



Quantifying Caregiver Change Across Early Autism Interventions Using the Measure of NDBI Strategy Implementation: Caregiver Change (MONSI-CC)

Bethany A. Vibert¹ · Sarah Dufek² · Claire B. Klein³ · Yeo Bi Choi⁴ · Jamie Winter³ · Catherine Lord⁵ · So Hyun Kim³ 

Published online: 10 January 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

This study aimed to provide initial validity and reliability of the *Measure of NDBI Strategy Implementation-Caregiver Change (MONSI-CC)*, a novel measure that captures changes in caregivers' implementation of NDBI strategies during early intervention. The MONSI-CC was applied to 119 observations of 43 caregiver-child dyads of preschoolers with autism spectrum disorders (ASD). The MONSI-CC showed high inter-rater and test-retest reliability and captured significant improvements in caregivers' implementation of NDBI strategies. Significant associations between improvements in caregiver NDBI implementation and improvements in the child's ASD symptoms also emerged. Our work shows promising evidence for the utility of the MONSI-CC to evaluate implementation of NDBI strategies by caregivers as a mediating and moderating factor for treatment effects on children with ASD.

Keywords NDBI · Autism spectrum disorder · Early intervention · Outcome measure · Parent-mediated intervention

Parent-mediated interventions (PMI) include treatments in which the clinician guides caregivers to use specific interactive skills to increase social communication, engagement, and adaptive functioning, and decrease disruptive behaviors in children and adolescents with developmental and/or behavioral challenges (Barkley 2013; Bearss et al. 2015; Eyberg and Funderburk 2011; Kasari et al. 2015; Kazdin

2005; Rogers et al. 2019b; Wetherby et al. 2014). Parent involvement in treatment for autism spectrum disorders (ASD) has long been recommended and supported by experts and federal policy because it significantly enhances generalization of skills attained during treatment for young children with ASD (IDEA 2004; Lovaas 1987; National Research Council 2001; Pickles et al. 2016; Schreibman et al. 2015; Wetherby et al. 2014). Along this line, teaching caregivers to apply interactive strategies to preexisting daily routines (e.g., meals or baths) has been shown to address the needs and goals of many families (Pickard et al. 2016). Consequently, most early intervention models for children with ASD, including Discrete Trial Teaching (Crockett et al. 2007), Developmental, Individual Differences, and Relationship-based (Solomon et al. 2007), as well as Naturalistic Developmental Behavioral Interventions (NDBIs; Gengoux et al. 2015; Kasari et al. 2015; Pickles et al. 2016; Rogers et al. 2019a; Wetherby et al. 2014), have involved PMI as a core component. Previous studies indicated improvements in receptive and expressive language, cognitive and adaptive functioning, social communication, restricted and repetitive behaviors (RRBs), and joint attention skills in children with ASD in response to interventions that include parent-mediated intervention sessions with or without clinician-mediated

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10803-019-04342-0>) contains supplementary material, which is available to authorized users.

✉ So Hyun Kim
sok2015@med.cornell.edu

¹ Autism Center, Child Mind Institute, New York, NY, USA

² University of California, Davis, MIND Institute, Sacramento, CA, USA

³ Center for Autism and the Developing Brain, Department of Psychiatry, Weill Cornell Medical College, New York-Presbyterian Hospital, 21 Bloomingdale Rd, White Plains, NY, USA

⁴ Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA

⁵ Semel Institute for Neuroscience and Behavior, University of California Los Angeles, Los Angeles, CA, USA

sessions (Dawson et al. 2010; Grahame et al. 2015; Kasari et al. 2014; Pickles et al. 2016; Wetherby et al. 2014).

Understanding Caregiver Behavior as a Mechanism of PMIs

Despite promising evidence and a strong rationale supporting PMIs for young children with ASD, the mechanisms of change, including the active ingredients, necessary dosage, duration, treatment intensity, and instructional methods that contribute to treatment outcomes for children with ASD are still unknown (Carter et al. 2011; Casagrande and Ingersoll 2017; Oono et al. 2013; Rogers et al. 2012; Rogers et al. 2019b). Better understanding of caregiver characteristics, specifically caregiver implementation of learned strategies, as mediating and moderating factors within PMIs can allow us to answer some of the remaining questions about the mechanisms of PMIs (Davlantist et al. 2019; Gulsrud et al. 2016; Pickles et al. 2015; Rogers et al. 2012, 2019b; Shire et al. 2015). For example, caregiver baseline interactive features, such as provision of learning opportunities (i.e., caregivers using behavioral learning principles to provide opportunities for children to learn social communication and play) have been found to predict improvements in child cognitive and language skills over the course of early interventions that incorporate PMIs (Davlantist et al. 2019). Caregiver behaviors and responsivity have also been observed to change significantly during early interventions involving (but not limited to) parent-mediated sessions within the first three to five months (Carter et al. 2011; Gulsrud et al. 2016; Rogers et al. 2012; Siller et al. 2013). Several recent studies examining changes in caregiver behaviors and their direct impact on specific child outcomes during PMIs found that improved parental synchrony and mirrored pacing mediate changes in the child's social initiations (Pickles et al. 2015) and joint attention (Gulsrud et al. 2016), respectively.

Most previous studies used study-specific fidelity of implementation (FOI) ratings (Carter et al. 2011; Green et al. 2010; Gulsrud et al. 2016; Pickles et al. 2016; Rogers et al. 2012, 2019b), or a focused measure targeting a single strategy (Davlantist et al. 2019) to measure caregiver behaviors. Despite the fact that these interventions teach caregivers similar NDBI techniques (Schreibman et al. 2015; see below for more details), various FOI scales use different terms which may share some similar concepts, and use different levels of measurement (e.g., coding if a specific behavior occurred within a given interval [e.g., a minute] and converting that to a percentage vs. giving a 1–5 quality rating of a general behavior). The psychometric properties (e.g., inter-rater reliability, test re-test reliability, validity) of these FOI scales have been also limited, making it challenging to synthesize the results. Additionally, FOI ratings

were typically created to measure if a specific intervention was being implemented to a predetermined standard (usually 80%) and have not been applied to other interventions. Single-strategy ratings are also limited because they cannot capture the variety of strategies that constitute at least NDBI-based PMIs and would fail to capture if an intervention aimed at one strategy (e.g., overall synchrony) actually improved another (e.g., manipulating the environment). The psychometric, conceptual, methodological, and terminology differences in existing caregiver behavior ratings and the importance of better understanding mechanisms of change in PMIs underscore the need for a valid, reliable outcome measure of caregiver change in NDBI strategy implementation during PMIs (Oono et al. 2013; Proctor et al. 2011). The development of such a measure applicable to many interventions would aid in evaluating how caregiver use of NDBI strategies moderate and/or mediate child outcomes across studies and which sets of NDBI strategies have the most significant impacts on child outcomes (Casagrande and Ingersoll 2017; Frost et al. 2019; Oono et al. 2013; Pickles et al. 2015).

The Development of a Common Caregiver Outcome Measure Based on Core Features of Naturalistic Developmental Behavioral Interventions (NDBIs; Schreibman et al. 2015)

In 2015, a seminal report was generated by a group of researchers “to bring parsimony to a field” by describing self-identified commonalities in theory and practice among different NDBIs (Schreibman et al. 2015, p. 2411). Beyond describing a shared, integrated theoretical foundation in both developmental science and behavioral learning theories, Schreibman et al. (2015) reported several common strategies used by NDBIs, including Three Part Contingency (antecedent-behavior-consequence), Child-Initiated Teaching Episodes (e.g., following the child's lead or interest and presenting teaching opportunities within those activities), Environmental Arrangement (e.g., structuring the environment to increase child attention to the social partner and desired materials), Natural Reinforcement and Related Methods for Enhancing Motivation of the Child (e.g., using reinforcement that is related to the child's goal rather than providing an unrelated reward), Use of Prompting and Prompt Fading (e.g., scaffolding or cuing to support new learning systematically), balanced turns within object or social play routines (e.g., increasing social reciprocity with social partner), Modeling, Adult Imitation of the Child's Language, Play, or Body Movements, and Broadening the Attentional Focus of the Child. There are differences in the focus on certain strategies and the way they are included within

specific interventions. For example, some interventions have explicit, clear guidelines within their fidelity of implementation ratings pertaining to how, when, and what to imitate (e.g., Dawson et al. 2010; Kasari et al. 2014) and others are less specific. However, the common elements described by Schreibman et al. (2015) make it possible to develop a comprehensive caregiver outcome measure that targets changes in these common, core ingredients of NDBIs.

