight into the generalization of learned social skills in situations where children need these skills. Based on knowledge from the literature, we therefore developed the Social skills Observation Measure (SOM) for observation at school, i.e. in a naturalistic, unstructured, and familiar situation with peers on the school playground and in the classroom. Before this instrument could be used in the evaluation of SST, more insight into its reliability, validity, and feasibility was needed. In chapter 5 I described the development of an independent observation instrument for social skills in daily life school situations, and its psychometric outcomes, based on the baseline measurement of 86 children in the RCT. The aim was to investigate whether it would be possible to feasibly, reliably, and validly observe naturalistic social communicative skills in a daily life situation that children regularly encounter. A benefit of the SOM compared to earlier studies on independent observation measures is that the SOM observed the children in two naturalistic situations to assess social skills with peers.

With the SOM, *specific behaviors*, the *quality of behavior*, and the *number of children* with whom the child interacted were observed in both situations at school. We only included behaviors that occurred in at least one-third of the participants to be able to properly investigate the reliability. The presence of *specific behaviors* differed between both situations, with more variation in the playground setting. In both situations the *specific behaviors* 'Verbal/physical aggression' and 'Reactive aggression' were not prevalent enough to include in the analyses.

The SOM had a relatively small convergent validity when compared to parent and teacher reported social skills. The largest (small to moderate) convergent validity existed compared to the teacher SSRS Assertion subscale. However, this coherence

was only found for the classroom observation and not on the playground. The SOM showed weak coherence with the ADOS and no coherence with parent reported measures (Vineland socialization domain and SSRS Assertion subscale). Most aspects of the SOM were reliably rated, except the specific behavior 'solitary engaged' on the playground. The findings indicate that the rating of the SOM is consistent and independent from the observer, and that the SOM measures something additional to the behavior reported by parents and teachers.

Since we involved parents in our training I was interested how their emotional well-being was related to behavior of the children with ASD. Therefore, in chapter 6, we investigated the association between parental stress and parental criticism towards the child in both mothers and fathers, and between these parental factors and ASD symptoms and externalizing and internalizing behaviors of the child.

Compared to the normal population the level of parental stress was above average to high in the current study. Compared to the clinical population it was below average to average. Only a limited number of parents was critical (9.6%). We found no differences between fathers and mothers in the level of parental stress or criticism. Parental stress was not related to ASD symptoms, but showed an association with comorbid internalizing and externalizing behaviors. Maternal stress was associated with child anxiety, depression, hyperactivity-impulsivity, inattention, and oppositional behaviors. Paternal stress was only associated with anxiety, hyperactivity, and oppositional behavior, and at a slightly lower level. When taking all behaviors into account, only maternal stress was related to child behaviors, i.e. anxiety, inattention, hyperactivity-impulsivity, and oppositional behaviors. Having two parents with higher levels of stress was not associated with severity of child behaviors. Also, parental criticism was not associated with any of the behaviors of the child.

Effectiveness of SST

The studies in this dissertation showed small to moderate evidence for the effectiveness of SST for children with ASD, especially for specific subgroups. The significant, small effect sizes of SST compared to the CAU condition as reported by parents were in line with the overall small to moderate effect sizes according to two recent meta-analyses (Gates et al., 2017; Wolstencroft et al., 2018). In addition to an overall combined effect size, Wolstencroft and colleagues reported separate effect sizes for two measurement instruments, i.e. SRS and SSRS. The SSRS effect size of $SDM = -.55$ corresponds with the SSRS findings in my study. Of all subscales of the SSRS, I only found a significant effect size on one (Cooperation). Remarkably, on the other subscales of the SSRS (Assertion, Self-Control, and Responsibility), I found no difference between children in SST, with or without parent and teacher involvement and CAU, while children in the CAU condition also showed significant improvement

on these subscales. This same pattern existed for perceived difficulty of the specific social skills that were trained during SST as reported by parents (ESTIA-TS). Against my expectations, I found only significant improvement between the conditions in more general and broader social skills (Vineland and SSRS) instead of the specific social skills learned during the SST (ESTIA-TS). A possible explanation is that the ESTIA-TS measured the perceived difficulty of these skills instead of the performance of the skills. The improvement in the CAU condition could reflect natural growth in time. In that case, SST did not seem to improve skills beyond natural development. Another explanation could be that children (and parents) in all conditions were motivated for SST. This may have led parents of children in the CAU condition to try to improve the social skills of their child otherwise, e.g. by organizing social situations for the child, coaching the child in such situations, or asking the teacher for help to improve social interaction in the classroom. In combination with the relatively high level of functioning of the children with ASD in my study, this may have resulted in improvement in the children in the CAU condition.

Unexpectedly, parent and teacher involvement did not lead to more improvement in social communicative skills in children, as reported by parents. This is in contrast with findings from a previous meta-analysis (Wolstencroft et al., 2018), which found large effect sizes for children in an SST with parental involvement compared to moderate effect sizes for children in an SST without parental involvement. This meta-analysis used several studies to perform group analyses for parent involvement (parent groups versus no parent group). All included studies compared two conditions, i.e., either SST with parental involvement versus a control group or SST without parental involvement versus a control group. Different SSTs (with and without parent involvement) were thus compared with a control group, which means that other factors than parental involvement can have influenced the effects that I found. My study was the first to directly compare the same SST with and without parent involvement to a CAU condition and to each other. Perhaps the motivation of the parents in my study may have caused them to support and practice with their child on their own initiative because they expected a possible added value of doing so, even when parents were not actively taking part in the intervention directly.

Notably, teachers reported more improvement in the condition in which they were involved in the intervention. This indicates that involving parents and teachers leads to generalization of social skills to the school situation. The teacher reported improvement is in contrast with other studies, which regularly found no teacher reported effect of SST (McMahon et al., 2013a; Gates et al., 2017). However, in these studies teachers were not involved in the intervention and may thus have been less aware of the learned social skills. In our SST-PTI condition the teacher was involved in the intervention directly and actively. Due to the difference in teacher involvement between the two SST conditions, we cannot rule out that the teacher-reported

improvement could also be affected by expectancy and allegiance effects (McMahon et al., 2013a; Gates et al., 2017; Wolstencroft et al., 2018). Also, teachers in the SST-PTI condition may have learned to better observe and interpret social behaviors. I think all in all that the involvement of teachers is important. Not only to improve the social skills of the children, but also for a better understanding of the problems of the child and to understand what a particular child needs to function in school and in peer groups. So that in addition to improving social skills, the child also feels good at school. Another method to reach this goal is psycho-education for teachers.

In my study, I found no further improvement of social skills after the end of the intervention in either SST condition reported by parents, even though children participated in three booster sessions between posttreatment and six months later, to repeat the learned social skills. This was against my expectation that SST would form the basis for further development of social skills after training. No other study evaluated the long-term effect from post treatment to six months follow-up. In the evaluation from pretreatment to follow-up, studies reported contradictory outcomes (Gates et al., 2017; Wolstencroft et al., 2018). Two studies into SST's consisting of 12 to 14 weekly sessions reported improvement between pretreatment and three months follow-up (Freitag et al., 2016; Deckers et al., 2016). Another study reported improvement between pretreatment and posttreatment, but this improvement did not longer exist three months after the intervention of 12 weekly sessions (Soorya et al., 2015). In the latter study, only approximately half of the participants participated in the follow-up measurements. Choque Olssen and colleagues (2017) found significant improvement in adolescents (13-17 years) but not children (7-12 years) in a similar SST condition (12 weekly sessions) compared to standard care alone from pretreatment to posttreatment and to three months follow-up. Although the four studies with three months follow-up did not evaluate the effects between post treatment and follow-up, they did report the scores of the measurements of each timepoint (sometimes in combination with a graph). From this it can be seen that in these studies (Choque Olssen et al., 2017; Freitag et al., 2016; Deckers et al., 2016; Soorya et al., 2015), this progress was mainly due to the period from pretreatment to post treatment. The combination of all findings indicates that a 12-18 week SST is not enough to teach children with ASD adequate social skills that enable them to further develop social skills on their own. Adding (three) booster sessions after the standard SST does not seem sufficient for children to further develop and generalize social skills to daily functioning in the long term either. It seems as if more is needed to help children develop social skills and use them in their daily life when meeting peers. Perhaps SSTs with more sessions (for example with more booster sessions), support in the daily situation (e.g. buddies at school; Sreckovic et al., 2017), or more intensive (but shorter) SSTs through summer camps (Lopata et al., 2010) could help to further develop and maintain develop social skills in children with ASD.

When investigating the role of participant characteristics in benefiting of SST, my finding that more anxiety was related to less improvement is in line with research that found that anxiety decreased the effect of SST (Pellecchia et al., 2016). Additionally, children of whom parents report lower perceived difficulty of social skills benefit more from SST. Possibly, children with ASD who report less anxiety or perceived difficulty of social skills, find it easier to practice the learned social skills during the SST and in daily life situations as part of their homework. This may result in benefiting more from SST. Therefore, anxiety and perceived difficulty are important to address before the start or in the beginning of the SST. Studies into the treatment of anxiety in children with ASD suggest that CBT-based interventions can be effective in decreasing anxiety (see for a review: Perihan et al., 2020). Such CBT-based interventions typically include anxiety management techniques and exposure to feared situations. In the context of ASD, these programs have been shown to be especially helpful if there is parent involvement and when treatments have a longer rather than a short duration (Perihan et al., 2020). For children with both ASD, and clinical levels of anxiety, it may be good to not only consider a social skills training, but to also add anxiety management techniques and exposure, and involve parents in these interventions.

