

Measuring Social Motivation in Autism Spectrum Disorder: Development of the Social  
Motivation Interview

Rebecca Elias, M.S.

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State  
University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy  
In  
Psychology

Susan W. White, Ph.D., Committee Chair  
Russell T. Jones, Ph.D.  
Jungmeen Kim-Spoon, Ph.D.  
Richard A. Winnett, Ph.D.

March 29, 2019  
Blacksburg, VA

Keywords: Autism spectrum disorder, social motivation, assessment, instrument  
development

Copyright 2019

# Measuring Social Motivation in Autism Spectrum Disorder: Development of the Social Motivation Interview

Rebecca Elias, M.S.

## ABSTRACT

Social motivation in individuals with ASD is currently derived from the observation of overt behaviors and neurological correlates, from which motivational processes are inferred. Motivation, however, is not the same as behavior, and most theoretical conceptualization of the construct assign primary importance to cognitive processes. Nevertheless, few studies have examined the cognitive processes that may influence goal-directed tasks involved in social interaction. Understanding internalized cognitive processes may distinguish underlying motivations which influence engagement in social behavior. This study aims to assess how beliefs, mindsets, and attitudes can influence one's motivation to engage in social interaction through the development of a novel interview. The Social Motivation Interview (SMI) assesses for internal cognitions as they relate to social motivation by determining levels of social desire, interest, and behaviors in children with ASD. The development of the SMI followed stringent criteria to create a unified measure that was methodologically sound and theoretically informed. SMI development followed guidelines to ensure item pool development was consistent with the proposed construct. Pilot testing suggested feasibility of administration, user satisfaction, and promising psychometric properties. Future examination of the SMI in large-scale field testing is warranted.

# Measuring Social Motivation in Autism Spectrum Disorder: Development of the Social Motivation Interview

Rebecca Elias, M.S.

## GENERAL AUDIENCE ABSTRACT

This study aims to assess how beliefs, mindsets, and attitudes can influence one's motivation to engage in social interaction through the development of a novel interview. The Social Motivation Interview (SMI) assesses for internal cognitions as they relate to social motivation by determining levels of social desire, interest, and behaviors in children with ASD. The development of the SMI followed stringent criteria to create a unified measure that was methodologically sound and theoretically informed. SMI development followed guidelines to ensure item pool development was consistent with the proposed construct. Pilot testing suggested feasibility of administration, user satisfaction, and promising psychometric properties. Future examination of the SMI in large-scale field testing is warranted.

## **Acknowledgements**

This work was supported by grants from the American Psychological Association (APA), PI: Elias and The Virginia Tech Center for Autism Research (VTCAR), PI: Elias. We gratefully thank the children and families who participated in this study. The expert review panel included Nirit Bauminger, Ph.D., Kirstin Greaves-Lord, Ph.D., Natasha Marrus, M.D., Ph.D., Carla Mazefsky, Ph.D., Thomas Ollendick, Ph.D., and Kate Sofronoff, Ph.D., who we appreciatively acknowledge for their role in providing feedback to the measure. I would also like to express my sincerest gratitude to my dissertation committee members including Russell T. Jones, Ph.D., Jungmeen Kim-Spoon, Ph.D., and Richard Winett, Ph.D. for their help and guidance in implementing this study. Lastly, I would like to express deep appreciation and admiration for my mentor, advisor, and committee chair, Susan W. White, Ph.D. She has served as an invaluable resource who has supported and guided me in becoming a competent and thoughtful researcher and clinician. Her feedback and training have been invaluable to my graduate school training.

## Table of Contents

Chapter 1 – Introduction .....	1
Motivation in ASD.....	1
Social Motivation: Where to from here? .....	7
Chapter 2 – Phase 1 Methods.....	10
Measures .....	12
Data Analysis .....	14
Participants.....	15
Chapter 3 – Phase 1 Results.....	17
Chapter 4 – Phase 2 Methods.....	18
Participants.....	18
Procedure .....	20
Measures .....	23
Data Analysis.....	27
Chapter 5 – Phase 2 Results.....	32
Chapter 6 – Discussion .....	37
Chapter 7 – References .....	46
Tables.....	60
Figures.....	72
Appendices.....	76

## List of Tables

Table 1: Expert Panel Content Validity Ratings.....	60
Table 2: Demographic characteristics of participants.....	63
Table 3: Descriptive Measure Data.....	64
Table 4: Inter-Item Correlation Matrix.....	65
Table 5: Item Analysis of the SMI.....	66
Table 6: Bivariate correlations among study measures. ....	70
Table 7: Correlations between the SMI and other validated measures.....	71

## List of Figures

Figure 1 .....	72
Figure 2 .....	73
Figure 3 .....	74
Figure 4 .....	75

## List of Appendices

Appendix 1.....	76
Appendix 2.....	77



## Chapter 1 – Introduction

### Motivation in ASD

Motivation is broadly regarded as goal-directed activity which is initiated, energized, and sustained (Schunk, Pintrich, & Meece, 2008). Motivation, under this conceptualization, has shaped the field of psychology and continues to be a primary focus of research (Ryan & Deci, 2000). Early conceptualizations (i.e., the first half of the 20<sup>th</sup> century) viewed motivation as solely a stimulus linked to primary or secondary reinforcement and reward (Deci & Moller, 2005; Freud, 1915; Hull, 1943; Schunk & Usher, 2012). Competing contemporary theories of motivation, largely in the educational psychology literature, center around ‘competence’ motivation, or the idea that the thoughts, feelings, and emotions can also influence goal-directed tasks which are initiated and sustained (Deci & Moller, 2005; Schunk & Usher, 2012).

Many specific facets within the broader framework of motivation have been examined throughout development, both in typically developing individuals and populations with psychopathology. One such construct is social motivation. Motivation to interact, and form relationships, with other people, has been examined as a unique construct given the evolutionary benefits of acting in collaboration (e.g., hunting, foraging for food, and sharing of knowledge; Kaplan, Hooper, & Gurven, 2009).

Despite demonstrated utility of examining social motivation, formal definitions of the construct of ‘social motivation’ are largely absent from the peer-reviewed literature base. Instead, broad definitions of motivation have traditionally been applied to social situations and/or stimuli. Studies of social motivation have lacked theoretical unity, consisting of varied emphasis, scope, and conceptual framework (Geen, 1991). Despite heterogeneous conceptual accounts of social motivation, this ambiguous construct has been widely applied to the study of

children with Autism Spectrum Disorder (ASD) (Chevallier, Kohls, Troiani, Brodtkin, & Schultz, 2012).

Diagnosis of ASD necessitates pervasive deficits in social communication and interaction (American Psychiatric Association, 2013). Etiological and mechanistic reasons for impaired social competence in ASD have long been hypothesized. One such hypothesis postulates that diminished social motivation, thought to be characteristic of ASD, reduces social opportunity and experience, thereby lowering overall social competence over time (Chevallier et al., 2012). The social motivation theory of ASD (Chevallier et al., 2012) integrates evidence from the behavioral, biological, and evolutionary domains to assert that social motivation is diminished in those with ASD. Within this framework, behavioral markers of social motivation include: 1) orientation to social stimuli, 2) use of interpersonal maintenance strategies, and 3) response to social rewards. Such reliance on behavior, though understandable, has obfuscated a more holistic conceptualization of social motivation in individuals with ASD.

The construct of social motivation is, at the present moment, most comprehensively outlined by two related theories of social motivation (Chevallier et al., 2012; Dawson, Webb, & McPartland, 2005). Specifically, these theories assert that there are three core behaviors that form our conceptualization of social motivation in ASD: 1) social attention, 2) response to reward, and 3) overt behavior strategies (e.g., approach and maintenance). Social attention has been a term broadly used to focus on attention to social stimuli either overtly or covertly produced by the individual (Salley & Colombo, 2015). It has been used as a proxy for understanding social motivation, largely because orienting to stimuli is thought to reflect social intent and needs/wants. Additionally, orientation towards socially relevant information is utilitarian for functioning in society at large. For example, social stimuli such as emotional

facial or body expressions may serve to communicate needs, desires, or possible threat (Miskovic & Keil, 2013). These stimuli are biologically relevant and salient to typically developing individuals, beginning early in infancy (Fletcher-Watson, Findlay, Leekam, & Benson, 2008).

Dysfunction in decoding and orienting to static and dynamic social stimuli is regarded as a core feature of ASD and has been robustly demonstrated among infants who later develop ASD, as well as in children who have received a formal diagnosis (e.g., Dawson et al., 2004; Hutman, Chela, Gillespie-Lynch, & Sigman, 2012; Jones & Klin, 2013; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Salley & Colombo, 2015). First seen at 6 months of age, perceptual atypicalities (e.g., viewing patterns) of social stimuli are apparent in infants with a later confirmed ASD diagnosis, versus typically developing individuals (Chawarska, Macari, & Shic, 2013; Shic, Macari, & Chawarska, 2014). One such abnormal viewing pattern, gaze to the eyes, appears to decline between 2 to 6 months of age (Jones & Klin, 2013), suggesting that social attention may begin at normative levels, but decline in specific domains as the infant with ASD develops (Jones & Klin, 2013). In childhood, individuals with ASD show a preference for non-social objects over human faces in naturalistic and laboratory-based designs (Klin et al., 2002; Nakano et al., 2010). This preference for non-social stimuli in children extends to the auditory domain as well (Kuhl, Coffey-Corina, Padden, & Dawson, 2005). Children with ASD also show reduced social orienting, joint attention (Jones & Carr, 2004), and attention to social distress (Dawson et al., 2004) compared to typically developing age-matched groups. These studies of reduced attention to socially relevant stimuli employ eye-tracking (e.g., Klin et al., 2002), measures of psychophysiology (Key & Corbett, 2014; Kuhl et al., 2005), and examination of

behavioral eye-gaze patterns and behavioral orientation (e.g., Dawson et al., 2004; Oruc, Shafai, & Iarocci, 2018).

Reduced value, or reward salience, given to social stimuli has been hypothesized to be the product of broad impairments in social attention (Chevallier et al., 2012; Dawson, Bernier, & Ring, 2012; Bottini, 2018; Dawson et al., 2005; Grelotti, Gauthier, & Schultz, 2002). That is, the desire to obtain social rewards and avoid social punishment is related to the presence or lack of orientation to social stimuli (Chevallier et al., 2012). The social motivation hypothesis proposed by Geraldine Dawson suggests that diminished attention to social stimuli in ASD produces blunted neural appraisal of rewards (Dawson et al., 2005). The presence or absence of reward has been measured primarily through measures of psychophysiology and behavior, including electroencephalography (EEG), functional magnetic resonance imaging (fMRI), go/no go tasks, and pupillometry.

In typically developing individuals, neurobiological research suggests that the ventral striatum and nucleus accumbens are implicated in both craving social incentives and avoiding punishment (Kohls et al., 2013). Additionally, neural sensitivity within reward-encoding regions toward the motivational salience of social stimuli is associated with social proficiency in typically developing adults (Gossen et al., 2014). Given the functional importance of neural reward pathways in normative social functioning in typically developing individuals, these processes have been explored in individuals with ASD (c.f., Clements et al., 2018). There is a robust finding of diminished neural responsiveness to social reward across development in ASD (Demurie, Roeyers, Baeyens, & Sonuga-Barke, 2011; Richey et al., 2014; Scott-Van Zeeland, Dapretto, Ghahremani, Poldrack, & Bookheimer, 2010). This neural dysfunction in the domain of social rewards that is associated with ASD, is apparent across visual and auditory modalities

(Abrams et al., 2013) and appears to be somewhat distinct from other diagnostic categories (Kohls et al., 2014; Richey et al., 2014)

Results from studies arguing for a pervasive domain-general altered activation within brain regions implicated in reward processing appear mixed, as there is incongruence in studies examining salience for monetary reward over social reward across varied measurement techniques. Studies of children, adolescents, and young adults show similar psychophysiological patterns (Feedback Related Negativity and P3) to those with ASD and typically developing controls engaged in tasks of monetary gain and loss (Larson, South, Krauskopf, Clawson, & Crowley, 2011). This finding suggests that responsiveness to reward is intact in paradigms involving non-social stimuli (i.e. money) when measured by event related potential. Although some research has shown a deficit in reward responsiveness to social stimuli specifically, other research has suggested a more global deficiency in reward as a construct (Richey et al., 2014). In typically developing individuals, facial expressions of happiness in others are intrinsically rewarding in a variety of contexts. Perceived reward is often operationalized via increased pupillary diameter in response to viewing happy facial expressions paired with direct eye gaze (Sepeta et al., 2012). When viewing happy versus other basic facial emotions, children and adolescents with ASD exhibited no significant pupillary diameter modulation (Sepeta et al., 2012). Taken together, dysfunction in reward processing has been demonstrated across multiple levels of analysis in those with ASD. However, the extent that these abnormalities are specific to social reward dysfunction remains unclear.

Overt behavioral strategies of social interaction have been used to infer that motivation underlies one's ability to approach or maintain a social situation/relationship. Engagement in maintenance strategies is implicated in the preservation of social relationships. Flattery,

reputation management, and maintenance of self-image are all components integral to sustaining lasting professional, personal, and romantic relationships with other peers (Chevallier, et al., 2012). Children and adults with ASD display diminished use of social maintenance strategies in a variety of contexts (Barbaro & Dissanayake, 2007; Hobson & Lee, 1998; Izuma, Matsumoto, Camerer, & Adolphs, 2011). For example, children with ASD engage in less non-verbal greetings such as a hand wave, display less laughter during verbal exchange, and engage in less feigned-flattery (Chevallier, Molesworth, & Happé, 2012; Hudenko, Stone, & Bachorowski, 2009). Differing mechanistic explanations for diminished social reputation strategies have been offered, including impaired diminished theory of mind (Cage, Bird, & Pellicano, 2016; Izuma et al., 2011). Alternate explanations including the diminished expectation of reduced reciprocal behavior from others (Cage, Pellicano, Shah, & Bird, 2013) imply that individuals with ASD can employ strong reputation management when it is beneficial to think about the opinion of others.

In addition to behavioral maintenance strategies, social motivation has been indexed by the amount of behavioral approach. When engaged in a task of free-play, 38 children aged 8-12 years of age with a diagnosis of ASD showed higher levels of cortisol in response to social engagement with peers (Corbett et al., 2014). Higher cortisol levels in the ASD group corresponded to less social motivation, as operationalized by fewer verbal interactions/initiations and increased self-play. Other behavioral measures used to derive social motivation include measures of social persistence (Garman et al., 2016). Behavioral approach (toward a target) appears to be used as an indicator of underlying motivation in those with ASD. However, there has yet to be evidence indicating a link between social engagement strategies and underlying desire in individuals with ASD across the lifespan.

Despite a lack of well-validated parent/self-report measures of social motivation in ASD, the Social Responsiveness Scale-2 (SRS-2; Constantino & Gruber, 2012) has commonly been used as a metric of social motivation in children with ASD (e.g., Factor, Condy, Farley, & Scarpa, 2016). The SRS-2 is a clinical assessment tool designed to assess social impairment in children and adults suspected of having ASD. The measure yields 5 distinct subscales, one of which is social motivation. The 10-item social motivation subscale relies only on the measurement of behavioral approach and maintenance to derive dimensional levels of social motivation. Using outwardly expressed behaviors, in the absence of consideration of cognitive processes, to index social motivation, has been criticized (e.g., Capriola, Maddox, & White, 2016; Swain, Scarpa, White, & Laugeson, 2015). Specifically, this commonly used subscale of the SRS-2 fails to integrate internal cognitive attributions (i.e., social interest, desire, anxiety) of an individual, and therefore, may misattribute origins of behavioral avoidance. Thus, the widely-used existing measures of social motivation in ASD do not probe clinically-relevant cognitive processes, and as such, are insufficient given the complexity of the construct.

### **Social Motivation: Where to from here?**

The current measurement techniques used to measure social motivation in those with ASD have primarily relied on overt behavioral actions to infer motivation. Specifically, social attention through orientation of eye-gaze is a valuable tool to understand salience within context. Additionally, this methodological consideration can be applied to individuals with a variety of cognitive abilities. The social motivation hypothesis (Chevallier et al., 2012; Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998; Dawson et al., 2004) asserts that through time, this decreased attentional salience for social stimuli beginning in infancy results in diminished neural specialization to social engagement. However, we contend that after the emergence of

observable social difficulty, there is a need to integrate cognitive factors related to the perception of the self in the conceptualization of social motivation. Additionally, there logically appears to be a developmental factor which may be associated with possible dynamic changes across the lifespan or within context. However, this approach appears most appropriate for high functioning individuals with ASD, as verbal abilities are typically implicated in the expression of internal perception of the self.