The Aims of the Current Study

The main objective of the current study is to describe the development of the Measure of NDBI Strategy Implementation—Caregiver Change ([MONSI-CC]; Kim et al. 2019) and to examine its initial validity and reliability. This measure is designed to capture changes in caregiver implementation of core NDBI strategies used in early interventions that implement parent mediated components (e.g., parent coaching sessions) over the course of treatment and to explore the relationship between changes in caregiver behavior and changes in child behaviors. The MONSI-CC is intended for caregivers of infants, toddlers, and preschoolers (with or without autism and with varying cognitive and developmental abilities) who are learning NDBI strategies. Specifically, we aimed (1) to identify items to be included in the final MONSI-CC coding scheme through analysis of item correlations, (2) explore the factor structure of the MONSI-CC items, (3) examine psychometric properties (e.g., inter-rater, test–retest, and split-half reliability, as well as validity) of the MONSI-CC, (4) provide initial evidence for the validity of the MONSI-CC in measuring changes in caregiver implementation of NDBI strategies over the course of several treatments, and (5) describe the relationships between changes in caregiver strategy implementation and child outcomes.

Method

Participants

Participants are detailed in Table 2 and consisted of subsets of children with ASD who were enrolled in RCTs of various NDBIs that included a component of PMI (Early Start Denver Model [ESDM] Rogers et al. 2019a; Joint Attention Symbolic Play, Engagement, and Regulation [JASPER], Kasari et al. 2014; Early Social Interaction [ESI], Wetherby et al. 2014) at the University of Michigan Autism and Communication Disorders Center (UMACC). Informed consent was obtained from all individual participants included in this study. The duration of our participants' treatment in the various RCTs ranged from 3 months to 2 years (Kasari et al. 2014; Rogers et al. 2019a; Wetherby et al. 2014). Because

previous research observed significant changes in caregiver behavior in 12 weeks while enrolled in the same (Rogers et al. 2012) or similar interventions (Gulsrud et al. 2016; Shire et al. 2015), we included children with a minimum of two recorded play samples from 43 dyads with the same caregiver that were at least 3–8 months apart with an average of 6 months apart ($M=6.35$ months, range = 3–8 months, $SD=0.88$) to ensure that the dyads had received sufficient PMI sessions to show changes. Twenty-six (10 control, 16 treatment) dyads were included from the Early Social Interaction (ESI) trial (Wetherby et al. 2014), 11 (3 control, 8 treatment) from the Joint Attention, Symbolic Play, Engagement, and Regulation (JASPER; Kasari et al., 2014), and 6 (2 control, 4 treatment) from the Early Start Denver Model trial (Rogers et al. 2019b). The control groups in each study were receiving group-based (compared to individual) playgroups and education sessions (Kasari et al. 2014; Wetherby et al. 2014), or community based, “as usual” interventions (Rogers et al. 2019a). Specific intervention effects are not evaluated in this study, as the primary focus is to examine the reliability and validity of the MONSI-CC in measuring changes in caregiver behaviors in various PMIs. All children (range = 12.5–55.1 months at baseline) had a best estimate clinical diagnosis of ASD based on diagnostic evaluations using the Autism Diagnostic Interview-Revised (ADI-R; Le Couteur et al. 2003), the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2000; Lord et al. 2012), developmental testing (Mullen Scales of Early Learning [MSEL]; Mullen 1995), as well as developmental and medical histories. Many of the children in the dyads were minimally verbal (simple phrase speech or less; ADOS Toddler Module, $n=31$; ADOS Module 1, $n=6$) with a small number of children with emerging two- to three-word phrases (ADOS Module 2, $n=4$).

Primary Measure (MONSI-CC)

The MONSI-CC coding scheme was applied to 10-min videos of free-play interactions between a caregiver and child using similar sets of toys across dyads. These caregiver-child interactions were gathered during participation in available intervention trials, and most interactions were collected in a clinic setting ($n=83$, 96.3%) while the remaining observations were collected in participants' homes. Scores did not vary significantly by context between lab or home settings in MONSI-CC Total Score ($t(84)=1.463$, $p=.15$). Caregivers were instructed to play “how you typically would” with the child.

The MONSI-CC item development was carried out by expert clinicians who had achieved fidelity of implementation in a variety of NDBIs (authors: BLINDED). They first reviewed and incorporated the common strategies defined in the original NDBI article (Schreibman et al.

2015). Additionally, the PIs of each treatment provided intervention specific fidelity measures to help identify primary mechanisms specific to their intervention which were considered in the development of the MONSI-CC. Many strategies overlapped across the three models (Dawson et al. 2010; Kasari et al. 2014; Wetherby et al. 2014). The fidelity ratings of other related intervention models were also considered when publicly available (Ingersoll and Dvortcsak 2010; Pickles et al. 2016). We took a comprehensive approach that included not only overlapping strategies across various NDBIs, but also some that were unique to a subset of interventions (e.g., specific imitation of play and social communication, management and modulation of child affect; Dawson et al. 2010; Kasari et al. 2015). We further characterized strategies as separate items when those strategies were identified as distinct strategies for some interventions (e.g., imitation of play versus social communication as two separate items) even though they were grouped together in other interventions. The intention was to avoid prioritizing different strategies a priori, but to develop a comprehensive list which could capture common strategies that cut across different NDBIs with parent-mediated components. Thus, the MONSI-CC is not designed as a parent fidelity rating of a particular intervention model, but as a potential monitoring and outcome measure to capture changes in a range of NDBI strategies taught to caregivers over the course of PMIs. The final MONSI-CC items are described in more detail in Table 1.

The original MONSI-CC coding scheme consisted of strategies described by 21 items coded on a 5-point scale based on the intersection of three aspects of the behavior: frequency/consistency, effectiveness, and the proportion of missed opportunities to implement each strategy. Frequency/consistency was defined by occurrences of a behavior. For ratings of effectiveness, coders observe if the caregiver is using the strategy as it is intended to be implemented (e.g., with developmentally appropriate expectations for the child's behavior), with variety (e.g., using all of the expected components of the strategy; using multiple modalities such as use of vocalizations and gestures to expand play routines), and with flexibility (e.g., not overusing a strategy repetitively). Missed opportunities were coded when the context allowed for the implementation of a strategy (e.g., a child demonstrates a desired behavior that the caregiver should reinforce) but the strategy was not implemented. Based on these definitions, each of the 21 items was assigned a code between 1 (strategy was rarely implemented; almost all opportunities were missed; lacking developmental appropriateness) and 5 (strategies were well-timed such that they were implemented frequently and consistently across activities/routines with variety and flexibility; opportunities were rarely missed). The item selection procedure is described in

Preliminary Analyses and the detailed item descriptions are provided in Table 1.

The MONSI-CC was coded by assigning codes to each of two 5-min segments (Segment A and B) of a 10-min video, watched twice, and averaging the codes from the two segments, following the model of the Brief Observation of Social Communication (BOSCC; Grzadzinski et al. 2016). Three codes have been created to describe specific aspects of the child's language and play level (Language Level, Play Level) and to rate the overall quality of the caregiver/child interaction (Dyadic Engagement), which are not included in the MONSI-CC totals because they are not intended to capture specific NDBI strategies. During the first viewing, coders take notes of these behaviors and assign the ratings on the three items, which are intended to facilitate scoring of strategies by identifying the abilities of the child and the context in which the strategies are being implemented and to evaluate developmental appropriateness (e.g., if the language being modeled by the parent is within the child's range of development). Language Level (rated from no words or directed communication [code of 1] to complex language [code of 8]) and Play Level (rated from sensory motor play [1] to imaginative play [6]) are assigned if the skill was consistently, spontaneously demonstrated ("mastered") rather than developing ("emerging"). Dyadic Engagement (rated from no example of dyadic engagement [1] to sustained dyadic engagement for more than 45 s across more than two activities [5]) is intended to be used as an additional outcome for researchers and clinicians particularly interested in measuring the dyad's time spent reciprocally engaged in a shared activity (e.g., the dyad sharing attention, engaging in a back and forth play, imitating each other's actions; both the child and caregiver initiating and maintaining the interaction during the activity).

On the second viewing of each segment, caregiver strategy implementation ratings are assigned. Item scores on each of the strategies are then averaged across the two segments and summed within five Domain Scores determined by the factor analyses described below (See *Preliminary Analyses*). The Domain Scores are then summed to produce a single MONSI-CC Total Score. Coding each 10-min interaction takes less than 30 min. A copy of the measure will be available upon request.

