

# Process Research in Early Intensive Intervention in Autism Spectrum Disorder: Sensitivity to Change of the Autism Behavior Coding System

Evelyn Herbrecht , Olga Lazari, Marianne Notter, Klaus Schmeck, and René Spiegel

The development of sensitive measures to capture changes in core autism symptoms is crucial in early intervention research. The study examines the sensitivity to change of the Autism Behavior Coding System (ABCS), a video-based instrument to assess core autism symptoms during therapist-child interaction. Video sequences of 40 young children treated in the Frühintervention bei Autistischen Störungen center were analyzed with regard to the question of whether short-term changes during an 18 day period of early intervention could be captured, and whether these results are reflected in an independent clinical assessment (Developmental Disorders-Child-Global Assessment Scale [DD-C-GAS]). ABCS results showed statistically significant improvements on behavioral domains such as “expression of wishes” and “social cooperative behavior” ( $P < 0.01$ ), less pronounced on “eye contact.” Improvements on the DD-C-GAS were highly significant on all subdomains. Both scales showed high correlations within their subdomains, yet no significant correlations between the changes in both instruments’ scores were found. An additional analysis between the DD-C-GAS scores at day 18 and the changes in the ABCS scores showed statistically significant associations in the expected direction between the changes in the variable “eye contact” and all DD-C-GAS subdomains. The correspondence of the two levels of assessment is low, but the specifics of this relationship deserve further study. The ABCS may prove useful in addition to standard assessment tools, especially in early intervention research settings, as it allows reliable analysis of core behavioral elements in young children with autism. *Autism Res* 2019, 12: 1817–1828. © 2019 The Authors. *Autism Research* published by International Society for Autism Research published by Wiley Periodicals, Inc.

**Lay Summary:** The study examined the sensitivity of an autism-specific video coding system (ABCS) in assessing changes after an 18 day period of intensive early intervention. Video sequences of therapist-child-interaction of 40 young children with autism spectrum disorder (ASD) were analyzed. Children’s behavior improved in expression of wishes, social cooperativity and eye contact. A therapist-based global assessment scale also showed important improvement after 18 days, yet both assessment instruments showed weak correlations between their respective changes. We showed that the ABCS may prove useful in capturing short-term changes in autism-related behaviors, especially in early intervention research.

**Keywords:** autism spectrum disorder; early intensive intervention; process research; ABCS; sensitivity to change

## Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder marked by impairments in social communication and interaction, and by pattern of restricted and repetitive interests and behaviors [American Psychiatric Association, 2013]. Impairment in social communication and responsiveness is a core feature of ASD and is already observed in very young children [Williams White, Keonig, & Scahill, 2007]. As they are considered an early indicator of ASD, improvement of restricted social communication skills is a highly important target of early intervention approaches, beyond improvement of parent-child interaction and social adaptive functioning [Ozonoff et al., 2010; Brian, Bryson, & Zwaigenbaum,

2015]. Early intensive interventions are aimed at enhancing precursors of social communication such as joint attention, play, and imitation. Repetitive behaviors represent another clinically meaningful target for early intervention and should be included in measures of outcome [Scahill et al., 2015].

Sensitive measures to capture behavioral changes and, more specifically, changes in core autistic symptoms, are rare [McConachie et al., 2015]. Improvement in social communication may occur subtly, and is thus difficult to capture by global measures of intervention responses. Anagnostou et al. [2015] evaluated existing social communication measures and defined key classification criteria such as clinical relevance and psychometric properties (validity, reliability, and sensitivity to change). Sensitivity

From the Department of Child and Adolescent Psychiatry, University Hospital for Psychiatry Basel, Basel, Switzerland (E.H., O.L., M.N., K.S., R.S.)

Received November 13, 2018; accepted for publication July 15, 2019

Address for correspondence and reprints: Evelyn Herbrecht, Department of Child and Adolescent Psychiatry, University Hospital for Psychiatry Basel, Kornhausgasse 7, CH-4051 Basel, Switzerland. E-mail: evelyn.herbrecht@upk.ch

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Published online 9 August 2019 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/aur.2188

© 2019 The Authors. *Autism Research* published by International Society for Autism Research published by Wiley Periodicals, Inc.

to change is considered an especially crucial criterion given that only few interventions have so far shown to alter core autism symptoms. Several social communication measures have been identified as being appropriate for use in evaluation studies, however, only few of them are suitable for use in studies of young children. The majority are caregiver or teacher ratings, and results of those are likely to be biased by subjective expectations [Anagnostou et al., 2015]. Another limitation of most existing measures is that they show sensitivity to change only for longer-term interventions, that is, lasting 1 year or longer. One particular metric, the Autism Diagnostic Observation Schedule (ADOS) [Lord et al., 2012; German version: Poustka et al., 2015] severity score is regularly used as an outcome measure in early intervention studies of autism, as it allows standardization and comparability across modules. However, it has not proved to be sufficiently sensitive to capture relevant behavioral changes [McConachie et al., 2015]. Grzadzinski et al. [2016] reported preliminary data on the Brief Observation of Social Communication Change, a video-based measure of treatment response for social communication behaviors derived from the ADOS-2 codes, providing global scores representing children's behavior over a given period of time. These authors describe promising preliminary results, although, to our knowledge, development of the instrument is not yet finalized. In conclusion, none of the existing measures can be considered as fully satisfactory, that is, all show some limitations, be it regarding their psychometric properties, preconditions of administering them or their sensitivity to change.

In order to address the need for sensitive measures to capture core autism-related behaviors focusing on the social communication domain as well as on repetitive behaviors, we started to work on a video-based assessment tool, called the Autism Behavior Coding System (ABCS) in 2011 and published preliminary data on its psychometric properties, as well as its sensitivity to change to proximal intervention effects [Dima et al., 2017]. The ABCS was originally developed to evaluate the effectiveness of the "Frühintervention bei Autistischen Störungen" (FIAS) early intervention approach. FIAS is a play-based, intensive early intervention approach for children with autism aged up to 5 years. It is intended to encourage children to take part in social interactions by increasing shared enjoyment and social motivation [see Herbrecht et al., 2015 for a detailed description]. FIAS treatment consists of an initial 18 day intervention period with the core family at the FIAS center, followed by 2 years of follow-up. During this follow-up period, a designated FIAS therapist supports the family, offers regular coaching, analyzes play interventions at home, and exchanges information with other involved institutions in order to follow the development of the child. Given the short initial period of very intensive intervention

compared to other approaches with lower but continuous intensities over 1 year, our goal was to quantitatively assess clinically evident changes following the 18 day FIAS intervention period. For this purpose, we needed an instrument that is sufficiently sensitive to changes over a relatively short period of time.

