

Web-based Assessment of Social–Emotional Skills in School-Aged Youth with Autism Spectrum Disorder

Nicole M. Russo-Ponsaran , Matthew D. Lerner , Clark McKown, Rebecca J. Weber, Ashley Karls, Erin Kang, and Samantha L. Sommer

Few tools are available to comprehensively describe the unique social–emotional skill profiles of youth with autism spectrum disorder (ASD). The present study describes the usability, reliability, and validity of SELweb, a normed, web-based assessment designed to measure four core social–emotional domains, when used to measure these skills in a sample of 57 well-characterized youth with ASD (ages 6–10 years with IQ \geq 80). SELweb measures facial emotion recognition, theory of mind, social problem solving, and self-control. SELweb was well tolerated and yielded scores with reliabilities comparable to those found in normative samples. SELweb scores showed good evidence of convergent and discriminant validity for three of the four skills it was designed to assess. Mean deficits were found for theory of mind, social problem solving, and self-control, whereas no mean deficits were found for emotion recognition. Individual profiles varied considerably, suggesting the sensitivity of SELweb to the within- and between-person individual differences among youth with ASD. Findings support the usefulness and accessibility of SELweb as a tool for measuring complex social–emotional skill profiles in youth with ASD. *Autism Res* 2019, 00: 1–12. © 2019 International Society for Autism Research, Wiley Periodicals, Inc.

Lay Summary: No single, simple, high-quality test exists that measures multiple social thinking skills directly in children with autism spectrum disorder (ASD). The present study suggests that SELweb, a web-based assessment system, is an effective and valid way to measure how children with ASD think about and understand social and emotional information, and is able to capture strengths and weaknesses experienced by children with ASD.

Keywords: autism; assessment; emotion recognition; theory of mind; social problem solving; self-control

Introduction

Decades of research suggest that difficulties in social–emotional skills (i.e., thinking about, understanding, and responding to others) are broadly characteristic of youth with autism spectrum disorder (ASD; Baron-Cohen, 2000). Recent work indicates that substantial heterogeneity within and across social–emotional domains may be the norm rather than the exception in ASD (e.g., Lombardo et al., 2016; Lozier, Vanmeter, & Marsh, 2014; Mendelson, Gates, & Lerner, 2016; Rice, Moriuchi, Jones, & Klin, 2012). As such, describing the unique social–emotional skill profiles of youth with ASD has proven challenging. This challenge is in large part due to the complex nature of the social deficits themselves. Such deficits are often context-specific and multifaceted, involving multiple social–emotional domains (e.g., perception and interpretation of social signals). Relatively few studies evaluate multiple domains in their samples. Rather, most studies evaluate skills in isolation or in

broad heterogeneous samples of children with ASD. For this reason, researchers and practitioners must piece together disparate data from different samples to try to describe social–emotional skills in ASD. In order to effectively understand the social–emotional skill profiles in ASD and better inform treatment development, a simple, usable, feasible, and technically sound system that assesses multiple domains is needed. To address this need, the present study evaluates SELweb as a tool to offer insight into the multifaceted nature of social–emotional skill profiles in youth with ASD.

SELweb

The web-based assessment SELweb (early elementary version) was designed in the context of a social–emotional learning framework (Lipton & Nowicki, 2009; McKown, Gumbiner, Russo-Ponsaran, & Lipton, 2009). This framework implicates emotion recognition, social perspective taking or theory of mind, social problem solving, and

From the Department of Psychiatry & Behavioral Sciences, Rush University Medical Center, Chicago, Illinois (N.M.R.-P., C.M., A.K.); Department of Psychology, Stony Brook University, Stony Brook, New York (M.D.L., R.J.W., E.K., S.L.S.)

Received November 23, 2018; accepted for publication April 21, 2019

Address for correspondence and reprints: Nicole M. Russo-Ponsaran, Rush NeuroBehavioral Center, 4711 Golf Road, Suite 1100, Skokie, IL 60076. E-mail: nicole_russo@rush.edu; Matthew D. Lerner, Department of Psychology Stony Brook University, Stony Brook, NY 11794-2500. E-mail: matthew.lerner@stonybrook.edu

Published online 00 Month 2019 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/aur.2123

© 2019 International Society for Autism Research, Wiley Periodicals, Inc.

Table 1. SELweb Modules, Domains, and Descriptions

Module	Domain	Description of task
Facial Emotion Recognition	Social Awareness	Individual child faces are shown ($n = 40$). Subjects indicate whether each face looks Happy, Sad, Angry, Scared, or Just Okay
Social Problem Solving	Social Reasoning	Illustrated and narrated vignettes of challenging social situations are presented ($n = 6$). Subjects must imagine themselves in that situation and answer questions indicating Problem Identification, Hostile Attribution, Goal Preference, and Solution Preference
Theory of Mind	Social Meaning	Illustrated and narrated vignettes are presented ($n = 12$). Subjects must infer a character's mental or emotional state (e.g., the reason they said something)
Delay of Gratification	Self-Control	Subjects send a total of 10 illustrated rockets to space. Subjects can choose between three rockets each time that vary in speed, with slower rockets earning more points (Task was previously referred to as Choice-Delay Task; McKown et al., 2016)
Frustration Tolerance	Self-Control	Pairs of shapes are shown, varying in both shape and color. Subjects must indicate whether the shapes in each pair are the same, regardless of color, answering as many correctly as they can in 90 s. Several items are programmed to become "stuck"

self-control as central in social functioning (McKown, Allen, Russo-Ponsaran, & Johnson, 2013). Accordingly, SELweb contains assessment modules for each of these areas (Table 1). It is a brief (35–40 min) assessment typically used in school settings to measure the social-emotional skills of general education (GenEd) students in kindergarten through third grades. Because SELweb is web-based and self-administered, it is scalable and easy to use. Although the skills measured by SELweb are not exhaustive of the challenges facing youth with ASD, these are noted areas of difficulty for individuals with ASD.

Three studies including 8,881 kindergarten through third grade GenEd students were conducted to establish SELweb's psychometric properties (McKown et al., 2013; McKown, Russo-Ponsaran, Johnson, Russo, & Allen, 2016). Score reliabilities for SELweb ranged from 0.78 to 0.85 in its first validation field trial, from 0.78 to 0.88 in its second validation field trial, and from 0.79 to 0.89 in its national norming study. Six-month temporal stabilities ranged from 0.60 to 0.70 in the first field trial, from 0.52 to 0.69 in the second field trial, and from 0.55 to 0.79 in the norming study. Across those studies, scores on the assessment modules fit a hypothesized four-factor model that includes emotion recognition, theory of mind, social problem solving, and self-control. Analyses also supported SELweb's criterion-related validity: controlling for age, IQ (in field trial one), sex, and ethnicity; overall performance on SELweb was positively associated with teacher-reported social skill, peer acceptance, and academic competence; and negatively associated with teacher-reported problem behavior. Analyses also supported the convergent and discriminant validity of SELweb's social-emotional composite score.

Although it was designed for use with GenEd students, the four distinct social-emotional domains SELweb measures are all often impaired in ASD (Baron-Cohen, 2000; Dawson, Webb, & McPartland, 2005; Mendelson et al., 2016). For example, many individuals with ASD demonstrate deficits in facial emotion recognition (e.g., Klin et al., 1999).