Some items require a caregiver response to a specific child behavior. However, not all children may exhibit the behavior (e.g., vocalization to request or undesired behavior). To account for caregivers not having opportunities to use these strategies within the 10-min observations, some items have an option to assign a "9" for "No opportunity." Two methods of conversion were used when calculating the Domain Score items assigned "9." First, "9"s on *Imitates Social Communication*, *Imitates Play Behavior*, *Reinforces Desired Behavior*, and *Reinforces Goal-directed Behavior*

Table 1 Measure of NDBI strategy implementation—caregiver change (MONSI-CC): domain and item description

Domain name	Items under each domain
<p>Environmental Set-up To optimize the child’s focus, these strategies involve adjusting the physical space and the environment. This includes removing distractions in the play area, the interactant physically positioning themselves directly across from the child with the play area directly in-between the interactant and child, and using a variety of strategies to get or maintain the child’s attention before or during teaching opportunities</p>	<p><i>Environmental Arrangement:</i> The environment is clear from overt distractions. If items are drawing the child’s attention away from the desired toy/play scheme, the caregiver identifies this and adjusts/removes those items</p> <p><i>Face to Face Seating:</i> The caregiver and child are seated directly across from each other, such that the play area, or items/objects of focus are directly in-between the child and the caregiver. The caregiver adjusts and moves to maintain this seating arrangement</p> <p><i>Gains Attention Before Learning Opportunities:</i> The caregiver attempts to get the child’s attention prior to teaching and when attempting to embed a learning opportunity in the play. The caregiver adjusts their behavior and varies strategies if an initial attempt is not successful</p> <p><i>Modulates Child’s Affect:</i> Caregiver selects and adjusts activities, as needed, to manage child’s affect (e.g., engaging in high-energy tasks to help increase energy/attention of the child or engaging the child in slow-tempo/soothing tasks if her/his energy is too high or not well-suited for optimal learning)</p> <p><i>Developmentally Appropriate Cues:</i> Caregiver adjusts language and play to be within an appropriate range for the child to be able to learn from and engage with. Language Level and Play Level assigned during the first viewing are considered in this item</p> <p><i>Follows Child’s Selection:</i> The caregiver uses what the child is already interested into incorporate teaching and learning opportunities. If a child shows interest in an object or toy, a caregiver will adjust their behavior to maximize the use of that object or toy</p> <p><i>Waits for Child to Lead:</i> The caregiver waits for the child to take actions within the play (either within a play scheme or in overall selection). The caregiver is not overly suggestive or directive during play, unless structure is required. The caregiver may take the lead, model, or expand play, but readily returns the floor to the child</p> <p><i>Balances Mastered/Acquisition Tasks:</i> Based on the child’s affect, motivation, and behaviors, the caregiver both encourages the child to try and practice new skills while also reinforcing and maintaining previously mastered skills. The caregiver balances these tasks appropriately such that the child remains interested and learning opportunities are maximized</p>
<p>Child-Guided Interactions Based on the child’s motivation, affect, and developmental abilities, caregivers join in the child’s interests and adjust their behaviors in response to the child. Caregivers introduce activities and utilize task demands to manage the child’s motivation and optimize arousal for learning. These strategies involve the caregiver recognizing the child’s needs or interests and adjusting their own behaviors to meet those needs</p>	

Table 1 (continued)

Domain name	Items under each domain
<p>Active Teaching and Learning These strategies involve being an active partner in the child’s learning through imitating, modeling, and expanding on social communication and play, and responding positively to the child. The caregiver shows the child an appropriate play or communicative behavior, moves the play or activity in a new direction, appropriately prompts the child to demonstrate that same skill, and guides the child through elaborating on that play or activity. If the child is unable to play or behave appropriately, these strategies are also used to intervene and demonstrate more appropriate play</p>	<p><i>Positive/Responsive to Social Communication/Behaviors:</i> Caregiver maintains positive regard toward and acknowledges the child’s social communication behaviors throughout the interaction with enthusiastic, animated, and positive responses</p> <p><i>Imitates Play Behavior:</i> Caregiver imitates the child’s appropriate play behavior within a timely fashion and while the child is attending</p> <p><i>Models and Expands Play:</i> Caregiver demonstrates new actions on objects that are developmentally appropriate, well-timed, and in the view of the child. The Play Level assigned during the first viewing is considered in this item</p> <p><i>Models and Expands Social Communication:</i> Caregiver uses language, gestures, and facial expressions appropriate to the context, in the child’s developmental level while the child’s attending. The Language Level assigned during the first viewing is considered in this item</p> <p><i>Manages Routines and Transitions:</i> Caregiver facilitates and minimizes adverse effects of moving between play items or activities. Caregiver provides transition supports and expectations (such as warnings or clean-up routines) to facilitate transition from one activity to another</p> <p><i>Prompts Social Communication and Play:</i> Caregiver uses a hierarchy of prompts appropriately (adjusting to the needs of the child) to promote social communication and play development in the child</p> <p><i>Manages Unwanted Behaviors:</i> When unwanted behaviors occur, parent manages them using appropriate and effective responses (e.g., ignoring, environmental control, etc.) to decrease the frequency of the behavior. Caregivers do not reinforce unwanted behaviors</p>
<p>Opportunities for Engagement Caregivers use specific developmentally appropriate strategies to promote interactions with the child. Choices are offered within and between play activities and routines. During interactions and routines, caregivers are using turn taking, blocking, and expectant waiting to promote communication and interaction with the child</p>	<p><i>Provides Choices:</i> Caregiver gives the child options of play materials and activities within and between activities. Child’s previously assigned Language Level from the first viewing is considered in determining the appropriateness and variety of choices provided by the caregiver (e.g., “yes/no” choices, holding up two objects for the child to indicate preference, asking a question with a choice but no visual supports)</p> <p><i>Uses Turn-Taking/Blocking/Expectant Waiting:</i> Caregiver uses a variety of strategies such as turn taking, blocking, and expectant waiting to promote the child’s use of communication within an interaction</p>
<p>Natural Reinforcement and Scaffolding Caregivers provide natural reinforcement for desired behavior (e.g., a child receives what s/he wants after an appropriate request) and goal-directed behavior (e.g., if a child is learning new words and s/he approximates a word rather than articulate it perfectly s/he is still naturally reinforced for the behavior). As an additional modality of reinforcing behavior, imitation of social communication is done immediately and promptly after the child uses appropriate social communication (e.g., vocalizations, gestures, and facial expressions)</p>	<p><i>Reinforces Desired Behaviors:</i> Caregiver provides natural reinforcement for desired behaviors in a timely and appropriate manner. This includes reinforcement for previously mastered skills and those the child is still acquiring</p> <p><i>Reinforces Goal-Directed Behaviors:</i> Caregiver recognizes the child’s attempts of social communication or play actions and provides natural reinforcement.</p> <p><i>Imitates Social Communication:</i> Caregiver immediately imitates and repeats the child’s social communication behaviors</p>

were converted to a “3” (representing the median score on the Likert scale), so as not to inflate or reduce caregiver scores. A “9” on *Manages Unwanted Behavior*, was converted to a “5” with the assumption that such a code implied that antecedents for unwanted behaviors were well-managed, providing no opportunities for undesired behavior.

Coders included a post-doctoral fellow and four research assistants who were able to achieve reliability under the supervision of licensed clinical psychologists who had achieved fidelity in parent-mediated NDBIs. Initial

reliability criteria required a coder to achieve (1) at least 80% reliability on all items within each segment (no more than four items could be different by more than one point); (2) 80% accuracy on each domain score within each segment; and (3) 90% accuracy for total scores. Domain scores were considered within 80% of the consensus codes based on the total number of points in that domain (e.g., for domains with only two items, the domain score had to be within two points of the agreed upon consensus code, but for domains with three items, the domain score had to be within three

points). Exact agreement was required for Language, Play, and Dyadic Engagement items. This reliability criteria had to be met for three consecutive observations. When there were multiple coders for each video, consensus codes were used for analysis; but consensus codes were not used for inter-rater reliability. All five coders became reliable on the measure with 10–20 practice videos.

Child Measures

Baseline Cognitive Functioning

All children were administered the MSEL (Mullen 1995) at baseline as a part of the original studies (Table 2). The MSEL produces age equivalents for expressive language, receptive language, fine motor, and visual reception, from which Full-Scale (FSIQ) and Ratio IQs (FRIQ) can be calculated (Bishop et al. 2011). All children were administered the MSEL on average 10 days before the first dyadic interaction occurred ($M = 10.7$ days, $SD = 20.2$).

Baseline Autism Symptoms

All children were administered the ADOS-2 at baseline by research reliable clinicians during the original studies (Table 2). We report ADOS-2 Comparison Scores for Social Affect (CS-SA) and Restricted and Repetitive Behavior (CS-RRB; Gotham et al. 2009; Hus et al. 2014). All children were

administered the ADOS-2 on average 12 days before the first dyadic interaction occurred ($M = 11.53$ days, $SD = 19.07$).