Regarding sex, my finding of no relation between sex and outcome of SST is comparable to most previous other studies (Gates et al., 2017). However, this finding may also be affected by the limited number of females and thereby limited power in my study. The one study that found a significant greater improvement over time (from pre to post treatment) for females but not males with SST compared to standard care alone suggests that SST is probably more effective in females than males (Choque Olsson et al., 2017), and more research including more females is needed. In line with former research I could not draw firm conclusion regarding age. This is due to the relatively small age range in my study. However, even in this small age range, I found a possible relation of younger age to improvement in the subgroup with lower levels of social skills at start. Also corroborating former research, children with higher verbal cognitive levels benefited more from SST than children with lower levels (Herbrecht et al., 2009; Stevens et al., 2000; Kim et al., 2016; Paynter et al., 2018). This outcome may be related to the verbal character of the SST in the current study, i.e. with regard to explanation of skills and homework assignments. Notably, the severity of ASD symptoms was not unequivocally related to improvement. The subgroup with the most improvement showed the least severe ASD symptoms, however the other subgroup that improved showed the most severe ASD symptoms. Probably the fact that my sample of relatively high-functioning children with ASD scored relatively low on the ADOS and ADI-R has led to a too small severity range of ASD symptoms, resulting in limited power. This complicates drawing conclusions on the role of ASD symptom severity.

In my study, ADHD symptoms seemed to relate to the baseline levels of social skills but not to outcome: children with more ADHD symptoms had lower levels of social skills at start. Probably social skills are more difficult for children with combined ASD and ADHD symptoms. In addition to a lack of social knowledge, they have difficulties applying social skills through poor impulse control, or they forget to use the skills because of their attention problems (Antshel et al., 2011). However, these symptoms did not seem to affect the effect of training in the current study. This corroborates some earlier research (Deckers et al., 2016), yet is in contrast with other findings (Antshel et al., 2011) that found that group SST was less effective in children with comorbid ASD and ADHD. An important difference with the latter study is that Antshel and colleagues (2011) included children with ASD with a comorbid ADHD diagnosis whereas an ADHD diagnosis was present in 20 children (20.4%) of my sample. The ADHD symptoms as reported by parents in my study were categorized as either below or above population mean scores. Scores above the population mean were used as indicating presence of ADHD symptoms (based on this, 69 children [70.4%] had Inattention symptoms and 59 [60.2%] Hyperactivity-Impulsivity symptoms). Therefore, also children with milder ADHD symptoms were included in my analysis, which has probably attenuated their influence on the effectiveness of the SST. It seems fair to conclude that also high functioning children with ADHD symptoms besides ASD, but not a full ADHD diagnosis benefit from the SST, and that this is also warranted, given the lower social skills in that group before the training. However, when ADHD symptoms reach the level that they justify a clinical diagnosis more may need to be done. In that case it would be important to know how SST should be adjusted, and what additional skills or methods it should address.

Regarding the intervention characteristic of the presence or absence of parent and teacher involvement, we found that children in the poor performance subgroup (i.e. lower baseline levels of social skills, higher perceived difficulty, and no change after SST) participated relatively more in the SST without parent and teacher involvement. This indicates that children with lower baseline levels of social skills and higher perceived difficulty probably benefit less from SST without parent and teacher involvement. Unfortunately, based on the current study it cannot be concluded that adding parental and teacher involvement to SST would be beneficial for children with low baseline social skills. This should be investigated more thoroughly. However, SST alone is certainly not sufficient for these children to improve their social skills. Therefore, I think that involvement of parents and teacher in the SST might be useful and of added value, especially for children with lower baseline levels of social skills who have more anxiety and/or higher perceived difficulties with social skills.

Based on all outcomes, our SST showed small but statistically significant effects in daily life for high-functioning pre-adolescent children with ASD. Additional involvement of parents and teachers did not show the extra improvement that was

expected. Notably, in general the children in all subgroups showed low levels of social skills at start and these levels remained relatively low, even in the subgroups that improved.

Observation in a naturalistic setting

For a long time, researchers have argued for independent observation of the child in a natural, social situation as an outcome measure, in order to investigate the effectiveness of SST on social performance in daily life.

In order to provide an answer to this appeal, I have developed and studied a naturalistic observation (called SOM; Social skills Observation Measure) of social skills in unstructured situations among peers at school. Such naturalistic observation in daily life social situations is expected to be complementary to parent or teacher questionnaires and to a semi-structured observational instrument such as the ADOS. This expectation was confirmed in the current study. First, I found that the observations performed in unstructured peer situations in school only showed significant correlations with the teacher reported social skills in the classroom but not on the playground. This is important, since this indicates that children's social behaviors at the playground are different from their social behaviors in a social situation in the classroom, and that teachers seem to base their report on the social behaviors they observe in the classroom. Since the correlations between teacher ratings and the SOM were small, even in the classroom situation, blind observations in school seem to contribute to a broader perspective on how children perform socially. Second, the study showed a weak coherence between the ADOS and the SOM. Although the ADOS encompasses an observation of social behavior by a trained and independent observer, the interaction in this semi-structured situation is with one adult, and not with one or more peers.

Regarding reliability, the rating of the SOM is consistent and independent from the observer. The interrater reliability resembles ratings of observational instruments in naturalistic settings in other studies (Hopkins et al., 2011, Frankel et al., 2011, Shih et al., 2016, Wood et al., 2014).

Even though the SOM was shown to provide specific information that was not measured with teacher information or the ADOS and which could be coded reliably, several feasibility issues associated with observing in naturalistic situations hindered use of the SOM in my study as an outcome measure, in line with previous studies (White et al., 2007; McMahon et al., 2013b). First, the scoring of the SOM could be improved. As recommended by Frankel and colleagues (2011), the SOM observed the number of children the child interacted with, to increase the insight in the acceptance of the participating child in the group. However, I only had information about the number of children the child interacted with and not on the duration, which complicated the understanding of the quality or superficiality of the interactions.

In the current form it was difficult to assess the quality of the interaction, because the observer could not always stand close enough to the child to understand what the child said. Obtaining audio-records or sitting closer to the child may improve the quality of the SOM assessments, but the practical implementation of these adjustments is difficult.

Second, the observation would benefit from further standardization of the observation situations and clearer instructions for teachers. Especially the observation in the classroom was difficult to standardize in a way that the conditions in which the children were observed were comparable enough to use the SOM as a standardized measure. The aim was to observe the children in an unstructured situation in the classroom in which they were allowed to interact. However, for some teachers it was difficult to create such a situation within the regular school situation. Examples of the situations created by teachers were working on a collaboration assignment (more opportunity for interaction, but not spontaneous), watching television (little opportunity for interaction), structured situations with limited possibilities to interact, and a regular work situation in which children were allowed to discuss their work. This resulted in a large variety of classroom situations leading to large differences in possibilities for the children to interact with peers. Another aspect that needs standardization and clear instructions, is the interaction of the child with the teacher. Some teachers initiated interaction with the child during the observation and sometimes children initiated interaction with the teacher. In those cases, the content of what the teacher says, or the fact that the child is less focused on peers affects the interaction with peers. If I excluded observations that were more structured than hoped for or influenced by interaction with the teacher, 43.6% of the observations would be lost, especially in the classroom (62.8%) but also on the playground (24.4%). For future research it is important to build upon the current experiences and to aim for more standardization. Foremost, standardization of the unstructured situations in the classroom is needed, in a way that is still feasible for teachers. In addition, clearer instructions for the teacher help decrease the variation between teachers in these situations.

After the suggested ameliorations to the scoring, clearer instructions, and standardization of the situations it is important to further investigate the usefulness of the SOM for both research and clinical purposes. It would then be necessary to observe typically developing children on this instrument and establish norm scores.

Parental burden

For optimally involving parents in SST it is important to understand their emotional well-being and its relation with the behavior of their child with ASD, as this may affect their engagement (Karst & van Hecke, 2012; Hickey et al., 2019a). Compared to other studies in children with ASD, parents in my study reported relatively low levels of parental stress (Estes et al., 2013; Hickey et al., 2019a). A possible explanation could

be that the factors that are known to cause stress, i.e. lower level of functioning (Karst & van Hecke, 2012), higher severity of ASD symptoms (Hickey et al., 2019a), and finding the right education (Whitaker, 2002) are not or no longer applicable in the sample in my study. All children were verbal, and most were high-functioning, had moderate ASD symptoms, and participated in regular education. Additionally, a considerable part of the parents already had received psycho-education and/or parent training preceding SST, which may possibly have improved their understanding of the behavior of their child, which may reduce parental stress (Patra et al., 2015). Another explanation may be the fact that children were eligible for SST, implying that a mental health professional saw opportunities for improvement of the child on social skills. This may indicate that if parents have found their way in health care and education (Whitaker, 2002) their parental stress can decrease (Nesbitt, 2000).