Past (e.g., Baron-Cohen, 2000; Yirmiya, Erel, Shaked, & Solomonica-Levi, 1998) and current (e.g., Livingston, Colvert, Bolton, & Happé, 2018; Livingston & Happé, 2017) research suggest that deficits in theory of mind, or making mental representations of others' emotions or thoughts, may be a core feature of a substantial proportion of the ASD population. Impairments are also often seen in the ability to identify a complex social problem, adapt to the social environments accordingly, and generate and select socially appropriate solutions (e.g., Channon, Crawford, Orlowska, Parikh, & Thoma, 2014; Russo-Ponsaran et al., 2018; Shulman, Guberman, Shiling, & Bauminger, 2012). Finally, individuals with ASD frequently exhibit difficulties with self-control, or self-regulation, such as the use of less adaptive self-regulation strategies (e.g., Barnard-Brak, Ivey-Hatz, Ward, & Wei, 2014; Cai, Richdale, Uljarevic, Dissanayake, & Samson, 2018; Hepburn & Wolff, 2013; Loveland, 2005; Mazefsky et al., 2013).

There are few well-validated direct assessment measures that address multiple domains of social-emotional learning and their relationship to social challenges in individuals with ASD. Existing assessments address some of these needs, but to date, none addresses all. Parent and teacher rating scales are well-suited for measuring general behavior but are not well-suited for measuring social-emotional comprehension skills. By definition, these skills involve mental events that do not have direct behavioral correlates and therefore require a high level of inference for an observer to rate (Epkins, 1994; van der Valk, van den Oord, Verhulst, & Boomsma, 2001). Furthermore, the most popular and well-validated rating scales for measuring social behavior in this population either do not capture social difficulties most relevant to ASD (Gresham & Elliott, 2008) or conform to a model of ASD symptomology rather than a broader theoretical framework for understanding social behavior (Bellini & Hopf, 2007; Constantino & Gruber, 2012). Behavior observation systems can measure specific actions but may not be treatment-sensitive or may have limited generalizability across contexts (Murphy, Martin, &

Garcia, 1982). Physiological indicators may measure neural or hormonal metrics of change, but they often do not reliably correlate with the behaviors they are meant to indicate (De Los Reyes et al., 2012; Nelson, Perlman, Hajcak, Klein, & Kotov, 2015; Patriquin, Lorenzi, Scarpa, Calkins, & Bell, 2015). Finally, a growing array of computer-based tools offer objective performance metrics of social, cognitive, or perceptual abilities (e.g., Canty, Neumann, Fleming, & Shum, 2017; Dadds, Hawes, & Merz, 2004; Dziobek et al., 2006; Golan, Baron-Cohen, & Golan, 2008; Klin, 2000; Sivaratnam, Cornish, Gray, Howlin, & Rinehart, 2012; Tanaka et al., 2012; Wirth & Klieme, 2003). Most of these either target a single construct, must be administered in a highly structured laboratory setting with a proctor's support, or lack broader population norms to which performance may be compared. The limitations of different methods and theoretical frameworks for assessing social processes have been of similar concern in application to other unique populations, such as in schizophrenia (Yager & Ehmann, 2006). Without the availability of a single assessment that addresses these gaps, previous studies seeking to measure social-emotional learning have often required lengthy batteries of assessments of varying formats (e.g., Demopoulos, Hopkins, & Davis, 2013; Denham et al., 2012; Didehbani, Allen, Kandalaf, Krawczyk, & Chapman, 2016; Sasson, Nowlin, & Pinkham, 2012; Stichter et al., 2010), which can limit comparability across studies and may be more likely to yield incomplete data. SELweb addresses many of the limitations of other assessments.

Despite its potential, no study to date has evaluated the usability, sensitivity, and psychometric properties of SELweb in youth with ASD. To address these gaps, we conducted a multisite study to evaluate the psychometric properties of SELweb in a well-characterized sample of verbal youth with ASD (ages 6–10 years). This study tested four main hypotheses: (a) SELweb would be usable by a sample of youth with ASD; (b) SELweb module scores would demonstrate evidence of internal consistency reliability; (c) SELweb would demonstrate criterion-related validity as reflected in expected associations between performance on SELweb with other measures of the same social-emotional skills, and with ASD symptoms; and (d) youth with ASD would exhibit overall lower performance on SELweb in comparison to established normative data from GenEd students.

Methods

University Institutional Review Boards at two sites approved all procedures. Participants were made aware of the second site and its investigators and were informed that only coded data without names or contact information would be shared with the other site to facilitate inclusive data analyses. After eligibility was established, participating youth completed SELweb and alternate measures of the same domains as SELweb that were also

used in our prior validation of SELweb in GenEd students (McKown et al., 2016).

Eligibility, Recruitment, and Informed Consent

Six- to 10-year old boys and girls who were verbal, met autism spectrum cutoffs on the Autism Diagnostic Observation Schedule-2 (ADOS-2; Lord et al., 2012), had a score ≥ 11 on parent-report on the Lifetime Social Communication Questionnaire (Norris & Lecavalier, 2010), and had average intellectual ability (≥ 80 on standardized test) were eligible for the study. The original assessment was designed for and validated in kindergarten through third grade GenEd students or roughly 5–8 year olds. Within the ASD population, the extant literature indicates greater variability in social-emotional skills up to higher ages and delayed mastery of skills relative to nonclinical populations (see Baron-Cohen, 2001; Happé, 1995, for reviews). Because of the expected developmental delay coincident with a diagnosis of ASD and to capture the variance in measures of social-emotional skills in this population, we extended our age range to 10 years old.

One hundred fifteen youth were recruited from site clinics, institutional review board-approved research registries, local schools and agencies in the Chicagoland and Stony Brook areas, word-of-mouth referrals, and through the use of electronic media, parent magazines, and information tables at area events. Research staff followed up on leads with parent phone calls to complete pre-enrollment.

After an initial phone call, verbal consent was obtained and parents were invited to complete the Social Communication Questionnaire online via Research Electronic Data Capture (Harris et al., 2009) as a preliminary eligibility screener. In instances where a parent preferred an alternate format or did not have access to the screening questionnaire electronically, the parent either completed a paper copy or was administered the questionnaire through a semi-structured interview over the phone or in person with a staff member. One hundred four youth whose score on the Social Communication Questionnaire was ≥ 11 (Norris & Lecavalier, 2010) were invited to complete further testing. Ninety-two families subsequently consented and assented to participate in the full study. Testing included both eligibility and experimental components (approximately 3 hours). Testing was offered in one or two sessions held either at university offices or at community partner facilities, according to the family's preference.

Procedures

The ADOS-2 was only administered by certified, independently research-reliable administrators (masters-level staff, doctoral candidates, and doctorate-level faculty) associated with the study. All other aspects of the study were administered by bachelors- and masters-level research staff and

university students who have experience working with clinical and pediatric populations.

At the first in-person eligibility session, the ADOS-2 (Module 2 or 3; Lord et al., 2012) and the Kaufman Brief Intelligence Test-2 (Kaufman & Kaufman, 2004) were administered to youth. The ADOS-2 is the gold standard instrument for characterizing symptoms of ASD. It involves administering and scoring a sequence of “social presses” to elicit normative social responses, if capacity is present in the participant. Eligible participants had an overall score ≥ 7 on Module 3 or ≥ 8 on Module 2 which is consistent with an autism spectrum classification (Lord et al., 2012). In addition, the ADOS-2-calibrated severity score was calculated to provide a standardized severity metric of autism symptomatology for other analyses. The Kaufman Brief Intelligence Test-2 provides standard scores of crystallized (verbal) and fluid (nonverbal) IQ for a range of individuals aged 4–90 years. (One child had scores from the Wechsler Preschool and Primary Scale of Intelligence—Fourth Edition (Wechsler, 2012), completed within the past year, and these scores were allowed in lieu of completing an additional IQ assessment.) Youth who met criteria for an ASD as determined by their scores on both the Social Communication Questionnaire and the ADOS-2 and who scored ≥ 80 on the intelligence test were invited to complete the experimental portion of testing ($n = 59$). Of those 59, two youth dropped out before completing the study, yielding a total of 57 participants.