As recently published by our group, results from a small patient sample suggested that the ABCS could be suitable for capturing short-term intervention effects, and could be useful to evaluate early interventions for ASD [Dima et al., 2017]. However, the changes detected using the ABCS were not statistically significant, most likely due to the small size of the patient sample. The inter-rater reliability (IRR) ranged from high to very high for most of the ABCS variables.

The objective of the current study is to examine the sensitivity to change of the ABCS within a relevant patient sample with regard to the following questions: (a) Is the ABCS able to capture short-term changes during an 18 day period of very intensive early intervention? (b) Is there a relationship between ABCS results and an independent, more general clinical assessment, as reflected in the Developmental Disorders-Child-Global Assessment Scale [DD-C-GAS; Wagner et al., 2007; German translation: Bölte, 2008]?

## Methods

### *Participants*

Participants were 40 young children with autism according to DSM-IV [American Psychiatric Association, 1994] and enrolled in the FIAS early intensive intervention approach [Herbrecht et al., 2015] between 2011 and 2017. They had been diagnosed using the ADOS-2 Module 1 and the Autism Diagnostic Interview-Revised [ADI-R; Rutter, LeCouteur, & Lord, 2003; German version: Bölte, Rühl, Schmötzer, & Poustka, 2006]. None of the children had relevant additional neurological or somatic problems. Families were all currently living in Switzerland, but had different nationalities. All parents spoke either German or English. At the start of the intervention, the children were classified as verbal or nonverbal according to the current language level score of item 29 on the ADI-R. Parents gave written informed consent to the anonymized use of the data.

### *Intervention*

During the initial 18 day period in the FIAS center, the child with autism receives about 100 hr of play-based intervention. Basic everyday activities such as having lunch or toilet training are also part of the intervention. FIAS focuses on the development of a reciprocal relationship between therapists and the child and offers social interaction opportunities without directly prompting behaviors or targeting isolated skills. Besides the reference

therapist working most with the child, up to three additional therapists hold play sessions with the child, so that communication and interaction behaviors can be generalized to different persons [see Herbrecht et al., 2015 for details]. The majority of play sessions, each lasting between 75 and 90 min, are recorded for documentation and intervention planning, as well as for video analysis together with the parents. FIAS was originally derived from the Mifne approach [Vorgraft, Farbstein, Spiegel, & Apter, 2007].

#### Assessment Procedures

**DD-C-GAS.** As in our previous study [Dima et al., 2017], we used the DD-C-GAS, an ASD-adapted version of the Children's Global Assessment Scale [Schaffer et al., 1983], as the reference measure of the children's level of functioning. DD-C-GAS results from a previous study showed important improvement during intervention [Dima et al., 2017]. In the current study, we used the DD-C-GAS as an independent external instrument to be able to determine whether positive changes on this scale after FIAS interventions would be reflected in the ABCS. The DD-C-GAS ratings were made at baseline and at the end of 18 days of intensive intervention (days 1 and 18) by the lead FIAS therapist working with the respective child. The DD-C-GAS comprises ten categories of functioning, using a 0 to 100 scale within four domains: everyday functioning, intellectual performance, communication, and social behavior. Scores below 70 indicate that a child has important special needs in a particular domain.

**ABCS.** The ABCS is a video-based observational instrument for assessing core autism symptoms during semistructured adult-child interaction [see Dima et al., 2017 for details]. In contrast to the majority of other observational tools available for use in evaluation studies of early intensive interventions, the ABCS is used by independent raters not involved in the intervention process. The ABCS enables an external evaluator to assess core autism-related behaviors and autism-specific interaction characteristics. It focuses on basic, autism-sensitive aspects of interaction and communication (Table 1 for ABCS variable description).

The ABCS is an event record-based coding system and provides a detailed, continuous protocol of the presence or absence of a particular behavior and its total duration (in seconds) or frequency within a given observation period. Higher scores indicate an improvement of the specified behavior, except for the duration of repetitive behavior where improvement will be reflected in lower scores. ABCS categories focus on core autism behavioral features and reflect the objectives of the FIAS intervention approach, as FIAS attempts to enhance social motivation and spontaneous communication and interaction

**Table 1. ABCS Behavioral Variables**

Variables	Scale unit	Description
Eye contact	Frequency	<ul style="list-style-type: none"> <li>Looks at therapist's eyes/face, no matter how short</li> </ul>
Joint attention	Frequency	<ul style="list-style-type: none"> <li>Child follows therapist's gaze/pointing (responding to joint attention)</li> <li>Child brings objects, points to objects, or uses sounds/words to direct therapist's attention (initiating joint attention)</li> </ul>
Expression of wishes	Frequency	<ul style="list-style-type: none"> <li>Spontaneous expression of wishes addressed to the therapist. Yes/no answers and making a choice are included</li> </ul>
Imitation	Frequency	<ul style="list-style-type: none"> <li>Child repeats an action</li> <li>Child repeats a sound/word/sentence (no echolalia)</li> </ul>
Social cooperative behavior	Duration (sec)	<ul style="list-style-type: none"> <li>Child is engaged in social interaction: helping the therapist to prepare the table for lunch, cleaning after lunch, preparing the play scene, choosing together games</li> </ul>
Functional play	Duration (sec)	<ul style="list-style-type: none"> <li>Child interacts with the therapist and participates actively to play</li> </ul>
Repetitive behavior	Duration (sec)	<ul style="list-style-type: none"> <li>Ritual play</li> <li>Sensory-motor stereotypes</li> <li>Vocal stereotypes</li> <li>Motor and vocal/verbal stereotypes combined</li> <li>Sensory-motor/other repetitive behaviors</li> <li>Verbal stereotypes</li> </ul>
Other	Duration (sec)	<ul style="list-style-type: none"> <li>Passage default code describing all behaviors that are not included in any of the previous categories, i.e., moving from one place to another, lying on the floor, eating/drinking</li> </ul>
Unscorable	Duration (sec)	<ul style="list-style-type: none"> <li>The child cannot be seen on the screen for at least 3 sec</li> </ul>

Abbreviation: sec, seconds.

behaviors. Therapist's behaviour is not coded separately, but therapists are constantly aiming at offering the most possible interaction opportunities to the child.

#### Video Data Collection and Coding

Continuous 60-min video sequences of therapist-child interaction were coded at three time points: at intervention start (Baseline, corresponding to day 4 of the intensive intervention, after the child could adapt to the situation and was separated from its parents during the day), at midintervention (day 9 or 10), and at the end of the intensive period (day 17). Based on findings of our earlier study [Dima et al., 2017], we now decided to

choose a longer time period for scoring, which should provide a more stable base for assessments across the treatment period. Additionally, we only chose morning sequences with the same reference therapist at all time points. All but one therapist were females. As play sessions last between 75 and 90 min each, we chose the first 60 min of the respective sequences. Video sequences were viewed once for coding. It takes approximately 90 min to complete the coding of one 60 min sequence.