An explanation for the low rate of parental criticism found in my study, compared to earlier studies in ASD is that some studies used a lower cut-off for criticism (Sher-Censor et al., 2018; Hickey et al., 2019b), which inevitably results in higher rates of criticism. I used the criteria as described in the manual of the instrument, the Five Minute Speech Sample (Magaña, 1986).

The relatively low rates of stress and criticism may also play a role in the absence of associations between parental stress and parental criticism, which was in contrast with previous research (Hickey et al., 2019a; Sher-Censor et al., 2018). The same may be true for the absence of an association between parental stress and ASD symptoms, for which our findings also contrasts with previous research (Rodriguez et al., 2019; Greenberg et al., 2006).

Another finding was that the association of maternal stress with the behaviors of the child was more consistent than with paternal stress, as only the relation with maternal stress persisted in the regression analysis (paternal stress did not). This may be a true difference, or it could be explained by the fact that mothers also rated the externalizing behaviors of the child, which may have led to reporter bias for the association between maternal stress and externalizing behaviors. However, such reporter bias cannot explain the association between parental stress and internalizing child behaviors and ASD symptoms, since internalizing child behaviors were rated by the child and the ASD symptoms by an independent observer.

Since my study was based on a cross-sectional design, I am not able to provide insight into the direction of the relationship between parental stress and the child behaviors. Longitudinal research is needed to investigate the direction of the association and get more insight into the larger or smaller influence of parental stress on child behaviors (Bader & Barry, 2014; Baker et al., 2019). As previous research mainly focused on mothers, my study contributed to understanding father's stress and criticism and on the difference between mothers and fathers.

Strengths and Limitations

A major strength of the studies into the effectiveness of SST for children with ASD in this dissertation is the RCT design with three conditions, including SST with and without parental and teacher involvement, and a care as usual condition. Participants were referred to mental health care clinics, implying high external validity of the sample and generalizability to mental health settings. The study was preregistered and CONSORT guidelines were adhered to when describing the results. Outcome was established on several levels of social behavior (directly related to the social skills training in the SST sessions, more general social skills, and broader social functioning) using multiple informants. Additionally, the follow-up assessment in all conditions can be considered as a strength, since this enabled me to compare change after the intervention to 'natural' changes in the CAU condition, which obviously lacks in designs that have a non-active or waitlist control condition.

In future studies it would be interesting to include follow-up measurements for teachers too, because teachers reported different outcomes than parents in the current study. In this study were follow-up data unavailable for teachers, due to change of class (and teacher) between posttreatment and follow-up.

When interpreting the results on the effectiveness of SST for children with ASD, note that the children with ASD in this study were a rather specific subgroup of all children with ASD. My sample was relatively high functioning, with moderate ASD symptoms, which affects the generalizability of the findings with regard to the effectiveness of SST. Additionally, although my study has shed some light on factors that may play a role in the effectiveness of SST for ASD, I could not address all factors that have been mentioned in earlier research, due to limited variability (age range, ASD symptom severity) and power (sex, level of functioning). Unfortunately, like all studies into the effectiveness of SST for ASD, the studies in this dissertation suffered from the limited availability of objective measures for changes in social skills. Although I tried to develop an instrument to blindly observe social skills in a naturalistic daily life situation with the SOM, the feasibility of this instrument showed to be limited. Therefore, the instrument could not be used properly as an effect measure in this study, and my outcomes were dependent on parent and teacher assessment of general social skills, informants who were not blind to the condition. As many studies have done before (McMahon et al., 2013a; Wolstencroft et al., 2018), I would recommend the use of instruments for social skills in a naturalistic situation and blind raters. However, although I made some steps in developing a reliable, valid, and complementary observation method, I have (again) found that doing so is not straightforward. Additionally, future studies should use ASD specific measures of social communication in order to complement measuring relevant changes after SST.

Regarding parental stress, my study was cross-sectional and therefore does not provide insight into the direction of the relationship between parental stress and child

behaviors. Future longitudinal research is needed to investigate how parental stress develops during SST and which role parental stress has on child behaviors over time, or the other way around.

A factor that affected all findings was the challenge it was to recruit children for the RCT, due to the chance that the child would be randomized to the CAU condition. Many parents were really motivated for an SST for their child and did not want to risk being in the CAU condition for a year, leading to the decision to decline participation. Other factors that led parents to decline participation were a specific preference for the SST with or without the involvement of parents and teacher, or the unfeasibility of SST with parent and teacher involvement. Therefore, although the children in my sample represent the children who are referred for SST in the participating clinics, the sample of parents may not fully represent parents of children with ASD who are eligible for SST.

Conclusions and implications

I found small to moderate but statistically significant effects of SST in high-functioning pre-adolescent children with ASD. Involving parents and teachers in the intervention made the training more burdensome for everyone involved, i.e. therapists, parents, and teachers. The therapists needed to provide more sessions and invest more in the meetings with parents and teachers and in discussing the assignments to practice with the child. Parents and teachers participated in additional sessions and were asked to actively support the child in practicing the trained social skills. However, compared to SST without parent and teacher involvement, I did not find the expected additional effect of directly and actively involving parents and teachers, as far as parents reported. Teachers did report an additional effect, but I cannot rule out that this was based on expectancy and allegiance. On the other hand, teacher involvement may have led the teacher to better understand the child's problems and their needs to function in school and in peer groups. Therefore, I think that involvement of the teacher in the treatment of the child is useful, especially when the child has lower levels of social functioning and higher perceived difficulty of performing these skills. Perhaps for this subgroup involvement of parents is of added value too.

Between post-treatment and follow-up, no further improvements were found on all measurements. In my opinion, a 12-18 week SST is probably not long enough for the children to improve their social skills beyond the duration of the training. Therefore, I think that SSTs with more sessions (for example with more booster sessions) and support in daily life situation (e.g. buddies at school) could help to maintain and further develop social skills in children with ASD. However, it is important to find the right balance between costs and benefits.

Additionally, an SST is probably more effective for children with less perceived difficulty of social skills as reported by parents, less general anxiety, and higher

verbal ability. Future research is needed to gain more insight in what works for whom, to further personalize SST for children with ASD. My findings indicate that, in order to increase the effect, it might be useful to modify SST for anxious children and add aspects such as psycho-education about anxiety and exposure to training in order to help them benefit. The same may hold true for adjustments of SST for children with ADHD symptoms in addition to ASD.

In order to optimize involvement of parents (especially through the interaction between parents' well-being and the behaviors of the child), stress and criticism are important to take into account. Even in my sample with relatively low levels of stress and criticism, parental stress was related to child internalizing and externalizing behaviors.

The instrument that I developed for observing social behavior in a naturalistic, unstructured situation at school seems of additional value for measuring social skills compared to other measures, but further standardization is needed, as well as better observation of the actual behavior and content of the interactions of children with peers. Additionally, norm scores would improve the measurement of (changes in) peer social interaction in naturalistic situations. That is important because improving peer social interaction in daily life is the ultimate goal of SST.

All in all, the current study has shown that the goal of SST is at least partially reached, providing knowledge on the effectiveness of SST immediately after training, on the longer term, and for specific subgroups of children. Additionally, developing an independent observation measure of peer interactions in a naturalistic situation has added to the desire for such instrument as an independent outcome measure of SST.