Primary Measure

SELweb. SELweb is web-based, self-administered, auto-scored direct assessment. It is easy to use with built-in, narrated instructions delivered by an animated talking

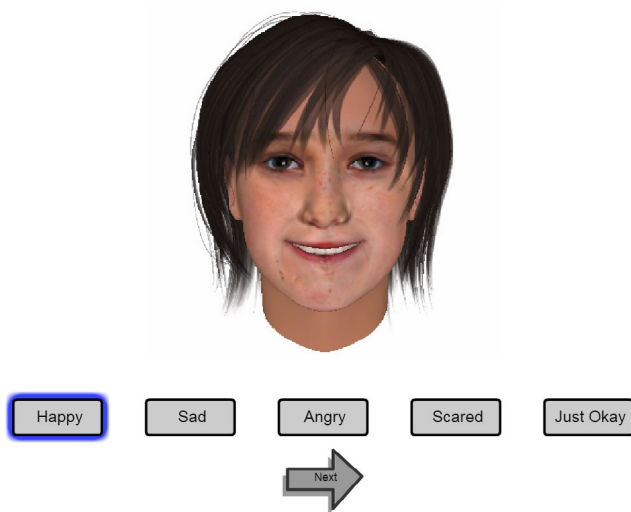


Figure 1. A sample screenshot from SELweb’s emotion recognition module. For this module, users view individual faces and indicate the emotion expressed.

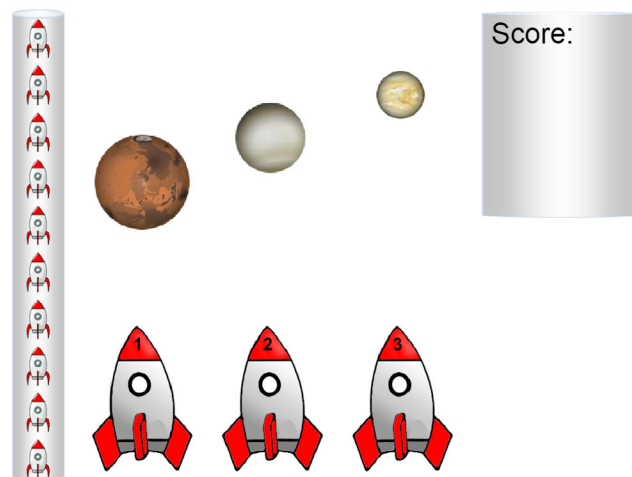


Figure 2. A sample screenshot from SELweb’s self-control (delay of gratification) module. For this module, users send animated rocket ships to space and earn points for selecting different rockets. More points are awarded when the user chooses the slower and more tedious rockets.

dog. SELweb modules are presented in the following order: emotion recognition (Fig. 1), self-control #1 (delay of gratification; Fig. 2), theory of mind (Fig. 3), social problem solving (Fig. 4), and self-control #2 (frustration tolerance; Fig. 5). Details about each module are presented in Table 1. All items are narrated and answer choices are repeated when the user hovers the cursor over



Figure 3. A sample screenshot from SELweb’s theory of mind module. For this module, users listen to illustrated, narrated vignettes and answer questions about why the story character did or said what he or she did.

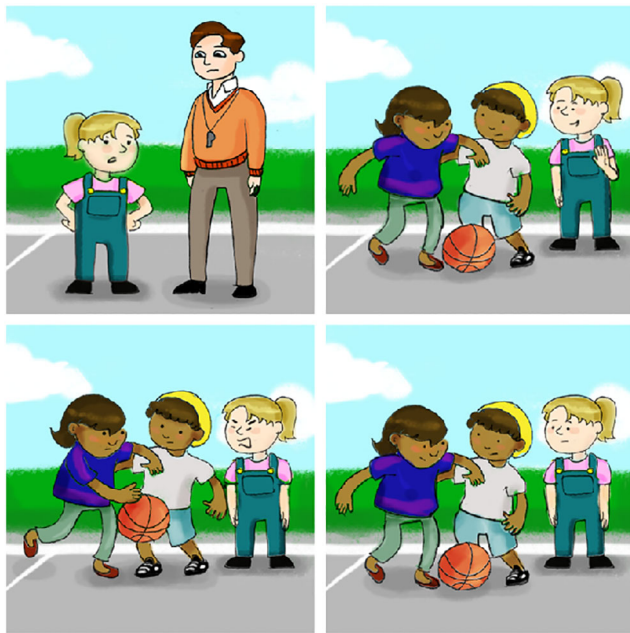


Figure 4. A sample screenshot from SELweb's social problem-solving module. For this module, users are asked to pretend that they are one of the story characters while they hear illustrated, narrated vignettes involving either ambiguous provocation or peer entry.

a response. There are no minimum reading level requirements, nor do any of the activities require a verbal or written response. A comprehensive description of stimuli, response options, and scoring for all modules are described in McKown et al. (2016). SELweb raw scores have been normed on large samples of youth and converted to standard scores ($M = 100$, $SD = 15$).

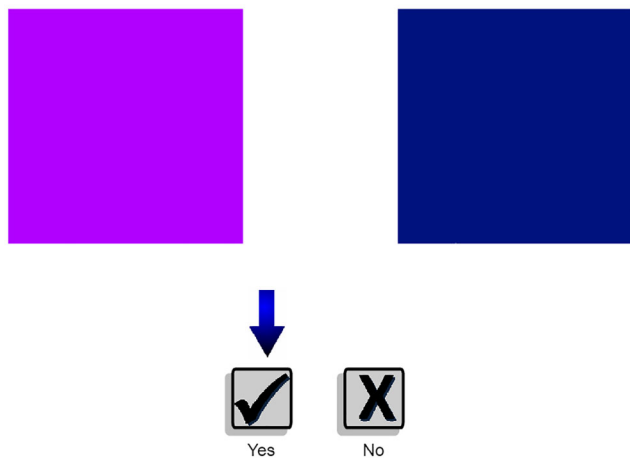


Figure 5. A sample screenshot from SELweb's self-control (frustration tolerance) module. For this module, users view pairs of shapes and indicate whether they are the same or different shapes. Several items are programmed to get stuck and induce frustration. Total score is the number correct and represents the ability of the user to stay focused and perform well despite frustration.

Validation Measures

The selected validation measures were chosen because they are widely used, have strong psychometric properties, and are designed to measure skills that parallel what SELweb is designed to measure.

Emotion recognition. For emotion recognition, a stimulus set from the University of California at Davis Set of Emotion Expressions (UCDSEE; Tracy, Robins, & Schriber, 2009) was administered via computer. Participants viewed four sets of 18 photographs of people and indicated whether the emotion expressed in each photograph matched a target emotion (happy, sad, angry, and scared). Photographs were presented in random order for each of the four target emotions. Final score was the number of correct items. Cronbach's α , an estimate of internal consistency, was 0.90 for the UCDSEE in our sample.