Video sequences were analyzed using Interact (Mangold International GmbH), a professional software package that provides multiple functionalities including the ability to watch a child's behavior in slow motion, automatic recording of the frequency and duration of specific behaviors, graphical displays of data, and the option to view behavior from two perspectives (e.g., recorded from opposite corners of the room) simultaneously. Depending on a child's movements in the room, it is possible that he or she could temporarily not be seen by either of the two cameras. This situation corresponds to the ABCS score of "unscorable." During the 60-min video sequence, we allowed "unscorable" scores of at most 5 min. Coding was performed by two clinical expert raters (both clinical psychologists, Olga Lazari, second author of this article, and Diana Dima) experienced in the treatment of children with ASD, but not involved in the FIAS intervention. Raters were blinded with regard to the children's intervention status by means of randomizing the three video sequences of each child before submitting them for assessment.

Based on our ongoing work experience with the ABCS, we added some specifications and redefined some ABCS variables. Most importantly, we replaced the variable "social positive affect" with "social cooperative behavior," defining this variable by more easily observable behavioral features than the former term of "affect" (see Methods and Table 1 for variable definitions). For all other variables, more precise coding definitions and examples have been elaborated in order to reduce ambiguity. Follow-up analysis of the IRR after revision of the ABCS variable definitions, using video sequences of 60-min duration, showed high to very high intra-class correlations (ICC) for the following behavioral variables: frequency of eye contact (ICC = 0.96), frequency of expression of wishes (ICC = 0.95), duration of functional play (ICC = 0.89), and duration of repetitive behavior (ICC = 0.86). IRR was moderate to high for frequency of joint attention (ICC = 0.72), frequency of imitation (ICC = 0.79), and duration of social cooperative behavior (ICC = 0.76). The variables are listed and described in Table 1. Details of the coding scheme are provided in the ABCS manual [Herbrecht, Lazari, Dima, Spiegel, & Schmeck, 2017] and are available from the authors upon request.

### *ABCS "Unscorable" Duration and Adjustment of ABCS Values*

In a total of 20 subjects, one or more assessments resulted in an "unscorable" duration of more than 5 min. For seven children, the maximum was between 10 and 20 min, and one child spent 21 min in "unscorable" because of a prolonged toilet procedure. In order to compensate for these high values of "unscorable," the values of the other ABCS variables were extrapolated to values corresponding to the total observation time of 3,600 s.

**Statistical methods.** Data analysis was performed using nonparametric methods. Short-term behavioral changes assessed by ABCS and DD-C-GAS were evaluated using Friedman's rank sum test for repeated measurements (three time points), and Wilcoxon signed rank tests for comparisons between two time points. If the Friedman test indicated significance, Wilcoxon signed rank tests were performed between each of two time points, with main interest in the comparisons between day 4 versus day 10 and day 4 versus day 17. The outcome measurements were the DD-C-GAS subdomains and the eight ABCS categories including "Other" (Table 1). In order to explore the extent of relationships within the DD-C-GAS subdomains, within the ABCS categories, and between the changes from baseline observed in the DD-C-GAS and the ABCS scores, analyses were performed using Spearman's rank correlation.

## **Results**

### *Background and Demographic Data*

We analyzed data of 40 children (32 males and 8 females) treated at the FIAS center between June 2011 and October 2017 (see Table 2 for patient characteristics at baseline). The sample was split into two equally sized groups according to the median age, that is,  $\leq 43$  months and  $> 43$  months, and two groups according to the time of intervention start, that is, up to May 04, 2015 and after May 04, 2015, and according to sex and language level. No relevant group differences were observed for the time of intervention start. Therefore, this criterion was not considered for further analysis. As only three children were rated as verbal, the language level category was not maintained either. The median ADOS severity score at baseline was significantly higher (Mann-Whitney-*U*-test,  $U = 124$ ,  $P = 0.034$ ) in participants older than 43 months (median value 8.0 vs. 6.0 for those  $\leq 43$  months). Six of the eight female participants were in the older age group. Given the small number of girls we describe baseline values for the two sexes, but did not analyze sex differences with regard to changes in outcome parameters.



**Table 2. Summary Statistics of Background and Demographic Data for All Participants and by Subgroups**

Parameter	All <i>N</i> = 40	Age		Sex	
		≤43 mo. <i>N</i> = 21	>43 mo. <i>N</i> = 19	Male <i>N</i> = 32	Female <i>N</i> = 8
Sex— <i>n</i> (%)					
Male	32 (80.0)	19 (90.5)	13 (68.4)	32	0
Female	8 (20.0)	2 (9.5)	6 (31.6)	0	8
Age (months)					
Mean (SD)	43.9 (7.04)	38.5 (4.14)	49.7 (4.31)	43.4 (7.74)	45.6 (2.50)
Median	43.0	39.0	48.0	43.0	46.0
Min–max	31–59	31–43	44–59	31–59	42–49
Start of intervention— <i>n</i> (%)					
Up to May 04, 2015	20 (50.0)	14 (66.7)	6 (31.6)	15 (46.9)	5 (62.5)
after May 04, 2015	20 (50.0)	7 (33.3)	13 (68.4)	17 (53.1)	3 (37.5)
ADOS total score					
Mean (SD)	19.8 (3.76)	18.9 (3.70)	20.7 (3.70)	19.8 (3.93)	19.8 (3.20)
Median	20.0	20.0	21.0	20.0	19.0
Min–max	14–28	14–24	16–28	14–28	17–26
ADOS severity score					
Mean (SD)	7.0 (1.54)	6.5 (1.40)	7.5 (1.54)	7.1 (1.58)	6.8 (1.39)
Median	6.0	6.0	8.0	7.0	6.0
Min–max	5–10	5–9	6–10	5–10	6–10

Abbreviations: mo., months; *n*, sample size; SD, standard deviation; ADOS, autism diagnostic observation schedule.

### ABCS Findings at Baseline

Summary statistics of baseline ABCS scores for all subjects and for subgroups classified by age and sex are presented in Table 3. Some of the variables that are assessed by frequency were observed more often (for example “eye contact”) than others such as “joint attention,” which occurred only rarely. Male children tended to score higher on “expression of wishes,” while females scored higher on “imitation.” However, there were no significant age- or sex-related group differences for any of the ABCS variables at baseline. A nonsignificant trend toward less time spent with “repetitive behavior” within the female group and within the older children’s group was observed.