References

- Antshel, K.M., Polacek, C., McMahon, M., Dygert, K., Spenceley, L., Dygert, L., Miller, L., & Faisal, F. (2011). Comorbid ADHD and anxiety affect social skills group intervention treatment efficacy in children with autism spectrum disorders. *Journal of Developmental and Behavioral Pediatrics*, 32(6), 439–446.
- Bader, S.H., & Barry, T.D. (2014). A longitudinal examination of the relation between parental expressed emotion and externalizing behaviors in children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44(11), 2820–2831.
- Baker, J.K., Fenning, R.M., Howland, M.A., & Huynh, D. (2019). Parental criticism and behavior problems in children with autism spectrum disorder. *Autism*, 23(5), 1249–1261.
- Choque Olsson, N., Flygare, O., Coco, C., Görling, A., Råde, A., Chen, Q., Lindstedt, K., Berggren, S., Serlachius, E., Tammimies, K., Kjellin, L., & Bölte, S. (2017). Social Skills Training for Children and Adolescents With Autism Spectrum Disorder: A Randomized Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(7), 585–592.
- Deckers, A., Muris, P., Roelofs, J., & Arntz, A. (2016). A Group-Administered social Skills Training for 8- to 12-Year-Old, high-Functioning Children With Autism Spectrum Disorders: An Evaluation of its Effectiveness in a Naturalistic Outpatient Treatment Setting. *Journal of Autism and Developmental Disorders*, 46(11), 3493–3504.
- Estes, A., Olson, E., Sullivan, K., Greenson, J., Winter, J., Dawson, G., & Munson, J. (2013). Parenting-related stress and psychological distress in mothers of toddlers with autism spectrum disorders. *Brain & Development*, 35(2), 133–138.
- Frankel, F.D., Gorospe, C.M., Chang, Y., & Sugar, C.A. (2011). Mothers' reports of play dates and observation of school playground behavior of children having high-functioning autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 52(5), 571–579.
- Freitag, C.M., Jensen, K., Elsuni, L., Sachse, M., Herpertz-Dahlmann, B., Schulte-Rüther, M., Hänig, S., Von Gontard, A., Poustka, L., Schad-Hansjosten, T., Wenzl, C., Sinzig, J., Taurines, R., Geißler, J., Kieser, M., & Cholemkery, H. (2016). Group-based cognitive behavioral psychotherapy for children and adolescents with ASD: the randomized, multicentre, controlled SOSTA-net trial. *Journal of Child Psychology and Psychiatry*, 57(5), 596–605.
- Gates, J.A., Kang, E., & Lerner, M.D. (2017). Efficacy of group social skills interventions for youth with autism spectrum disorder: A systematic review and meta-analysis. *Clinical Psychology Review*, 52, 164–181.
- Greenberg, J.S., Seltzer, M.M., Hong, J., & Orsmond, G.I. (2006). Bidirectional effects of expressed emotion and behavioral problems and symptoms in adolescents and adult with autism. *American Journal of Mental Retardation*, 111(4), 229–249.
- Herbrecht, E., Poustka, F., Birnkammer, S., Duketis, E., Schlitt, S., Schmötzer, G., & Bölte, S. (2009). Pilot evaluation of the Frankfurt social skills training for children and adolescents with autism spectrum disorder. *European Child & Adolescent Psychiatry*, 18(6), 327–335.
- Hickey, E.J., Hartley, S.L., & Papp, L. (2019a). Psychological well-being and parent-child relationship quality in relation to child autism: an actor-partner modeling approach. *Family Process*, 59(2), 636–650.
- Hickey, E.J., Nix, R.L., & Hartley, S.L. (2019b). Family emotional climate and children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 49(8), 3244–3256.
- Hopkins, I.H., Gower, M.W., Perez, T.A., Smith, D.S., Amthor, F.R., Wimsatt, F.C., & Biasini, F.J. (2011). Avatar Assistant: improving social skills in students with an ASD through a computer-based intervention. *Journal of Autism and Developmental Disorders*, 41(11), 1543–1555.
- Karst, J.S., & Van Hecke, A.V. (2012). Parent and family impact of autism spectrum disorders: A review and proposed model for intervention evaluation. *Clinical Child and Family Psychology review*, 15(3), 247–277.
- Kim, S.H., Macari, S., Koller, J., & Chawarska, K. (2016). Examining the phenotypic heterogeneity of early autism spectrum disorder: subtypes and short-term outcomes. *Journal of Child Psychology and Psychiatry*, 57(1), 93–102.
- Lopata, C., Thomeer, M.L., Volker, M.A., Toomey, J.A., Nida, R.E., Lee, G.K., Smerbeck, A.M., & Rodgers, J.D. (2010). RCT of a manualized social treatment for high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 40(11), 1297–1310.

- Magaña, A.B., Goldstein, M.J., Karno, M., Miklowitz, J.J., Jenkins, J., & Falloon, I.R.H. (1986). A brief method for assessing expressed emotion in relatives of psychiatric patients. *Psychiatry Research*, 17, 203–212.
- McMahon, C.M., Lerner, M.D., & Britton, N. (2013a). Group-based social skills interventions for adolescents with higher-functioning autism spectrum disorder: a review and looking to the future. *Adolescent Health, Medicine and Therapeutics*, 4, 23–38.
- McMahon, C.M., Vismara, L.A., & Solomon, M. (2013b). Measuring changes in social behavior during a social skills intervention for higher-functioning children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(8), 1843–1856.
- Nesbitt, S. (2000). An evaluation of multi-agency service provision for children with autistic spectrum disorders. *The British Journal of Development Disabilities*, 46(90), 43–50.
- Patra, S., Arun, P., & Chavan, B.S. (2015). Impact of psychoeducation intervention module on parents of children with autism spectrum disorders: A preliminary study. *Journal of Neurosciences in Rural Practice*, 6(4), 529–535.
- Paynter, J., Trembath, D., & Lane, A. (2018). Differential outcome subgroups in children with autism spectrum disorder attending early intervention. *Journal of Intellectual Disability Research*, 62(7), 650–659.
- Pellecchia, M., Connell, J.E., Kerns, C.M., Xie, M., Marcus, S.C., & Mandell, D.S. (2016). Child characteristics associated with outcome for children with autism in a school-based behavioral intervention. *Autism*, 20(3), 321–329.
- Perihan, C., Burke, M., Bowman-Perrott, L., Bicer, A., Gallup, J., Thompson, J., & Sallese, M. (2020). Effects of cognitive behavioral therapy for reducing anxiety in children with high functioning ASD: A systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, 50(6), 1958–1972.
- Rodríguez, G., Hartley, S.L., & Bolt, D. (2019). Transactional relations between parenting stress and child autism symptoms and behavior problems. *Journal of Autism and Developmental Disorders*, 49(5), 1887–1898.
- Sher-Censor, E., Shulman, C., & Cohen, E. (2018). Associations among mothers' representations of their relationship with their toddlers, maternal parenting stress, and toddlers' internalizing and externalizing behaviors. *Infant Behavior & Development*, 50, 132–139.
- Shih, W., Patterson, S.Y., & Kasari, C. (2016). Developing an adaptive treatment strategy for peer-related social skills for children with autism spectrum disorders. *Journal of Clinical Child & Adolescent Psychology*, 45(4), 469–479.
- Soorya, L.V., Siper, P.M., Beck, T., Soffes, S., Halpern, D., Gorenstein, M., Kolevzon, A., Buxbaum, J., & Wang, A.T. (2015). Randomized comparative trial of a social cognitive social skills group for children with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 54(3), 208–216.
- Sreckovic, M.A., Hume, K., & Able, H. (2017). Examining the Efficacy of Peer Network Interventions on the Social Interactions of High School Students with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 47(8), 2556–2574.
- Stevens, M.C., Fein, D.A., Dunn, M., Allen, D., Waterhouse, L.H., Feinstein, C., & Rapin, I. (2000). Subgroups of children with autism by cluster analysis: a longitudinal Examination. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39 (3), 346–52.
- Whitaker, P. (2002). Supporting families of preschool children with autism: What parents want and what helps. *Autism*, 6(4), 411–426.
- White, S.W., Koenig, K., & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: a review of the intervention research. *Journal of Autism and Developmental Disorders*, 37(10), 1858–1868.
- Wolstencroft, J., Robinson, L., Srinivasan, R., Kerry, E., Mandy, W., & Skuse, D. (2018). A Systematic Review of Group Social Skills Interventions, and Meta-analysis of Outcomes, for Children with High Functioning ASD. *Journal of Autism and Developmental Disorders*, 48(7), 2293–2307.
- Wood, J.J., Fujii, C., Renno, P., & van Dyke, M. (2014). Impact of cognitive behavioural therapy on observed autism symptom severity during school recess: a preliminary randomized, controlled trial. *Journal of Autism and Developmental Disorders*, 44(9), 2264–2276.

Nederlandse samenvatting

Nederlandse samenvatting

Een autismespectrumstoornis (ASS) is een neurobiologische ontwikkelingsstoornis die gekenmerkt wordt door beperkingen in de sociale communicatie en interactie en door beperkte en repetitieve patronen van gedrag, interesse of activiteiten. Kinderen met ASS leren minder automatisch en langzamer dan zich normaal ontwikkelende kinderen als het gaat om sociaal-communicatieve vaardigheden. Sociaal-communicatieve beperkingen leiden vaak tot kleinere sociale netwerken en een grotere kans op pesten of sociale isolatie vergeleken met leeftijdsgenoten. Om de sociale vaardigheden te verbeteren zijn trainingen ontwikkeld. Socialevaardigheidstrainingen beogen het sociale functioneren van een kind te verbeteren en gedragsproblemen en negatieve gevolgen van de sociaal-communicatieve beperkingen te verminderen. Belangrijke aannames daarbij zijn dat kinderen de geleerde sociale vaardigheden kunnen generaliseren naar situaties buiten de training en dat de geleerde sociale vaardigheden na afloop van de training behouden blijven en zo een basis vormen voor verdere ontwikkeling van sociaal-communicatieve vaardigheden. Socialevaardigheidstrainingen zijn in de klinische praktijk een belangrijk onderdeel van de behandeling van kinderen met ASS. Toch is het bewijs voor de effectiviteit ervan voor deze kinderen nog niet overtuigend. Het zou van toegevoegde waarde kunnen zijn om ouders en leerkrachten uitgebreider bij de training te betrekken om de geleerde vaardigheden te generaliseren naar de school- en thuissituatie. Zo'n uitgebreide variant van een socialevaardigheidstraining zou effectiever kunnen zijn, met name op de lange termijn. Over deze langetermijneffecten met en zonder betrokkenheid van ouders en leerkrachten is nog weinig bekend. Verder is het belangrijk dat de uitkomstmaten zich niet alleen richten op de geleerde vaardigheden, maar ook op sociale vaardigheden op school en thuis en op het algemene en brede sociaal functioneren van het kind. Om dit in kaart te kunnen brengen is het noodzakelijk dat er bij meerdere informanten informatie wordt verzameld om zo een goed beeld te krijgen van de effecten van de training. In dit proefschrift is dan ook de effectiviteit van een socialevaardigheidstraining onderzocht voor hoog-functionerende kinderen met ASS in de laatste 2,5 jaar van de basisschool. Er werd gekeken of een socialevaardigheidstraining met of zonder ouder- en leerkrachtbetrokkenheid effectiever was dan reguliere zorg direct na de training en op langere termijn. Ook onderzocht ik of er een mogelijk verschil is in effectiviteit voor specifieke subgroepen. Daarnaast heb ik een observatie-instrument voor sociaal-communicatief gedrag in een naturalistische situatie ontwikkeld en onderzocht op psychometrische kwaliteiten. Tot slot beschrijft het proefschrift een studie naar ouderlijke stress en kritiek van vaders en moeders in relatie tot elkaar en tot verschillende types gedragsproblemen van het kind. Ik heb ervoor gekozen om de kinderen met ASS te werven bij jeugd-GGZ-instellingen. Hierdoor zijn de uitkomsten van mijn studie te generaliseren naar

kinderen met ASS (en comorbide problematiek) in de reguliere praktijk. In *hoofdstuk 1* staat een uitgebreidere beschrijving van de achtergrond van deze studies beschreven.