Theory of mind. Participants completed five vignettes from the Strange Stories (Happé, 1994). In each vignette, a character states one thing but intends something else. Participants listened to recordings of the stories while looking at printouts of the stories with minimal accompanying pictures. They were then asked why the character said what he or she said. Participants received one point for each correct inference of the speaker's intention. The final score on Strange Stories was the summed item scores. Cronbach's α was 0.81 for these Strange Stories vignettes in our sample.

Social problem solving. Participants completed four video-vignettes from the Social Information Processing Application (SIP-AP; Kupersmidt, Stelter, & Dodge, 2011), a web-based measure of social problem solving. Videos are from a first-person perspective, and all the main actors are males. Following each vignette, participants answered a selection of questions assessing problem identification, goals, and solution preference. These questions were grouped into five different categories: hostile intent attribution ($\alpha = 0.50$), aggressive goal ($\alpha = 0.84$), prosocial goal ($\alpha = 0.69$), overtly aggressive solution ($\alpha = 0.92$), and relationally aggressive solution ($\alpha = 0.87$). Total score for each category was the average of scores across vignettes. The average score across all categories was also examined; Cronbach's α was 0.89 for the overall average score on the SIP-AP in our sample.

Self-control. Participants completed the Distractibility and Go/No Go subtests from the computer-based Test of Attentional Performance for Children (KiTAP; Zimmermann, Gondan, & Fimm, 2005). Scores for Distractibility were based on the number of misses with (0.69–0.80) and without (0.66–0.72) a distractor, and scores for Go/No Go were based on the number of false reactions (0.65–0.66) and the

median reaction time (0.77). The total number of correct responses was also considered for each subtest. Because individual item-level data were unavailable, published split-half reliabilities were reviewed, which ranged from 0.66 to 0.74.

Data Analytic Plan

To test the first hypothesis that SELweb is usable in our target population of 6- to 10-year-old youth with ASD who do not have an intellectual deficit, we collected tester records and observations of each SELweb administration and assessed the data qualitatively.

To test the second hypothesis that SELweb maintains internal consistency reliability in our target population, two methods were used. For single-module scores, Cronbach's alpha values were calculated. For composite scores made up of individual scores, internal consistency reliabilities were calculated using a formula described by Nunnally and Bernstein (1994). This calculation estimates the reliability of composite scores using an algorithm that takes into account the internal consistency of each component score, the strength of the factor loadings, and the covariance between the component scores. The current sample was too small to estimate factor loading with confirmatory factor analysis. Instead, estimated factor loadings from the SELweb norming sample were used (McKown, 2018).

To test the third hypothesis that SELweb is a valid assessment tool of social-emotional skills in our target population, we examined the association between SELweb performance and performance on alternate measures of the same constructs (criterion measures). For these analyses, as well as hypothesis 2, a multitrait-multimethod matrix is presented to visualize aspects of convergent and discriminant validity. To show evidence of criterion-related validity, we ran regression models with the SELweb score as the predictor, the score on the alternate measure of the same construct as the criterion, and with age and estimated IQ entered as covariates. (We calculated regressions with and without site as a covariate. The magnitude did not vary greatly and significance did not change; therefore, we maintained the simpler model in our presentation of methods and results.) This method provides a stricter test of convergent validity as it controls for potential confounding by developmental and cognitive factors.

To test the fourth hypothesis that performance on SELweb among our target population indicates social-emotional deficits compared to the normative sample as expected for youth with ASD, we ran one-sample *t*-tests, comparing the standard scores on SELweb and its subscales to the normative mean of 100.

Finally, we calculated bivariate correlations between all SELweb scores and the ADOS-2 calibrated symptom severity indices.

Results

Fifty-seven rigorously characterized youth ($M_{age} = 8.67$, $SD_{age} = 1.32$; 44 male) who met eligibility cutoffs based on the ADOS-2 ($M = 7.46$, $SD = 1.93$), Social Communication Questionnaire ($M = 21.18$, $SD = 5.81$), and intelligence testing ($IQ \geq 80$; $M = 103.72$, $SD = 14.33$) completed study procedures. Demographic data are provided in Table 2.

Usability

Overall, SELweb was well received and easily completed by participants. On average, participants needed 33.7 ± 4.36 min to complete SELweb. Seven participants required one short break of less than 5 min. The majority (82.5%) of participants completed SELweb without needing assistance.

Qualitative observations of usability. Behavioral observations and comments during testing suggested that participants were actively engaged with and largely enjoyed SELweb throughout testing. More than half of the participants verbally expressed engagement or enjoyment at some point during testing (e.g., "This part is

Table 2. Participant Demographics by Site and Combined

Measure	Site 1 (Stony Brook) <i>n</i> (%)	Site 2 (Rush) <i>n</i> (%)	Total <i>n</i> (%)
Age, <i>M</i> (<i>SD</i>)	6.5 (1.2)	6.1 (1.5)	6.1 (1.3)
IQ, <i>M</i> (<i>SD</i>)	81 (14.5)	80 (12.2)	80 (14.1)
Sex: male	26 (76.5)	18 (78.3)	44 (77.2)
Income			
<\$30,000	2 (5.9)	1 (4.3)	3 (5.3)
\$30,000–\$60,000	2 (5.9)	2 (8.7)	4 (7.0)
\$60,000–\$90,000	1 (2.9)	5 (21.7)	6 (10.5)
\$90,000–\$120,000	7 (20.6)	5 (21.7)	12 (21.1)
>\$120,000	21 (61.8)	9 (39.1)	30 (52.6)
Unknown	1 (2.9)	1 (4.3)	2 (3.5)
Caregiver 1 education			
HS graduate or GED	3 (8.8)	0	3 (5.3)
Some college or post-HS	7 (20.6)	3 (13)	10 (17.5)
College graduate	8 (23.5)	13 (56.5)	21 (36.8)
Advanced degree	16 (47.1)	7 (30.4)	23 (40.4)
Caregiver 2 education			
HS graduate or GED	4 (11.5)	3 (13)	7 (12.3)
Some college or post-HS	7 (20.6)	4 (17.4)	11 (19.3)
College graduate	10 (29.4)	6 (26.1)	16 (28.1)
Advanced degree	10 (29.4)	8 (34.8)	18 (31.6)
Unknown	3 (8.8)	2 (8.7)	5 (8.8)
Ethnicity			
White	30 (88.2)	14 (60.9)	44 (77.2)
Black	1 (2.9)	0	1 (1.8)
Asian	0	1 (4.3)	1 (1.8)
Mixed	3 (8.8)	6 (26.1)	9 (15.8)
Unknown	0	2 (8.7)	2 (3.5)
Total	34	23	57

HS, High School; GED, General Education Diploma.

awesome!"; "I'm getting so many right"; "It has to be one of these two"). Six participants requested task-specific clarification during the delay of gratification task (e.g., "Is this what we're supposed to do?") and the social problem-solving and theory of mind modules (e.g., "This is my opinion?"; "What are these? Things I can do?"). Two participants asked for basic clarification during the emotion recognition module, including asking research staff to read the emotion words aloud (despite availability of voiceover for an item whenever the computer cursor hovers over it) and asking research staff to define the phrase "just okay." Thirteen participants received some level of redirection. For example, participants became fixated on aspects of the activities or the computer itself (e.g., the order of answer options, the computer mouse, etc.) or became distracted by objects in the room. Two participants utilized "fidget" objects periodically to maintain focus and manage behavior. Fifteen participants expressed fatigue or boredom through either verbal or nonverbal indications, particularly during the emotion recognition module.