### DD-C-GAS Findings: Baseline and Changes from Baseline

Mean subdomain scores at baseline (data not shown) ranged from  $31 \pm 14.2$  (communication) to  $42 \pm 17.8$  (intellectual skills). Self-care had a mean score of  $32 \pm 16.2$ , while social behavior scored  $26 \pm 12.8$ . Subdomain median scores for self-care and communication were higher in the female group (data for subgroups not shown). There were statistically significant changes (i.e., improvements) in all four DD-C-GAS domains, most importantly in the communication and social behavior subdomains corresponding to the large effect sizes (Table 4). Improvements were of similar magnitude in both sexes (data not shown).

### Changes in ABCS Scores from Baseline

There were positive changes from baseline in some ABCS categories, most pronounced in “expression of

wishes” (frequency) and “social cooperative behavior” (duration), and somewhat less for “eye contact” (frequency). The Friedman test indicated significant to highly significant changes in “expression of wishes” and “social cooperative behavior,” “other,” and a trend for “eye contact” (Table 5). The change in “social cooperative behavior” was statistically significant between day 10 and baseline, and highly significant between day 17 and baseline (Wilcoxon test) corresponding to a large effect size. For “expression of wishes,” the change was significant between day 17 and baseline with a moderate effect size. Numbers were initially high for the variable “other” and decreased significantly from baseline to day 17, while the “unscorable” scores remained high throughout. The statistical analysis after adjustment for the durations of “unscorable” did not reveal any relevant differences from the original analysis (data not shown).

### Correlations of Changes Within the DD-C-GAS and ABCS Findings, and Between DD-C-GAS and ABCS Findings

There were positive and highly significant correlations between the changes within all four DD-C-GAS subdomains (Table 6). The ABCS-categories “joint attention” and “imitation” were excluded from correlation analysis because of their low frequency. There were positive and highly significant correlations between changes in the categories “eye contact” and “expression of wishes,” as well as between changes in “eye contact” and “social cooperative behavior.” Changes in “functional play” correlated significantly with changes in “eye contact” and

**Table 3. Summary Statistics of Baseline ABCS Scores for All Participants and by Subgroups**

	All <i>N</i> = 40	Age		Sex	
		≤43 mo. <i>N</i> = 21	>43 mo. <i>N</i> = 19	Male <i>N</i> = 32	Female <i>N</i> = 8
Eye contact—frequency					
Mean (SD)	38.9 (28.62)	39.2 (23.25)	38.6 (34.26)	38.9 (29.55)	39.0 (26.40)
Median	31.0	38.0	26.0	31.0	30.5
Joint attention—frequency					
Mean (SD)	2.3 (4.27)	3.1 (5.61)	1.4 (1.71)	2.5 (4.70)	1.4 (1.60)
Median	1.0	1.0	1.0	1.0	1.0
Expression of wishes—frequency					
Mean (SD)	5.8 (6.28)	5.9 (7.06)	5.7 (5.5)	6.4 (6.58)	3.4 (4.44)
Median	4.0	3.0	5.0	4.0	0.5
Imitation—frequency					
Mean (SD)	1.9 (3.16)	2.4 (4.14)	1.3 (1.37)	1.5 (2.99)	3.4 (3.54)
Median	1.0	1.0	1.0	1.0	3.5
Functional play—duration (sec)					
Mean (SD)	1,006.3 (576.21)	954.8 (609.00)	1,063.1 (548.44)	980.9 (553.74)	1,107.6 (690.80)
Median	972.5	957.2	1,034.6	972.5	942.1
Repetitive behavior—duration (sec)					
Mean (SD)	758.7 (752.58)	924.3 (867.81)	575.7 (568.35)	769.5 (713.98)	715.4 (946.19)
Median	543.6	558.3	425.7	552.0	305.3
Social cooperative behavior—duration (sec)					
Mean (SD)	254.3 (216.64)	246.9 (205.25)	262.6 (233.97)	255.8 (216.13)	248.4 (233.57)
Median	232.1	229.7	234.4	237.8	167.6
Other—duration (sec)					
Mean (SD)	1,379.6 (728.37)	1,323.9 (731.30)	1,441.2 (740.00)	1,412.7 (738.23)	1,247.4 (719.26)
Median	1,294.1	1,260.1	1,343.8	1,294.1	1,205.1
Unscorable—duration (sec)					
Mean (SD)	149.2 (195.84)	103.7 (139.71)	199.4 (237.36)	126.5 (180.54)	239.9 (239.95)
Median	64.1	40.4	71.9	50.4	243.5

Note. There are no significant group differences for any of the variables.  
Abbreviations: mo., months; *n*, sample size; SD, standard deviation; sec, seconds.

“expression of wishes.” Significant negative correlations were found between “repetitive behavior” and the variable “other,” as well as between “repetitive behavior” and

“functional play” (Table 6). None of the correlations between the changes in any of the DD-C-GAS subdomains and the ABCS categories were statistically

**Table 4. Summary Statistics of DD-C-GAS Findings by Time Point and Change from Baseline and Wilcoxon Signed Rank Test Results**

Score	Baseline <i>N</i> = 40	Day 18 <i>N</i> = 40	Day 18—baseline	
			<i>N</i> = 40	Effect size <sup>a,b</sup>
Self-care				
Mean (SD)	32.0 (16.17)	44.6 (15.82)	12.6 (11.17)	
Median	30.0	45.0	10.0	0.76***
Communication				
Mean (SD)	30.9 (14.22)	43.8 (15.15)	12.9 (8.27)	
Median	26.5	41.0	15.0	0.80***
Social behavior				
Mean (SD)	26.2 (12.79)	39.9 (13.18)	13.7 (9.90)	
Median	24.0	41.0	11.0	0.82***
Intellectual skills				
Mean (SD)	41.9 (17.83)	50.8 (18.58)	9.0 (7.79)	
Median	41.0	50.0	8.5	0.75***

<sup>a</sup>Effect size (*r*), 0.1, small; 0.3, medium; 0.5, large.

<sup>b</sup>Wilcoxon signed rank test.

Abbreviations: *n*, sample size; SD, standard deviation.

\*\*\**P*-value <0.001.