Distilling social motivation to the study of social attention, social reward, and behavioral engagement and maintenance poses difficulty for our understanding of how the field of psychology conceptualizes what social motivation is. Specifically, behavioral and psychophysiological correlates are being used as proxies to infer motivation, which has been particularly advantageous in assessing all individuals with ASD, irrespective of cognition level. However, the current state of measurement has treated social motivation as a static, intrinsic deficit pervasive across all individuals with ASD. That is, diminished social motivation in ASD has been conceptualized as a trait marker (or endophenotype), representing the biological and behavioral indicators that influence the emergence and maintenance of the diagnostic symptomatology of the disorder. The field has neglected to consider the role of development as a dynamic process that may alter motivational processes implicated in social engagement and approach. Consequently, the competence view of motivation (Deci & Moller, 2005; Schunk & Usher, 2012) has yet to be applied to the construct of social motivation in ASD.

This is noteworthy as recent studies have begun to challenge the assumption that those with ASD exhibit inherent deficits in social motivation. Jaswal and Akhtar (2018) proposed that future research examining social motivation in ASD specifically target the ways in which individuals with ASD may show their desire to engage with others, as this may be



unconventional in nature. Indeed, the authors assert that, “Research is also urgently needed to identify and characterize the range of behaviors that can signal social interest” (Jaswal & Akhtar, 2018, p.44). For the purposes of this study, we define social motivation as interest or desire to engage in social situations.

The primary aim of the present study was to create and evaluate a novel measure of social motivation for individuals with ASD - the Social Motivation Interview (SMI). The SMI is intended to be a theoretically grounded tool applicable for communicative individuals with ASD aged 8-17. Additionally, the SMI is designed to tap the internal cognitive processes that influence an individual’s expressed social behavior. As depicted in Figure 1, during Phase 1 we created an interview item pool based on expert consultation and statistical item refinement. We hypothesized that the item-level content validity index (I-CVI) across each individual question would be at least 0.78 (Polit & Beck, 2006; Polit, Beck, Owen, 2007).

Phase 2 of the study involved pilot testing to establish psychometric properties of the measure. To establish preliminary construct validity, we examined individual item performance, internal consistency, and associations between derived SMI scores and other indices of social motivation using different modalities (convergent validity), as well as theoretically unrelated variables (discriminant validity). We hypothesized that convergent validity would be established by high and significant correlations. It was hypothesized that concurrent validity would be demonstrated by moderate and significant correlations. Lastly, discriminant validity was hypothesized to be characterized by correlations which were not significant. The following exploratory hypothesis was proposed: interview parental participant involvement would be associated with child participant age and ASD severity level.

Lastly, to determine feasibility and acceptability, we gathered stakeholder satisfaction feedback and examined consensus among multiple raters (inter-rater reliability). We hypothesized that trained observers would code the interviews with a high level of inter-rater agreement ( $Kappa > 0.80$ ). Additionally, we hypothesized that at least 85% of caregivers would report ‘high’ or ‘very high’ acceptability ratings.

### **Phase 1: Creation of the SMI**

#### **Chapter 2 – Phase 1 Methods**

Approval for both phases of this study was granted by the Virginia Tech Institutional Review Board. A systematic literature search was executed to determine the presence or absence of need for a measure specifically tapping social motivation in youth with ASD. It was expected that an understanding of the limitations of practitioners’ ability to assess social motivation would make the resulting measure more relevant and, ultimately, more likely to be adopted in practice. The systematic search was conducted to identify all existing instruments used to assess social motivation (or related constructs) in child or caregiver formats. This search spanned a variety of research foci including psychopathology, education, personality, and severe mental illness.

Results of the literature search indicated paucity of available, relevant, and normed assessment measures. Specifically, few standardized measures of motivation toward social interaction were identified. Of those identified, most standardized assessment tools targeted motivation broadly in typically developing populations and were administered in a self-report format (i.e., Intrinsic Motivation Inventory; Deci, Eghrari, Patrick, & Leone, 1994; McAuley, Duncan & Tammen, 1989; Social Reward Questionnaire; Foulkes, Viding, McCrory, & Neumann, 2014; Motivation and Pleasure Scale - Self-Report; Llerena et al., 2013; Conte Social Interest Questionnaire; unpublished; and The Schizotypal Personality Questionnaire; Raine,

1991). No stand-alone measures of motivation were specifically created for and targeted toward children or adults with ASD. Two subscales of social motivation in youth with ASD were identified, both of which were normed for parent-report formats (i.e., Anxiety Disorders Interview Schedule-Autism Spectrum Addendum [ADIS/ASA] social motivation subscale; Kerns, 2015; Kerns, Renno, Kendall, Wood & Storch, 2017 and the SRS-2, social motivation subscale; Constantino & Gruber, 2012). Given the results of the literature search, we proceeded with the creation of a novel interview measure.

Two frameworks guided methodological considerations in creating the interview. The construct modeling approach to measurement is particularly advantageous in conceptualizing units which make up the content and intended audience of a measure (Wilson, 2005). The construct modeling approach encompasses four building blocks of measurement development including the construct map, item design, the outcome space, and the measurement model (Wilson, 2005). The Instrument Development and Validation Scientific Standards as proposed by the Patient Reported Outcome Measurement Information System (PROMIS®) were similarly used as a model for item development and validation (PROMIS, 2013). PROMIS® guidelines are empirically supported and provide a model for the development of measures which can be used in clinical research trials and practice (PROMIS, 2013). Both frameworks for instrument development were integrated at multiple stages of SMI development.

Prior to item development, a measure-specific statement of purpose was created, for use as a benchmark in evaluation of the external validity of the measure. Additionally, a purpose statement dictates that all forthcoming steps of development and psychometric testing are consistent with the measure's intended purpose (Wilson, 2005). The purpose of the SMI is to measure the degree to which children aged 8-17 with ASD are interested in socially interacting

with peers. The semi-structured interview is not diagnostic in nature, but rather, designed to measure social motivation on a continuum. The intended use of the SMI is to provide a picture of social motivation over time, to strengthen the research base of social motivation in ASD, and to guide and evaluate developmentally appropriate psychological interventions.

Developing theoretically based content is regarded as the first step to establishing the content validity of a novel instrument (Lynn, 1985). A construct map and internal model of the SMI was established to facilitate a visual specification of the measure content. A construct map is comprised of: 1) a coherent definition for the content of the construct and 2) the idea that the construct is composed of a latent continuum (Wilson, 2005). Self-Determination Theory (Ryan & Deci, 2000) and the Social Motivation Hypothesis of ASD (Chevallier et al., 2012) informed the theoretical underpinning for the SMI's construct map (see Figure 2 for construct map). The three tenants of social motivation that guided development of this interview were social cognition, social interest/desire, and social awareness (see Figure 3 for internal model). The construct map and internal model were the foundations for creating an initial item pool for the SMI.

### **Measures.**

Thirty-one items were drafted and assessed for face validity (DeVon et al., 2007; Hardesty & Bearden, 2004; Lynn, 1985). Face validity is the soundness of a newly created measure, determined by qualitative judgements of a group of laypeople (Lynn, 1985). It is not psychometric in nature, but rather, a manner of streamlining an instrument prior to obtaining individual item content validity. Three individuals (Ph.D. candidates in Clinical Psychology uninvolved with the current study) edited clarity, grammar, structure, relevance, and user-

acceptability prior to distribution to an expert panel. Fourteen changes were initiated [Clarity: 7, Grammar: 5, User-acceptability: 2].

The intent of item development was to create individual interview items, each of which targeted the construct of social motivation broadly in a clear and precise manner. PROMIS® guidelines advocate for literacy-level analyses to judge the accessibility of the written measure items (PROMIS, 2013). The revised measure had adequate readability statistics. Specifically, the Flesch-Kincaid ease level was 81.7, the Flesch-Kincaid grade level was 4.3, and the measure contained 0.0% passive sentences. Flesch-Kincaid readability ease levels are rated on a 0-100 scale where scores rated above ‘70’ are considered acceptable (e.g., D’Alessandro, Kingsley, & Johnson-West, 2001).

The minimally refined thirty-one items were next subject to expert review. Each expert received an electronic copy of the measure after the invitation to participate was accepted. Expert review was completed in Qualtrics, a subscription software with capabilities of collecting data in a secure online format (Qualtrics, Provo, UT). First, the survey queried respondents to provide demographic information. Demographic variables included: gender, occupation, employment setting, number of years of experience working with individuals with ASD, and the number of clients with ASD they personally worked with over the past calendar year. Next, experts rated each individual item on the measure.

Item content relevance was queried as follows: “The following questions ask you to rate the content relevance of each item of the survey. Content relevance refers to how well each item of the interview corresponds to the defined content domain that the item was written to reflect (e.g., social motivation)”. Item importance was also assessed: “On the next page, you will see the same item questions. However, you will be asked to rate the item *importance*.” Experts

were also instructed to provide input regarding specific instructions, the overall measure, the examiner manual, age-relevance, and strengths/weaknesses of the measure. Item content and importance ratings were on a 4-point Likert scale: Content (1 = *Not Relevant*, 2 = *Somewhat Relevant*, 3 = *Quite Relevant*, 4 = *Highly Relevant*) and Importance (1 = *Not Important; Remove*, 2 = *Of Little Importance; Keep on Survey If Room*, 3 = *Of Average Importance; Keep on Survey*, 4 = *Very Important; Do Not Remove From Survey*) based on the recommendations of Lynn, 1985. Original item wording sent to the expert reviewers can be viewed in Table 1.

### **Data Analysis.**

Content validity is defined as the extent to which the items of a measure accurately capture the construct the scale is intended to measure (Polit & Beck, 2006; Wynd, Schmidt, & Schaefer, 2003). Despite expectations that the construct of a measure must be described, little consistency remains in how to derive meaningful quantitative or qualitative input (Polit & Beck, 2006). For the purposes of this study, the content validity index (CVI) was derived based on the methodology set forth by Lynn (1985). CVI was calculated at the individual and scale levels. Item content validity (I-CVI) was the primary metric of interest since the scale is novel and was subject to revision based on feedback from the expert panel (Polit & Beck, 2006; Polit et al., 2007). Scale content validity (S-CVI) was calculated by determining the average of I-CVIs for each item on the interview (Lynn, 1985; Polit & Beck, 2006). An alternative method of S-CVI relies on universal agreement among experts (i.e., the proportion of items on the measure that received agreement by every expert). The S-CVI average method was employed because universal agreement was deemed overly stringent given the number of expert raters included in the sample (Polit & Beck, 2006).

Item content and item importance were rated on a 4-point ordinal Likert scale such that ratings of '3' or '4' indicated valid content and ratings of '1' or '2' indicated invalid content. Ratings were dichotomized into a proportion agreement. Chance agreement was minimized by using a minimum of five and a maximum of ten expert raters (Lynn, 1985). The proportion of the number of experts who might agree out of the total acquired as part of the expert panel is first established, and then the standard error of the proportion is set (Lynn, 1985). This allows for specific cut-offs which dictate chance versus actual agreement (Lynn, 1985). If five or fewer experts are acquired, all must agree on the content validity. For the present study, six experts provided ratings. In order to achieve significance at the 0.05 level, the CVI for each item was set at 0.83 or greater (Lynn, 1985). Thus, five out of six experts must deem each individual item as content valid in order for the item to be retained. Chance agreement was further mitigated by eliminating the option of a neutral response through a 4-point forced response (Lynn, 1985). To ensure that each member of the expert panel offered critical feedback, it was determined that if any expert evaluated all SMI items positively and offered no known areas of improvement, that individual's responses would be removed from analyses (Lynn, 1985). A visual examination of quantitative and qualitative responses was conducted by the principal investigator. It was deemed that each expert participant offered critical feedback and a spread of responses that was not positively biased. Thus, all expert feedback was integrated into the final analyses.

### **Participants.**

Expert consultants were procured to respond to the survey based on the guidelines of Lynn (1985). Specifically, recommendations guide that a minimum of three and a maximum of ten experts engage in the judgement-quantification state of content validity (Lynn, 1985). Ten individuals with scientific specialties in test construction, ASD symptomology, and/or social

motivation in ASD were contacted to review the draft items. A non-probability sampling approach was determined appropriate to solicit the advice of authorities in the field, a gold-standard feature of pretesting methodology (Lynn, 1985). Experts were chosen based on the aforementioned comprehensive literature search and relevant scientific contributions (i.e., at least 3 publications related to the target area of scientific specialty) per the recommendations of Grant & Davis, 1996.

Participation in the expert review phase was closed to the public. Of the 10 individuals that were initially solicited, one individual declined without explanation and one individual did not respond to the request. Therefore, eight individuals agreed to review the measure and respond to the survey. Of the eight who agreed to review, six individuals completed the survey in the designated one-month timeframe. Expert reviewers understood that their participation was not anonymous. Expert reviewers were offered a \$20 Amazon gift card honorarium for their participation.

The six participants comprised an international expert consultation panel with varied demographic characteristics and academic expertise. Three individuals resided in the United States and three individuals resided internationally (Australia, the Netherlands, and Israel). The expert panel was predominantly female (83.33%) and ranged between 37-72 years of age. Ethnic and racial diversity was limited. Specifically, all participants identified as White. With respect to ethnic identity, one individual identified as Hispanic or Latino. Individuals endorsed working in a variety of employment settings, with academic medical centers and university settings evenly represented. Two individuals endorsed joint employment (i.e., 1. medical center and university; and 2. medical center and mental health facility). The remaining participants only endorsed one employment setting. All participants had lengthy professional experience



working with individuals with ASD over the course of their career ( $M_{years} = 20.33$ ,  $SD = 10.89$ ,  $Range = 7-40$ ). When queried about the number of individuals with ASD between the ages of 8-17 they have worked with in a clinical capacity during this calendar year, 3 participants indicated 'less than 10' and 3 participants indicated '30-40'.

### **Chapter 3 – Phase 1 Results**

Participants took an average of 40.03 minutes ( $SD = 21.97$ ,  $Range = 15.43-69.47$ ) to complete the survey. Each participant answered every item on the survey and completed the survey in its entirety. All responses were retained for data analysis as there was no evidence of careless responding. I-CVI ratings are detailed in Table 1 along with SMI author decisions regarding inclusion in the revised measure. I-CVI ratings for content relevance ranged from 0.50 to 1. I-CVI ratings for item importance ranged from 0.50 to 1. At the scale level, 25 items were deemed content valid out of the possible 31. S-CVI for content relevance was 0.81. The intra-class correlation coefficient (ICC) was additionally used at the scale level to mitigate concern that high proportion ratings might be due to random chance. The Scale ICC was 0.439 indicating poor agreement among raters (Koo & Li, 2016). The S-CVI was below 0.83 (Lynn, 1985) and the ICC was below what is recommended for reliability research (Koo & Li, 2016) indicating that several items should be subject to revision or deletion prior to psychometric evaluation among a pilot sample of participants.

Refinement of the item pool was conducted by revision of the individual items, instructions, and scoring scheme. Revision of individual interview items was guided by I-CVI ratings (Table 1). Sixteen items were retained without modification in the final interview. Six items were revised, and three additional items were revised and collapsed into a single question. Six items were discarded. This totaled 23 items retained as original or modified in the final

version of the measure. Qualitative input guided fine-grained modifications of the interview and measure instructions (see Appendix 1 for qualitative results).

Quantitative CVI ratings and qualitative feedback were both considered in the refinement of the SMI prior to distribution in a pilot sample. The structure of the interview was significantly modified to assist the examiner and examinee in a seamless administration. In order to standardize the semi-structured interview in a more systematic fashion, optional probes were added to several interview items for the examiner to query if the initial prompt was not understood or if an insufficient response was received. A visual rating scale was added to encourage responses from minimally verbal participants, allowing them to answer nonverbally by pointing. The subdomains of social interest and social desire were combined into one part of the interview, remaining distinct from the social behavior subdomain.