Changes in Autism Symptoms

Changes in autism symptoms in children were coded using the Brief Observation of Social Communication (BOSCC; Grzadzinski et al. 2016). The BOSCC has been found to be more sensitive in detecting changes in social communication over short periods of time than the CS or raw algorithm totals of the ADOS (Grzadzinski et al. 2016; Kim et al. 2018). The BOSCC has two ASD symptom-related totals (Social Communication [SC] domain and Restricted and Repetitive Behaviors [RRB] domain) as well as an overall ASD Core domain, which sums the two symptom domains. Lower scores on these domains indicate better social communication, fewer RRBs, or lower overall ASD symptoms. BOSCC scores were rated by coders blind to time points of the videos, intervention model, and treatment condition; scores were previously reported for some of this sample in a previous study (Grzadzinski et al. 2016).

Data Analyses

Preliminary Analyses

An inter-item correlation analysis of the MONSI-CC items was conducted. Inter-item correlations ranged from minimal ($r = .088$, $p > .05$) to significant ($r = .752$, $p < .01$). One

Table 2 Demographics/baseline characteristics ($N = 43$)

Child	Mean (<i>SD</i>) or <i>n</i> (%)	Range	Interactant	Mean (<i>SD</i>) or <i>n</i> (%)	Range
Age (months)	25.4 (11.3)	12.5–55.1	Age (years) ^b	35.0 (5.8)	23.8–49.3
Sex (males)	30 (69.8)		Sex (females)	41 (95.3)	
Race			Race ^b		
Caucasian	33 (76.7)		Caucasian	29 (80.6)	
African American	3 (7.1)		African American	3 (8.3)	
Other	7 (16.2)		Other	4 (11.1)	
MSEL			Relationship to Child		
Verbal IQ	62.1 (24.6)	26.3–139.0	Mother	39 (90.6)	
Nonverbal IQ	84.9 (19.1)	36.5–130.9	Father	2 (4.7)	
FSIQ ^a	72.3 (20.4)	33.2–128.0	Grandmother	2 (4.7)	
ADOS-2			Education ^b		
CS-SA	7.6 (1.9)	4.0–10.0	BA/BS or above	23 (60.5)	
CS-RRB	7.2 (1.7)	1.0–10.0	Below BA/BS	15 (39.5)	

MSEL mullen scales of early learning; *FSIQ* full-scale IQ; *CS-SA* comparison score social affect; *CS-RRB* comparison score restricted and repetitive behaviors

^aFSIQ, $n = 30$; FRIQ full-scale ratio IQ, $n = 13$

^bAge, education level, and race only calculated for parents who reported the information (age, $n = 35$; education, $n = 38$; race, $n = 36$), grand-mother ages were unavailable ($n = 2$)

item (*Appropriately Animated*) showed the highest correlations with two other items (with *Positive/Responsive to Social Communication/Behaviors*, $r = .727$; with *Modulates Child's Affect*, $r = .752$). This item, *Appropriately Animated* was, therefore, subsequently omitted from further analyses to ensure that these items are specific, targeting non-overlapping behavioral constructs. See Online Resource Fig. 1a–e for item distributions.

To examine the factor structure of items and derive Domain Scores, an exploratory factor analysis (EFA) was conducted for the remaining 20 MONSI-CC items (*Appropriately Animated* removed, see above). EFA was conducted on the full sample ($N = 43$, 86 observations). Analysis was conducted in Mplus (Muthén and Muthén 2012) using a geomin oblique rotation because of multiple observations per child using a complex survey adjustment with the child as the cluster-level unit. As the number of factor solutions increased from 1 to 5, the fit changed from a comparative fit index (CFI) of 0.814, 0.885, 0.887, 0.993, to 0.967 with root mean square errors of approximation (RMSEA) values of 0.110, 0.092, 0.097, 0.080 to 0.059.

The 5-factor solution was chosen based on a goodness-of-fit rating of a CFI of 0.967 (CFI between 0.9 and 1 indicating good fit; Skrondal and Rabe-Hesketh 2004) and an RMSEA value of 0.059 (RMSEA of 0.07 or less is considered a satisfactory fit; Browne and Cudeck 1993). Also, a χ^2 test of model fit was significant for the 5-factor model ($\chi^2 = 129.57$, $df = 100$, $p < 0.05$); eigen values decreased down to 0.9 for the 5-factor model, and with the 6-factor model, it was even lower at 0.7. One item, *Waits for Child to Lead* loaded similarly on two factors (0.284 in factor 5, and 0.279 in factor 2) and was placed in factor 2 (Child-Guided Interactions) because it was more theoretically related to other items in factor 2 (focused on the caregivers' ability to allow the children to lead the play when appropriate opportunities allow). See Online Resource Table 1 for factors and loadings.

Inter-Rater, Test–Retest, and Split-Half Reliability

Fifteen MONSI-CC videos were coded by two different coders for inter-rater reliability. Thirty videos from 15 dyads gathered on two occasions and recorded less than 1.5 months apart ($M = 1.02$, $SD = 0.31$) were coded for test–retest reliability. All 119 observations were used for split-half reliability comparing codes from segment A and B. For all these analyses, two-way Random Absolute Intraclass Correlation Coefficients (ICCs) were computed for the MONSI-CC Total Score and the five Domain Scores.

Validity

To assess validity of the MONSI-CC as a measure of change over time, 43 dyads' MONSI-CC scores from

the first and last observation (86 videos) were compared using paired t tests. Effect sizes (Cohen's d) were used to examine the magnitude of changes over time. Based on the changes we observed across all caregivers, post hoc linear regressions were conducted to examine how treatment condition (treatment vs. control) predicted changes in MONSI-CC Total and Dyadic Engagement scores while controlling for baseline child features (symptom severity using the ADOS CS and FSIQ), baseline caregiver features (education and age), and baseline MONSI-CC and Dyadic Engagement scores respectively. In addition, using paired t -tests and effect sizes (Cohen's d), we also examined the magnitudes of change over time in the MONSI-CC Total Score, Domain Scores, and Dyadic Engagement. Independent samples t tests were also used to examine if the changes made by the treatment vs control groups over time differed significantly.

Preliminary Analyses: Evaluating the Relationship Between Change in Caregiver Strategy Implementation and Child Outcome

Pearson's r correlational analyses were used to examine associations between changes in caregiver behavior in the MONSI-CC Total Score/Domain Scores and changes in child symptoms using the BOSCC Domains. Linear regressions were conducted to examine how changes in caregiver strategy implementation (MONSI-CC scores) predict changes in child outcome (BOSCC scores) and Dyadic Engagement while controlling for baseline features (baseline outcome variable score, FSIQ, CS-SA, and CS-RRB).

Results

Inter-Rater, Test–Retest, and Split-Half Reliability

The inter-rater ICC for the MONSI-CC Total Score was high at $r = .80$ [95% Confidence Interval (CI) .39–.93]. Inter-rater ICCs were also strong for all MONSI-CC Domains ($r = .69$ –.81; Online Resource Table 2). The test–retest ICC for the MONSI-CC Total Score was high at $r = .92$ [95% CI (.76–.97)]. Test–retest ICCs for Domain Scores were high, ranging from $r = .80$ to .93 (Online Resource Table 3) except for the Opportunities for Engagement domain ($r = 0.34$), potentially due to limited variability (more than 50% scoring at 1). The split-half ICC for MONSI-CC Total Score was high at $r = .85$ [95% CI (.78–.89)]. Domain Scores also showed strong split-half reliability ($r = .83$ –.90; Online Resource Table 4).

Validity

Results of paired *t* tests demonstrated that MONSI-CC Total Scores increased significantly from the first to last observation, indicating improvements in caregiver implementation of strategies with a large effect size ($M = 11.85$, $SD = 14.22$, $t(42) = 5.47$, $p < .01$; Cohen's $d = .86$). As seen in Fig. 1, improvements on all Domain Scores were statistically significant with effect sizes ranging from .45 to .93: Environmental Set-up ($t(42) = 3.83$, $p < .01$; Cohen's $d = .69$), Child-Guided Interactions ($t(42) = 5.99$, $p < .01$; Cohen's $d = .93$), Active Teaching and Learning ($t(42) = 3.43$, $p < .01$; Cohen's $d = .58$), Opportunities for Engagement ($t(42) = 2.37$, $p < .01$; Cohen's $d = .45$), and Natural Reinforcement and Scaffolding ($t(42) = 4.50$, $p < .01$; Cohen's $d = .82$). The BOSCC SC and ASD Core domains for children also decreased significantly [Fig. 1; SC [$t(42) = -2.24$, $p < .03$; Cohen's $d = .30$], ASD Core [$t(42) = 2.18$, $p < .03$; Cohen's $d = .31$]], indicating reduction in child symptom levels with effect sizes ranging from .30 to .31. This was also consistent with the previous study that reported significant changes in social communication symptoms in a subset of the children included in our sample (Grzadzinski et al. 2016).