Onderzoeksopzet

Het onderzoek in dit proefschrift (*Efficacy of Social Skills Training In Autism*; ESTIA) betreft een gerandomiseerde gecontroleerde studie naar de effectiviteit van een socialevaardigheidstraining voor hoog-functionerende kinderen met ASS. In totaal zijn er 122 kinderen (103 jongens en 19 meisjes; gemiddelde leeftijd van 11 jaar (9,5-13,0 jaar); gemiddeld totaal IQ van 101,5 (72-135); 32% had één comorbide diagnose en 4.1% had twee comorbide diagnoses) per toeval verdeeld over drie verschillende condities. De drie condities waren een SoVa groep ($n=47$), een SoVa-PLUS groep ($n=51$) en een groep die gebruikelijke zorg kreeg (*care as usual*; CAU; $n=24$). De socialevaardigheidstraining bestond uit 15 sessies met 3 booster sessies van ieder 1,5 uur. Bij de SoVa-PLUS was er sprake van 8 aanvullende sessies voor ouders en leerkrachten volgden een startbijeenkomst en hadden 5 keer telefonisch contact met één van de therapeuten gedurende de eerste 15 sessies van de training. Tijdens drie verschillende meetmomenten werden vragenlijsten en interviews afgenomen bij ouders, kinderen en/of leerkrachten gericht op verschillende niveaus van sociale vaardigheden. Als eerste ging het om de specifieke sociale vaardigheden die de kinderen geleerd hebben tijdens de socialevaardigheidstraining (ESTIA-TS, ervaren moeilijkheid van deze vaardigheden gerapporteerd door ouders). Ten tweede werden meer algemene sociale vaardigheden gemeten met de Social Skills Rating Scale (SSRS; subschalen 'Cooperation', 'Self-Control', 'Assertion' en 'Responsibility'), beoordeeld door ouders en leerkrachten. Als laatste werden brede sociale vaardigheden in kaart gebracht aan de hand van het ouderinterview voor adaptief functioneren Vineland (socialisatie domein). Om de effectiviteit goed te onderzoeken bestudeerden we niet alleen veranderingen direct na afloop van de training, maar ook zes maanden daarna. In *hoofdstuk 2* staat gedetailleerd beschreven hoe het onderzoek is opgezet. De onderzoeken in de verschillende hoofdstukken zijn uitgevoerd bij de kinderen die deelnamen aan ESTIA. Zoals genoemd waren dit kinderen die in de praktijk waren aangemeld voor een socialevaardigheidstraining. De studie werd vooraf geregistreerd in het Dutch Trial Register (NTR2405), zodat voorafgaand aan de studie al duidelijk werd welke uitkomstmaten vooral belangrijk werden geacht.

Effectiviteit direct na afloop van de training en op lange termijn

Voor inzicht in de effectiviteit van de socialevaardigheidstraining zijn de training met (SoVa-PLUS) en zonder (SoVa) ouder- en leerkrachtbetrokkenheid vergeleken met CAU en met elkaar. Door beide vormen van training te vergelijken met reguliere zorg zonder training (CAU) kreeg ik inzicht in de effectiviteit van de socialevaardigheids-

trainingen ten opzicht van de reguliere zorg zonder training. Het vergelijken van de training met en zonder ouder- en leerkrachtenbetrokkenheid gaf inzicht in de toegevoegde waarde van het betrekken van ouders en leerkrachten bij de socialevaardigheidstraining.

In *hoofdstuk 3* heb ik de uitkomsten van deze vergelijkingen beschreven. In alle condities verbeterden kinderen evenveel op de door ouders gerapporteerde ervaren moeilijkheid van de specifieke sociale vaardigheden die de kinderen hadden geleerd tijdens de socialevaardigheidstraining (ESTIA-TS) en de meer algemene sociale vaardigheden (SSRS, subschalen 'Self-Control', 'Assertion' en 'Responsibility', behalve de subschaal 'Cooperation'). Kinderen die meededen aan een socialevaardigheidstraining waren qua socialisatie (gemeten met de Vineland) en de SSRS subschaal 'Cooperation' significant meer verbeterd na de training, vergeleken met kinderen die alleen regulier zorg hadden gekregen. Ook kinderen in de SoVa-PLUS conditie verbeterden significant meer dan kinderen in de CAU conditie. Tegen mijn verwachting waren er op geen van de ouderrapportages (ESTIA-TS, alle subschalen van de SSRS en Vineland socialisatie) verschillen tussen de SoVa en SoVa-PLUS conditie.

De bevindingen bij leerkrachten lieten wat anders zien dan de bevindingen uit de ouderrapportages. Leerkrachten rapporteerden een significante verbetering bij kinderen die meededen aan de SoVa-PLUS ten opzichte van de SoVa, met betrekking tot algemene sociale vaardigheden (SSRS 'Cooperation' subschaal en de 'Self-Control' subschaal). Daarnaast verbeterden kinderen die de SoVa-PLUS hadden gevolgd significant meer dan kinderen in de CAU conditie op alle SSRS leerkracht subschalen ('Cooperation', 'Self-Control' en 'Assertion'). Leerkrachten rapporteerden geen verschillen na de training tussen kinderen in de SoVa conditie en kinderen in de CAU conditie.

De verwachting was dat de betrokkenheid van ouders en leerkrachten bij de socialevaardigheidstraining de langetermijneffecten zou versterken, omdat ouders hun kind na afloop van de training kunnen blijven stimuleren om de geleerde sociale vaardigheden te gebruiken, zodat het kind sociale vaardigheden kan blijven oefenen en ontwikkelen. Het onderzoek wees echter anders uit. Op basis van oudergegevens bleek er in de periode direct na afloop van de training tot een half jaar later geen verschil in verandering in sociale vaardigheden, in welke vorm dan ook. De verwachte voortgang op lange termijn bleef dus uit. Leerkrachtgegevens waren alleen beschikbaar direct na de interventieperiode, en niet een half jaar na afloop vanwege het wisselen van klas en meestal ook leerkracht.

Wat werkt voor wie?

In *hoofdstuk 4* van dit proefschrift is onderzocht of en zo ja welke specifieke subgroepen van kinderen met ASS baat hebben bij een socialevaardigheidstraining en welke niet of in mindere mate. Door de kind- en trainingskenmerken van de verschillende subgroepen te onderzoeken kunnen we factoren identificeren die

kunnen helpen voorspellen welke kinderen met ASS kunnen profiteren van een socialevaardigheidstraining.

Onder de totale groep kinderen die meededen aan een SoVa of SoVa PLUS ($n=98$; 9-12 jaar; 17 meisjes) waren vier subgroepen te onderscheiden, die verschilden in niveau van sociaal-communicatieve vaardigheden aan het begin van de training en in hoe ze reageerden op de training (direct en zes maanden na de training). Twee subgroepen (*'best performers'*, 28,6% van de kinderen en *'improvers'*, 29,2% van de kinderen) verbeterden direct na afloop van de socialevaardigheidstraining: hun sociaal-communicatieve vaardigheden namen toe en de ervaren moeilijkheid van sociaal-communicatieve vaardigheden nam af (gerapporteerd door ouders). De andere twee subgroepen (*'stable class'*, 24,8% van de kinderen en *'poor performers'*, 17,4% van de kinderen) veranderden niet na de socialevaardigheidstraining. Verandering in sociaal-communicatieve vaardigheden lijkt niet gerelateerd aan het niveau van de sociaal-communicatieve vaardigheden bij start van de training. Zowel verbetering als gelijk blijven komen voor in de subgroepen met hogere en met lagere niveaus van sociale vaardigheden aan het begin van de socialevaardigheidstraining. De ervaren moeilijkheid van de specifieke sociale vaardigheden die werden getraind in de training (gerapporteerd door ouders) aan het begin van de socialevaardigheidstraining lijkt meer invloed te hebben op de verbetering van sociale vaardigheden. Kinderen in de twee subgroepen die verbeterden vonden sociale vaardigheden minder of gemiddeld moeilijk vergeleken met het gemiddelde over alle kinderen voorafgaand aan de training (gerapporteerd door ouders), terwijl kinderen in de twee subgroepen die niet verbeterden startten met meer ervaren moeilijkheid van sociale vaardigheden (gerapporteerd door ouders). Zes maanden na afloop van de socialevaardigheidstraining verbeterde of verslechterde geen van de subgroepen vergeleken met direct na de training.