Reliability

There were notes from staff suggesting possible perseverative response patterns for six participants on one of the self-control modules. These participants appeared overly focused on the ordering or numbering of response options (1, 2, 3, 1, 2, 3, etc.). Whether these cases were truly perseverative or merely coincidental could not be confirmed, and hence internal consistency was evaluated with and without the questionable data. Analyses yielded largely equivalent results, so all data from that module were included.

Cronbach's α for SELweb's social-emotional composite score was 0.89. Individual modules demonstrated Cronbach's alphas between 0.73 and 0.88 and were similar in magnitude to internal consistency reliabilities found in two large studies

with GenEd students (McKown, 2018; McKown et al., 2016; Table 3).

Validity

Associations with alternate measures of the same constructs. Evidence of convergent and discriminant validity is presented in Tables 4 and 5. Table 4 presents the multitrait-multimethod matrix, which demonstrates that, with the exception of self-control, monotrait-heteromethod correlations for each SELweb scale (a) were significant, denoting convergent validity; (b) were greater than heterotrait-heteromethod correlations in the same row or column, denoting discriminant validity, in that SELweb scores were indeed less correlated with measures of other social-emotional skills than with measures of the same skill; and (c) were greater than heterotrait-monomethod correlations, such that the variance shared by common SEL constructs was greater than shared method variance. Table 5 summarizes the relationship between SELweb scores and scores on alternate measures of the same constructs after controlling for age and IQ. The overall SELweb score was associated with scores on alternate measures of emotion recognition, theory of mind, and social problem solving, but not self-control. Performance on the SELweb emotion recognition module was significantly associated with performance on the UCDSEE, the alternate measure of emotion recognition. Similarly, performance on the SELweb theory of mind module was significantly and robustly associated with performance on Strange Stories, the alternate measure of theory of mind. Performance on the SELweb social problem-solving module was significantly and robustly associated with performance on the SIP-AP, the alternate measure of social problem solving. Performance on the SELweb self-control modules was not associated with the KiTAP, the alternate measure of self-control. Performance on SELweb emotion recognition, theory of mind, and

Table 3. SELweb Internal Consistency in the ASD and GenEd Samples

SELweb module and score	ASD vs. Mean GenEd		GenEd samples		
	ASD sample	Mean GenEd samples	Field trial 1	Field trial 2	Norming study
Emotion Recognition	0.74	0.84	0.78	0.84	0.86
Theory of Mind	0.78	0.79	0.80	0.78	0.79
Social Problem Solving	0.93	0.86	0.82	0.88	0.88
Positive Attribution	0.81	0.73	—	0.72	0.73
Positive Social Goal	0.84	0.72	0.71	0.72	0.72
Positive Solution Selection	0.88	0.82	0.81	0.83	0.82
Self-Control	0.82	0.83	0.85	0.80	0.86
Delay of Gratification	0.73	0.73	0.71	0.74	0.74
Frustration Tolerance	0.76	0.85	0.92	0.77	0.85

Cronbach's alphas are reported for all modules, with the exception of social problem-solving and self-control composite scores. Methods for these values are in accordance with a strategy outlined by Nunnally and Bernstein.

Table 4. Multitrait–Multimethod Matrix for SELweb Overall Scale and Subscales

	<i>M (SD)</i>	¹ SELweb					² Validation tasks		
		SE ₁	ER ₁	ToM ₁	SPS ₁	SC ₁	ER ₂	ToM ₂	SPS ₂
SE ₁	89.1 (14.2)								
ER ₁	98.7 (12)	0.50***							
ToM ₁	87 (16.8)	0.72***	0.08						
SPS ₁	92.2 (20.4)	0.68***	0.22	0.26*					
SC ₁	90.6 (14.3)	0.58***	0.17	0.41**	0.00				
ER ₂	60.4 (8.7)	<u>0.26</u>	<u>0.27*</u>	<u>0.08</u>	<u>0.09</u>	<i>0.27*</i>			
ToM ₂	2.4 (1.8)	<i>0.45***</i>	<i>−0.06</i>	<i>0.59***</i>	<i>0.12</i>	<i>0.46***</i>	0.26		
SPS ₂	2.2 (0.9)	<i>−0.53***</i>	<i>−0.20</i>	<i>−0.16</i>	<i>−0.81***</i>	<i>0.01</i>	−0.02	−0.16	
SC ₂	18 (5.2)	<i>−0.15</i>	<i>−0.10</i>	<i>−0.02</i>	<i>−0.20</i>	<u>−0.03</u>	0.25	−0.06	0.12

Pearson correlation coefficients between convergent and discriminant constructs and methods. The matrix is formatted to reflect patterns of correlation interpreted within a multitrait–multimethod matrix, wherein bold values denote the heterotrait–monomethod triangles, underlined values denote monotrait–heteromethod correlations (i.e., validity coefficients), and italicized values denote heterotrait–heteromethod triangles. A pattern of correlations in which validity coefficients are sufficiently high to warrant further investigation, of greater magnitude than heterotrait–heteromethod correlations in the same row or column, and of greater magnitude than heterotrait–monomethod correlations provides evidence of convergent and discriminant validity for SELweb emotion recognition, theory of mind, and social problem solving.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

SE₁, Social–Emotional/Overall SELweb standard score; ER₁, SELweb emotion recognition standard score; ToM₁, SELweb theory of mind standard score; SPS₁, SELweb social problem solving standard score; SC₁, SELweb self-control standard score; ER₂, University of California at Davis Set of Emotion Expressions (subset) Total Correct; ToM₂, Strange Stories (subset) Total Correct; SPS₂, Social Information Processing–Application (subset) Overall score; SC₂, KiTAP Go/NoGo Total Correct.

social problem-solving modules was generally not associated, or only weakly associated, with alternate measures of dissimilar constructs.

Table 5. Relationship between SELweb Domain Scores and Validation Measures

Variable	Criterion			
	UCDSEE	Strange stories	SIP-AP overall	KiTAP Go/NoGo
Age	0.36*	0.46***	−0.34**	0.06
IQ	−0.15	0.44***	0.35*	0.09
SELweb Overall	0.45**	0.36**	−0.83***	−0.18
Age	0.35*	0.36***	−0.13	0.10
IQ	0.10	0.65***	−0.12	−0.01
SELweb Emotion Recognition	0.40**	0.08	−0.25	−0.06
Age	0.23	0.48***	−0.08	0.13
IQ	0.04	0.32**	−0.02	−0.04
SELweb Theory of Mind	0.10	0.50***	−0.16	0.03
Age	0.22	0.34***	−0.16*	0.10
IQ	0.10	0.64***	0.01	0.02
SELweb Social Problem Solving	0.10	0.07	−0.84***	−0.19
Age	0.25	0.36***	−0.02	0.12
IQ	−0.09	0.53***	−0.20	−0.01
SELweb Self-Control	0.35*	0.20	0.13	−0.00

This table shows standardized regression coefficients; bold-face coefficients are associations between variables in the same SEL domain; the SIP-AP is scored such that a higher score reflects lower functioning, which accounts for the negative correlations.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

UCDSEE, University of California at Davis Set of Emotion Expressions; SIP-AP, Social Information Processing Application; KiTAP, Test of Attentional Performance for Children.

Associations with autism severity. A moderate, yet significant, correlation existed between the overall SELweb standard score and the ADOS-2 calibrated severity score ($r = -0.36$, $p = 0.01$). Further analysis examined the association between the four SELweb scores reflecting different skill areas. These analyses revealed a significant correlation between theory of mind ($r = -0.30$, $p = 0.02$) and ADOS-2 severity scores; associations between the other three SELweb domain scores and the ADOS-2 severity score were nonsignificant ($r = -0.26$ – -0.12 , $p = 0.06$ – 0.36). This pattern of associations was also observed in correlations between the ADOS-2 severity scores and performance on validation measures.