Post hoc analyses indicated that changes in MONSI-CC Total Score were significantly predicted by treatment condition ($\beta = .44$, $p = .02$) while controlling for the baseline child symptom severity and cognitive levels, and caregiver age

and education, as well as the baseline MONSI-CC Total Score (Table 3; placement below this paragraph). Additionally, changes in Dyadic Engagement were significantly predicted by treatment condition ($\beta = .37$, $p = .05$) while controlling for baseline child and caregiver characteristics and baseline Dyadic Engagement score (Table 3). *T* tests and effect sizes indicated that participants in the treatment condition increased their MONSI-CC Total Scores significantly from first to last observation (Fig. 2; $M = 16.32$, $SD = 12.89$, $t(27) = 6.70$, $p < .01$; Cohen's $d = 1.5$); those in the control condition did not improve their MONSI-CC Total Score significantly (Fig. 2; $M = 3.50$, $SD = 13.10$, $t(14) = 1.02$, $p = .32$; Cohen's $d = .20$). Additionally, the treatment group's Dyadic Engagement increased significantly from the first to last observation (Fig. 2; $M = .89$, $SD = 1.23$, $t(27) = 3.85$, $p < .01$; Cohen's $d = .77$) whereas the control condition did not significantly improve their Dyadic Engagement (Fig. 2; $M = .07$, $SD = .96$, $t(14) = .27$, $p = .80$, Cohen's $d = .06$). Changes in Domain Scores also showed similar patterns; the treatment group showed significant changes in all Domain Scores whereas the control group did not, except for one domain, Opportunities for Engagement, for which neither groups showed significant changes (See Online Resource Table 5). *T*-tests and also indicated that the treatment condition made significantly greater changes than the control group in MONSI-CC Total Score ($t(41) = -3.10$, $p < .01$) and Dyadic Engagement ($t(41) = -2.43$, $p = .03$).

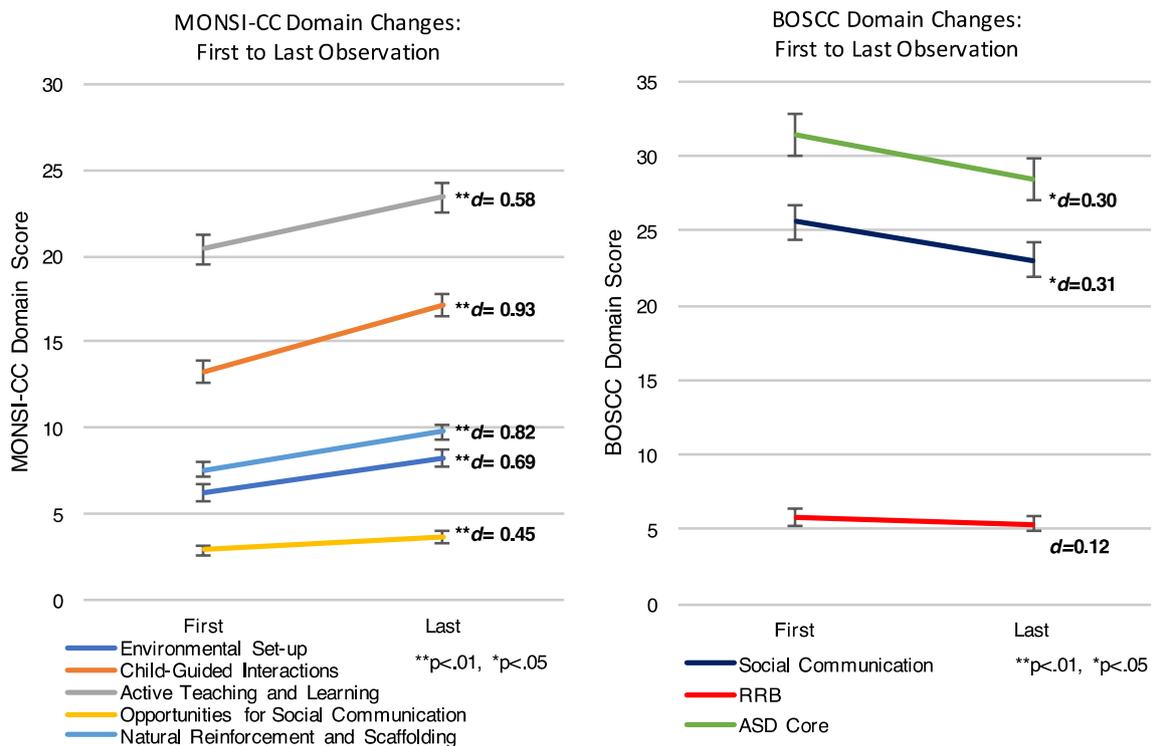


Fig. 1 MONSI-CC and BOSCC domain changes: first to last observation

Table 3 Regression analyses examining group assignment (treatment vs control) predicting changes in MONSI-CC total score and dyadic engagement

Outcome variable	Predictors in model	<i>B</i>	SE	<i>t</i>	<i>Sig</i>	<i>R</i> ²
MONSI-CC total score (changes)						0.42
	CS- total (baseline)	− 0.17	1.21	− 1.08	0.29	
	FSIQ (baseline)	0.25	0.11	1.50	0.15	
	MONSI-CC total score (baseline)	− 0.29	0.18	− 1.79	0.09	
	Caregiver education	0.02	4.26	0.11	0.91	
	Caregiver age (baseline)	− 0.07	0.43	− 0.39	0.70	
	Treatment condition	0.44	5.22	2.42	0.02	
MONSI-CC dyadic engagement (changes)						0.51
	CS- total (baseline)	− 0.19	0.10	− 1.31	0.20	
	FSIQ (baseline)	0.40	0.01	2.44	0.02	
	MONSI-CC DE (baseline)	− 0.57	0.18	− 3.34	<0.001	
	Caregiver education	0.02	0.37	0.15	0.88	
	Caregiver age (baseline)	− 0.09	0.04	− 0.53	0.60	
	Treatment condition	0.37	0.44	2.09	0.05	

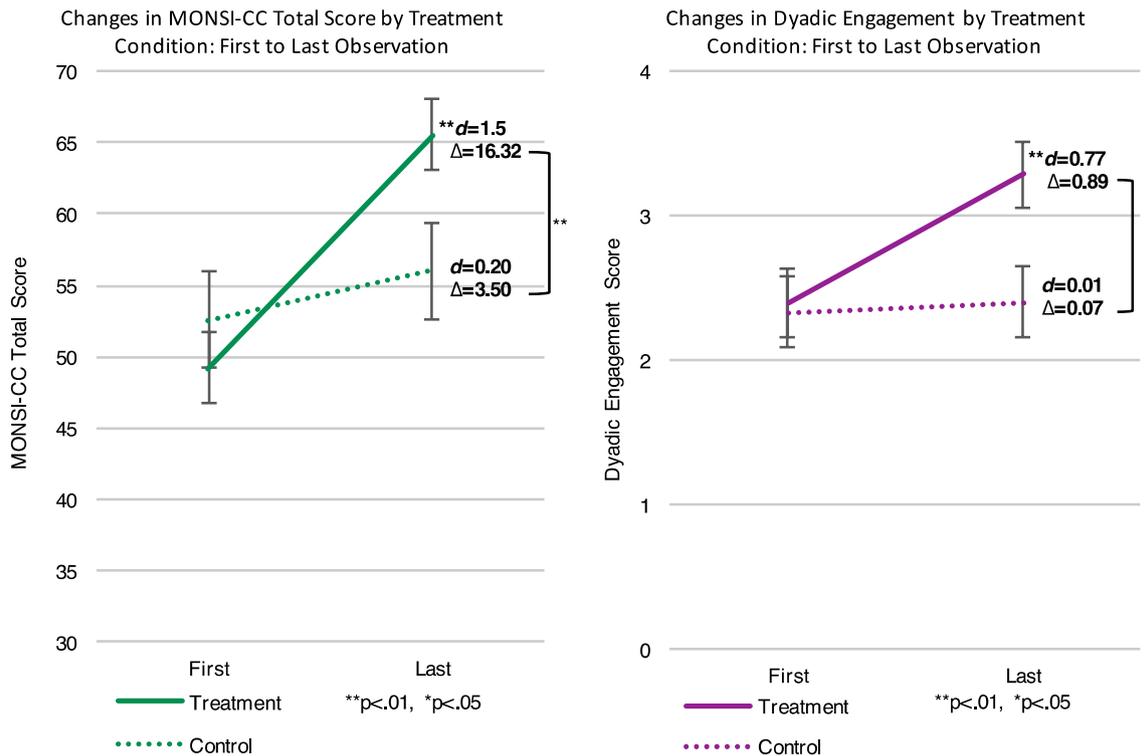


Fig. 2 Changes in MONSI-CC total score and dyadic engagement by treatment condition: first to last observation

Preliminary Analyses: Evaluating the Relationship Between Change in Caregiver Strategy Implementation and Child Outcome

Significant associations emerged between increases in MONSI-CC Total Score and decreases in children’s ASD symptoms as measured by the BOSCC SC, RRB, and ASD

Core domains ($r = - .383$ to $- .485, p < .05$), and improvements in Dyadic Engagement as measured by the MONSI-CC ($r = .657, p < .01$). Improvements in MONSI-CC domains were also significantly correlated with reductions in ASD symptoms ($r = - .313$ to $- .446, p < .05$) and improvements in Dyadic Engagement ($r = .361$ to $.693, p < .05$). See Table 4.