Vervolgens zijn verschillende kind- en trainingskenmerken van de subgroepen in kaart gebracht. Hieruit kwam naar voren dat de twee subgroepen met relatief hogere niveaus van sociale vaardigheden bij start van de training een hogere verbale intelligentie hadden, minder ASS symptomen en minder sociale fobie symptomen. Aanvullend daarop waren de subgroepen die vooruitgingen na training jonger en verbaal sterker en was er sprake van een lagere ervaren moeilijkheid van sociale vaardigheden (gerapporteerd door ouders) en mildere angstsymptomen. Van de overige onderzochte kindkenmerken werd geen relatie gevonden tussen geslacht en de effectiviteit van training. Symptomen van ADHD bleken samen te hangen met het niveau van sociale vaardigheden tijdens de start van de socialevaardigheidstraining, maar niet met de effectiviteit van de training. Ten aanzien van het trainingskenmerk ouder- en leerkrachtbetrokkenheid bleek dat kinderen in de subgroep *'poor performers'* met lagere sociale vaardigheden aan het begin en die niet vooruitgingen vaker deelnamen aan de socialevaardigheidstraining zonder ouder- en leerkrachtbetrokkenheid vergeleken met alle andere subgroepen.

Observatie van sociale vaardigheden in een naturalistische situatie op school

In de literatuur pleit men al geruime tijd voor een onafhankelijke observatie van het kind in een natuurlijke, dagelijkse sociale situatie als uitkomstmaat voor onderzoek naar de effectiviteit van een socialevaardigheidstraining. In dit proefschrift heb ik een observatie-instrument (*Social skills Observation Measure*; SOM) ontwikkeld om het gedrag van de kinderen met ASS in een naturalistische en ongestructureerde situatie te observeren. In *hoofdstuk 5* zijn de psychometrische eigenschappen van het observatie-instrument beschreven, zoals onderzocht bij 86 kinderen (15 meisjes) met ASS (9,8 tot 13,1 jaar), op basis van de voormeting van het gerandomiseerde behandel-effectonderzoek zoals beschreven in hoofdstuk 2, 3 en 4. De SOM betreft een observatie van specifieke gedragingen, de kwaliteit van gedrag en het aantal kinderen waar het kind mee interacteert in twee situaties op school (op het plein en tijdens een ongestructureerde situatie in de klas). Alleen specifieke gedragingen die bij ten minste een derde van de kinderen voorkwamen zijn meegenomen in de analyses, om de betrouwbaarheid goed te kunnen onderzoeken. De aanwezigheid van de gedragingen verschilde tussen beide situaties, met meer variatie op het plein. De interbeoordelaars-betrouwbaarheid was goed tot uitstekend. De convergente validiteit was relatief laag in vergelijking met de door ouders en leerkrachten gerapporteerde sociale vaardigheden. Voor leerkrachten was de convergente validiteit het grootst (klein tot matig) voor de SSRS subschaal 'Assertion', maar dit werd alleen gevonden voor observatie in de klas en niet op het plein. Dit geeft aan dat leerkrachten hun beoordeling lijken te baseren op het sociale gedrag dat ze in de klas zien en dat dit mogelijk verschilt van het sociale gedrag op het plein. Deze bevindingen lijken uit te wijzen dat een onafhankelijke observatie informatie toevoegt over de aard, frequentie en kwaliteit van sociaal gedrag in dagelijkse situaties boven rapportage van ouders en leerkrachten. Daarnaast lijken de scores op de SOM consistent en onafhankelijk van de observator. Het gebruik in de praktijk werd echter beperkt door een gebrek aan praktische haalbaarheid, bijvoorbeeld omdat het lastig bleek de situatie in de klas op zo'n manier te standaardiseren dat de situaties waarin de kinderen worden geobserveerd voldoende overeenkomen om de SOM te kunnen gebruiken als gestandaardiseerde maat. Duidelijkere instructies en afspraken over de standaardisatie van de observatiesituatie zijn nodig om de SOM te kunnen gebruiken als uitkomstmaat. Vervolgens is het van belang om zich normaal ontwikkelende kinderen van dezelfde leeftijd te observeren, zodat normscores ontwikkeld kunnen worden.

Ouderlijke stress en kritiek in relatie tot gedragsproblemen van het kind

Niet alleen kinderen met ASS ervaren moeilijkheden en uitdagingen, maar dit geldt ook voor ouders die een kind met ASS opvoeden. Ouders ervaren stress en de emotionele kwaliteit van de relatie met hun kind staat onder druk. Deze factoren lijken

op hun beurt weer samen te hangen met verschillende gedragsproblemen van het kind. Meer zicht op de samenhang tussen ouderlijke stress, het kritisch zijn ten aanzien van het kind en gedragsproblemen van het kind, is van belang voor de invulling van de behandeling van kinderen met ASS en hun ouders, omdat deze aspecten mogelijk een rol spelen in hoe ouders betrokken kunnen zijn bij een sociaalvaardigheidstraining. Het meeste onderzoek dat is uitgevoerd, heeft alleen de moeder als informant gebruikt. In dit proefschrift heb ik me in *hoofdstuk 6* gericht op de samenhang tussen ouderlijke stress en ouderlijke kritiek richting het kind met ASS ($n=122$, 9-13 jaar) bij vaders en bij moeders, en op de samenhang tussen ouderparen in stress en kritiek. Hiervoor werd een vragenlijst gebruikt om ouderlijke stress in kaart te brengen (PSI-SF; Parental Stress Index - Short Form) en een kort interview om te meten of er sprake was van overmatige kritiek bij de ouder (FMSS; Five Minute Speech Sample). Daarnaast onderzocht ik de samenhang tussen deze ouderfactoren en ASS symptomen (ADOS; Autism Diagnostic Observation Schedule, ADI-R; Autism Diagnostic Interview-Revised) en externaliserend (SNAP-IV; Swanson, Nolan, and Pelham questionnaire) en internaliserend gedrag van het kind (RCADS; Revised Child Anxiety and Depression Scale). Vergeleken met de normale populatie was het niveau van ouderlijke stress in dit onderzoek gemiddeld tot hoog. Vergeleken met de klinische populatie was de ouderlijke stress beneden gemiddeld tot gemiddeld. Een beperkt aantal ouders werd getypeerd als kritisch over hun kind (9,6%). Er werden geen verschillen tussen vaders en moeders gevonden. Ouderlijke stress hing niet samen met ASS symptomen, maar wel met internaliserend en externaliserend gedrag. Moederlijke stress hing samen met angst- en depressiesymptomen (gerapporteerd door het kind) en hyperactiviteit-impulsiviteit, onoplettendheid en oppositioneel gedrag gerapporteerd door de ouder (voornamelijk door moeders). Vaderlijke stress hing samen met angst en hyperactiviteit-impulsiviteit en oppositioneel gedrag van het kind. Wanneer alle gedragingen werden samengenomen vertoonde alleen moederlijke stress een relatie met een aantal kindkenmerken (angst, hyperactiviteit-impulsiviteit, onoplettendheid en oppositioneel gedrag). Ouderlijke stress bij beide ouders hing niet significant samen met ernstigere gedragsproblemen van het kind. Ouderlijke kritiek kende geen enkele significante relatie met de onderzochte gedragingen van het kind.

Conclusies

Mijn onderzoeken tonen aan dat sociaalvaardigheidstraining resulteert in kleine tot matige verbeteringen van het sociaal functioneren van kinderen met ASS, gemeten direct na afloop van de training. Het verwachte extra effect van het direct en actief betrekken van ouders en leerkrachten bij de training bleef uit op basis van ouder-rapportage maar werd wel gerapporteerd door leerkrachten. Het valt niet uit te sluiten dat verwachtingen van de leerkrachten en de eigen inzet mogelijk een rol spelen bij

deze uitkomsten. Aan de andere kant kan de betrokkenheid van de leerkracht ertoe geleid hebben dat de leerkracht meer zicht kreeg op de problemen van het kind op school en in de omgang met leeftijdsgenoten en meer begrip heeft gekregen. Het is belangrijk na te gaan voor wie het betrekken van ouders en leerkrachten zinvol is, want dit vraagt veel tijd en inzet van alle betrokkenen. De betrokkenheid van ouders en leerkrachten lijkt vooral nuttig wanneer het kind op een lager sociaal niveau functioneert en het moeilijker vindt om de vaardigheden uit te voeren voordat het aan de training begint. Hiervoor is verder onderzoek nodig.