Scoring was similarly refined to encourage consistency across raters (i.e., inclusion of a 4-point Likert Scale for the rater to complete after each question). The interview manual was significantly modified to provide several examples and specific instructions to guide raters on how to make a scoring determination. Wording was modified to facilitate participant understanding and to foster cultural sensitivity.

## **Phase 2**

### **Chapter 4 – Phase 2 Methods**

#### **Participants.**

We sought a sample of 20 participants to evaluate preliminary psychometric properties. However, due to attrition between phone screen and consent, eighteen youth with ASD ( $M_{age} = 12.84$ ,  $SD = 2.61$ ,  $Range = 8-17$ ) and their caregivers consented to jointly participate in the development and initial evaluation of the SMI. This sample size was appropriate for a pilot

study as sample size recommendations for field testing of an instrument typically ranges from 100-1000 participants (Anthoine, Moret, Regnault, Sbille, & Hardouin, 2014). The small sample was judged sufficient to demonstrate proof of concept, initial feasibility, and acceptable preliminary psychometric properties (reliability and validity) prior to large-scale evaluation of the instrument.

Child participants were between the ages of 8-17, inclusive. This fairly large age range was sought because of the need to assess social motivation over time longitudinally as well as in the context of intervention. As such, the SMI needs to be appropriate for use in school-age children through adolescence. Further, the social motivation hypothesis applies to childhood through adulthood, despite the fact that there has been little research in this area in school-age and older individuals (e.g., Jones & Klin, 2013; Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Mendelson, Gates, & Lerner, 2016).

Both male and female children and caregiver participants were eligible for the study. Child participants must have met DSM-5 diagnostic criteria for ASD as determined by a clinical evaluation and supported by a research-reliable administration of the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) within the past 4 years. In the domain of communication skills, child participants were required to have adequate receptive language in order to understand the assessment battery prompts. However, instead of putting emphasis on verbal fluency, we asked that children be communicative instead. This was defined as expressive communication through either verbal or non-verbal modalities which could be reflective of the use of communication devices or note-writing. Preliminary determination of inclusion was assessed through a brief phone screen described in detail below.

Recruitment targeted caregivers and youth with ASD via flyers and online resources. Local advertising occurred in Southwest Virginia and was disseminated through web-based research portals. Online listservs, participant databases of the Virginia Tech Department of Psychology, Virginia Tech ASD registries (Virginia Tech Center for Research and Treatment, and Virginia Tech Autism Clinic), and parent support groups were targeted in the New River Valley region of Virginia. Caregivers were provided with a brief summary of the assessment findings by mail within 2-weeks of the assessment session. Additionally, participants were compensated for their time. Caregivers received a \$50 cash honorarium and child participants received a \$25 cash honorarium, totaling \$75 per family.

Demographic characteristics of the study participants can be found in Table 2. Cognitive abilities were derived from the full scale intelligence quotient (FSIQ) as estimated by the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II). The WASI-II administration requires verbal expressive language. One child participant did not have verbal expressive language and was administered the Leiter International Performance Scale, 3<sup>rd</sup> Edition (Leiter-3). An IQ score was not obtained for this individual due to testing demands and time considerations. Two out of the four Leiter-3 fluid intelligence subtests were administered (i.e., Figure Ground and Form Completion; Scaled Scores = 8 & 5 respectively).

## **Procedure**

To determine initial eligibility, caregiver (s) completed a brief phone screen with the principal investigator. The phone screen assessed communication of the child participant, information related to the child's ASD diagnosis, and availability with regards to scheduling an in-person laboratory assessment. The General Language Screen (GLS; Merricks, Bolton, & Goodyer, 2002) was modified and administered with the child's parent over the phone in order to

get a standardized assessment of parent-reported expressive and receptive language functioning. The screen is 12-items and response choices are YES/NO. Sample items include: *When your child [communicates] can he or she be understood by you?*, *When your child [communicates] can he or she be understood by other members of your family?*, *Is what your child [communicates] usually meaningful and relevant to the ongoing conversation or situation?*, *Can your child string three or more words together in a meaningful way? If not, can your child communicate in alternate ways?* Although the GLS was normed in typically developing children, younger than the sample recruited for this study, the items assess for a basic level of understanding and responses relevant to the tasks assessed during the assessment. If the parent endorsed YES to all of the items queried, they were invited to participate in the in-person assessment.

Twenty-six individuals expressed interest in the study and a desire to complete the requisite phone screen. Of those, five participants did not respond to investigator attempts to contact. Twenty-one individuals completed a phone screen. All individuals who completed a phone screen were found eligible to participate in the study. All 21 eligible child/caregiver dyads were scheduled to participate in the assessment session. However, 3 families never enrolled in the study due to parent-reported illness, life stressors, and/or difficulties with attendance due to substantial commute.

Upon arrival at the clinic, the child and caregiver each signed assent and consent documents, respectively. Children participated in one in-session assessment, jointly with at least one identified caregiver, which lasted approximately 3 hours. Assessing the caregiver and child jointly allows the interviewer to cross-validate responses to minimize bias. This is particularly advantageous as codes informed by more than one respondent allow for data triangulation (De

Los Reyes et al., 2015; Leech & Onwuegbuzie, 2007; Storch et al., 2012; Van der Ende, Verhulst, & Tiemeier, 2012). Additionally, integrating caregivers as respondents allows for the assessment of non-verbal participants. One such example of assessing children with ASD and their parents jointly is the Pediatric Anxiety Rating Scale (PARS; cf., RUPP 2002; Storch et al., 2012).

After consent, the assessment battery was administered. The order of self and other-report assessments was predetermined such that each participant engaged in the same assessment experience and order. The order of tasks was as follows: 1) Consent/assent, 2) phenotyping measures (ADOS, WASI-II), 3) The SMI, 4) Child eye-tracking while parent completes the parent-report measures, 5) ADIS/ASA interview modules, and 6) Child self-report measures. The SMI and ADIS/ASA interview module were jointly administered to the caregiver and child participants.

Nineteen caregivers completed the SMI for 18 youth participants with ASD. Fifteen biological mothers (78.95%), 1 biological father (4.16%), and 1 adoptive mother (4.16%) participated in the SMI and ADIS/ASA administration jointly in a dyad format (1 caregiver and 1 youth with ASD). All administrations of the SMI and ADIS/ASA were administered jointly unless individual administration was explicitly requested by the family. One grandmother (4.16%) and grandfather (4.16%), a couple, participated in the interviews, without their grandchild present. The interviews for that one participant were separately administered, with the youth independently completing items without caregivers present. Self/other-report assessments were administered electronically through Research Electronic Data Capture (REDCap; Harris, Taylor, Thielke, Payne Gonzalez, & Conde, 2009).

During the eye-tracking task, participants were seated in a chair which was arranged 70 centimeters from an 18-inch Tobii T60 XL eye-tracker monitor (Tobii Technology AB, Falls Church VA). The Tobii eye-tracker collects raw eye-movement data and transports them to the Tobii Pro Studio software where filters are applied. A five-point calibration procedure was initiated at the beginning of the task. Specifically, a calibration dot was presented at the four corners and center of the eye-tracking monitor. If calibration was not attained for all five points, re-calibration was attempted up to three times. After the calibration procedure, participants were instructed to look at the screen and keep their head still. They were informed that they would be watching videos, and that they could do this without any guidelines.

**Measures.**

***Social motivation.***

*Social Motivation Interview (SMI) for individuals with ASD.* The SMI is a semi-structured interview designed to measure levels of social motivation in youth aged 8-17 with ASD. The interview protocol contains 23 questions which, correspond to the following two subsections: 1) social wanting and 2) social behavior. Questions are asked verbally by the interviewing clinician to youth and caregiver dyads. No more than 2 caregivers/parents are encouraged to participate in the interview to ensure efficiency in administration. The domains of social cognition, social interest/desire, social awareness, and social behavior are queried. A visual analog scale is provided to participants during administration to facilitate understanding of response options. Clinicians administering the interview score each item on a 0-3 rating scale (0 = *None at all*, 1 = *A Little*, 2 = *A Moderate Amount*, 3 = *A Lot*). At the termination of the interview, subscale and total scores are derived based on individual item summation. Item scores are based on verbalized responses and observed behaviors. As such, the SMI is not simply respondent report as the final scores reflect verbalizations (responses) and observed behaviors.

The SMI seeks to assess current thoughts, behaviors, and emotions. As such, item responses and corresponding scoring are limited to the past 7 days. Given this timeframe, the interview may be repeated as frequently as every 2 weeks. The final iteration of the SMI can be found in Appendix 2. The principal investigator administered all of the SMI administrations.

*Anxiety Disorder Interview Schedule- Autism Spectrum Addendum (ADIS/ASA; Kerns, 2015; Kerns, Renno, Kendall, Wood & Storch, 2017).* The ADIS/ASA is a set of queries assessed in addition to the traditional ADIS –Parent Version (Albano & Silverman, 2016). The Interpersonal Relationships section of the ADIS/ASA was administered to parent and child jointly as part of this assessment battery. Five subsections comprise the Interpersonal Relationships section of the ADIS/ASA: 1) Friends, 2) Social Motivation, 3) Bullying/Peer Rejection, 4) Theory of Mind/Awareness of Social Opinion, and 5) Social Opportunity. A clinician rating is holistically given after receiving answers to the interview queries in each subsection. For the purposes of this study, particular interest was paid to the social motivation subsection, which includes five queries. The clinician rating is scored on a 0 to 3 Likert scale with higher scores indicating more impairment. Interrater reliability is sufficient ( $\kappa = 0.76$ ; Kerns et al., 2017). For the purposes of this study, an independent coder (someone other than the principal investigator) administered the ADIS/ASA to both parent and child jointly. During a training phase prior to administration, each independent coder observed via videotape review at least 1 administration of the measure given by an external party and was also trained at the VT Child Study Center on diagnostic interview administration. The principal investigator (R. Elias), who trained the independent coders, was trained on the ADIS/ASA by the developer (C. Kerns).

***Demographic characteristics.***



*Wechsler Abbreviated Scale of Intelligence – second edition (WASI-II; Wechsler, 2011).*

The WASI-II provides an estimate of general cognitive ability (i.e., FSIQ). For the purposes of this study, the 2-subtest version was used, which acquires FSIQ from the Vocabulary and Matrix Reasoning subtests. The measure has strong reliability in individuals 6 to 90 years of age with expressive language and can be completed in 15 minutes.

*Loneliness Questionnaire (Asher, Hymel, & Renshaw, 1984).* The Loneliness Questionnaire is a self-report measure that assesses feelings of loneliness in children. This study utilized a modified version to assess feelings of loneliness in children with ASD (Bauminger, Shulman, & Agam, 2003). The modified scale is 24 items and has high internal consistency ( $\alpha = 0.93$ ). The scale derives two subscales, emotional loneliness defined as feelings of isolation, and social loneliness, defined as perceived lack of involvement with peers.

***ASD symptom presentation.***

*Autism Diagnostic Observation Schedule (ADOS-2; Lord et al., 2012).* The ADOS-2 is a 45-60-minute observation and diagnostic test to assess deficits associated with Autism Spectrum Disorders in various domains, including Communication, Reciprocal Social Interaction, Imagination/Creativity, and Restricted and Repetitive Behaviors. It consists of a series of structured and unstructured situations that allow for the observation of spontaneous social-communicative behaviors and the ability to respond to social cues. Five modules of the ADOS-2 are available, each corresponding to a distinct communication and developmental level. For the purposes of this study, the ADOS-2 was administered by a clinician with established research reliability. The majority of participants were administered Module 3 of the ADOS-2 ( $n = 16$ , 88.89%), followed by equal distribution of Module 2 ( $n = 1$ , 5.56%) and Module 4 ( $n = 1$ , 5.56%).

*Social Responsiveness Scale, Second Edition, Parent Version (SRS-2; Constantino & Gruber, 2012).* The SRS-2 is a 65-item parent-report measure of ASD-related social impairments, including social awareness, social information processing, reciprocal social communication, social motivation, and restricted interests/repetitive behaviors. The SRS-2 provides a total T-score and subscale T-scores about the degree of interference in everyday life situations. The SRS-2 has a T-score range of 59 or less (normal range), 60 to 65 (mild range), 66 to 75 (moderate range), and 76 or greater (severe range). Psychometric properties for this scale are strong.

***Anxiety symptom presentation.***

*Anxiety Disorder Interview Schedule for DSM-5 - Parents/Children (ADIS-P/C; Albano & Silverman, 2016).* The ADIS-P/C is a semi-structured clinical interview administered to a parent and child by a trained clinician. The ADIS-P/C assesses for a wide range of psychopathology in children. The Interpersonal Relationships section of the ADIS-P/C, rather than the whole interview, was administered as part of this study's assessment battery. This module of the ADIS assesses preference for spending time with other kids and expressed desire to have more friends. Responses on this module are derived from Yes/No forced-choice and clinical severity ratings ranging from 0 to 8.

*Screen for Childhood Anxiety Related Emotional Disorders (SCARED; Parent and Child version; Birmaher et al., 1999).* The SCARED is a parent- or child-report measure used to screen for childhood anxiety disorders, including Panic Disorder or Significant Somatic Symptoms, Generalized Anxiety Disorder, Separation Anxiety Disorder, Social Anxiety Disorder, and Significant School Avoidance. The SCARED consists of 41 items, each rated on a 3-point scale regarding how well the item describes the child over the last 3 months. Although the whole

measure was administered, the Social Anxiety Disorder subscale was of specific interest for the purposes of this study. Both caregivers and children completed respective versions of the SCARED.

### ***Gaze Patterns.***

Eye-tracking was derived from child participants viewing the Interactive Visual Exploration (IVE) task, which has been previously implemented in children with ASD (Chevallier et al., 2015). The paradigm presents 22 dynamic video clips of children engaged in social and non-social (parallel) play activities. The children are engaging in activity in a naturalistic environment so that light switches, toys, posters, etc. are visible in the screenshot. The child actors participate in one of two conditions: social interaction or parallel play. In the social interaction condition, both children engage in a game together (e.g., card game). In the parallel condition, the children individually participate in their own task without interacting with the other individual in the screenshot. The stimuli involve sibling pairs who are of both genders and are school-aged. See Figure 4 for a screenshot of the paradigm. The entire paradigm lasts under 7 minutes and consists of 15 second clips followed by a white crosshair on a black background which is presented for 1 second.

Descriptive measure data for youth and parent participants can be found in Table 3. Measure bivariate correlations can be found in Table 6.

### **Data Analysis.**

Descriptive statistics were calculated for all administered assessment measures to characterize the participant sample. They were also calculated to determine acceptability of and satisfaction with the SMI. Participant acceptability ratings were derived from a 5-point scale (*1 = not acceptable to 5 = very acceptable*). At the end of the interview, parent and child

participants provided verbal feedback on aspects of interview that they liked and disliked.

Variables were assessed for missing values, distribution, and variance. Data were analyzed with IBM SPSS Statistics 25 (IBM Corp, 2017). The SMI was analyzed at the item and construct level.

Item analysis was adopted to analyze the participant's performance on individual measure items and their performance on the overall construct of social motivation through the SMI. Item-level psychometric data was examined with jMetrik™ (version 4.1.1; Meyer, 2014), an open-source software. Specifically, the following methods were acquired with jMetrik™ to assess the psychometric properties of the SMI including item analysis and reliability estimation. The data were analyzed with pairwise deletion of missing data so that as much of the examinee's data as possible was preserved (Meyer, 2014). Across all participants, two items of the SMI from separate participants were unable to be scored and thus considered 'missing' for the purposes of analyses. Missing response items included: social wanting subscale item 1 and social wanting subscale item 11.

Item difficulty (referred to as item endorsability, given the nature of the SMI) is the average item score. It measures the degree to which the highest response option is endorsed. For the SMI, item scores ranged from 0 to 3. The closer the mean is to 0, the less participants were assigned the highest category (a score of 3). The closer the mean is to 3, the more participants obtained scores of the highest category (Furr, 2017; Meyer, 2014). Item endorsability scores in the middle of the range are most desired as items with means hovering over one of the extremes will have low variances (DeVellis, 2017). An item may not sensitively detect certain values of the construct if item endorsability scores are near the highest or lowest scores of the range. Additionally, item endorsability scores with a narrow range will not

correlate well with other items (DeVellis, 2017). Item endorsability was also transformed to a 0 to 1 scale by dividing the produced number by the maximum possible item score (3).