Comparison to GenEd youth. Youth with ASD scored significantly lower ($p < 0.006$) than the GenEd normative sample on the SELweb social–emotional composite score ($M_{SS} = 89$), theory of mind ($M_{SS} = 87$), social problem solving ($M_{SS} = 92$), and self-control ($M_{SS} = 91$), but not emotion recognition ($M_{SS} = 99$; $p = 0.41$).

Discussion

Overall, this study suggests that SELweb is a usable and feasible task for youth with ASD and shows promise as a tool for measuring multiple social–emotional domains in this population. Because social–emotional skills are not directly observable and rating scales require a high level of inference, SELweb’s direct assessment format provides an ideal alternative to frequently used third-party reports. The fact that the social domains SELweb assesses are implicated in both contemporary and longstanding

models of ASD also supports its use in this population (Baron-Cohen, 2000; Beauchamp & Anderson, 2010; Cai et al., 2018; Dawson et al., 2005; McMahon, Lerner, & Britton, 2013; Mendelson et al., 2016). Because challenges in emotion recognition, theory of mind, social problem solving, and self-control may be present early in development and persistent over time, being able to evaluate performance with the same measure at different ages is another benefit of SELweb. The usability of SELweb in youth with ASD is likely augmented by its portability across settings, as well as the fact that administration does not require shifting across materials and modalities.

Generally, participants were interested in and engaged with the activities within SELweb. Although some participants exhibited instances of distraction or off-task behavior, this was not observed to be of especially greater intensity or frequency than typically observed during administration of comparable tasks. The emotion recognition module was at times fatiguing for participants, but this finding is consistent with what was observed in GenEd populations.

Overall, our validation results suggest that most SELweb scores are reliable and measure what they were designed to measure in youth with ASD in a laboratory setting, and they do so in an integrated instrument that assesses multiple interrelated constructs on a straightforward, usable, and scalable platform. Contrary to expectations, the SELweb self-control module did not demonstrate expected convergent and discriminant validity. The fact that performance on the self-control modules was related to performance on the both the emotion recognition and theory of mind validation measures suggests that this subscale may be capturing a construct other than traditional impulse control. Disentangling those data is of interest and represents a future direction. A promising result from this study is the moderate and significant association between overall SELweb performance and the ADOS-2 calibrated severity score. This finding supports the contention that social-emotional skills may be related to symptom expression in youth with ASD (Beauchamp & Anderson, 2010; Lipton & Nowicki, 2009; Loveland, 2005; Rice et al., 2012; Russo-Ponsaran et al., 2015). The fact that this effect was driven largely by the psychometrically sound theory of mind subscale is consistent with current thinking that this capacity represents a vector vital to understanding and explaining heterogeneity in ASD symptomatology (Livingston & Happé, 2017; Lombardo et al., 2016; Rice et al., 2012).

The deficit levels in the various domains measured by SELweb are consistent with literature showing heterogeneity in social-emotional impairment in ASD (Lozier et al., 2014). Most of the standard scores are significantly below GenEd population norms. Such scores do *not* reflect severe levels of impairment or global impairment, but they *do* reflect a wide range of functioning ($SD_{score} = 12.02\text{--}20.42$). That performance on the emotion

recognition module was not significantly impaired is not entirely surprising because the existence, magnitude, and nature of facial emotional recognition deficits in ASD is mixed and has not been fully established (e.g., Klin et al., 1999; Tracy, Robins, Schriber, & Solomon, 2011). Recent work has suggested that emotional deficits in ASD may be attributable to alexithymia, rather than ASD per se (e.g., Bird & Cook, 2013). Although the underlying mechanisms of these deficits require further exploration, the fact that SELweb was sensitive to deficits in ASD is important. Future studies should include measurement of alexithymia as a means of understanding variability in emotion recognition deficits.

Although exploring the specific phenotypic heterogeneity present in the social-emotional skill profiles of individuals with ASD (Kim et al., 2018) is outside the scope of this initial pilot study, informal analyses did reveal several distinct patterns of ability without demonstrating the sort of floor effects, or range restriction, often seen in other measures that are also usable in non-ASD populations. For example, some domain scores within individuals varied by more than two SDs from each other—with strengths in certain domains and challenges in others. Others scored high or low on all domains, and others demonstrated an area of strength or deficit in all *but one* domain. Larger studies should evaluate these profiles using latent class analyses.

Often, scientists and practitioners try to understand the phenotypes in ASD by piecing together findings in disparate heterogeneous samples across multiple studies. SELweb offers a unique opportunity to describe complex social-emotional skill profiles—both of clinically distinct subgroups of youth with ASD (Lerner, De Los Reyes, Drabick, Gerber, & Gadow, 2017) and of individuals—with a single assessment tool. SELweb offers the opportunity to potentially better characterize, and ultimately understand, the subgroups of the ASD population.

This study is a first step in understanding the utility of SELweb for assessment in verbal youth with ASD who have average intellect. Although the present findings support the feasibility of SELweb for this population, future work should examine feasibility in larger samples and in youth with ASD with a wider range of cognitive and attentional capabilities. It should also be noted that language level was not explicitly tested as part of the eligibility criteria or experimental battery. Receptive language-level would be an important variable to consider for future testing and to establish whether a language-level cutoff should be implemented, over and above an IQ cutoff. Although we consider SELweb as a tool that reflects a wide range of skill and impairment levels, we recognize that it may not be appropriate for youth with ASD who, for example, have co-occurring severe intellectual, social, or receptive language challenges or are unable to sit and attend during a structured

computerized activity. Another important consideration is that SELweb does not capture social-emotional abilities via interactive activities or in real time (Mundy, 2018). Although outside the scope of the present study, validation of SELweb with behavioral observations may contribute to our understanding of the utility of SELweb as a measurement tool. From a design perspective, developers of future assessments should explore ways to better capture the ecological validity in interactive ways. Another future direction to consider is exploring the sensitivity of SELweb to behavioral or pharmacological interventions in ASD and other populations.

Dissemination of tools for the assessment of social-emotional skills in ASD has been hampered by administrative and interpretive requirements of most contemporary instruments. Because SELweb provides a total score and modular scores, it increases the capacity to generate individualized performance profiles within ASD, providing a rich picture of the autism spectrum. Thus, this study sheds light on the presence of phenotypic heterogeneity in social-emotional skill profiles in this population. Once further validated in larger samples, SELweb may provide a platform to standardize performance on social-emotional assessments across research, educational, and clinical settings, effectively bridging the research-practice assessment gap.

ACKNOWLEDGMENTS

This work was supported by a grant from the Simons Foundation (SFARI# 381283, MDL, NRP), as well as by the Institute of Education Sciences through Grant R305A110143 to CM. The opinions expressed are those of the authors and do not represent views of the funding agencies. We would like to thank the families who participated in this study. We would also like to thank Ms. Irene Wu who assisted with the preparation of this manuscript and Mr. Jason Johnson (deceased) who contributed to the early development of this project. Drs. Russo-Ponsaran and McKown have financial interests in/relative to xSEL-Labs, Inc., which could potentially benefit from the outcomes of this research.