In de zes maanden na afloop van de training blijven de vaardigheden stabiel maar treedt geen extra vooruitgang op. Een socielevaardigheidstraining van 12-18 weken lijkt niet voldoende voor kinderen met ASS om hun sociale vaardigheden ook na de training verder te verbeteren. Socielevaardigheidstrainingen met meer sessies (bijvoorbeeld met meer boostersessies) en ondersteuning in het dagelijks leven (bijvoorbeeld met leeftijdsgenoten op school) zouden mogelijk wél kunnen helpen om sociale vaardigheden bij kinderen met ASS verder te ontwikkelen. Het is echter wel belangrijk om de juiste balans te vinden tussen kosten en baten en die goed te onderzoeken.

Kindkenmerken die gerelateerd zijn aan de betere effecten zijn lagere ervaren moeilijkheid van sociale vaardigheden (gerapporteerd door ouders), jongere leeftijd, hogere verbale capaciteiten en mildere angstsymptomen. Deze uitkomsten zijn des te opvallender omdat de groep als geheel bestond uit hoog-functionerende kinderen met ASS. Deze bevindingen geven aanleiding voor vervolgonderzoek naar mogelijke verbetering van socielevaardigheidstraining voor specifieke subgroepen van kinderen om de behandeling verder te personaliseren. Mogelijk is het bijvoorbeeld zinvol om bij kinderen met ASS die angstig zijn de training aan te passen en aspecten zoals psycho-educatie over angst, angst regulatie strategieën en blootstelling aan de situaties die angst oproepen (*exposure*) toe te voegen aan de training om deze kinderen mogelijk beter te laten profiteren van een socielevaardigheidstraining. Het effect van dergelijke aanpassingen zou moeten worden onderzocht.

Het instrument dat ik ontwikkelde voor het observeren van sociaal-communicatief gedrag in een naturalistische situatie op school lijkt van aanvullende waarde voor het meten van sociale vaardigheden in vergelijking met andere meetinstrumenten, maar het bleek in het huidige onderzoek onvoldoende haalbaar om het instrument voldoende te standaardiseren en te gebruiken als uitkomstmaat. Voor vervolgonderzoek is het belangrijk dat de naturalistische situaties op school meer worden gestandaardiseerd en dat de kwaliteit van de interacties met andere kinderen beter in kaart wordt gebracht. Vervolgens is het zinvol om zich normaal ontwikkelende kinderen van dezelfde leeftijd te observeren om normscores te ontwikkelen. Op deze manier kan een dergelijke observatiemethode het meten van verbeteringen in interactie met leeftijdsgenoten in het dagelijks leven bevorderen. Dat is belangrijk omdat dát het ultieme doel is dat we met een socielevaardigheidstraining willen bereiken.

Ten slotte impliceert mijn onderzoek naar ouderlijke stress dat het van belang is om dit in kaart te brengen en hiermee rekening mee te houden in de behandeling van kinderen met ASS. Zelfs in een groep ouders van hoog-functionerende kinderen met ASS aan het einde van de basisschool en met een voor een klinische populatie laag tot gemiddeld stressniveau, blijkt het stressniveau samen te hangen met internaliserend en externaliserend gedrag van het kind.

Samenvattend kunnen kinderen met ASS in enige mate profiteren van een socialevaardigheidstraining en zijn er daarnaast aanknopingspunten om de training verder te personaliseren waardoor sommige subgroepen kinderen mogelijk meer kunnen profiteren van de training in vergelijking met de gevonden uitkomsten in mijn onderzoek.

Dankwoord

Dankwoord

Wat heb ik er lang naar uitgekeken om dit laatste hoofdstuk te mogen schrijven en nu is het dan eindelijk zo ver. Tegelijkertijd is het ook bijzonder om te beseffen dat dit, na zo'n lange tijd, echt het laatste hoofdstuk is dat ik voor mijn proefschrift ga schrijven. Het einde is in zicht!

Er zijn ontzettend veel mensen die ik wil bedanken voor hun betrokkenheid en hun bijdrage aan de totstandkoming van dit proefschrift. Zowel bij de opzet van het onderzoek, de praktische uitvoering, de dataverzameling en het schrijven van de papers zijn veel mensen betrokken geweest.

Allereest wil ik Annelies de Bildt, mijn copromotor, bedanken. Wat hebben wij veel overlegd en samen afgestemd. Tijdens alle fasen kon ik bij je terecht met welke vraag dan ook. Je nam altijd de tijd om mee te denken en te zoeken naar een oplossing. In de hectiek van alle metingen heb je weleens een ADOS afgenomen. Ook toen ik eigenlijk had besloten om te stoppen, heb je met me meegedacht in welke vorm het voor mij wel haalbaar zou zijn om door te gaan. Zie hier het eindresultaat! In de laatste periode heb je vooral inhoudelijk meegedacht, ontzettend fijn dat je zo toegankelijk, flexibel en snel met reageren was. Ik waardeer de manier waarop je me begeleidt enorm.

Daarnaast wil ik mijn promotoren Maaïke Nauta en Pieter Hoekstra bedanken voor hun bijdrage. Maaïke, jij was vanaf het begin betrokken en hebt in deze eerste periode een belangrijke bijdrage geleverd aan de opzet van de studie. Maar ook in de laatste fase, tijdens het schrijven van de papers heb ik veel geleerd van je inhoudelijke feedback. Pieter, jij bent op een later moment betrokken geraakt bij mijn promotietraject. Ik heb gebruik kunnen maken van je kennis en ervaring tijdens het schrijven van de papers, bedankt hiervoor.

Ook Erik Mulder wil ik bedanken voor zijn zeer diverse bijdrage. In eerste instantie was je betrokken bij de opzet van de studie, daarna vanuit je functie als clustermanager van toenmalig cluster 1 en later je inhoudelijke bijdrage bij de papers.

Voor statistische vragen en het meedenken over passende analyses kon ik altijd terecht bij Marieke Timmerman. Dit heeft onder andere een prachtig artikel opgeleverd over welke specifieke subgroepen kinderen met een autismespectrumstoornis wel baat hebben bij een sociaalvaardigheidstraining en welke in mindere mate of juist niet. Ik heb veel gehad aan je uitleg over de verschillende statistische programma's waarmee ik de analyses voor de verschillende papers heb uitgevoerd en vervolgens natuurlijk de uitleg van de interpretatie van de uitkomsten. Dank voor het vertrouwen dat je me hierin gaf.

Ook co-auteur Sjoerd Sytema wil ik bedanken voor de fijne samenwerking, het meedenken in het analyseren van de schoolobservatie en het schrijven van het paper hierover.

De socialevaardigheidstraining (SoVa) die ik tijdens mijn promotietraject onderzocht heb (het ESTIA-onderzoek), is gebaseerd op het SoVa protocol zoals ontwikkeld door Barbara van den Hoofdakker en Lianne van der Veen-Mulders. Met dank aan Leonieke Vet en Sjoukje van Warners is dit protocol vervolgens aangepast voor het ESTIA-onderzoek. Daarnaast wil ik Sjoukje en Leonieke bedanken voor het trainen van alle behandelaren in het ESTIA-protocol en het verzorgen van de intervisie tijdens de SoVa's. Voor de uitvoering van de SoVa wil ik alle behandelaren bedanken. Doordat het een gerandomiseerd gecontroleerde studie was, hoorden jullie vaak pas op het laatste moment welke training je moest geven of dat je helemaal geen training hoefde te geven. Dank voor jullie flexibiliteit hierin en natuurlijk jullie betrokkenheid bij het geven van de training zelf: Erika Bakker, Kristel Bindels, Marjolein ten Cate, Lineke Davids, Simone Hein, Deranda Hoekzema-de Boer, Kim Guldenaar, Judith Havinga, Mara van den Heuvel, Marianne van der Horn, Jolanda Koster, Anneke Memelink, Harma Moorlag, Marjan Neef, Dorien Rump, Marjan Sikkema, Marina Spekken, Ilona Visser, Lieneke Voerman, Annelies Vonk, Irene de Vries, Sjoukje van Warners en Lara van der Zwaag. Daarnaast wil ik de poli's van Accare (Groningen en Drachten) en van het Autisme Team Noord-Nederland (Groningen en Drachten) bedanken dat de SoVa in het kader van het ESTIA-onderzoek gegeven kon worden. En natuurlijk gaat mijn dank uit naar alle medewerkers van de poli's die alert zijn geweest op mogelijke deelnemers en het ESTIA-onderzoek met ouders en kinderen hebben besproken.

Zonder de medewerking van alle ouders en kinderen had er geen socialevaardigheidstraining gegeven kunnen worden en was dit proefschrift er nooit gekomen. Mijn dank aan jullie is groot! Jullie hebben geaccepteerd dat door middel van loting bepaald zou worden welke socialevaardigheidstraining (of alleen reguliere zorg) jullie kind zou krijgen. Dat was niet altijd makkelijk en heeft ook weleens tot teleurstelling geleid. Daarnaast zijn jullie extra naar de poli gekomen voor meetmomenten binnen het ESTIA-onderzoek en hebben jullie vragenlijstboekjes ingevuld. Deze extra investering waardeer ik enorm.