Table 4 Correlations of changes in MONSI-CC and child outcomes/dyadic engagement

Domain (changes)	RRB	SC	ASD core symptoms	Dyadic engagement
MONSI-CC total score	– .394**	– .383*	– .485**	.657**
Environmental set-up	– .112	– .351*	– .341*	.502**
Child-guided interactions	– .290	– .301	– .373*	.452**
Active teaching and learning	– .418**	– .313*	– .436**	.693**
Opportunities for engagement	– .353*	– .184	– .301	.371
Natural reinforcement and scaffolding	– .260	– .241	– .309*	.282

** $p < .01$ (2-tailed)* $p < .05$ (2-tailed)

As shown in Table 5, results from linear regression analyses indicated that improvements in specific strategies (Environmental Set-up, Child-Guided Interactions) and overall improvements in MONSI-CC Total Score significantly predicted BOSCC SC improvement while controlling for baseline features. Improvement in specific strategies (Environmental Set-up, Child-Guided Interactions, Active Teaching and Learning) and overall improvement in the MONSI-CC Total Scores also significantly predicted improvement in MONSI-CC Dyadic Engagement while controlling for baseline features. All p values were < 0.05 .

Discussion

The results of the current study show promising initial evidence for high test–retest, inter-rater, and split-half reliability of a new measure, the MONSI-CC, and for its validity in detecting changes in caregivers' implementation of NDBI strategies with their children with ASD while participating in various PMIs. Although intervention-specific fidelity ratings have been used to measure caregiver changes (Gulsrud et al. 2016; Pickles et al. 2016; Shire et al. 2015), the MONSI-CC provides a single measure to document behavioral changes in caregivers across different NDBIs involving components of PMIs. Our preliminary results showed that over an average of six months, the MONSI-CC detected improvements in caregivers' implementation of NDBI strategies with medium to large effect sizes. Rogers et al. (2012) detected changes in 12 weeks in caregiver strategies (using their FOI) over the course of their intervention with effect sizes ranging from .36 to .57 for both treatment and control groups (Rogers et al. 2012), similar in magnitude to changes measured with the MONSI-CC across caregivers who participated in three different intervention models. Our effect sizes are also consistent with changes in caregiver behavior found in other NDBIs (Green et al. 2010; Gulsrud et al. 2016; Shire et al. 2015). Post hoc analyses revealing that the treatment condition was a significant predictor of changes in NDBI strategy

implementation and Dyadic Engagement, with those in the treatment group showing significantly larger changes than those in the control group, further support the validity of the MONSI-CC. To our knowledge, the MONSI-CC is one of the first caregiver outcome measures to validly and reliably detect significant changes in caregivers' use of NDBI strategies across different interventions. The significant changes in caregiver behavior through various PMIs further reinforce the impact of caregiver interactions in early interventions for ASD (Gulsrud et al. 2016; Kasari et al. 2015; Pickles et al. 2016; Rogers et al. 2019b; Siller et al. 2013; Wetherby et al. 2014).

The comprehensiveness of the MONSI-CC and the analytical approach used to derive factor structures provided us the unique opportunity to move beyond theoretically groupings to evaluate empirically validated clusters of interactive strategies. In this sample, caregivers in PMIs improved in all measured interactive strategy domains grounded in NDBIs, although the magnitude of improvements varied across specific domains. This is consistent with other findings using conceptually grouped strategies within intervention-specific fidelity measures (Gulsrud et al. 2016; Shire et al. 2015). In addition, the empirically derived domains of strategies allowed us to examine how caregiver implementation of *overall* and *specific* interactive strategies are associated with changes in child outcomes in both domains of autism symptoms, social communication, and RRBs, as well as dyadic engagement between the child and the caregiver. Like others using intervention-specific fidelity measures (Gulsrud et al. 2016; Pickles et al. 2015; Rogers et al. 2012; Rogers et al. 2019b; Shire et al., 2015), we found that while accounting for baseline characteristics, improvement in *overall* caregiver implementation of strategies and *specific* domains (Environmental Set-up and Child-Guided Interactions) were significantly related to improvements in children's social communication. Furthermore, we also found that improvements in caregivers' implementation of *overall* strategies and improvements in *specific* domains of strategies (Environmental Set-up, Child-Guided Interactions,

Table 5 Preliminary regression analyses examining changes in parent strategy implementation predicting changes in child outcomes

Outcome variable	Predictors in model	<i>B</i>	SE	<i>t</i>	<i>Sig</i>	<i>R</i> ²	
BOSCC social communication (SC) (changes)	Environmental set-up (changes)	− 0.64	0.25	− 2.51	0.02	0.57	
	FSIQ (baseline)	− 0.05	0.04	− 1.22	0.23		
	CS-SA (baseline)	1.40	0.47	2.95	0.01		
	CS-RRB (baseline)	0.31	0.48	0.64	0.53		
	BOSCC SC (baseline)	− 0.60	0.10	− 5.80	<0.001		
	Child-guided interactions (changes)	Child-guided interactions (changes)	− 0.42	0.21	− 2.05	0.05	0.55
		FSIQ (baseline)	− 0.07	0.04	− 1.50	0.14	
		CS-SA (baseline)	1.54	0.48	3.23	<0.001	
		CS-RRB (baseline)	0.14	0.49	0.27	0.79	
		BOSCC SC (baseline)	− 0.59	0.11	− 5.54	<0.001	
	MONSI-CC total score (changes)	MONSI-CC total score (changes)	− 0.15	0.07	− 2.17	0.04	0.55
		FSIQ (baseline)	− 0.05	0.05	− 1.01	0.32	
		CS-SA (baseline)	1.33	0.50	2.67	0.01	
		CS-RRB (baseline)	0.03	0.50	0.06	0.95	
		BOSCC SC (baseline)	− 0.57	0.11	− 5.38	<0.001	
MONSI-CC dyadic engagement (DE) (changes)	Environmental set-up (changes)	0.12	0.04	3.07	<0.001	0.60	
	FSIQ (baseline)	0.02	0.01	2.91	0.01		
	CS-SA (baseline)	− 0.13	0.07	− 1.74	0.09		
	CS-RRB (baseline)	− 0.08	0.07	− 1.04	0.31		
	MONSI-CC DE (baseline)	− 0.59	0.12	− 4.96	<0.001		
	Child-guided interactions (changes)	Child-guided interactions (changes)	0.09	0.03	2.78	0.01	0.58
		FSIQ (baseline)	0.02	0.01	3.25	<0.001	
		CS-SA (baseline)	− 0.15	0.07	− 2.08	0.04	
		CS-RRB (baseline)	− 0.04	0.08	− 0.56	0.58	
		MONSI-CC DE (baseline)	− 0.59	0.12	− 4.80	<0.001	
	Active teaching and learning (changes)	Active teaching and learning (changes)	0.11	0.03	4.31	<0.001	0.66
		FSIQ (baseline)	0.01	0.01	1.98	0.06	
		CS-SA (baseline)	− 0.10	0.07	− 1.55	0.13	
		CS-RRB (baseline)	0.06	0.07	0.75	0.46	
		MONSI-CC DE (baseline)	− 0.45	0.12	− 3.88	<0.001	
MONSI-CC total score (changes)	MONSI-CC total score (changes)	0.04	0.01	4.18	<0.001	0.66	
	FSIQ (baseline)	0.02	0.01	2.38	0.02		
	CS-SA (baseline)	− 0.09	0.07	− 1.26	0.21		
	CS-RRB (baseline)	− 0.01	0.07	− 0.10	0.92		
	MONSI-CC DE (baseline)	− 0.53	0.11	− 4.69	<0.001		

and Active Teaching and Learning) significantly predicted changes in children's engagement level while controlling for child baseline characteristics. Improvements in strategies, specifically embedding Active Teaching and Learning and providing Opportunities for Engagement, were also significantly correlated with RRB symptom reduction.