Transformed item endorsability values close to 0 indicate an item which participants are rarely assigned the highest scoring option, and values close to 1 indicate an item where participants easily obtain the highest scoring option.

Item discrimination (referred to as item-test correlation for the purposes of this study) discerns how representative the individual item is of the entire measure. Item discrimination is the Pearson correlation coefficient between the item score and the total test score and indicates the degree to which the relevance of the item relates to the construct of social motivation being measured in the holistic interview. If the correlation is positive and high, the item strongly converges with the total score. If an item discriminates well, the item-test correlation has produced a result which narrowly captures the construct of interest without ambiguity (DeVellis, 2016).

Fixed item guidelines are not well established for the retention of polytomous items into a final measure. However, high item discrimination, high standard deviation, and moderate difficulty are typically recommended guidelines (Meyer, 2014). For the purposes of this study, adequate item transformed difficulty and item discrimination was set at a minimum of 0.30.

Internal consistency detects how well individual items group together and reflect the same underlying construct. Cronbach's Alpha was derived for both the measure as a whole and each individual subtest. To determine if individual item values may be driving internal consistency values, Cronbach's Alpha was also calculated for each individual item should it be omitted from the final measure. The standard error of measurement (SEM) indicates how

accurate each estimate is. SEM was also derived and calculates the error variability in the measure as a whole.

Group differences were examined with the Mann Whitney U Test since the group variables of interest were not normally distributed. Total scores of the SMI were compared with gender. Racial/ethnic groups were not able to be examined due to little variability in the racial and ethnic makeup of the sample.

Convergent validity, divergent validity, and concurrent validity were calculated using Pearson correlation coefficients. Convergent validity is the degree to which the SMI is correlated with measures of the same construct. Discriminant validity is the degree to which the SMI is uncorrelated with measures of unrelated constructs. Lastly, concurrent validity is the degree to which the SMI is correlated with measures of related constructs (Furr, 2018). Establishing correlation coefficients are the standard in determining measure validity (DeVellis, 2017). Unfortunately, psychometric standards of correlation statistics are not widely accepted in the instrument development literature (Devon et al., 2007). However, it is accepted that correlations above 0.50 are infrequent, and that the qualitative description of “high” or “moderate” varies widely among validated instruments (e.g., support in quality of life instrument ranged from  $r = 0.11$  to 0.88; Devon & Ferrans, 2003). To combat this discrepancy among reported statistics, it is recommended that narrow confidence intervals be additional support for validity in addition to the correlation obtained. Devon and colleagues (2007) suggest the marker of  $r \geq 0.45$  be accepted as “substantial and high” for tests of validity including convergent validity and  $r \leq 0.45$  for discriminant validity. Moderate correlation coefficients were defined for the purposes of this study as  $r > 0.30, < 0.45$ . Bivariate correlations were run among the following measures: 1) SMI, 2) ADIS/ASA Social Motivation differential diagnosis score, 3)

eye-tracking social prioritization score, 4) Loneliness Questionnaire, 5) WASI-II, and 6) parent-reported participant chronological age.

Convergent validity was determined by correlating the SMI with the ADIS/ASA Social Motivation differential diagnosis score (Kerns, 2015). Concurrent validity was established by correlating the SMI with the following assessments: social prioritization score (i.e., gaze fixation to socially relevant information), the Loneliness Questionnaire, and the SCARED social anxiety score. Discriminant validity was detected by correlating the SMI with IQ, ASD severity level, and participant age (see Figure 1).

Pearson correlation coefficients and corresponding confidence intervals were calculated in IBM SPSS Statistics 25 to establish preliminary validation. Confidence intervals are often used to aid interpretation of patterns of association. Confidence intervals for correlation coefficients were calculated by applying a Fisher's z-transformation to derive the standard normal distribution from which confidence intervals are computed.

Inter-rater reliability was calculated by having the SMI clinician administrator (R. Elias) and a Ph.D.-level clinician who was trained on administration and scoring (S.W. White) co-code 27.88% ( $n = 5$ ) of SMI administrations. Co-coding occurred either via live or videotaped recordings which were selected at random. Inter-rater reliability was derived from both percent agreement and a Kappa coefficient. Percent agreement codes responses dichotomously into 0 (no agreement) and 1 (agreement achieved). Obtaining a Kappa coefficient accounts for random chance agreement.

Gaze pattern was assessed via a standardized metric of fixation duration (FD). FD is characterized as the length that a participant fixates on an area of interest (AOI) of the stimulus which is presented. AOIs were drawn to capture individual faces and background objects in real-

time as they move throughout the visual video stimulus. Following methodology proposed by Chevallier et al. (2015), AOIs were differentiated into ‘social AOI’ (fixation to a human face) and ‘object AOI’ (fixation to background objects). Fixations were defined as gaze within a 30-pixel radius for at least 100 milliseconds. The following data export parameters were selected based on the literature and initiated within Tobii Pro Studio: Gap fill-in: 75 milliseconds maximum gap length, eye selection: average, noise reduction: disabled, velocity calculator and I-VT classifier: 20 milliseconds and 30 degrees per second, and merge adjacent fixations: max angle between fixations at 0.5 degrees. To draw accurate conclusions of overall looking patterns, individual participant fixation to at least 30% of each stimulus task was used as an inclusion benchmark. The methodological procedure outlined in Chevalier et al. (2015) was applied to derive a social prioritization FD score. Specifically, the social prioritization score represents a standardized FD to the social AOIs minus the standardized FD to object AOIs.

## **Chapter 5 – Phase 2 Results**

Variables were defined as continuous and missing values were identified. The skewness (1.21) and kurtosis (1.24) of the SMI measure total score was not greater than the absolute value of 2 indicating approximate normal distribution of the scale. There was a wide range of SMI scores ( $M_{total\ score} = 19.89$ ,  $SD = 13.63$ ,  $Range = 4-53$ ,  $Median = 17.00$ ). At the subscale level, variability in scores was also observed ( $M_{SocialWanting} = 12.56$ ,  $SD = 8.60$ ,  $Range = 3-32$ ,  $Median = 10.50$ ) and ( $M_{SocialBehavior} = 7.44$ ,  $SD = 5.76$ ,  $Range = 0-22$ ,  $Median = 7.50$ ). There were significant differences between the two subscale total scores ( $Z = -3.65$ ,  $p = 0.001$ ). SMI administration time was fairly brief ( $M_{mins} = 45.00$ ,  $Range = 21-60$ ). There were no significant group differences among scores obtained by male and female participants ( $U = 18.5$ ,  $p = 0.33$ ).

Participant level of relative engagement within the participant dyad, for the 17 families



who completed the SMI as a youth-caregiver dyad, was gauged after the interview on a 1-9 scale. Scores closer to 1 indicated that the child participated more in the interview, whereas higher scores indicated that the parent participated more in the interview. Scores hovering around 5 indicated equal participation from child and parent participants. This was true if both parent and child participants participated quite a bit or barely at all. Although not statistically significant, parental involvement was moderately correlated with participant age ( $r = -0.40, p = 0.10$ ), such that parents of young children were more engaged than the child. Similarly, parental involvement was moderately correlated with autism severity ( $r = -0.35, p = 0.15$ ). Participants, on average, participated equally in the interview ( $M = 4.50, SD = 0.92, Range = 3-6$ ). Similarly, child and parent participants primarily agreed upon responses ( $M = 4.78, SD = 0.65, Range = 3-6$ ).

At the construct level, the measure had strong overall reliability. Cronbach's alpha provided an internal consistency indicator of 0.96. The standard error of measurement was 2.98. Respondents had a 95% confidence interval of  $19.89 \pm 2.98 = (16.91, 22.87)$ . That means that assigning a 95% confidence interval rating, the true score of respondents lies between 16.91 and 22.87 if they obtained a score of the mean (19.89). This is a fairly narrow range meaning that the total score is relatively precise. Two hypothesized subscales were created during the development stage and analyzed independently as well. Specifically, the internal consistency of each subscale, when analyzed separately, was also strong (Social Wanting Subscale,  $\alpha = 0.93$ ; Social Behavior Subscale,  $\alpha = 0.89$ ).

Percent agreement was 80.87% aggregated among all items for the 5 selected administrations. The raters disagreed on scoring for a maximum of five items and agreed on scoring for a minimum of two questions. Cohen's Kappa was also calculated to account for

random chance agreement ( $\kappa = 0.69$ ). Inter-rater reliability was judged as substantial (Landis & Koch, 1977). In healthcare research, any Kappa coefficient above 0.60 is considered adequate (McHugh, 2012). Thus, the measure supported consistency among individuals assigning scores to the SMI.

Next, a correlation matrix amongst all individual items was generated to detect relatedness and extreme multicollinearity (Table 4). No specific guidelines exist to guide the retention or omission of items based on inter-item correlations for small sample sizes in pilot studies. As such, a visual examination of correlation matrices with set cutoff values is advised to guide the researcher in determining if exclusion prior to data analysis is appropriate (see Bodenhorn & Skaggs, 2005 for this procedure). For the present study, the following cutoff values were selected:  $r = 0.2$  to  $0.8$ . Specifically, if the majority of responses (defined as  $>80\%$ ) fell within this range, data are considered non-problematic and cleaning of specific variables should not be engaged at this stage of data analysis.

Most items (99.2%) were positively correlated. Responses on social behavior item 6 were negatively correlated with two items: social wanting subscale item 4 ( $r = -0.05$ ) and item social behavior item 7 ( $r = -0.06$ ). Although correlation coefficients are small, the directionality of the association suggests that the items should be considered for possible omission. A majority of the items (92.09%) were correlated with correlation coefficients ranging between 0.2 and 0.8. Specifically, 6.32% of item correlation coefficients fell below the 0.2 correlation cutoff (range =  $-0.06 - 0.20$ ), inclusive of the two negatively correlated values, and 1.58% of item correlation coefficients fell above the cutoff of 0.80 (range =  $0.81 - 0.87$ ). The values resting above the cutoff values suggest that: a) collinearity may exist or b) a factor analysis should be conducted in large-scale trials to understand the degree to which latent factors exist within the measure. Since

only 7.90% of item correlation coefficients fell above or below the pre-determined cutoff values, we proceeded with item analysis of all SMI items and considered these correlations non-problematic for the purposes of the pilot study analyses.

Item level analyses included item endorsability, standard deviation, item discrimination, and reliability coefficients if the item were dropped from the measure (Table 5). Eleven items obtained item endorsability scores less than 0.30 indicating that, on almost half of the items, the administering clinician rarely assigned the highest rating of a 3. This indicates that the spread of scores was narrow and favored lower item ratings. No items obtained overall discrimination scores less than 0.30. Overall item discrimination scores on the SMI indicated that the strength of the relationship between the item and the total test is strong. Similarly, the SMI scoring was able to distinguish among participants with similar total scores on the measure. One item (social behavior item 7) only obtained the scores '0' or '1' on a 4-point scale. The standard deviation of this item (0.43) was much smaller than the average standard deviation indicating that the individual item responses were not adequately discriminatory.

Correlation strength informed convergent and divergent validity of the SMI (see Table 7 for correlation coefficients). SMI scores were not significantly associated with participant age or FSIQ, although SMI subscales and FSIQ were moderately correlated at a strength higher than hypothesized. Two outliers were identified through a visual boxplot identifying values greater than or less than 1.5 times the interquartile range for FSIQ. These values resided on both the upper and lower bounds (FSIQ = 127 & FSIQ = 63) of the spectrum. With or without the inclusion of the outliers, results remained almost identical. Additionally, autism severity was moderately, though non-significantly, correlated with the SMI as judged by both the SRS-2 Total Score and the ADOS Comparison Score (see Table 7).

Convergent validity was examined by comparing the SMI to the ADIS/ASA (social motivation subscale) and gaze fixation to socially relevant information (see Table 7).

Correlation coefficients for the ADIS/ASA Social Motivation subscale ranged from -0.43 to -0.49 and resulted in significant correlations for the Social Behavioral subscale and the total score. The social prioritization score of the IVE task and the total score of the SMI was moderately positively correlated, although not statistically significant ( $r = 0.35, p = 0.16$ ).

Concurrent validity was further analyzed by comparing scores on the SMI to degree of loneliness and levels of social anxiety. Correlation strength between degree of loneliness and SMI ranged from -0.34 to -0.27 (Table 7). Scores on the SCARED were broken down by child and caregiver informants for the variables of interest (i.e., the social anxiety subscale and the total score). The association profiles for the SCARED social anxiety subscale and total score and the SMI were variable (Table 7). Specifically, as hypothesized, the SMI moderately correlated with self-reported social anxiety and loneliness. The confidence intervals for the correlation coefficients between the SMI and other measures were large for all variables assessed. Narrow confidence intervals typically suggest additional support for validity (Devon et al., 2007).

Participant acceptability ratings were high ( $M = 4.50, SD = 0.62$ ). 27.78% of parent participants indicated that they enjoyed hearing their child's responses and opinions about social interactions. The same proportion of child participants (27.78%) stated that the interview was too lengthy. Additional feedback pertained to the timeframe of the interview (i.e., one week of retrospective reporting was too narrow of a snapshot because of variability in children's social interests), the repetitiveness of questions, and satisfaction with the interview administration structure (i.e., approval of the open-ended response format). Several children indicated that they

would rather be engaging in a preferred activity (i.e., video games or reading) or be alone in a different setting, than completing the interview.

## **Chapter 6 - Discussion**

The primary aim of this study was to develop an interview measure of social motivation in youth with ASD that was methodologically sound and theoretically informed. The SMI fills a needed gap in the literature by creating a measure of social motivation which directly assesses both internal social processes and social behaviors. The creation of the SMI followed rigorous and systematic standards of instrument development (Meyer, 2005; PROMIS, 2013). Through the adopted two-phase approach, results of the study indicate stringent item development and promising initial psychometric properties. The study demonstrated “proof of concept” through user satisfaction ratings, feasibility, and initial psychometrics (reliability and validity) in a small pilot sample. When examining the SMI from the PROMIS® Instrument Maturity Model Framework, the instrument has undergone developmental examination (stages 1A-2B) and has necessitated further large-scale testing (public release stages 3A-5; PROMIS, 2013).

SMI development followed established guidelines to ensure the item pool was consistent with the statement of purpose and the definition of the construct of social motivation. Expert panel consultation indicated that the SMI required revision at both the item and interview level prior to field testing. This feedback informed refinement of the measure content, structure, scoring, and administration. Integrating a construct map, internal model, and expert feedback into the initial stages of measure development likely aided in a measure which was unified in presentation of the construct of interest and easy to administer in a diverse sample.

Despite structured gathering of quantitative and qualitative feedback in the first phase of the study to develop a content valid measure, expert raters may have significantly differed on

their perception of problematic interview items. Calculating percent agreement amongst the raters through CVI (Lynn, 1985) has been criticized for possible inflation of random chance agreement or disagreement (Lynn, 1985; Polit et al., 2006). Potential factors relating to expert discrepancies include: expectations of involvement, time spent on the task, judgements about what makes a strong or weak item, and their prior experience in evaluating assessment measures (DeMaio & Landreth, 2004; Olson, 2010). To combat some variability due to chance, expert raters were given instructions and a unified schema of soliciting feedback (Olson, 2010). In the future, a modified Delphi approach may be implemented to minimize expert discrepancy (PROMIS, 2013).

Further, although expert review was guided by prompts of content validity, there are several domains which could additionally be assessed at the measure development level. Expert consultants did not evaluate the appropriateness of items for racially/culturally diverse populations or the potential for individual items to be translated into a language other than English (PROMIS, 2013). Three international expert reviewers were intentionally integrated into the expert review panel to potentially mitigate the chance of potential cultural bias in item wording and to promote generalizability across diverse populations. Although designed as a strength, the use of an international expert panel may also be viewed as a weakness. Specifically, although ASD impacts individuals globally; health service availability, treatment, and sociocultural factors related to ASD may vary across geographic regions. Thus, incorporating both domestic and international scholars may have produced heterogeneity with regards to perceptions of individual item content and importance. Alternatively, the racially homogeneous sample may have limited adaptation for ethnically and racially-diverse youth with ASD.