Het inplannen van alle afspraken en het verzamelen van alle gegevens was ook een zeer grote klus en was nooit gelukt zonder de medewerking van alle onderzoeks-medewerkers. Annemaaike van der Wal, José Hoogervorst, Karijne Schaap, Kerstin Wolters, Marieke Messchendorp, Marjan Houwing, Neeltje Valkhof en Wya Rijpkema bedankt voor jullie inzet, flexibiliteit en harde werken. Marjan en Neeltje, zonder jullie access-bestand weet ik niet hoe we alles administratief zo goed hadden kunnen bijhouden. Daarnaast was dit een hele effectieve manier voor het versturen van de meeste brieven. Ook de bijdrage van alle studenten met betrekking tot het doen van schoolobservaties, het scoren van FMSS fragmenten, het afnemen van de interviews en het invoeren van data was erg fijn. Bedankt voor jullie inzet en enthousiasme:

Ananda, Anna, Bart, Berber, Cris, Denise, Dorian, Elise, Fardau, Femke, Floor, Grieke, Julia, Lana, Laura, Louise, Maartje, Marcella, Mariska, Mirjam, Nel, Nicole, Petri, Rob, Solveig, Sophie en Wesley. Daarnaast wil ik Coby Gerlsma bedanken voor haar FMSS-training aan de studenten en haar bijdrage aan de consensusbijeenkomsten van de FMSS.

De Accare researchgroep wil ik ook bedanken voor hun hulp, een luisterend oor of gewoon een ontspannende lunch. Anna, Andrea, Anne-Flore, Annelies, Annemaaïke, Anne-Marie, Barbara, Catharina, Christina, Denise, Dennis, Djûke, Elise, Ellen, Elly, Eke, Esther, Florianne, Grieke, Halewijn, Hyun, Jessica, José, Julie, Karijne, Kerstin, Lianne, Lizanne, Lotte, Maartje, Marieke, Mariken, Marian, Marjan, Mark-Peter, Mascha, Mohammed, Monika, Natalie, Neeltje, Nel, Netty, Pieter, Sanne, Simone, Rianne, Thaira, Yvonne G., Yvonne M. en Wya, bedankt.

Tijdens de laatste jaren van mijn proefschrift werkte ik tevens als psycholoog bij Molendrift. Een heel fijne werkplek waar ook werd meegedacht met betrekking tot wat ik nodig had om mijn proefschrift af te ronden. Met daarnaast fijne collega's die zorgden voor steun en motivatie tijdens de laatste loodjes. Ik vond het erg fijn dat ik deze ruimte heb gekregen en mede daardoor is het gelukt om mijn proefschrift af te ronden naast mijn reguliere baan.

Als laatste (maar zeker niet het minst belangrijk), wil ik mijn thuisfront bedanken. Allereerst mijn twee paranimfen, Dorien en Sanne. Ik ben blij dat jullie aan mijn zijde staan tijdens de dag waar ik al lang naar uitkijk. Alle jaren tijdens mijn promotietraject heb ik veel aan jullie gehad.

Ook vriendinnen en familie wil ik bedanken. Jullie hebben naar mij geluisterd, mijn gemopper aangehoord en mij gesteund waar nodig. Of juist voor ontspanning gezorgd, zodat ik daarna weer met frisse moed verder kon. Heel fijn dat jullie er altijd voor mij waren. De laatste periode ben ik erg druk geweest om alles af te ronden en was het wat lastiger om af te spreken, daar gaat vanaf nu verandering in komen!

Leon, Bas en Anne, wat ben ik ontzettend blij en dankbaar dat jullie in ons leven zijn gekomen. Vanaf dat moment werd het wel iets lastiger om naast mijn werkzaamheden bij Molendrift ook de papers voor mijn proefschrift te schrijven. Dankzij alle momenten dat opa's, oma's en pake hebben opgepast, is het gelukt om het nu af te ronden. Ontzettend bedankt, de jongens vinden het altijd ontzettend leuk om jullie te zien. Desondanks ben ik heel blij dat dit nu niet meer in deze mate nodig zal zijn.

Lieve Johan, het is uiteindelijk een langdurig traject geworden en ondanks dat je niet altijd precies wist waar ik mee bezig was, heb je me altijd gesteund en je waardering uitgesproken over datgene waar ik mee bezig was (en het doorzettingsvermogen dat daarvoor nodig was ;)). Het laatste jaar speelde dit in toenemende mate, aangezien je weer onregelmatig ging werken en overdag veel bij de jongens was. Dit heeft

er wel voor gezorgd dat ik nu (hoogzwanger) dit dankwoord aan het schrijven ben en dat we straks samen kunnen genieten van ons mooie gezin, zonder dat ik hoeft te zeggen: 'ik ga niet mee of ik kan nu niet, want ik moet nog even wat voor mijn proefschrift doen'. Want dat is nu klaar!

About the author

About the author

Vera Dekker was born in Assen on February 28, 1986. In 2004 she started to study Psychology at the University of Groningen. In 2009 she received her Master's degree in Clinical and Developmental Psychology. After graduation, she worked as a researcher at the University Center of Psychiatry (UCP) in Groningen, studying winter depression. In January 2010 Vera started with her PhD project about the effectiveness of social skills training for children with autism spectrum disorder. During the first 4.5 years she mainly focused on the coordination and data collection of the study. In the next years she analyzed the data and wrote the papers. During her PhD trajectory Vera has supervised several students. Since August 2015, Vera combined her PhD with her job as a psychologist at Molendrift, a mental health institution for children, adolescents, and adults with behavioral problems, parenting problems, learning problems, and problems related to work. Vera is involved in both diagnosis and treatment of mostly, children and adolescents with ADHD, autism spectrum disorder, anxiety, behavioral problems or learning problems. In the last year, Vera also worked in a general practice and treated children and adolescents, and their parents with onset psychological problems. Vera lives in Assen with Johan and their three children Leon (4 years), Bas (2 years), and Anne (0 years).

Publications

Publications

- Meesters, Y., Dekker, V., Schlangen, L.J.M., Bos, E.H., & Ruiter, M.J. (2011). Low-intensity blue-enriched white light (750 lux) and standard bright light (10 000 lux) are equally effective in treating SAD. A randomized controlled study. *BMC Psychiatry*, 11, 17.
- De Bildt, A., Oosterling, I.J., van Lang, N.D.J., Kuijper, S., Dekker, V., Sytema, S., Oerlemans, A.M., van Steijn, D.J., Visser, J.C., Rommelse, N.N., Minderaa, R.B., van Engeland, H., van der Gaag, R., Buitelaar, J.K., de Jonge, M.V. (2013). How to use the ADI-R for classifying autism spectrum disorders? Psychometric properties of criteria from literature in 1,204 Dutch children. *Journal of Autism and Developmental Disorders*, 43(10), 2280-2294.
- Dekker, V., Nauta, M.H., Mulder, E.J., Timmerman, M.E., & de Bildt, A. (2014). A randomized controlled study of a social skills training for preadolescent children with autism spectrum disorders: generalization of skills by training parents and teachers? *BMC Psychiatry*, 14, 189.
- Dekker, V., Nauta, M.H., Mulder, E.J., & de Bildt, A. (2015). ESTIA: Effectiviteit van sociale vaardigheidstraining bij kinderen met een autismespectrumstoornis. *Kind en Adolescent Praktijk*, 14(1), 26-29.
- Dekker, V., Nauta, M.H., Mulder, E.J., Sytema, S., & de Bildt, A. (2016). A fresh pair of eyes: A blind observation method for evaluating social skills of children with ASD in a naturalistic peer situation in school. *Journal of Autism and Developmental Disorders*, 46(9), 2890-2904.
- Dekker, V., Nauta, M.H., Timmerman, M.E., Mulder, E.J., van der Veen-Mulders, L., van den Hoofdakker, B.J., van Warners, S., Vet, J.J., Hoekstra, P.J., & de Bildt, A. (2016). Als wetenschap iets anders zegt dan intuïtie. Het geringe effect van sociale vaardigheidstraining bij kinderen met ASS. *Kind & Adolescent Praktijk*, 15, 14-21.
- Dekker, V., Nauta, M.H., Timmerman, M.E., Mulder, E.J., van der Veen-Mulders, L., van den Hoofdakker, B.J., van Warners, S., Vet, J.J., Hoekstra, P.J., & de Bildt, A. (2019). Social skills group training in children with autism spectrum disorder: a randomized controlled trial. *European Child & Adolescent Psychiatry*, 28(3), 415-424.
- Montazeri, F., de Bildt, A., Dekker, V., & Anderson, G.M. (2019). Network analysis of anxiety in the autism realm. *Journal of Autism and Developmental Disorders*, 49, 2219-2230.
- Dekker, V., Nauta, M.H., Timmerman, M.E., Mulder, E.J., Hoekstra, P.J., & de Bildt, A. (2021). Application of Latent Class Analysis to identify subgroups of children with Autism Spectrum Disorders who benefit from Social Skills Training. *Journal of Autism and Developmental Disorders*, 51, 2004-2018

Dekker, V., Nauta, M.H., Hoekstra, P.J., & de Bildt, A. (2021). *Parental stress and criticism of fathers and mothers of children with autism spectrum disorder in association with co-occurring internalizing and externalizing problems*. Manuscript submitted for publication.