**Table 5. Summary Statistics of ABCS Variables by Time Point and Change from Baseline and Results of Friedman Test of ABCS Scores**

	Baseline N = 40	Day 10 N = 40	Day 17 N = 40	Friedman test P-value	Day 17—baseline N = 40	Effect size <sup>a,b</sup>
Eye contact—frequency						
Mean (SD)	38.9 (28.62)	42.1 (36.29)	45.4 (33.83)		6.4 (25.40)	
Median	31.0	27.0	38.0	0.072	6.5	0.22
Joint attention—frequency						
Mean (SD)	2.3 (4.27)	2.2 (4.32)	3.3 (5.97)		1.0 (3.76)	
Median	1.0	1.0	0.0	0.479	0.0	0.21
Expression of wishes—frequency						
Mean (SD)	5.8 (6.28)	7.4 (7.55)	7.6 (6.98)		1.8 (6.00)	
Median	4.0	5.0	5.0	0.006**	2.0	0.32*
Imitation—frequency						
Mean (SD)	1.9 (3.16)	2.1 (2.46)	1.7 (1.56)		-0.2 (3.37)	
Median	1.0	1.0	1.5	0.593	0.0	0.07
Functional play—duration (sec)						
Mean (SD)	1,006.3 (576.21)	1,091.9 (503.85)	1,095.6 (582.26)		89.4 (618.97)	
Median	972.5	1,081.5	1,085.4	0.792	-28.0	0.01
Repetitive behavior—duration (sec)						
Mean (SD)	758.7 (752.58)	697.9 (675.13)	672.4 (699.62)		-86.3 (589.51)	
Median	543.6	526.8	418.3	0.744	2.7	0.11
Social cooperative behavior—duration (sec)						
Mean (SD)	254.3 (216.64)	358.6 (266.40)	394.3 (213.21)		140.0 (223.83)	
Median	232.1	303.5	369.4	0.010*	116.4	0.55***
Other—duration (sec)						
Mean (SD)	1,379.6 (728.37)	1,150.2 (631.40)	1,140.3 (693.83)		-239.3 (587.65)	
Median	1,294.1	1,015.3	924.9	0.037*	-199.1	0.38**
Unscorable—duration (sec)						
Mean (SD)	149.2 (195.84)	259.3 (210.92)	251.7 (290.51)		102.5 (273.01)	
Median	64.1	198.0	134.0	n.a.	73.5	n.a.

<sup>a</sup>Effect size (*r*), 0.1, small; 0.3, medium; 0.5, large.

<sup>b</sup>Wilcoxon signed rank test.

Abbreviations: *n*, sample size; SD, standard deviation; sec, seconds; n.a., not applicable.

\**P*-value <0.05; \*\**P*-value <0.01; \*\*\**P*-value <0.001.

significant. There was also no pattern of consistently positive or negative, albeit not significant, correlations (Table 6).

Scatterplots of changes in ABCS categories versus changes in DD-C-GAS subdomains of the 40 children illustrate the heterogeneity of the individual results (see Fig. 1 as an example). However, the vast majority of data points are consistently found in the upper right quadrant of all scatterplots, indicating positive changes of varying magnitude in both areas investigated (DD-C-GAS and ABCS).

#### *Correlations Between DD-C-GAS Subdomain Scores at Day 18 and ABCS Categories Change Scores*

In view of the lack of correlation between the changes in the ABCS categories and the DD-C-GAS scores, and in order to further explore the relationship between the two assessment instruments, an additional analysis was performed between DD-C-GAS scores at day 18 and the changes in ABCS category scores. DD-C-GAS scores showed numerically high and statistically significant inter-correlations (upper part of Table 7), indicating that

the four DD-C-GAS subdomains are not independent from one another. Changes in the ABCS category “eye contact” showed positive and statistically significant correlations with all four DD-C-GAS subdomains (lower part of Table 7; Figs. 2 and 3). Changes in the ABCS category “expression of wishes” showed consistently positive correlations with all four DD-C-GAS subdomains, but only the one with “self-care” was statistically significant.

## Discussion

For therapeutic approaches to ASD that use a relatively short initial period of highly intensive intervention, as exemplified by FIAS, it is important to have an instrument available that will objectively capture behavioral changes within this period. The ABCS was designed for this purpose and was shown, in a first study, to provide adequate IRR. Compared to our previous publication [Dima et al., 2017], the video observation periods used for analysis have now been extended to 60 min, and for

**Table 6. Correlations Between ABCS Scores and DD-C-GAS Findings—Change from Baseline (*N* = 40)**

Spearman- $\rho$	DD-C-GAS—change day 18—baseline					ABCS—change from baseline				
	Self-care	Communication	Social behavior	Intellectual skills	Eye contact	Expression of wishes	Functional play	Repetitive behavior	Social coop. behavior	
DD-C-GAS—change day 18—baseline	1									
Communication	0.697 <sup>b</sup>	1								
Social behavior	0.613 <sup>b</sup>	0.672 <sup>b</sup>	1							
Intellectual skills	0.423 <sup>b</sup>	0.485 <sup>b</sup>	0.403 <sup>a</sup>	1						
Eye contact	-0.071	0.031	0.087	0	1					
Expression of wishes	-0.193	-0.099	0.03	-0.164	0.547 <sup>b</sup>	1				
Functional play	-0.167	-0.059	-0.068	0.17	0.314 <sup>a</sup>	0.370 <sup>a</sup>	1			
Repetitive behavior	0.061	0.045	-0.065	-0.287	-0.202	-0.169	-0.281	1		
Social coop. behavior	-0.202	-0.109	-0.111	-0.15	0.407 <sup>b</sup>	0.145	-0.101	-0.211	1	
Other	0.069	0.003	0.105	-0.047	-0.095	0.018	-0.431 <sup>b</sup>	-0.429 <sup>b</sup>	-0.045	

<sup>a</sup>Correlation significant at level 0.05 (two-sided).

<sup>b</sup>Correlation significant at level 0.01 (two-sided).

Abbreviations: DD-C-GAS, developmental disorder-child-global assessment; ABCS, autism behavior coding system; coop., cooperative; *n*, sample size.

each child only time segments spent with his or her individual reference therapist were chosen for analysis.