Taken together, results of the first phase of the study suggest the need and potential utility of a measure of social motivation for children with ASD which is theoretically informed and content valid. The second phase of the study proved the measure to be feasible with respect to administration and scoring as well as consumer acceptability. Most notably, participants completed the SMI in under 1 hour, which is fairly brief given the scope of the interview measure and the number of respondents present during administration. Equal participation of youth and caregiver respondents indicated that both parties do contribute substantially in the presence of one another. Further, the convergence of responses between youth and caregiver participants strengthened the credibility of data and richness of information for which the examiner could make an appropriate rating. The high perception of the acceptability of the interview indicates that parents and youth judged the interview content to be relevant and appropriate, and administration to be efficient and acceptable.

A 7-day reporting window was chosen to increase applicability of the SMI for clinical trial research. Specifically, a 7-day recall window is regarded as optimal to capture clinically-relevant symptoms without introduction of bias (Chella et al., 2010; PROMIS, 2013). Additionally, a narrow span of time is advantageous for psychological clinical trials in which outcome measures are subject to multiple assessments occurring at several timepoints. Despite past research indicating that measurement error of past experiences is minimal within a week timeframe, parent respondents reported qualitative difficulty answering SMI items under a time restriction. Future research may consider the two-time frame method (Bhandari & Wagner, 2006) in which respondents are asked to respond to queries over a short and long recall period (Clark, Fiebig, & Gerdtham, 2008). Alternatively, future iterations of the SMI may incorporate better anchors for eliciting meaningful information in a clear and precise manner. The diverse

sample characteristics serve both as a strength and weakness of the study. The gender makeup of the study was consistent with estimates of sex differences observed in ASD (Loomes, Hull, & Mandy, 2017) and the socioeconomic composition of the participating families was varied. Age and IQ were intentionally broad with primary diagnostic categorization of ASD unifying the sample. Limitations related to the size and scope of the sample exist. First, the sample size of the study was small indicating that results should be interpreted with greater reliance on association strength than statistical significance of values. Therefore, it is encouraged that future examination of the developed measure utilize a larger sample. Secondly, the broad age and IQ range contributed to a developmentally diverse sample which should also be examined at a deeper level when large sample sizes are acquired.

Given the aforementioned characteristics with the overall sample size, psychometric data of the SMI at this stage of analysis should be cautiously interpreted. When analyzed from a unidimensional perspective, the SMI had excellent internal consistency indicating that items complementary measured the same construct. However, internal consistency ratings of this strength can signal that items may be redundant in nature or measuring a limited and specific facet of the construct repeatedly (Clark & Watson, 1995). However, only 1.58% of item correlations fell above a correlation coefficient of 0.80 which may slightly mitigate this concern.

The standard error of measurement and item endorsability ratings indicate a narrow, although precise, range of scores on the SMI. Specifically, it appears that in this ASD sample, scores typically clustered in the lower range of social motivation. This scatter profile was unsurprising given the broad coverage intended during the item development phase. Further refinement of the measure can address the narrow range of total scores in one of two ways. First, the measure can be informed by this data by condensing the range of possible scores to



encompass only lower trait levels of social motivation in a sample of youth with ASD. However, this approach appears fairly restrictive given the overall spread of total scores on the SMI (*Range = 4-53, out of a possible 69 points*). Alternatively, validating the measure across different psychopathology may yield informative data. Future evaluation of this measure should contain comparison samples, inclusive of typically developing youth to determine if intrinsic social motivation deficits assessed from this framework exist. Further, a diverse sample may allow for discrimination of ASD and typically developing youth on the basis of the SMI. Alternatively, this may also allow for social motivation to be captured transdiagnostically from a NIMH Research Domain (RDoC) perspective (Insel et al., 2010) with applications to those with and without ASD.

Item-level analyses also identified a few individual items which may need revision or omission prior to the next stage of field testing (social behavior item 6 and 7) and social wanting item 4. Specifically, these items weakly negatively correlated with other SMI items indicating that these items may not contribute to the construct of social motivation in meaningful ways.

Validity estimates obtained in this study determined the extent in which variation in the SMI scores are due to changes in overall levels of social motivation. Specifically, classical test theory methodology guided preliminary estimates of the association between the SMI and other related (convergent) and unrelated (discriminant) measures. However, the small sample size limits the credibility of these results. Specifically, confidence intervals are recommended as a supplement to coefficient estimates when examining the validity of a new measure (Devon et al., 2007). However, confidence intervals for validity coefficients may not be accurately estimated for samples containing less than 100 participants (Mendoza, Stafford, & Stauffer, 2000).

Partial support for discriminant validity is indicated. Congruent with hypotheses, the age of participants did not correlate with SMI total scores suggesting that applicability of the measure may be warranted in a wide range of developmental levels. Contrary to hypotheses, although not statistically significant is the strength of the SMI's association with IQ ( $r = 0.40$ ) and ASD symptomatology ( $r = -0.39$ ). Previous literature supports ASD symptomatology as being related to overt social behavior (Dawson et al., 2005), including aloofness (Pallathra et al., 2018).

Convergent validity was consistent with hypotheses for the ADIS/ASA, which – like the SMI, is an interview. The ADIS/ASA social motivation subscale is currently the only existing metric of social motivation conceptualized as interest or desire in social situations. Overt looking behavior toward social stimuli, an alternative methodology for indexing social motivation, was moderately correlated with the SMI. This suggests that the current conceptualization of social motivation (Chevallier et al., 2012) can perhaps be expanded to encompass internal cognitions related to social interaction. In other words, results of this study suggest that a broader, unified, and more expansive definition of social motivation is warranted and that the SMI may be a useful tool in capturing this.

Several methodological barriers exist to comparing the social prioritization FD in the present manuscript to that of previous research (Chevallier et al., 2015). First, in the present study, individual participant fixation to at least 30% of the entire stimulus task was used as a benchmark for inclusion. Demarcating 'valid' and 'invalid' trials of participant gaze fixation through data cleaning is typically regarded as essential, because ensuring an adequately large proportion of looking-time increases trust in the accuracy of the data (Venker & Kover, 2015). However, no guidelines exist with respect to what constitutes invalid trials by gaze fixation time

(Venker & Kover, 2015). Missing fixation data, however, was not deemed invalid (at any ratio) and included into the final analyses of Chevallier and colleagues (2015). Additionally, in the present study fixations over 100ms were considered a fixation, whereas in the only published study using the IVE task (Chevallier et al., 2015), fixation of 30ms was used as the inclusion criteria. These methodological differences may hinder comparisons across the two studies using the same video paradigm.

It is difficult to assert the degree to which SMI scores and social anxiety may be inter-related, particularly due to parent and child differences. Specifically, child SCARED social anxiety ratings correlated much stronger to overall SMI scores than parent-rated scores did. However, child and parent total SCARED scores were not correlated low with SMI total scores. Self-reported social motivation, captured under the Chevallier et al. 2012 definition, has been shown to correlate with social anxiety in previous recent studies of adults with ASD (Pallathra et al., 2018).

The SMI manual was comprehensive and provided detailed suggestions for clinical administration of the interview. Although inter-rater reliability was fairly high, limits to this measurement approach exist. Namely, semi-structured interviews lend themselves to a certain degree of clinical judgment when making appropriate scoring estimations at the item level (McIntosh & Morse, 2015). Advocating for a minimum threshold of reliability for research administrations (i.e., 85%, see Appendix 2) may mitigate deviations from the SMI's intended scoring scheme.

Results of pilot psychometric examination of the measure from a classical test theory perspective suggest that a second, in-depth evaluation phase of the study is warranted. Specifically, incremental validity is one facet that should be further examined. The measure may

be appropriate for integration into large-scale characterization or treatment trials where incremental validity and predictive validity can be measured. Incremental validity addresses the performance of a measure relative to others (Haynes & Lench, 2003). Specifically, the currently existing measures of social motivation cover some (e.g., reward, behavior) but not all (e.g., social cognitions) facets of the proposed construct. Improved content validity should be reflected in greater predictive validity (Haynes & Lench, 2003). Other psychometric domains of interest in large-scale field testing include: reliability and validity in an ethnic and socioeconomically diverse sample inclusive of stability, predictive validity, and test-retest reliability of the measure.

The clinical implications of this study are meaningful and relevant to youth with ASD. At the present moment, there are few measurement tools to capture social motivation in youth with ASD that involve stakeholder input. To our knowledge, the SMI serves as the first assessment to evaluate intrinsic motivation toward social interaction in children with ASD. This is particularly noteworthy given the number of social skills interventions available which target an observable social behavior impairment. The availability of an interview to target internal cognitions (e.g., social desire and interest) can serve as a valuable tool in enhancing social behavior in meaningful ways. Examination of thoughts, attitudes, and beliefs are a central tenant across many psychological disorders and treatments (e.g., depression; Gotlib & Jorman, 2010). Extension of this psychological process to social situations in youth with ASD is logical given the transdiagnostic nature of motivation in humans (Ryan & Deci, 2000). Initial psychometric properties, including item functioning, feasibility of administration, consumer acceptability, and construct and criterion validity are promising, and suggest that this measure should be further evaluated.



## Chapter 7 – References

- Abrams, D.A., Lynch, C. J., Cheng, K. M., Phillips, J., Supekar, K., Ryali, S., . . . Menon, V. (2013). Underconnectivity between voice-selective cortex and reward circuitry in children with autism. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *110*(29), 12060-12065. doi:10.1073/pnas.1302982110
- Albano, A.M. & Silverman, W.K. (2016). *Anxiety disorders interview schedule for DSM-5: Child and Parent version*. Oxford University Press.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5<sup>th</sup> ed.). Washington, DC: Author.
- Anthoine, E., Moret, L., Regnault, A., Sbille, V., & Hardouin, J. B. (2014). Sample size used to validate a scale: A review of publications on newly-developed patient reported outcomes measures. *Health and Quality of Life Outcomes*, *12*(1), 1–10.  
<http://doi.org/10.1186/s12955-014-0176-2>
- Asher, S. R., Hymel, S., & Renshaw, P. D. (1984). Loneliness in children. *Child development*, *55*, 1456-1464.
- Barbaro, J., & Dissanayake, C. (2007). A comparative study of the use and understanding of self-presentational display rules in children with high functioning autism and Asperger's disorder. *Journal of Autism and Developmental Disorders*, *37*(7), 1235-1246.  
doi:10.1007/s10803-006-0267-y
- Bauminger, N., Shulman, C., & Agam, G. (2003). Peer interaction and loneliness in high-functioning children with autism. *Journal of autism and developmental disorders*, *33*(5), 489-507.

- Bhandari, A., & Wagner, T. (2006). Self-reported utilization of health care services: improving measurement and accuracy. *Medical Care Research and Review*, 63(2), 217-235.
- Birmaher, B., Brent, D. A., Chiappetta, L., Bridge, J., Monga, S., & Baugher, M. (1999). Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): A replication study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(10), 1230–6.
- Bodenhorn, N., & Skaggs, G. (2005). Development of the School Counselor Self-Efficacy Scale. *Measurement and Evaluation in Counseling and Development*, 38, 14–29.
- Bottini, S. (2018). Social reward processing in individuals with autism spectrum disorder: A systematic review of the social motivation hypothesis. *Research in Autism Spectrum Disorders*, 45, 9–26. <http://doi.org/10.1016/j.rasd.2017.10.001>
- Cage, E., Bird, G., & Pellicano, E. (2016). Reputation management in children on the autism spectrum. *Journal of Autism and Developmental Disorders*, 1-14.
- Cage, E., Pellicano, E., Shah, P., & Bird, G. (2013). Reputation management: Evidence for ability but reduced propensity in autism. *Autism Research*, 6(5), 433-442.  
doi:10.1002/aur.1313
- Capriola, N. N., Maddox, B. B., & White, S. W. (2016). No Offense Intended : Fear of Negative Evaluation in Adolescents and Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*. <http://doi.org/10.1007/s10803-016-2827-0>
- Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biological Psychiatry*, 74(3), 195–203. <http://doi.org/10.1016/j.biopsych.2012.11.022>

- Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E.S., & Schultz, R.T. (2012). The Social Motivation Theory of Autism. *Trends in Cognitive Sciences*, 16(4), 231-239.  
doi:10.1016/j.tics.2012.02.007
- Chevallier, C., Molesworth, C., & Happé, F. (2012). Diminished social motivation negatively impacts reputation management: autism spectrum disorders as a case in point. *PloS One*, 7(1), e31107. <http://doi.org/10.1371/journal.pone.0031107>
- Chevallier, C., Parish-Morris, J., McVey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. T. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research*, 8(5), 620-628.
- Clarke, P. M., Fiebig, D. G., & Gerdtham, U. G. (2008). Optimal recall length in survey design. *Journal of health economics*, 27(5), 1275-1284.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological assessment*, 7(3), 309.
- Clements, C. C., Zoltowski, A. R., Yankowitz, L. D., Yerys, B. E., Schultz, R. T., & Herrington, J. D. (2018). Evaluation of the social motivation hypothesis of autism a systematic review and meta-analysis. *JAMA Psychiatry*, 75(8), 797–808.  
<http://doi.org/10.1001/jamapsychiatry.2018.1100>
- Constantino, J.N., & Gruber, C.P. (2012). *Social Responsiveness Scale (SRS-2)*. Los Angeles, CA: Western Psychological Services.
- Corbett, B. A, Swain, D.M., Newsom, C., Wang, L., Song, Y., & Edgerton, D. (2014). Biobehavioral profiles of arousal and social motivation in autism spectrum disorders. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 55(8), 924–34.  
<http://doi.org/10.1111/jcpp.12184>



- D'Alessandro D.M., Kingsley P, Johnson-West J. (2001). The Readability of Pediatric Patient Education. *Arch Pediatr Adolesc Med*, 155(7):807–812. doi:[10.1001/archpedi.155.7.807](https://doi.org/10.1001/archpedi.155.7.807)
- Dawson, G., Bernier, R., & Ring, R.H. (2012). Social attention: a possible early indicator of efficacy in autism clinical trials. *Journal of Neurodevelopmental Disorders*, 4(1), 11. doi:10.1186/1866-1955-4-11
- Dawson, G., Meltzoff, A.N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *J Autism Dev Disord*, 28.
- Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental Psychology*, 40(2), 271–283. <http://doi.org/10.1037/0012-1649.40.2.271>
- 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. doi:10.1207/s15326942dn2703\_6
- Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of personality*, 62(1), 119-142.
- Deci, E.L., & Moller, A.C. (2005). The Concept of Competence: A Starting Place for Understanding Intrinsic Motivation and Self-Determined Extrinsic Motivation. In *Handbook of Competence and Motivation* (Vol. 1, pp. 579–597). <http://doi.org/10.1017/S000748530002229X>
- Deci, E.L., & Ryan, R.M. (2000). The “ What ” and “ Why ” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227–268. [http://doi.org/10.1207/S15327965PLI1104\\_01](http://doi.org/10.1207/S15327965PLI1104_01)

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The Validity of the Multi-Informant Approach to Assessing Child and Adolescent Mental Health. *Psychol Bull*, *141*(4), 858–900.

<http://doi.org/10.1037/a0038498>

Demaio, T., & Landreth, A. (2004). Examining expert reviews as a pretest method.

In *Conference on Questionnaire Evaluation Standards-QUEST 2003*, *9*, 60-73.

Demurie, E., Roeyers, H., Baeyens, D., & Sonuga-Barke, E. (2011). Common alterations in sensitivity to type but not amount of reward in ADHD and autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *52*(11), 1164-1173. doi:10.1111/j.1469-7610.2010.02374.x

DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26). Sage publications.

Devon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., ...

Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability.

*Journal of Nursing Scholarship*, *39*(2), 155–164. <http://doi.org/10.1111/j.1547-5069.2007.00161.x>

Devon, H. A., & Ferrans, C. E. (2003). The psychometric properties of four quality of life instruments used in cardiovascular populations. *Journal of Cardiopulmonary Rehabilitation and Prevention*, *23*(2), 122-138.

Factor, R.S., Condy, E.E., Farley, J.P., & Scarpa, A. (2016). Brief report: Insistence on sameness, anxiety, and social motivation in children with autism spectrum disorder.

*Journal of Autism and Developmental Disorders*, *46*(7), 2548-2554.

doi:10.1007/s10803-016-2781-x

Fletcher-Watson, S., Findlay, J.M., Leekam, S.R., & Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception*, 37(4), 571–583.

<http://doi.org/10.1068/p5705>

Foulkes, L., McCrory, E. J., Neumann, C. S., & Viding, E. (2014). Inverted social reward: Associations between psychopathic traits and self-report and experimental measures of social reward. *PloS one*, 9(8), e106000.