REFERENCES

- Barnard-Brak, L., Ivey-Hatz, J., Ward, A. K., & Wei, T. (2014). Self-regulation and social interaction skills among children with autism across time. *Advances in Mental Health and Intellectual Disabilities*, 8(4), 271–279. <https://doi.org/10.1108/AMHID-12-2012-0007>
- Baron-Cohen, S. (2000). Theory of mind and autism: A fifteen year review. In S. Baron-Cohen, H. Tager-Flusberg, D. J. Cohen, S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*. 3–20). New York: Oxford University Press.
- Baron-Cohen, S. (2001). Theory of mind in normal development and autism. *Prisme*, 34(1), 74–183.
- Beauchamp, M. H., & Anderson, V. (2010). SOCIAL: An integrative framework for the development of social skills. *Psychological Bulletin*, 136(1), 39–64. <https://doi.org/10.1037/a0017768>
- Bellini, S., & Hopf, A. (2007). The development of the autism social skills profile: A preliminary analysis of psychometric properties. *Focus on Autism and Other Developmental Disabilities*, 22(2), 80–87. <https://doi.org/10.1177/10883576070220020801>
- Bird, G., & Cook, R. (2013). Mixed emotions: The contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry*, 3, e285. <https://doi.org/10.1038/tp.2013.61>
- Cai, R. Y., Richdale, A. L., Uljarevic, M., Dissanayake, C., & Samson, A. C. (2018). Emotion regulation in autism spectrum disorder: Where we are and where we need to go. *Autism Research*, 11(7), 962–978. <https://doi.org/10.1002/aur.1968>
- Canty, A. L., Neumann, D. L., Fleming, J., & Shum, D. H. (2017). Evaluation of a newly developed measure of theory of mind: The virtual assessment of mentalising ability. *Neuropsychological Rehabilitation*, 27(5), 834–870. <https://doi.org/10.1080/09602011.2015.1052820>
- Channon, S., Crawford, S., Orlowska, D., Parikh, N., & Thoma, P. (2014). Mentalising and social problem solving in adults with Asperger's syndrome. *Cognitive Neuropsychiatry*, 19(2), 149–163. <https://doi.org/10.1080/13546805.2013.809659>
- Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale, second edition (SRS-2)*. Torrance, CA: Western Psychological Services.
- Dadds, M., Hawes, D., & Merz, S. (2004). *The UNSW Facial Emotion Task*. Sydney: Department of Psychology, University of New South Wales.
- Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Developmental Neuropsychology*, 27(3), 403–424. https://doi.org/10.1207/s15326942dn2703_6
- De Los Reyes, A., Aldao, A., Thomas, S. A., Daruwala, S., Swan, A. J., Van Wie, M., ... Lechner, W. V. (2012). Adolescent self-reports of social anxiety: Can they disagree with objective psychophysiological measures and still be valid? *Journal of Psychopathology and Behavioral Assessment*, 34(3), 308–322. <https://doi.org/10.1007/s10862-012-9289-2>
- Demopoulos, C., Hopkins, J., & Davis, A. (2013). A comparison of social cognitive profiles in children with autism spectrum disorders and attention-deficit/hyperactivity disorder: A matter of quantitative but not qualitative difference? *Journal of Autism and Developmental Disorders*, 43(5), 1157–1170. <https://doi.org/10.1007/s10803-012-1657-y>
- Denham, S. A., Bassett, H., Mincic, M., Kalb, S., Way, E., Wyatt, T., & Segal, Y. (2012). Social-emotional learning profiles of preschoolers' early school success: A person-centered approach. *Learning and Individual Differences*, 22(2), 178–189. <https://doi.org/10.1016/j.lindif.2011.05.001>
- Didehbani, N., Allen, T., Kandalaf, M., Krawczyk, D., & Chapman, S. (2016). Virtual reality social cognition training for children with high functioning autism. *Computers in Human Behavior*, 62, 703–711. <https://doi.org/10.1016/j.chb.2016.04.033>

- Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., ... Convit, A. (2006). Introducing MASC: A movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders*, 36(5), 623–636. <https://doi.org/10.1007/s10803-006-0107-0>
- Epkins, C. C. (1994). Peer ratings of depression, anxiety, and aggression in inpatient and elementary school children: Rating biases and influence of rater's self-reported depression, anxiety, and aggression. *Journal of Abnormal Child Psychology*, 22(5), 611–628. <https://doi.org/10.1007/bf02168941>
- Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The 'reading the mind in films' task [child version]: Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 38(8), 1534–1541. <https://doi.org/10.1007/s10803-007-0533-7>
- Gresham, F., & Elliott, S. N. (2008). *Social skills improvement system: Rating scales*. Bloomington, MN: Pearson Assessments.
- Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, 24(2), 129–154. <https://doi.org/10.1007/BF02172093>
- Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, 66(3), 843–855.
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 42(2), 377–381.
- Hepburn, S. L., & Wolff, B. C. (2013). Self-regulation in persons with autism spectrum disorders. In K. C. Barrett, N. A. Fox, G. A. Morgan, D. J. Fidler, & L. A. Daunhauer (Eds.), *Handbook of Self-Regulatory Processes in Development: New Directions and International Perspectives* (pp. 429–452). New York, NY, US: Psychology Press.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test—Second Edition Manual*. Circle Pines, MN: American Guidance Service.
- Kim, H., Keifer, C. M., Rodriguez-Seijas, C., Eaton, N. R., Lerner, M. D., & Gadow, K. D. (2018). Quantifying the optimal structure of the autism phenotype: A comprehensive comparison of dimensional, categorical, and hybrid models. *Journal of the American Academy of Child & Adolescent Psychiatry* 10.1016/j.jaac.2018.09.431
- Klin, A. (2000). Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: The social attribution task. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41(7), 831–846. <https://doi.org/10.1111/1469-7610.00671>
- Klin, A., Sparrow, S. S., De Bildt, A., Cicchetti, D. V., Cohen, D. J., & Volkmar, F. R. (1999). A normed study of face recognition in autism and related disorders. *Journal of Autism and Developmental Disorders*, 29(6), 499–508.
- Kupersmidt, J. B., Stelter, R., & Dodge, K. A. (2011). Development and validation of the social information processing application: A web-based measure of social information processing patterns in elementary school-age boys. *Psychological Assessment*, 23(4), 834–847. <https://doi.org/10.1037/a0023621>
- Lerner, M. D., De Los Reyes, A., Drabick, D. A. G., Gerber, A. H., & Gadow, K. D. (2017). Informant discrepancy defines discrete, clinically useful autism spectrum disorder subgroups. *Journal of Child Psychology and Psychiatry*, 58(7), 829–839. <https://doi.org/10.1111/jcpp.12730>
- Lipton, M., & Nowicki, S. (2009). The social emotional learning framework (SELF): A guide for understanding brain-based social emotional learning impairments. *Journal of Developmental Processes*, 4(2), 99–115.
- Livingston, L. A., Colvert, E., Bolton, P., & Happé, F. (2018). Good social skills despite poor theory of mind: Exploring compensation in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 60, 102–110. <https://doi.org/10.1111/jcpp.12886>
- Livingston, L. A., & Happé, F. (2017). Conceptualising compensation in neurodevelopmental disorders: Reflections from autism spectrum disorder. *Neuroscience & Biobehavioral Reviews*, 80, 729–742. <https://doi.org/10.1016/j.neubiorev.2017.06.005>
- Lombardo, M. V., Lai, M.-C., Auyeung, B., Holt, R. J., Allison, C., Smith, P., ... Baron-Cohen, S. (2016). Unsupervised data-driven stratification of mentalizing heterogeneity in autism. *Scientific Reports*, 6, 35333. <https://doi.org/10.1038/srep35333>
- Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule Second Edition (ADOS-2) Manual (Part 1): Modules 1–4*. Torrance, CA: Western Psychological Services.
- Loveland, K. A. (2005). Social-emotional impairment and self-regulation in autism spectrum disorders. In J. Nadel & D. Muir (Eds.), *Emotional Development: Recent Research Advances* (pp. 365–382). New York: Oxford University Press.
- Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Developmental Psychopathology*, 26(4 Pt 1), 933–945. <https://doi.org/10.1017/S0954579414000479>
- Mazefsky, C. A., Herrington, J., Siegel, M., Scarpa, A., Maddox, B. B., Scahill, L., & White, S. W. (2013). The role of emotion regulation in autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(7), 679–688. doi:10.1016/j.jaac.2013.05.006
- McKown, C. (2018). Reliability, factor structure, and measurement invariance of a web-based assessment of children's social-emotional comprehension. *Journal of Psychoeducational Assessment*, 073428291774968. <https://doi.org/10.1177/0734282917749682>
- McKown, C., Allen, A. M., Russo-Ponsaran, N. M., & Johnson, J. K. (2013). Direct assessment of children's social-emotional comprehension. *Psychological Assessment*, 25(4), 1154–1166. <https://doi.org/10.1037/a0033435>
- McKown, C., Gumbiner, L. M., Russo-Ponsaran, N. M., & Lipton, M. (2009). Social-emotional learning skill, self-regulation, and social competence in typically developing and clinic-referred children. *Journal of Clinical Child & Adolescent Psychology*, 38(6), 858–871.
- McKown, C., Russo-Ponsaran, N. M., Johnson, J. K., Russo, J., & Allen, A. (2016). Web-based assessment of children's social-emotional comprehension. *Journal of Psychoeducational Assessment*, 34(4), 322–338. <https://doi.org/10.1177/0734282915604564>