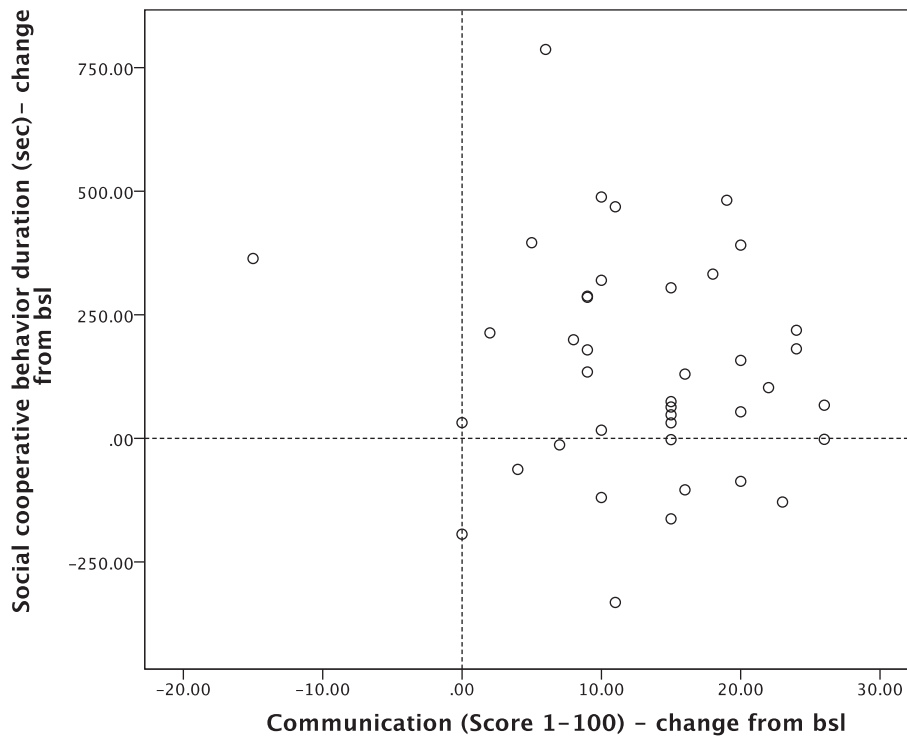
The primary objective of the current study was to examine the sensitivity of the ABCS with regard to its ability to capture short-term changes – from baseline to end of the treatment (day 4 to day 17)—in young children with autism during the period of very intensive early intervention. We also investigated whether and how the changes observed in any of the ABCS categories corresponded with assessments made by the children’s reference therapists, who used a more comprehensive and metrically different scale, called DD-C-GAS.

A first observation in the current study of 40 young children with autism, treated according to the FIAS approach, concerns the different frequencies of occurrence of the four event-based behavioral categories of the ABCS. While the frequency of “eye contact” within 60 min was relatively high, the category “expression of wishes” was scored less frequently, and instances of “joint attention” as well as “imitation” were rare. This general pattern did not change during the intervention period. Although this numerical inequality could be seen as a metric weakness of the ABCS categorical system, one should consider the likelihood that the frequency of occurrence of the desirable behaviors, such as “joint attention,” will increase as children develop more adaptive behavior with continued treatment. The inequality was less pronounced for the four time-based ABCS categories: “functional play,” “repetitive behavior,” “social cooperative behavior,” and “other.”

Another interesting observation was the lack of strong associations between the ABCS categories and age and sex in our patient sample. This is paralleled by a lack of strong association between these demographic factors and the severity of autism, as expressed by the ADOS scores in our sample. Within the ABCS outcomes, a negative correlation was found between “repetitive behavior,” as well as “functional play,” and the variable “other.” The variable “other” summarizes a number of behavioral elements which are not sufficiently distinctive to be identified as one of the specified variables. The duration of this variable declined during intervention. We interpret this finding, as well as the negative correlation with other variables, as a development toward behavioral differentiation throughout the intensive intervention, implying that unspecific behavior is reduced.

The third, and probably the most important, finding of this study concerns the sensitivity to change of the ABCS categories. The frequency of “expression of wishes” increased from a mean of  $5.8 \pm 6.28$  to  $7.6 \pm 6.98$  ( $P < 0.01$ ), and the duration of periods of “social cooperative behavior” from  $254.3 \pm 216.64$  to  $394.30 \pm 213.21$  s ( $P < 0.01$ ). “Eye contact” showed a similar change, albeit not statistically significant. In accordance with these decreases the duration of “other” periods decreased from





**Figure 1.** Scatterplots of correlations between changes in Autism Behavior Coding System social cooperative behavior versus Developmental Disorders-Child-Global Assessment Scale communication ( $N = 40$ ). Abbreviations: Sec, seconds. Bsl, baseline.

1,379.6 ± 728.37 to 1,140.3 ± 693.83 ( $P < 0.05$ ). Given the short duration of the intensive FIAS intervention, we consider this as a rather remarkable finding, both for the FIAS therapeutic approach as well as for the ABCS. If compared to the highly significant improvements noted by means of the DD-C-GAS, the changes seen on the ABCS variables may appear modest. However, the DD-C-GAS assessments were made both at baseline and end of intervention by the children’s reference therapists, thus a strong bias on their part, resulting in overly positive

assessments on this essentially subjective scale, cannot be excluded.

So far, the findings discussed are in line with the clinical focus of the FIAS intervention, as FIAS attempts to enhance social motivation and spontaneous communication and interaction behaviors. The child is almost continuously encouraged to express needs, social intentions, and emotions spontaneously and to take part in social play with the therapist. Important changes of the variables “expression of wishes” and “social cooperative

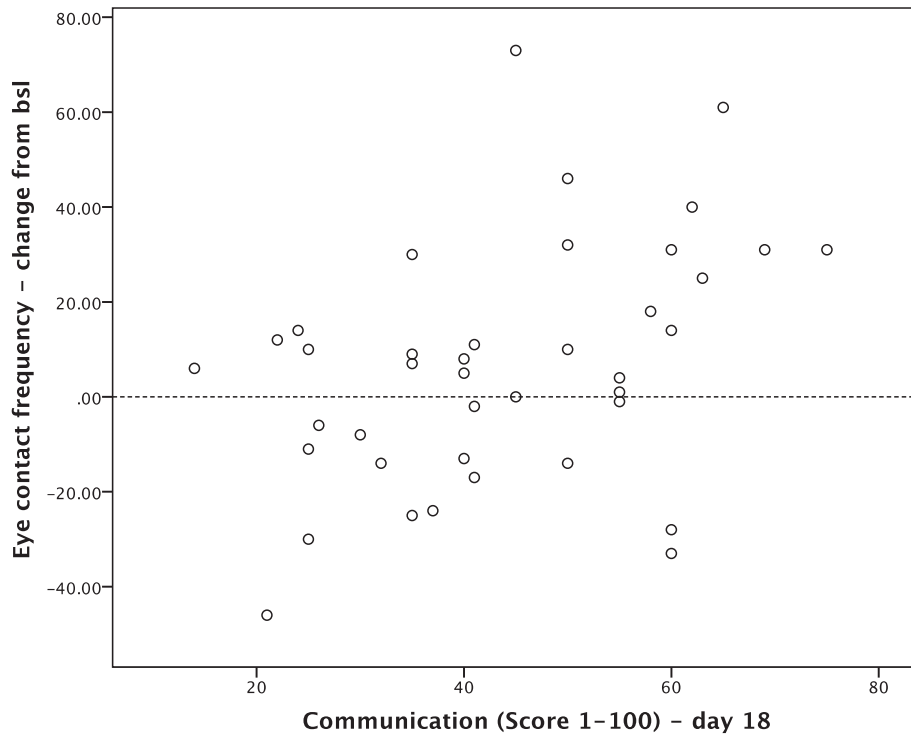
**Table 7. Correlations Between ABCS Change Scores and DD-C-GAS Findings Day 18 ( $N = 40$ )**