Freud, S. (1915). Instincts and their Vicissitudes. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIV (1914-1916): On the History of the Psycho-Analytic Movement, Papers on Metapsychology and Other Works, 109-140

Furr, R. M. (2017). *Psychometrics: an introduction*. Sage Publications.

Garman, H. D., Spaulding, C. J., Webb, S. J., Mikami, A. Y., Morris, J. P., & Lerner, M. D. (2016). Wanting it Too Much: An Inverse Relation Between Social Motivation and Facial Emotion Recognition in Autism Spectrum Disorder. *Child Psychiatry and Human Development*, 47(6), 1–13. <http://doi.org/10.1007/s10578-015-0620-5>

Geen, R.J. (1991). Social Motivation. *Annu Rev Psychol*, 42, 377-399.

Gossen, A., Groppe, S.E., Winkler, L., Kohls, G., Herrington, J., Schultz, R.T., . . .

Spreckelmeyer, K.N. (2014). Neural evidence for an association between social proficiency and sensitivity to social reward. *Social Cognitive and Affective Neuroscience*, 9(5), 661-670. doi:10.1093/scan/nst033

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual review of clinical psychology*, 6, 285-312.

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in nursing & health*, 20(3), 269-274.

- Grelotti, D.J., Gauthier, I., & Schultz, R.T. (2002). Social interest and the development of cortical face specialization: What autism teaches us about face processing. *Developmental Psychobiology*, 40(3), 213-225. doi:10.1002/dev.10028
- Hardesty, D. M., & Bearden, W. O. (2004). The use of expert judges in scale development. Implications for improving face validity of measures of unobservable constructs. *Journal of Business Research*, 57(2), 98–107. [http://doi.org/10.1016/S0148-2963\(01\)00295-8](http://doi.org/10.1016/S0148-2963(01)00295-8)
- 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.
- Haynes, S. N., & Lench, H. C. (2003). Incremental validity of new clinical assessment measures. *Psychological Assessment*, 15(4), 456.
- Hobson, R.P., & Lee, A. (1998). Hello and goodbye: A study of social engagement in autism. *Journal of Autism and Developmental Disorders*, 28(2), 117-127.  
doi:10.1023/A:1026088531558
- Hudenko, W.J., Stone, W., & Bachorowski, J.A. (2009). Laughter differs in children with autism: an acoustic analysis of laughs produced by children with and without the disorder. *Journal of Autism and Developmental Disorders*, 39(10), 1392–400.  
<http://doi.org/10.1007/s10803-009-0752-1>
- Hull, C.L. (1943). Principles of behavior: An introduction to behavior theory. *Journal of the American Medical Association*, 124(2), 131-131. doi:10.1001/jama.1944.02850020061030
- Hutman, T., Chela, M. K., Gillespie-Lynch, K., & Sigman, M. (2012). Selective visual attention at twelve months: Signs of autism in early social interactions. *Journal of Autism and*

- Developmental Disorders*, 42(4), 487–498. <http://doi.org/10.1007/s10803-011-1262-5>
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... & Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders.
- Izuma, K., Matsumoto, K., Camerer, C.F., & Adolphs, R. (2011). Insensitivity to social reputation in autism. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 108(42), 17302-17307. doi:10.1073/pnas.1107038108
- Jaswal, V. K., & Akhtar, N. (2018). Being vs. appearing socially uninterested: challenging assumptions about social motivation in autism. *Behavioral and Brain Sciences*, 1-84.
- Jones, E.A., & Carr, E.G. (2004). Joint Attention in Children with Autism: Theory and Intervention. *Focus on Autism and Other Developmental Disabilities*, 19(1), 13–26.
- Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, 504(7480), 427–31.  
<http://doi.org/10.1038/nature12715>
- Kaplan, H.S., Hooper, P.L., & Gurven, M. (2009). The evolutionary and ecological roots of human social organization. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364(1533), 3289–99. <http://doi.org/10.1098/rstb.2009.0115>
- Kerns C.M. (2015). The Autism Spectrum Addendum to the Anxiety Disorders Interview Schedule. Unpublished manual.
- Kerns, C. M., Renno, P., Kendall, P. C., Wood, J. J., & Storch, E. A. (2017). Anxiety disorders interview schedule–autism Addendum: Reliability and validity in children with autism spectrum disorder. *Journal of Clinical Child & Adolescent Psychology*, 46(1), 88-100.

- Key, A.P., & Corbett, B.A. (2014). ERP responses to face repetition during passive viewing: A nonverbal measure of social motivation in children with autism and typical development. *Developmental Neuropsychology*, 39(6), 474-495. doi:10.1080/87565641.2014.940620
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, 59(9), 809-816.  
doi:10.1001/archpsyc.59.9.809
- Kohls, G., Perino, M.T., Taylor, J.M., Madva, E.N., Cayless, S. J., Troiani, V., . . . Schultz, R. T. (2013). The nucleus accumbens is involved in both the pursuit of social reward and the avoidance of social punishment. *Neuropsychologia*, 51(11), 2062-2069.  
doi:10.1016/j.neuropsychologia.2013.07.020
- Kohls, G., Thönessen, H., Bartley, G.K., Grossheinrich, N., Fink, G. R., Herpertz-Dahlmann, B., & Konrad, K. (2014). Differentiating neural reward responsiveness in autism versus ADHD. *Developmental Cognitive Neuroscience*, 10, 104-116.  
doi:10.1016/j.dcn.2014.08.003
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, 15(2), 155-163.
- Kuhl, P.K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science*, 8(1), F1-F12.  
doi:10.1111/j.1467-7687.2004.00384.x
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

- Larson, M.J., South, M., Krauskopf, E., Clawson, A., & Crowley, M.J. (2011). Feedback and reward processing in high-functioning autism. *Psychiatry Research, 187*(1–2), 198–203. <http://doi.org/10.1016/j.psychres.2010.11.006>
- Leech, N. L., & Onwuegbuzie, A. J. (2007). An Array of Qualitative Data Analysis Tools: A Call for Data Analysis Triangulation. *School Psychology Quarterly, 22*(4), 557–584. <http://doi.org/10.1037/1045-3830.22.4.557>
- Llerena, K., Park, S. G., McCarthy, J. M., Couture, S. M., Bennett, M. E., & Blanchard, J. J. (2013). The Motivation and Pleasure Scale–Self-Report (MAP-SR): Reliability and validity of a self-report measure of negative symptoms. *Comprehensive psychiatry, 54*(5), 568-574.
- Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(6), 466-474.
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule: ADOS*. Los Angeles, CA: Western Psychological Services.
- Lynn, M. R. (1985). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–385. <http://doi.org/10.1097/00006199-198611000-00017>
- McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport, 60*(1), 48-58.
- Mchugh, M. L. (2012). Lessons in biostatistics Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282.
- McIntosh, M. J., & Morse, J. M. (2015). Situating and constructing diversity in semi-structured

- interviews. *Global qualitative nursing research*, 2, 2333393615597674.
- 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*, *In Press*.
- Mendoza, J. L., Stafford, K. L., & Stauffer, J. M. (2000). Large-sample confidence intervals for validity and reliability coefficients. *Psychological Methods*, 5(3), 356.
- Merricks, M. J., Bolton, P. F., & Goodyer, I. M. (2002). Screening for Speech and Language Disorders: the reliability, validity and accuracy of the General Language Screen. *International Journal of Language and Communication Disorders.*, 37(2), 133–151.  
<http://doi.org/10.1080/13682820110116785>
- Meyer, J. P. (2014). *Applied measurement with jMetrik*. New York, NY: Routledge.
- Miskovic, V., & Keil, A. (2013). Perceiving threat in the face of safety: excitation and inhibition of conditioned fear in human visual cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 33(1), 72–8.  
<http://doi.org/10.1523/JNEUROSCI.3692-12.2013>
- Nakano, T., Tanaka, K., Endo, Y., Yamane, Y., Yamamoto, T., Nakano, Y., . . . Kitazawa, S. (2010). Atypical gaze patterns in children and adults with autism spectrum disorders dissociated from developmental changes in gaze behaviour. *Proceedings of the Royal Society B: Biological Sciences*, 277(1696), 2935.
- Olson, K. (2010). An examination of questionnaire evaluation by expert reviewers. *Field Methods*, 22(4), 295-318.



- Oruc, I., Shafai, F., & Iarocci, G. (2018). Link Between Facial Identity and Expression Abilities Suggestive of Origins of Face Impairments in Autism: Support for the Social-Motivation Hypothesis. *Psychological Science, 29*(11), 1859-1867.
- Pallathra, A. A., Calkins, M. E., Parish-Morris, J., Maddox, B. B., Perez, L. S., Miller, J., ... Brodtkin, E. S. (2018). Defining behavioral components of social functioning in adults with autism spectrum disorder as targets for treatment, *Autism Research, 11*, 488–502.  
<http://doi.org/10.1002/aur.1910>
- Polit, D.F., & Beck, C.T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health, 29*(5), 489–497. <http://doi.org/10.1002/nur.20147>
- Polit, D.F., Beck, C.T., & Owen, S.V. (2007). Is the CVI an Acceptable Indicator of Content Validity? Appraisal and Recommendations. *Research in Nursing & Health, 30*, 459–467.  
<http://doi.org/10.1002/nur>
- PROMIS® Validity Standards Committee on behalf of the PROMIS® Network (2013). PROMIS® instrument development and psychometric evaluation scientific standards. 1-72. Retrieved from  
<http://www.nihpromis.org/science/methodology?AspxAutoDetectCookieSupport=1>.
- Raine, A. (1991). The SPQ: a scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophrenia bulletin, 17*(4), 555-564.
- Research Units on Pediatric Psychopharmacology Anxiety Study Group. (2002). The pediatric anxiety rating scale (PARS): Development and psychometric properties. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(9), 1061-1069.

- Richey, J.A., Rittenberg, A., Hughes, L., Damiano, C. R., Sabatino, A., Miller, S., . . . Dichter, G.S. (2014). Common and distinct neural features of social and non-social reward processing in autism and social anxiety disorder. *Social Cognitive and Affective Neuroscience*, 9(3), 367-377. doi:10.1093/scan/nss146
- Ryan, R.M., & Deci, E.L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25, 54–67.  
<http://doi.org/10.1006/ceps.1999.1020>
- Salley, B., & Colombo, J. (2015). Conceptualizing Social Attention in Developmental Research. *Social Development*, 25(4), 687–703. <http://doi.org/10.1111/sode.12174>
- Schunk, D.H., Pintrich, P.R., & Meece, J.L. (2008). *Motivation in education: Theory, research, and applications* (3rd ed.). Upper Saddle River, N.J: Pearson/Merrill Prentice Hall.
- Schunk, D.H., & Usher, E.L. (2012). *Social cognitive theory*. In R.M. Ryan. *Oxford Handbook of Human Motivation* (13-27). New York, New York: Oxford University Press.
- Scott-Van Zeeland, A.A., Dapretto, M., Ghahremani, D.G., Poldrack, R.A., & Bookheimer, S.Y. (2010). Reward processing in autism. *Autism Res*, 3(2), 53-67.
- Sepeta, L., Tsuchiya, N., Davies, M. S., Sigman, M., Bookheimer, S.Y., & Dapretto, M. (2012). Abnormal social reward processing in autism as indexed by pupillary responses to happy faces, 1–9.
- Shic, F., Macari, S., & Chawarska, K. (2014). Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biological Psychiatry*, 75(3), 231–237.  
<http://doi.org/10.1016/j.biopsych.2013.07.009>
- Storch, E. A., Wood, J. J., Ehrenreich-May, J., Jones, A. M., Park, J. M., Lewin, A. B., & Murphy, T. K. (2012). Convergent and discriminant validity and reliability of the pediatric

- anxiety rating scale in youth with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 42(11), 2374–2382. <http://doi.org/10.1007/s10803-012-1489-9>
- Swain, D., Scarpa, A., White, S., & Laugeson, E. (2015). Emotion Dysregulation and Anxiety in Adults with ASD: Does Social Motivation Play a Role? *Journal of Autism and Developmental Disorders*, (Mc 0436). <http://doi.org/10.1007/s10803-015-2567-6>
- Van der Ende, J., Verhulst, F. C., & Tiemeier, H. (2012). Agreement of informants on emotional and behavioral problems from childhood to adulthood. *Psychological Assessment*, 24(2), 293–300. <http://doi.org/10.1037/a0025500>
- Venker, C. E., & Kover, S. T. (2015). An open conversation on using eye-gaze methods in studies of neurodevelopmental disorders. *Journal of Speech, Language, and Hearing Research*, 58(6), 1719-1732.
- Wechsler, D. (2011). WASI II: Wechsler Abbreviated Scale of Intelligence. 2nd. *Psychological Corporation*.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. New York, NY: Psychology Press.
- Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research*, 25(5), 508-518.

Table 1

*Expert Panel Content Validity Ratings*

<b>Subdomain: Social Interest</b>				
<b>Item</b>	<b>I-CVI Rele- vance</b>	<b>I-CVI Impor- tance</b>	<b>Hypothesized Component</b>	<b>Decision</b>
1. Do you tend to notice what kids your age are doing?	0.83	0.83	Social Awareness	Keep
2. Do you ever try to figure out what someone is thinking or feeling? Can you give me an example?	1	1	Social Cognition	Keep
3. What annoys you about hanging with peers? Why?	0.83	1	Social Cognition	Keep
4. Do you want other kids your age to be your friends? Tell me more about that.	1	1	Social Desire/ Interest	Keep
5. What makes someone friendly?	0.50	0.67	Social Cognition	<b>Revise &amp; Combine with 6</b>
6. What makes someone your age seem approachable (e.g., classmate, teammate)?	0.67	0.83	Social Cognition	<b>Revise &amp; Combine with 5</b>
7. Tell me about some of the trends that people your age are interested in ( <i>clinician inquire about: clothing, television shows, hobbies, music</i> )?	0.83	0.50	Social Cognition	<b>Discard</b>
8. Is there anyone in your life that you dislike? Do you hang out with them even though you don't like them very much? Why?	0.50	0.50	Social Behavior	<b>Discard</b>
9. What do you expect to happen if someone asks you, "Can I play [that] with you" ( <i>clinician can insert individual interests such as video games, Legos, music, etc.</i> )?	0.83	0.67	Social Behavior	<b>Revise</b>
10. If you took a trip to the park, what types of things would you be interested in doing? Would you like to do them alone or with other kids your age?	0.83	0.67	Social Desire	<b>Revise</b>
11. Rank the following activities in order of preference: [base off of youth's interest] A) [act which is solitary], B) Act which has remote involvement, (i.e., online video game partner or chat), C) Act with in-person	1	1	Social Desire	Keep

involvement.