- McMahon, C. M., Lerner, M. D., & Britton, N. (2013). Group-based social skills interventions for adolescents with higher-functioning autism spectrum disorder: a review and looking to the future. *Adolescent Health, Medicine and Therapeutics*, 4, 23–38. <https://doi.org/10.2147/AHMT.S25402>
- Mendelson, J. L., Gates, J. A., & Lerner, M. D. (2016). Friendship in school-age boys with autism spectrum disorders: A meta-analytic summary and developmental, process-based model. *Psychological Bulletin*, 142(6), 601–622.
- Mundy, P. (2018). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*, 47(6), 497–514. <https://doi.org/10.1111/ejn.13720>
- Murphy, K. R., Martin, C., & Garcia, M. (1982). Do behavioral observation scales measure observation? *Journal of Applied Psychology*, 67(5), 562–567. <https://doi.org/10.1037//0021-9010.67.5.562>
- Nelson, B. D., Perlman, G., Hajcak, G., Klein, D. N., & Kotov, R. (2015). Familial risk for distress and fear disorders and emotional reactivity in adolescence: An event-related potential investigation. *Psychological Medicine*, 45(12), 2545–2556.
- Norris, M., & Lecavalier, L. (2010). Screening accuracy of level 2 autism spectrum disorder rating scales: a review of selected instruments. *Autism*, 14(4), 263–284. <https://doi.org/10.1177/1362361309348071>
- Nunnally, J. C., & Bernstein, I. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Patriquin, M. A., Lorenzi, J., Scarpa, A., Calkins, S. D., & Bell, M. A. (2015). Broad implications for respiratory sinus arrhythmia development: Associations with childhood symptoms of psychopathology in a community sample. *Developmental Psychobiology*, 57(1), 120–130.
- Rice, K., Moriuchi, J. M., Jones, W., & Klin, A. (2012). Parsing heterogeneity in autism spectrum disorders: Visual scanning of dynamic social scenes in school-aged children. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(3), 238–248. doi:10.1016/j.jaac.2011.12.017
- Russo-Ponsaran, N. M., McKown, C., Johnson, J., Russo, J., Crossman, J., & Reife, I. (2018). Virtual environment for social information processing: Assessment of children with and without autism spectrum disorders. *Autism Research*, 11(2), 305–317. <https://doi.org/10.1002/aur.1889>
- Russo-Ponsaran, N. M., McKown, C., Johnson, J. K., Allen, A. W., Evans-Smith, B., & Fogg, L. (2015). Social-emotional correlates of early stage social information processing skills in children with and without autism spectrum disorder. *Autism Research*, 8(5), 486–496. <https://doi.org/10.1002/aur.1463>
- Sasson, N. J., Nowlin, R. B., & Pinkham, A. E. (2012). Social cognition, social skill, and the broad autism phenotype. *Autism*, 17(6), 655–667. <https://doi.org/10.1177/1362361312455704>
- Shulman, C., Guberman, A., Shiling, N., & Bauminger, N. (2012). Moral and social reasoning in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 42(7), 1364–1376. <https://doi.org/10.1007/s10803-011-1369-8>
- Sivaratnam, C. S., Cornish, K., Gray, K. M., Howlin, P., & Rinehart, N. J. (2012). Brief report: Assessment of the social-emotional profile in children with autism spectrum disorders using a novel comic strip task. *Journal of Autism and Developmental Disorders*, 42(11), 2505–2512. <https://doi.org/10.1007/s10803-012-1498-8>
- Stichter, J. P., Herzog, M. J., Visovsky, K., Schmidt, C., Randolph, J., Schultz, T., & Gage, N. (2010). Social competence intervention for youth with asperger syndrome and high-functioning autism: An initial investigation. *Journal of Autism and Developmental Disorders*, 40(9), 1067–1079. <https://doi.org/10.1007/s10803-010-0959-1>
- Tanaka, J. W., Wolf, J. M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., ... McPartland, J. C. (2012). The perception and identification of facial emotions in individuals with autism spectrum disorders using the Let's Face It! Emotion Skills Battery. *Journal of Child Psychology and Psychiatry*, 53(12), 1259–1267.
- Tracy, J. L., Robins, R. W., & Schriber, R. A. (2009). Development of a FACS-verified set of basic and self-conscious emotion expressions. *Emotion*, 9(4), 554–559. <https://doi.org/10.1037/a0015766>
- Tracy, J. L., Robins, R. W., Schriber, R. A., & Solomon, M. (2011). Is emotion recognition impaired in individuals with autism spectrum disorders? *Journal of Autism and Developmental Disorders*, 41(1), 102–109. <https://doi.org/10.1007/s10803-010-1030-y>
- van der Valk, J. C., van den Oord, E. J. C. G., Verhulst, F. C., & Boomsma, D. I. (2001). Using parental ratings to study the etiology of 3-year-old twins' problem behaviors: Different views or rater bias? *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(7), 921–931. <https://doi.org/10.1017/S0021963001007648>
- Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence—Fourth Edition*. San Antonio, TX: Psychological Corporation.
- Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy & Practice*, 10(3), 329–345. <https://doi.org/10.1080/0969594032000148172>
- Yager, J. A., & Ehmann, T. S. (2006). Untangling social function and social cognition: A review of concepts and measurement. *Psychiatry: Interpersonal and Biological Processes*, 69(1), 47–68. <https://doi.org/10.1521/psyc.2006.69.1.47>
- Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin*, 124(3), 283–307.
- Zimmermann, P., Gondan, M., & Fimm, B. (2005). *Test of attention performance for children (KITAP)*. Würselen, Germany: Psytest.