Spearman-rho		DD-C-GAS—day 18			
		Self-care	Communication	Social behavior	Intellectual skills
DD-C-GAS—day 18	Self-care	1			
	Communication	0.844 <sup>b</sup>	1		
	Social behavior	0.775 <sup>b</sup>	0.755 <sup>b</sup>	1	
	Intellectual skills	0.795 <sup>b</sup>	0.906 <sup>b</sup>	0.785 <sup>b</sup>	1
ABCS—change from baseline	Eye contact	0.422 <sup>b</sup>	0.405 <sup>b</sup>	0.339 <sup>a</sup>	0.470 <sup>b</sup>
	Expression of wishes	0.411 <sup>b</sup>	0.296	0.218	0.265
	Functional play	0.101	0.188	0.06	0.201
	Repetitive behavior	−0.067	−0.096	−0.177	−0.149
	Social coop. behavior	0.06	0.026	0.037	−0.022
	Other	−0.216	−0.301	−0.163	−0.209

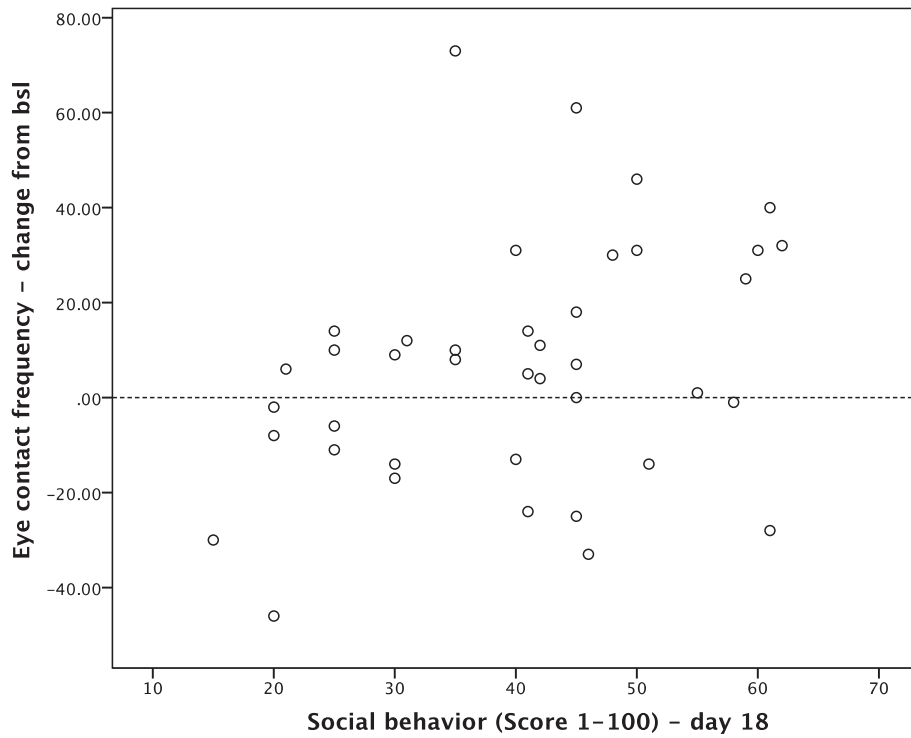
<sup>a</sup>Correlation significant at level 0.05 (two-sided).

<sup>b</sup>Correlation significant at level 0.01 (two-sided).

Abbreviations: DD-C-GAS, developmental disorder-child-global assessment; ABCS, autism behavior coding system; coop., cooperative;  $n$ , sample size.



**Figure 2.** Scatterplots of correlations between changes in Autism Behavior Coding System eye contact versus Developmental Disorders-Child-Global Assessment Scale communication scores at day 18 ( $N = 40$ ). Abbreviation: Bsl, baseline.



**Figure 3.** Scatterplots of correlations between changes in Autism Behavior Coding System eye contact versus Developmental Disorders-Child-Global Assessment Scale social behavior scores at day 18 ( $N = 40$ ). Abbreviation: Bsl, baseline.

behaviors” reflect these primary goals of the FIAS intervention. As behaviors such as eye-contact, joint attention or imitation are not directly prompted by the therapists, but their occurrence is associated with the child’s motivation for spontaneous social communication, we interpret the less important changes in these domains also as being in line with the FIAS intervention focus.

The DD-C-GAS subdomains were highly correlated among themselves, both with regard to the changes from baseline to day 18 and, even higher, with regard to the assessments made on day 18. This suggests that the correlations calculated between the ABCS categories and the DD-C-GAS subdomain scores must be regarded with caution. Against our expectations, none of the correlations between the two instruments, with regard to changes from baseline, were statistically significant, although the DD-C-GAS as well as the ABCS indicated improvement in the majority of cases. This finding is most likely due to the fact that the two assessment methods focus on different aspects of behavior, and differ even more with regard to the source and the time window considered in the respective assessments: three times 1 hr for the ABCS, daily interaction of several hours duration over 18 days for the DD-C-GAS. It should also be noted that the baseline ratings for the DD-C-GAS were made at Day 1 of the intervention, whereas the baseline assessment of the ABCS corresponds to Day 4 of treatment, that is, a time when the child has adapted to the novel circumstances of the FIAS environment. As a consequence, the potential for improvement was probably higher for the DD-C-GAS than for the ABCS assessments.

Given the lack of significant correlations between the changes in the ABCS and the DD-C-GAS scores, it was of interest to further explore the relationship between the two assessment instruments. An additional analysis was performed between the DD-C-GAS scores provided by the therapists at day 18 and the changes observed in the ABCS category scores. In contrast to the analysis of after-minus-before differences, there were statistically significant associations in the expected direction between the changes in some of the ABCS categories and the DD-C-GAS assessments made on day 18. While these relationships should not be over-rated, they do suggest that the two levels of assessment, despite many differences, are not independent of one another: therapists’ final assessments were obviously not only determined by their own expectations but also by the changes in a child’s observable behavior. The ABCS variable “eye contact” may be particularly important in this respect.