<b>Subdomain: Social Behavior</b>				
12. Do you spend time with other kids? Why? What do you like to do together?	1	1	Social Behavior	Keep
13. What do you do during free or unstructured times during the school-day (e.g., recess)? Do you like doing _____ at [recess]? Have you ever tried doing something else at [recess]? What was that like?	1	0.67	Social Behavior	Keep
14. Do you do things to fit in? Do you do things to get kids to like or notice you ( <i>clinician inquire about types of clothes, social media use, slang, trends in media, afterschool activities</i> )?	1	0.83	Social Behavior	Keep
15. Do you sit with other kids at lunch? Tell me more about that.	0.83	0.50	Social Behavior	<b>Revise</b>
16. Tell me about your friends? What do you like to do with your friends?	1	0.83	Social Behavior	Keep
17. Do you ever adjust your behavior to fit in with the other kids? What types of things do you do to not stand out?	1	1	Social Behavior	Revise
18. Are you involved in any clubs, sports, or activities outside of school? Why do you participate in these activities? Are other people involved in them as well? <i>If youth is involved in extracurricular activities: Would you participate in these activities if your parent did not make you, or encourage you to participate?</i>	0.83	0.83	Social Behavior	Keep
19. Do you interact with people online (e.g., video game online multiplayer, chat rooms, instant messaging)? Do you have a preference for online or in-person social interactions?	0.83	0.83	Social Behavior	<b>Revise</b>
20. Do you use social media (e.g., Facebook, Twitter, and Instagram)? Tell me more about that.	0.67	0.67	Social Behavior	<b>Discard</b>
21. If a group of kids are doing something you're interested in, would you want to be included? Why? Give me some reasons why you choose to approach or avoid groups of other	1	1	Social Behavior	Keep

kids.				
22. Do you do things to help you fit in with other kids? Why?	1	1	Social Behavior	<b>Discard (Duplicate)</b>
<b>Subdomain: Social Desire</b>				
23. Do you like chatting with other kids your age? Why? Do you like having a conversation with other kids your age, even if it is about things you're not too interested in?	1	1	Social Desire/Interest	Keep
24. Would you like to spend more time with other kids, if given the opportunity? Why?	1	1	Social Reward	Keep
25. Would you like more friends? Why?	1	0.67	Social Desire	Keep
26. Would you like to spend more time with your current friends or peers? Would you like to become closer to any kids you know?	1	1	Social Desire/Interest	Keep
27. A best friend is someone you can really trust. Do you have friends that you can trust with your secrets? Do you have any friends that you are really close to? What makes them different from your other friends?	0.67	0.67	Social Desire/Interest	<b>Revise</b>
28. Everyone has things that they would like to work toward in their life. These things are called goals. For example, some kids have the goal of getting a good grade on a test they have studied hard for. Are having good friends that you can count on a goal of yours? How important is this to you? How likely is it that this will happen in the future? How much have you already attained this goal?	0.83	0.67	Social Behavior	<b>Discard</b>
29. What makes someone "cool"? Do you do anything to make yourself seem "cool" from time to time?	0.83	0.67	Social Cognition	Keep
30. If you saw someone who appeared friendly, would you like to get to know them?	1	0.83	Social Awareness	<b>Revise &amp; Combine with 5 &amp; 6</b>
31. If I was creating a television program all about you, and I asked some kids to describe what you're like, what would they say? What would you want them to say?	0.50	0.50	Social Cognition	<b>Discard</b>

Table 2

*Demographic characteristics of participants*

<b><i>n</i> = 18</b>	
Male	14 (77.78%)
Age	<i>M</i> = 12.84 ( <i>Range</i> = 8-17, <i>SD</i> = 2.61)
IQ	<i>M</i> = 100.12 ( <i>Range</i> = 63-127, <i>SD</i> = 18.00) <sup>a</sup> <i>M</i> = 100.80 ( <i>Range</i> = 72-119, <i>SD</i> = 14.82) <sup>b</sup>
Family Income	
Less than \$24,999	4 (22.2%)
\$25,000-\$49,999	4 (22.2%)
\$50,000-\$99,999	9 (50%)
Over \$100,000	1 (5.56%)
Ethnicity	
Hispanic/Latino	0
Race <sup>c</sup>	
African-American	0
American Indian or Alaska Native	1 (5.56%)
Asian-American	0
White	16 (88.89%)
Other	2 (11.11%)
Prefer not to indicate	1 (5.56%)
Current Psychiatric Medication	8 (44.44%)
Ever Been in Therapy	15 (83.33%)

*Note.* <sup>a</sup> Entire Sample, <sup>b</sup> With outliers excluded, <sup>c</sup> Participants were instructed to choose all that apply

Table 3

*Descriptive Measure Data*

Measure	Children ( <i>M, SD</i> )	Caregivers ( <i>M, SD</i> )	<i>P</i> -Value of Pairwise Comparison
ADOS-2 ( <i>Comparison Score</i> )	7.44 (1.98) <sup>a</sup>	-	-
SRS-2 Total Score ( <i>T-Scores</i> )	-	78.78 (7.71) <sup>a</sup>	-
Social Awareness	-	75.78 (10.63) <sup>a</sup>	-
Social Cognition	-	73.44 (10.10) <sup>a</sup>	-
Social Communication	-	75.89 (6.47) <sup>a</sup>	-
Social Motivation	-	69.89 (11.62) <sup>a</sup>	-
Restricted Interest and Repetitive Behavior	-	82.17 (9.67) <sup>a</sup>	-
Loneliness Questionnaire Emotional Subscale ( <i>Raw Score</i> )	21.00 (7.35)	-	-
Loneliness Questionnaire Social Subscale ( <i>Raw Score</i> )	27.53 (11.16)	-	-
ADIS/ASA Social Motivation Subscale ( <i>Raw Score</i> )	1.29 (0.77) <sup>b</sup>	-	-
SCARED Total Score ( <i>Raw Scores</i> )	22.35 (12.09)	26.00 (14.09) <sup>a</sup>	-0.83 (p = 0.41)
Panic/Somatic	4.59 (3.34)	5.44 (5.07)	-0.37 (p = 0.71)
GAD	5.24 (3.67)	7.17 (3.93)	-1.83 (p = 0.70)
Separation Anxiety	4.35 (2.57)	4.44 (3.68)	-0.55 (p = 0.59)
Social Anxiety	6.24 (3.53)	6.83 (4.41)	-0.43 (p = 0.67)
School Avoidance	1.94 (2.11)	2.11 (2.40)	-0.57 (p = 0.57)
Social Prioritization ( <i>Raw Score, ms.</i> )	2.9 (0.12)	-	-

Note: \*p < 0.05, <sup>a</sup> = clinical significance, <sup>b</sup> = jointly administered to child and parent



Table 4

<b>Inter-Item Correlation Matrix</b>																						
	SW1	SW2	SW3	SW4	SW5	SW6	SW7	SW8	SW9	SW10	SW11	SW12	SW13	SB1	SB2	SB3	SB4	SB5	SB6	SB7	SB8	SB9
SW1	1.00																					
SW2	0.63	1.00																				
SW3	0.63	0.81	1.00																			
SW4	0.12	0.34	0.51	1.00																		
SW5	0.42	0.67	0.77	0.53	1.00																	
SW6	0.16	0.55	0.66	0.65	0.79	1.00																
SW7	0.62	0.77	0.71	0.42	0.80	0.69	1.00															
SW8	0.43	0.39	0.53	0.30	0.68	0.34	0.59	1.00														
SW9	0.52	0.77	0.71	0.42	0.57	0.49	0.64	0.26	1.00													
SW10	0.56	0.58	0.49	0.08	0.28	0.17	0.56	0.61	0.42	1.00												
SW11	0.54	0.57	0.48	0.25	0.63	0.24	0.67	0.74	0.31	0.57	1.00											
SW12	0.54	0.59	0.57	0.26	0.73	0.46	0.65	0.51	0.46	0.27	0.55	1.00										
SW13	0.28	0.60	0.65	0.71	0.71	0.87	0.59	0.30	0.52	0.12	0.17	0.53	1.00									
SB1	0.51	0.70	0.57	0.44	0.54	0.44	0.62	0.46	0.39	0.56	0.60	0.49	0.60	1.00								
SB2	0.67	0.61	0.62	0.41	0.74	0.59	0.84	0.48	0.55	0.30	0.60	0.55	0.57	0.51	1.00							
SB3	0.69	0.38	0.51	0.40	0.67	0.43	0.63	0.50	0.32	0.16	0.52	0.54	0.47	0.48	0.79	1.00						
SB4	0.66	0.51	0.52	0.31	0.55	0.34	0.83	0.48	0.52	0.52	0.56	0.54	0.33	0.58	0.66	0.63	1.00					
SB5	0.52	0.44	0.34	0.17	0.57	0.42	0.64	0.46	0.13	0.28	0.52	0.37	0.39	0.47	0.70	0.79	0.44	1.00				
SB6	0.53	0.56	0.37	-0.05	0.33	0.34	0.47	0.18	0.39	0.27	0.19	0.58	0.50	0.32	0.41	0.25	0.28	0.30	1.00			
SB7	0.50	0.45	0.61	0.35	0.73	0.47	0.68	0.73	0.37	0.48	0.70	0.39	0.28	0.45	0.71	0.74	0.56	0.68	-0.06	1.00		
SB8	0.46	0.60	0.66	0.28	0.61	0.63	0.63	0.55	0.43	0.51	0.41	0.40	0.55	0.33	0.60	0.25	0.27	0.36	0.49	0.41	1.00	
SB9	0.63	0.61	0.73	0.24	0.47	0.24	0.52	0.39	0.61	0.58	0.57	0.48	0.20	0.52	0.36	0.29	0.60	0.02	0.18	0.45	0.35	1.00
SB10	0.41	0.62	0.65	0.51	0.64	0.76	0.72	0.26	0.64	0.19	0.25	0.36	0.68	0.22	0.75	0.40	0.38	0.40	0.45	0.38	0.79	0.25

*Inter-Item Correlation Matrix*

Table 5

*Item Analysis of the SMI*

Item	Option (Score)	Difficulty/ Endorsability (0-3)	Difficulty/ Endorsability Transformed (0-1)	SD	Discrimination (Item-Test Correlation)	Reliability if Item Deleted ( $\alpha$ )
SW1	Overall	0.389	0.130	0.608	0.687	0.953
	0.0	0.611	0.204	0.502	-0.622	
	1.0	0.278	0.093	0.461	0.303	
	2.0	0.056	0.019	0.236	0.578	
	3.0	0.000	0.000	0.000	NaN	
SW2	Overall	0.611	0.204	0.778	0.813	0.951
	0.0	0.556	0.185	0.511	-0.683	
	1.0	0.278	0.093	0.461	0.036	
	2.0	0.167	0.056	0.384	0.792	
	3.0	0.000	0.000	0.000	NaN	
SW3	Overall	0.778	0.259	0.878	0.834	0.951
	0.0	0.444	0.148	0.511	-0.638	
	1.0	0.389	0.130	0.502	-0.021	
	2.0	0.111	0.037	0.323	0.544	
	3.0	0.056	0.019	0.236	0.524	
SW4	Overall	1.500	0.500	0.786	0.492	0.955
	0.0	0.056	0.019	0.236	-0.280	
	1.0	0.500	0.167	0.515	-0.310	
	2.0	0.333	0.111	0.485	0.075	
	3.0	0.111	0.037	0.323	0.451	
SW5	Overall	1.000	0.333	1.138	0.865	0.950
	0.0	0.444	0.148	0.511	-0.732	
	1.0	0.278	0.093	0.461	-0.010	
	2.0	0.111	0.037	0.323	0.188	
	3.0	0.167	0.056	0.384	0.735	
SW6	Overall	1.167	0.389	0.985	0.689	0.953
	0.0	0.278	0.093	0.461	-0.555	
	1.0	0.390	0.130	0.502	-0.179	
	2.0	0.222	0.074	0.428	0.360	
	3.0	0.111	0.037	0.323	0.451	
SW7	Overall	0.889	0.296	0.900	0.921	0.950
	0.0	0.333	0.111	0.485	-0.654	
	1.0	0.556	0.185	0.511	0.047	
	2.0	0.000	0.000	0.000	NaN	
	3.0	0.111	0.037	0.323	0.812	

SW8	Overall	1.444	0.481	0.984	0.641	0.953
	0.0	0.167	0.056	0.384	-0.219	
	1.0	0.389	0.130	0.502	-0.497	
	2.0	0.278	0.093	0.461	0.018	
	3.0	0.167	0.056	0.384	0.735	
SW9	Overall	0.890	0.297	0.900	0.662	0.953
	0.0	0.389	0.130	0.502	-0.367	
	1.0	0.389	0.130	0.502	-0.359	
	2.0	0.167	0.056	0.384	0.520	
	3.0	0.056	0.019	0.236	0.524	
SW10	Overall	1.444	0.481	0.922	0.531	0.955
	0.0	0.111	0.037	0.323	0.123	
	1.0	0.500	0.167	0.515	-0.672	
	2.0	0.222	0.074	0.428	-0.036	
	3.0	0.167	0.056	0.384	0.746	
SW11	Overall	1.056	0.352	0.938	0.627	0.953
	0.0	0.278	0.093	0.461	-0.431	
	1.0	0.333	0.111	0.485	-0.302	
	2.0	0.278	0.093	0.461	0.340	
	3.0	0.056	0.019	0.236	0.578	
SW12	Overall	0.500	0.167	0.707	0.698	0.951
	0.0	0.611	0.204	0.502	-0.678	
	1.0	0.278	0.093	0.461	0.294	
	2.0	0.111	0.037	0.323	0.544	
	3.0	0.000	0.000	0.000	NaN	
SW13	Overall	0.778	0.259	0.943	0.679	0.953
	0.0	0.500	0.167	0.515	-0.696	
	1.0	0.278	0.093	0.461	0.073	
	2.0	0.167	0.056	0.384	0.701	
	3.0	0.056	0.019	0.236	0.092	
SB1	Overall	0.556	0.185	0.856	0.679	0.953
	0.0	0.611	0.204	0.502	-0.662	
	1.0	0.278	0.093	0.461	0.349	
	2.0	0.056	0.019	0.236	0.021	
	3.0	0.056	0.019	0.236	0.578	
SB2	Overall	1.333	0.444	0.908	0.829	0.951
	0.0	0.167	0.056	0.384	-0.488	
	1.0	0.444	0.148	0.511	-0.321	
	2.0	0.278	0.093	0.461	0.100	
	3.0	0.111	0.037	0.323	0.812	

SB3	Overall	0.611	0.204	0.850	0.680	0.953
	0.0	0.556	0.185	0.511	-0.620	
	1.0	0.333	0.111	0.485	0.268	
	2.0	0.056	0.019	0.236	0.074	
	3.0	0.056	0.019	0.236	0.578	
SB4	Overall	0.389	0.130	0.850	0.699	0.953
	0.0	0.778	0.259	0.428	-0.527	
	1.0	0.111	0.037	0.323	-0.175	
	2.0	0.056	0.019	0.236	0.524	
	3.0	0.056	0.019	0.236	0.578	
SB5	Overall	0.890	0.297	0.900	0.581	0.954
	0.0	0.390	0.130	0.390	-0.481	
	1.0	0.390	0.130	0.390	0.029	
	2.0	0.167	0.056	0.167	0.119	
	3.0	0.056	0.019	0.056	0.578	
SB6	Overall	0.844	0.281	0.786	0.456	0.955
	0.0	0.389	0.130	0.502	-0.441	
	1.0	0.389	0.130	0.502	0.004	
	2.0	0.222	0.074	0.428	0.410	
	3.0	0.000	0.000	0.000	NaN	
SB7	Overall	0.222	0.074	0.428	0.715	0.954
	0.0	0.778	0.259	0.428	-0.743	
	1.0	0.222	0.074	0.428	0.715	
	2.0	0.000	0.000	0.000	NaN	
	3.0	0.000	0.000	0.000	NaN	
SB8	Overall	1.056	0.352	0.938	0.666	0.953
	0.0	0.333	0.111	0.485	-0.687	
	1.0	0.333	0.111	0.485	0.206	
	2.0	0.278	0.093	0.461	0.146	
	3.0	0.056	0.019	0.236	0.524	
SB9	Overall	0.611	0.204	0.778	0.586	0.954
	0.0	0.556	0.185	0.511	-0.375	
	1.0	0.278	0.093	0.461	-0.316	
	2.0	0.167	0.056	0.384	0.792	
	3.0	0.000	0.000	0.000	NaN	
SB10	Overall	0.944	0.315	0.802	0.686	0.953
	0.0	0.278	0.093	0.461	-0.642	
	1.0	0.556	0.185	0.511	0.212	
	2.0	0.111	0.037	0.323	0.084	
	3.0	0.056	0.019	0.236	0.524	