In summary, our results show promising evidence for the utility of the MONSI-CC. The results support the use of the MONSI-CC as a valid, reliable measure of change in caregivers' behavior during various PMIs, which will allow the aggregation of samples across different studies to obtain greater power to examine effects and mechanisms

of PMI treatments. Although we were able to demonstrate that caregivers are in fact changing over time, we recognize there are many other factors and variables that impact the amount of change that occurs in both the caregiver and the child, which need to be explored further. Additionally, use of empirically derived domains to identify how implementation of *specific* strategies can affect child outcomes is critical for more effective programming of early interventions to meet the individual needs of the children with ASD and their caregivers. Therefore, with replication and further validation studies, we hope the MONSI-CC can be useful to select and prioritize both general and specific NDBI strategies for the varying needs of families and children.

These results have several theoretical and clinical implications. Like previous intervention studies that used study specific fidelity measures (Carter et al. 2011; Green et al. 2010; Gulsrud et al. 2016; Pickles et al. 2016; Rogers et al. 2012; Rogers et al. 2019b), based on a more broadly derived measure, we were able to show that caregivers can successfully learn and implement multiple interactive strategies taught in early treatments that incorporate parent-mediated components to promote social communication and to increase engagement in young children with ASD. Some strategies might improve more readily than others, although caregivers showed improvements across all domains of strategies measured by the MONSI-CC. These findings contribute to our understanding of how and when caregiver behavior changes, and eventually may help us understand which strategies may be more challenging for caregivers to learn and they are consistent with previous research showing that certain strategies might improve more than others (Gulsrud et al. 2016; Shire et al. 2015). For example, like others (Shire et al. 2015), play-based strategies such as expanding and developing play routines, represented by the Active Teaching and Learning domain, improved the least among caregivers as a whole. Caregivers, even those in active treatment conditions, also showed fewer improvements in their provision of Opportunities for Engagement (e.g., providing choices, using blocking, expectant waiting). It may be important for interventionists to spend more time, additional resources, or attention when teaching caregivers about play, expansion, developing routines, offering choices, and creating contexts to elicit communication.

Limitations and Future Directions

Although our preliminary findings are promising, they should be considered in the context of limitations. Our preliminary analyses examining the correlations between changes in implementation and changes in child outcomes does not equate causation. To evaluate causation would require random assignment across the whole sample and

identification of a sequence of change (e.g. parent change yields child change, rather than accompanies). Another limitation is that this sample of caregivers is predominantly white and female, so that more information is needed about more diverse populations and about male caregivers. Additionally, reliability was strong for MONSI-CC Total Scores and most of the Domain Scores, except for test–retest for Opportunities for Engagement, which showed lower ICC's, potentially due to the limited variability in item distribution (more than 50% scoring at 1 for test and re-test) and a small sample size ($n = 15$). Because the test–retest observations were 4–6 weeks apart, future studies should explore ICCs with test–retest observations with shorter intervals and a larger sample size. Moreover, there are other important areas of child development (e.g., language development, motor skills, adaptive functioning), besides the core ASD symptoms measured by the BOSCC that were not specifically evaluated in the current study. Although the BOSCC has been found to align well with other clinically meaningful developmental changes (Grzadzinski et al. 2016), additional analyses of the impact of strategy implementation on other child outcomes will be important. In addition, there may be other aspects of PMIs that mediate/moderate change that we did not capture with the MONSI-CC, which merit further exploration (e.g., family and caregiver characteristics, goals, and resources; (Shalev et al. 2019). Finally, this preliminary investigation was limited to caregiver strategies; however, we hope to extend similar coding to other social partners and settings (e.g., teachers in school settings) in the near future.

From a different point of view, while studies have shown that the involvement of caregivers maximizes the generalization of skills (Pickles et al. 2016; Strauss et al. 2012), and that teaching caregivers to apply NDBI strategies to preexisting daily routines (e.g., meals or baths) addresses the needs and goals of many families (Pickard et al. 2016), it is possible that adherence to professional or researcher-dictated strategies may not always be appropriate or helpful for all families (Nock et al. 2006). Thus, many researchers have underscored the importance of understanding the ways in which implementation and adherence impacts child and family outcomes (Kasari et al. 2014; Wainer and Ingersoll 2013). In other areas of psychology and mental health, this has led others to examine how much adherence and/or variation is detrimental to outcomes, leading to mixed findings regarding adherence to an arbitrary standard, i.e., 80% which has been widely accepted in the field (Durlak and DuPre 2008). The MONSI-CC would allow direct testing of different levels and patterns of change across different families, as well as across different interventions. We also recognize that the caregiver changes using MONSI-CC and child changes using BOSCC were rated based on the same segments of videos. Therefore, to measure the impact of PMI on the generalization of child's skill attainment and

symptom improvement across different contexts, measuring child behaviors with a non-caregiver as a social partner would be necessary.

Conclusion

The initial psychometric properties of a newly developed measure of changes in caregiver implementation of NDBI strategies, the MONSI-CC, are promising. The MONSI-CC was able to detect improvements in frequency, effectiveness, and appropriateness of caregiver strategy usage across several interventions with strong reliability and validity. The MONSI-CC also enabled us to find significant associations between increases in caregiver implementation and improvement in child behaviors with the caregiver. The MONSI-CC provides an opportunity for future research to evaluate the contribution of broadly defined changes in caregiver implementation to treatment effects for young children with ASD, furthering the development of individualized early interventions.

Acknowledgement This work was supported by a grant awarded to author S.H.K National Institute of Mental Health (NIMH; 1R01MH114925-01). Work for this project was also supported by Grants awarded to author C.L. from NIMH (R01MH081757, 1RC1MH089721, R01RFAMH14100, R01MH078165), Autism Speaks (5766), and Health Resources and Service Administration (HRSA; UA3MC11055). The views expressed in this work are those of the authors and not necessarily those of the NIMH, Autism Speaks or HRSA. The authors would like to thank Nurit Benrey-Grinberg and Hannah Thomas for their data coding. The authors would also like to thank Connie Kasari, Sally Rogers, and Amy Wetherby for the data and support of this project.

Author Contributions BV participated in the design of the study, statistical analysis, interpretation of data, and drafted the manuscript; SD conceived of the study, participated in the design of study, interpretation of the data and manuscript preparation; CK helped to draft the manuscript and performed the statistical analysis; YB helped to draft the manuscript and performed the statistical analysis; JW participated in the design of the study, and helped to draft the manuscript; CL conceived of the study, assisted in interpretation of the data, statistical analysis, and helped to draft the manuscript; SK conceived of the study; participated the design of the study, interpretation of data, statistical analysis, and manuscript writing. All authors read and approved the final manuscript.

Compliance with Ethical Standards

Conflict of interest C.L. receives royalties from the sale of the ADI-R and the ADOS-2. All royalties related to the research were donated to a non-profit organization. No other authors have conflicts of interest regarding this study.