Recent findings from a simulated 8-week clinical trial protocol for ASD [Jones, Carberry, Hamo, & Lord, 2017] showed placebo-like (positive) responses in caregiver ratings even in absence of treatment, while no significant changes in clinical ratings of ASD symptoms were noted.

This study indicates that a placebo-like effect is less likely to influence independent expert ratings, whereas therapists and caregivers will tend to overestimate intervention effects because they believe in the intervention they administer and in their therapeutic orientation. To answer the challenge of using sensitive tools alongside standard assessments to evaluate psychosocial interventions for ASD, observational coding (OC) by independent, blind assessors is considered to not only be an effective method [Hawkes, 2013], but also to be more sensitive to changes in outcome variables than other forms of assessment, especially with respect to short-term, post-treatment changes. The ABCS, as an observational video-based evaluation instrument, was designed to provide an objective, microanalytical assessment of children’s ASD-related behaviors, and to capture heterogeneity in expression of the disorder, as well as subtle changes, during a short-term intensive intervention. The results of the current study do not really support the idea of superior sensitivity of OC approaches as claimed by Hawkes, Dadds, and Pasalich [2013], but they do reflect significant improvement of important autism-related behavioral domains within a substantial patient sample.

In conclusion, the ABCS proved to be a sensitive assessment instrument to capture short-term changes in autism-related behaviors on a microanalytical level. Changes assessed with a more global clinical scale such as the DD-C-GAS, which is typically used by immediately involved therapists, may be numerically more impressive, but are more likely biased by personal expectations than a relatively objective instrument such as the video-based ABCS. Depending on the circumstances, the correspondence of the two levels of assessment may be low, but the specifics of this relationship deserve further study. The ABCS may prove useful in addition to standard assessment tools, especially in early intervention research settings, as it allows reliable analysis of core behavioral elements and their changes in young children with autism.

## Acknowledgments

We thank all parents and their children who participated in this evaluation. We also thank our former research colleague Diana Dima for video coding.

## Conflict of Interest

Olga Lazari was partially funded by the Universitäre Psychiatrische Kliniken (UPK) Basel Forschungsfonds and by the Botnar foundation. René Spiegel is the President of the FIAS Foundation, a charitable nonprofit organization. No other potential conflict of interest was reported.

## References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.
- Anagnostou, E., Jones, N., Huerta, M., Halladay, A. K., Wang, P., Scahill, L., ... Dawson, G. (2015). Measuring social communication behaviors as a treatment endpoint in individuals with autism spectrum disorder. *Autism*, 19(5), 622–636.
- Bölte, S. (2008). *Children's Global Assessment Scale für Entwicklungsstörungen (DD-C-GAS)*. J.W. Goethe-Universität Frankfurt/M.
- Bölte, S., Rühl, D., Schmötzer, G., & Poustka, F. (2006). *Diagnostisches Interview für Autismus-revidiert (ADI-R)*. Bern, Switzerland: Huber.
- Brian, J. A., Bryson, S. E., & Zwaigenbaum, L. (2015). Autism spectrum disorder in infancy: Developmental considerations in treatment targets. *Current Opinion in Neurology*, 28(2), 117–123.
- Dima, D., Lazari, O., Schmeck, K., Brönnimann, A., Goth, K., & Herbrecht, E. (2017). Assessing symptoms during comprehensive interventions for young children with autism spectrum disorders: Development and preliminary analysis of the Autism Behaviour Coding System (ABCS). *Psychiatry Research*, 247, 63–67.
- 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*, 46, 2464–2479.
- Hawkes, D. J., Dadds, M. R., & Pasalich, D. S. (2013). Observational coding strategies. In J. S. Colmer & P. C. Kendall (Eds.), *The Oxford handbook of research strategies for clinical psychology* (pp. 120–141). Oxford, UK: Oxford University Press.
- Herbrecht, E., Kievit, E., Spiegel, R., Dima, D., Goth, K., & Schmeck, K. (2015). Become related: FIAS, an intensive early intervention for young children with autism spectrum disorders. *Psychopathology*, 48, 162–172.
- Herbrecht, E., Lazari, O., Dima, D., Spiegel, R., & Schmeck, K. (2017). *The autism behavior coding system manual*. Basel, Switzerland: Department of Child and Adolescent Psychiatry, Psychiatric University Clinics (UPK) in preparation.
- Jones, R. M., Carberry, C., Hamo, A., & Lord, C. (2017). Placebo-like response in absence of treatment in children with autism. *Autism Research*, 10, 1567–1572.
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Luyster, R. J., Gotham, K., ... Guthrie, W. (2012). *ADOS-2 autism diagnostic observation schedule* (2nd ed.). Torrance, CA: Western Psychological Services.
- McConachie, H., Parr, J. R., Glod, M., Hanratty, J., Livingstone, N., Oono, I. P., ... Williams, K. (2015). Systematic review of tools to measure outcomes for young children with autism spectrum disorder. *Health Technology Assessment*, 19(41), 1–506.
- Ozonoff, S., Iosif, A. M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., ... Young, G. S. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(3), 256–266.
- Poustka, L., Rühl, D., Feineis-Matthews, S., Poustka, F., Hartung, M., & Bölte, S. (2015). *ADOS-2, Diagnostische Beobachtungsskala für Autistische Störungen-2. Deutschsprachige Fassung der Autism Diagnostic Observation Schedule-2 von C. Lord, M. Rutter, P.C. DiLavore, S. Risi, K. Gotham und S.L. Bishop (Module 1–4) und C. Lord, R.J. Luyster, K. Gotham und W. Guthrie (Kleinkind-Modul)*. Bern, Switzerland: Huber.
- Rutter, M., LeCouteur, A., & Lord, C. (2003). *Autism diagnostic interview, revised*. Torrance, CA: Western Psychological Services.
- Scahill, L., Aman, M. G., Lecavalier, L., Halladay, A. K., Bishop, S. L., Bodfish, J. W., ... Dawson, G. (2015). Measuring repetitive behaviors as a treatment endpoint in youth with autism spectrum disorder. *Autism*, 19(1), 38–52.
- Schaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry*, 40, 1228–1231.
- Vorgraft, Y., Farbstein, I., Spiegel, R., & Apter, A. (2007). Retrospective evaluation of an intensive method of treatment for children with pervasive developmental disorder. *Autism*, 11, 413–424.
- Wagner, A., Lecavalier, L., Arnold, L. E., Aman, M. G., Scahill, L., Stigler, K. A., ... Vitiello, B. (2007). Developmental disabilities modification of the Children's Global Assessment Scale. *Biological Psychiatry*, 61, 504–511.
- Williams White, S., Keonig, 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, 1858–1868. <https://doi.org/10.1007/s10803-006-0320-x>