References

- Kim, S.H., Dufek, S., Vibert B., & Lord, C. (2019). *Measure of NDBI Strategy Implementation - Caregiver Change* (MONSI-CC). Copyright in Progress.
- Barkley, R. A. (2013). *Defiant children: A clinician's manual for assessment and parent training*. New York: Guilford Press.
- Bearss, K., Burrell, T. L., Stewart, L., & Scahill, L. (2015). Parent training in autism spectrum disorder: What's in a name? *Clinical Child and Family Psychology Review*, 18(2), 170–182.
- Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the Differential Ability Scales in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, 116(5), 331–343.
- Browne, M., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen (Ed.), *Testing structural equation models* (Vol. 154, pp. 136–162). Newbury Park, CA: Sage.
- Carter, A. S., Messinger, D. S., Stone, W. L., Celimli, S., Nahmias, A. S., & Yoder, P. (2011). A randomized controlled trial of Hanen's 'More Than Words' in toddlers with early autism symptoms. *Journal of Child Psychology and Psychiatry*, 52(7), 741–752.
- Casagrande, K., & Ingersoll, B. (2017). Parent-mediated interventions for social communication in young children with ASD. In J. B. Leaf (Ed.), *Autism and Child Psychopathology. Handbook of Social Skills and Autism Spectrum Disorder* (pp. 285–312). Cham: Springer.
- Crockett, J. L., Fleming, R. K., Doepke, K. J., & Stevens, J. S. (2007). Parent training: Acquisition and generalization of discrete trials teaching skills with parents of children with autism. *Research in Developmental Disabilities*, 28(1), 23–36.
- Davilantis, K., Estes, A., Dawson, G., & Rogers, S. (2019). A novel method for measuring learning opportunities provided by parents to young children with autism spectrum disorder. *Autism*, 23(6), 1563–1574.
- Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., ... Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: The Early Start Denver Model. *Pediatrics*, 125(1), e17–e23.
- Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and factors affecting implementation. *American Journal of Community Psychology*, 41(3–4), 327–350.
- Eyberg, S., & Funderburk, D. (2011). *Parent-child interaction therapy protocol*. Gainesville, FL: PCIT International.
- Frost, K., Koehn, G. N., Russell, K. M., & Ingersoll, B. (2019). Measuring child social communication across contexts: Similarities and differences across play and snack routines. *Autism Research*, 12(4), 636–644.
- Gengoux, G. W., Berquist, K. L., Salzman, E., Schapp, S., Phillips, J. M., Frazier, T. W., ... Hardan, A. Y. (2015). Pivotal response treatment parent training for autism: Findings from a 3-month follow-up evaluation. *Journal of Autism and Developmental Disorders*, 45(9), 2889–2898.
- Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39(5), 693–705.
- Grahame, V., Brett, D., Dixon, L., McConachie, H., Lowry, J., Rodgers, J., ... Le Couteur, A. (2015). Managing repetitive behaviours in young children with autism spectrum disorder (ASD): Pilot randomised controlled trial of a new parent group intervention. *Journal of Autism and Developmental Disorders*, 45(10), 3168–3182.
- Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., ... Pickles, A. (2010). Parent-mediated

- communication-focused treatment in children with autism (PACT): A randomised controlled trial. *The Lancet*, 375(9732), 2152–2160.
- Grzadzinski, R., Carr, T., Colombi, C., McGuire, K., Dufek, S., Pickles, A., et al. (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(7), 2464–2479.
- Gulsrud, A. C., Helleman, G., Shire, S., & Kasari, C. (2016). Isolating active ingredients in a parent-mediated social communication intervention for toddlers with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 57(5), 606–613.
- Hus, V., Gotham, K., & Lord, C. (2014). Standardizing ADOS domain scores: Separating severity of social affect and restricted and repetitive behaviors. *Journal of Autism and Developmental Disorders*, 44(10), 2400–2412.
- Individuals with Disabilities Education Act (IDEA). 20 U.S.C. § 1400 (2004).
- Ingersoll, B., & Dvortcsak, A. (2010). *Teaching social communication to children with autism: A manual for parents*. New York: Guilford Press.
- Kasari, C., Gulsrud, A., Paparella, T., Helleman, G., & Berry, K. (2015). Randomized comparative efficacy study of parent-mediated interventions for toddlers with autism. *Journal of Consulting and Clinical Psychology*, 83(3), 554–563.
- Kasari, C., Lawton, K., Shih, W., Barker, T. V., Landa, R., Lord, C. (2014). Caregiver-mediated intervention for low-resourced preschoolers with autism: An RCT. *Pediatrics*, 134(1), e72–e79.
- Kazdin, A. E. (2005). *Parent management training: Treatment for oppositional, aggressive, and antisocial behavior in children and adolescents*. New York: Oxford University Press.
- Kim, S. H., Grzadzinski, R., Martinez, K., & Lord, C. (2018). Measuring treatment response in children with autism spectrum disorder: Applications of the Brief Observation of Social Communication Change to the Autism Diagnostic Observation Schedule. *Autism*, 23(5), 1176–1185.
- Le Couteur, A., Lord, C., & Rutter, M. (2003). *The autism diagnostic interview-revised (ADI-R)*. Los Angeles, CA: Western Psychological Services.
- Lord, C., Rutter, M., DiLavore, P., & Risi, S. (2000). *ADOS: Autism diagnostic observation schedule*. Los Angeles, CA: Western Psychological Services.
- Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule: ADOS-2*. Los Angeles, CA: Western Psychological Services.
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, 55(1), 3–9.
- Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines, MN: AGS.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide 1998–2012* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- National Research Council. (2001). *Educating children with autism*. Washington, DC: National Academies Press.
- Nock, M., Ferriter, C., & Holmberg, E. (2006). Parents beliefs about treatment credibility and effectiveness: Assessment and relation to subsequent treatment participation. *Journal of Child and Family Studies*, 16(1), 27–38.
- Oono, I. P., Honey, E. J., & McConachie, H. (2013). Parent-mediated early intervention for young children with autism spectrum disorders (ASD). *Evidence-Based Child Health: A Cochrane Review Journal*, 8(6), 2380–2479.
- Pickard, K. E., Kilgore, A. N., & Ingersoll, B. R. (2016). Using community partnerships to better understand the barriers to using an evidence-based, parent-mediated intervention for autism spectrum disorder in a medicaid system. *American Journal of Community Psychology*, 57(3–4), 391–403.
- Pickles, A., Harris, V., Green, J., Aldred, C., McConachie, H., Slonims, V., ... the PACT Consortium. (2015). Treatment mechanism in the MRC preschool autism communication trial: Implications for study design and parent-focussed therapy for children. *Journal of Child Psychology and Psychiatry*, 56(2), 162–170.
- Pickles, A., Le Couteur, A., Leadbitter, K., Salomone, E., Cole-Fletcher, R., Tobin, H., ... Green, J. (2016). Parent-mediated social communication therapy for young children with autism (PACT): Long-term follow-up of a randomised controlled trial. *The Lancet*, 388(10059), 2501–2509.
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., ... Hensley, M. (2011). Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health*, 38(2), 65–76.
- Rogers, S., Estes, A., Lord, C., Munson, J., Rocha, M., Winter, J., ... Talbot, M. (2019). A multisite randomized controlled two-phase trial of the Early Start Denver Model compared to treatment as usual. *Journal of the American Academy of Child & Adolescent Psychiatry*, Advance online publication.
- Rogers, S., Estes, A., Lord, C., Vismara, L., Winter, J., Fitzpatrick, A., ... Dawson, G. (2012). Effects of a brief Early Start Denver Model (ESDM)-based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(10), 1052–1065.
- Rogers, S., Estes, A., Vismara, L., Munson, J., Zierhut, C., Greenson, J., ... Talbot, M. (2019). Enhancing low-intensity coaching in parent implemented early start denver model intervention for early autism: A randomized comparison treatment trial. *Journal of Autism and Developmental Disorders*, 49(2), 632–646.
- Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., ... Halladay, A. (2015). Naturalistic developmental behavioral interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 45(8), 2411–2428.
- Shalev, R., Lavine, C., & Di Martino, A. (2019). A systematic review of the role of parent characteristics in parent-mediated interventions for children with autism spectrum disorder. *Journal of Developmental and Physical Disabilities*, 1–21.
- Shire, S. Y., Goods, K., Shih, W., Mucchetti, C., Kaiser, A., Wright, C., ... Kasari, C. (2015). Parents' adoption of social communication intervention strategies: Families including children with autism spectrum disorder who are minimally verbal. *Journal of Autism and Developmental Disorders*, 45(6), 1712–1724.
- Siller, M., Hutman, T., & Sigman, M. (2013). A parent-mediated intervention to increase responsive parental behaviors and child communication in children with ASD: A Randomized Clinical Trial. *Journal of Autism and Developmental Disorders*, 43(3), 540–555.
- Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models* (1st ed.). Boca Raton: Chapman & Hall/CRC.
- Solomon, R., Necheles, J., Ferch, C., & Bruckman, D. (2007). Pilot study of a parent training program for young children with autism: The PLAY Project Home Consultation program. *Autism*, 11(3), 205–224.
- Strauss, K., Vicari, S., Valeri, G., D'Elia, L., Arima, S., & Fava, L. (2012). Parent inclusion in Early Intensive Behavioral Intervention: The influence of parental stress, parent treatment fidelity and parent-mediated generalization of behavior targets on child outcomes. *Research in Developmental Disabilities*, 33(2), 688–703.
- Wainer, A., & Ingersoll, B. (2013). Intervention fidelity: An essential component for understanding ASD parent training research

and practice. *Clinical Psychology: Science and Practice*, 20(3), 335–357.

Wetherby, A. M., Guthrie, W., Woods, J., Schatschneider, C., Holland, R. D., Morgan, L., et al. (2014). Parent-implemented social intervention for toddlers with autism: An RCT. *Pediatrics*, 134(6), 1084–1093.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.