Table 6. Bivariate correlations among study measures

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Age	1										
2. FSIQ <sup>a</sup>	-1.52 ( <i>p</i> =.59)	1									
3. SRS Total Score	0.16 ( <i>p</i> =.53)	-0.39 ( <i>p</i> =.15)	1								
4. ADOS Comparison Score	0.05 ( <i>p</i> =.83)	-0.44 ( <i>p</i> =.10)	0.42 ( <i>p</i> =.08)	1							
5. Caregiver SCARED Social Anxiety Subscale	0.10 ( <i>p</i> =.71)	-0.23 ( <i>p</i> =.41)	0.30 ( <i>p</i> =.23)	0.35 ( <i>p</i> =.16)	1						
6. Caregiver SCARED Total Score	0.15 ( <i>p</i> =.56)	-0.16 ( <i>p</i> =.58)	0.62 ( <i>p</i> =.00*)	0.39 ( <i>p</i> =.11)	0.66 ( <i>p</i> =.00*)	1					
7. Child SCARED Social Anxiety Subscale	0.24 ( <i>p</i> =.35)	-0.16 ( <i>p</i> =.57)	0.40 ( <i>p</i> =.12)	0.46 ( <i>p</i> =.07)	0.10 ( <i>p</i> =.70)	0.42 ( <i>p</i> =.09)	1				
8. Child SCARED Total Score	0.17 ( <i>p</i> =.52)	-0.35 ( <i>p</i> =.20)	0.28 ( <i>p</i> =.28)	0.60 ( <i>p</i> =.01*)	0.17 ( <i>p</i> =.52)	0.43 ( <i>p</i> =.09)	0.75 ( <i>p</i> =.00*)	1			
9. Loneliness Emotional Subscale	0.03 ( <i>p</i> =.92)	0.01 ( <i>p</i> =.97)	0.17 ( <i>p</i> =.52)	-0.08 ( <i>p</i> =.75)	0.11 ( <i>p</i> =.67)	0.09 ( <i>p</i> =.74)	0.23 ( <i>p</i> =.38)	0.27 ( <i>p</i> =.30)	1		
10. Loneliness Social Subscale	0.30 ( <i>p</i> =.25)	0.24 ( <i>p</i> =.40)	0.02 ( <i>p</i> =.95)	0.22 ( <i>p</i> =.40)	-0.01 ( <i>p</i> =.97)	0.04 ( <i>p</i> =.88)	0.44 ( <i>p</i> =.08)	0.42 ( <i>p</i> =.10)	0.71 ( <i>p</i> =.00*)	1	
11. ADIS/ASA Social Motivation Subscale	0.25 ( <i>p</i> =.33)	-0.21 ( <i>p</i> =.45)	0.36 ( <i>p</i> =.16)	0.20 ( <i>p</i> =.45)	-0.01 ( <i>p</i> =.98)	0.05 ( <i>p</i> =.84)	0.50 ( <i>p</i> =.04*)	0.33 ( <i>p</i> =.20)	0.60 ( <i>p</i> =.01*)	0.70 ( <i>p</i> =.00*)	1
12. Social Prioritization FD	0.36 ( <i>p</i> =.15)	0.34 ( <i>p</i> =.22)	-0.46 ( <i>p</i> =.06)	0.11 ( <i>p</i> =.67)	0.08 ( <i>p</i> =.77)	-0.12 ( <i>p</i> =.96)	-0.08 ( <i>p</i> =.77)	-0.02 ( <i>p</i> =.93)	-0.45 ( <i>p</i> =.07)	0.00 (0.99)	-0.26 ( <i>p</i> =.31)

Note: \* =  $p < 0.05$ , <sup>a</sup> Adjusted for outliers

Table 7. Correlations between the SMI and other validated measures

Variable	SMI Social Wanting Score ( <i>r</i> , <i>p</i> -value)	SMI Social Behavior Score ( <i>r</i> , <i>p</i> -value)	SMI Total Score ( <i>r</i> , <i>p</i> -value)	95% Confidence Interval of SMI Total Score ( <i>r</i> )	
				Lower	Upper
Age	0.12 ( <i>p</i> = .65)	-0.13 ( <i>p</i> = .65)	0.02 ( <i>p</i> = 0.94)	-0.29	0.61
FSIQ <sup>a</sup>	0.38 ( <i>p</i> = 0.17)	0.41 ( <i>p</i> = 0.13)	0.40 ( <i>p</i> = 0.14)	-0.14	0.76
SRS Total Score	-0.39 ( <i>p</i> = 0.11)	-0.37 ( <i>p</i> = 0.13)	-0.39 ( <i>p</i> = 0.11)	-0.72	0.09
ADOS Comparison Score	-0.46 ( <i>p</i> = 0.06)	-0.48 ( <i>p</i> = 0.04*)	-0.48 ( <i>p</i> = 0.05*)	-0.77	-0.12
Caregiver SCARED Social Anxiety Subscale	-0.30 ( <i>p</i> = 0.91)	-0.12 ( <i>p</i> = 0.65)	-0.07 ( <i>p</i> = 0.80)	-0.87	-0.30
Caregiver SCARED Total Score	-0.19 ( <i>p</i> = 0.46)	-0.31 ( <i>p</i> = 0.21)	-0.24 ( <i>p</i> = 0.34)	-0.64	0.26
Child SCARED Social Anxiety Subscale	-0.47 ( <i>p</i> = 0.06)	-0.65 ( <i>p</i> = 0.01*)	-0.55 ( <i>p</i> = 0.02*)	-0.82	-0.10
Child SCARED Total Score	-0.34 ( <i>p</i> = 0.19)	-0.46 ( <i>p</i> = 0.07)	-0.39 ( <i>p</i> = 0.12)	0.73	0.11
Loneliness Emotional Subscale	-0.33 ( <i>p</i> = 0.20)	-0.34 ( <i>p</i> = 0.18)	-0.34 ( <i>p</i> = 0.18)	-0.71	0.17
Loneliness Social Subscale	-0.27 ( <i>p</i> = 0.29)	-0.31 ( <i>p</i> = 0.23)	-0.29 ( <i>p</i> = 0.26)	-0.68	0.22
ADIS/ASA Social Motivation Subscale	-0.43 ( <i>p</i> = 0.08)	-0.56 ( <i>p</i> = 0.02*)	-0.49 ( <i>p</i> = 0.04*)	-0.79	-0.01
Social Prioritization FD	0.34 ( <i>p</i> = 0.17)	0.35 ( <i>p</i> = 0.16)	0.35 ( <i>p</i> = 0.16)	-0.14	0.70

Note: \* =  $p < 0.05$ , <sup>a</sup> Adjusted for outliers

Figure 1. Visual Schema of Study Design

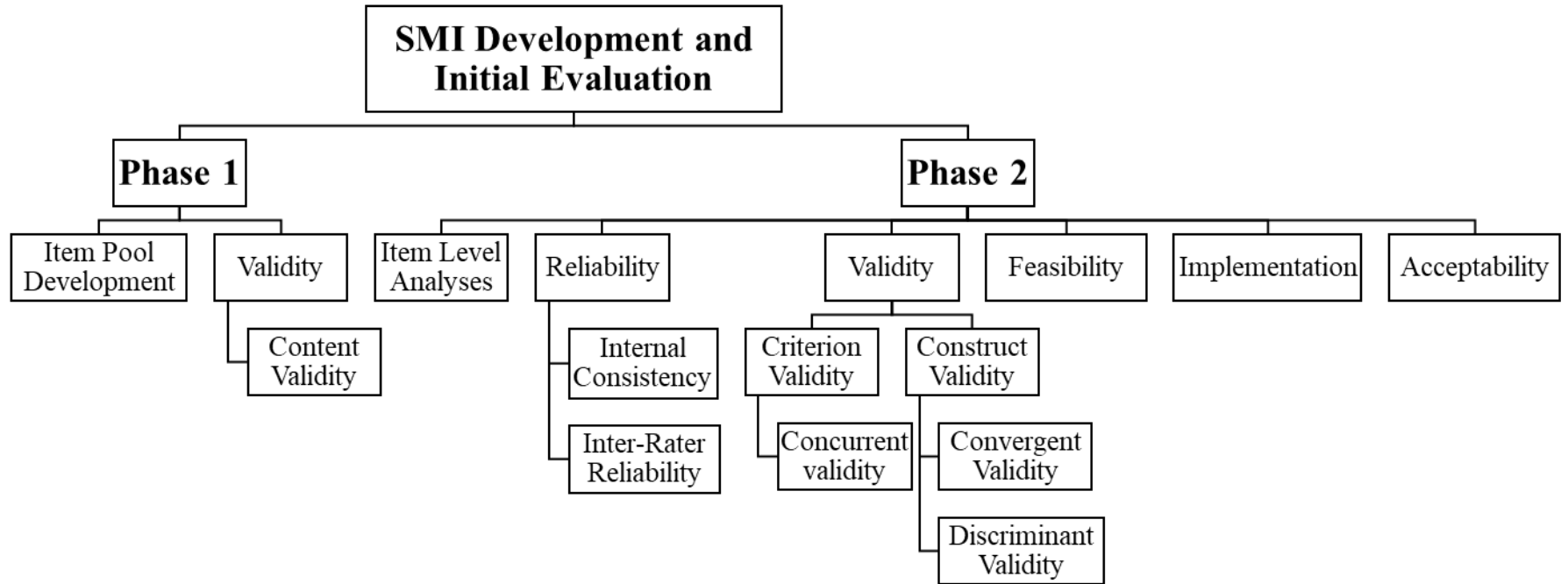




Figure 2. Construct Map

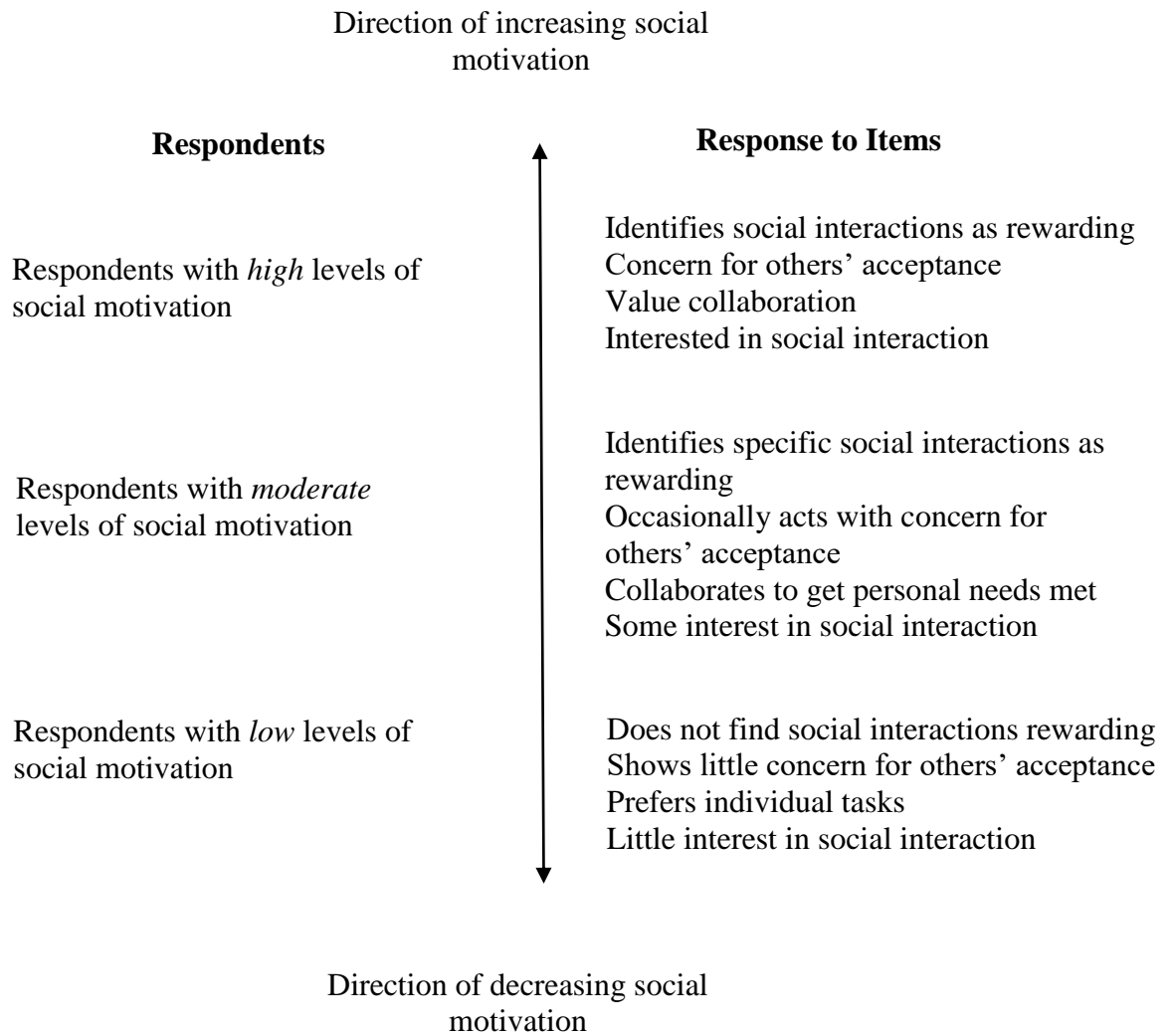


Figure 3. Internal Model

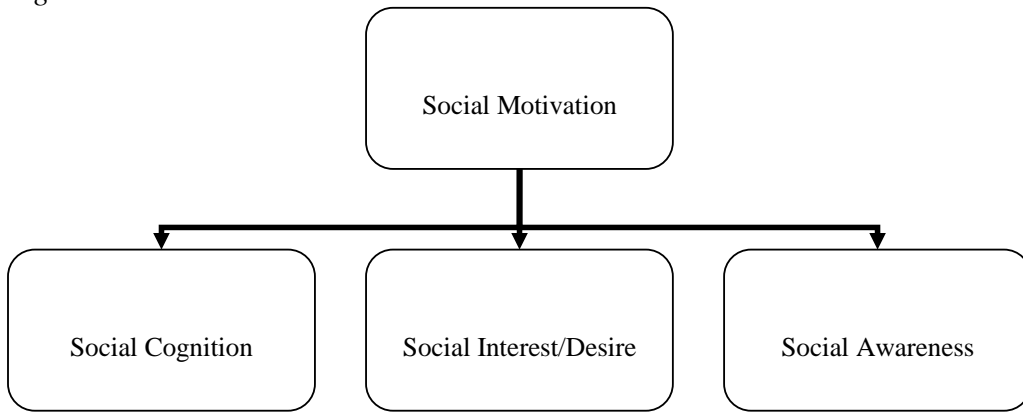


Figure 4. Screenshot of eye-tracking paradigm. Image taken from Chevallier et al., 2015

C.



## Appendix 1. Qualitative Expert Feedback

Three primary themes of the qualitative responses were evident including: (1) appropriateness of certain items for child audiences (e.g., “*the use of [the term] social preferences might be confusing for a child*”), (2) the structure of the interview (e.g., “*I am wondering if you really want to spilt out [the subdomains] desire and interest. I'm not sure you'd really get those two as separate factors in a factor analysis of social motivation though I see why behavior is separate. Maybe if the distinction was made clear, but I thought some of the interest questions definitely could be called desire*” and “*the headings were useful. I thought it was very helpful to lay out how parents and the child would contribute information and how to weigh the information*”), and (3) scoring considerations (e.g., “*the process that is described has the capacity to be quite subjective and this could lead to high rates of inconsistency between raters or even between occasions by the same rater. It seems to me that quite rigorous training would be required to achieve concordance and I would also think that you need greater than 80%*”). The utility of the SMI was overwhelmingly positive, “*the survey is a great idea...*”, “*really nice work! Would love to work with this!!*”, and “*sounds like an interesting and useful measure.*”

## Appendix 2. Final Version of the Social Motivation Interview

### **Social Motivation Interview (SMI)**

#### Manual

Overview: The Social Motivation Interview (SMI) is a semi-structured interview designed to measure levels of social motivation in youth aged 8-17 with Autism Spectrum Disorder (ASD). The interview is not diagnostic in nature. For the purposes of this interview, social motivation is conceptualized as an interest or desire to engage in social situations. The domains of social cognition, social interest/desire, social awareness, and social behavior will all be queried within the context of the interview. Youth who are invited to participate in this interview should have adequate verbal or non-verbal communication. The interview should not exceed 20 minutes in length, inclusive of scoring. The assessor should jointly survey the youth with ASD (aged 8-17) and a caregiver who is familiar with the youth's daily behavior. The purpose of jointly interviewing the youth and caregiver is to get multiple informant perspectives on several domains of social motivation.

Purpose: The purpose of this document is to instruct the clinician assessor how to administer the interview, derive meaningful scores, and interpret interview results. The SMI can be administered by a clinician with minimal training. The primary function of the administering clinician is to obtain detailed descriptions, examples, and behaviors exhibited by the interviewed youth as they relate to the individual item queries. Clinicians administering the interview will subsequently score the interview based on verbalized responses and observed behaviors. To encourage consistency in administration and scoring (inter-rater reliability), those who wish to use the SMI for research purposes may be provided additional training materials. Specifically, assessors may be asked to co-code two training videos to make sure there is at least 85% agreement on ratings assigned. A training session offered by the interview developer may also be available, per request. The primary goal of the training session will be to instruct novice evaluators how to, within the context of a semi-structured interview, obtain rich data. More specifically, evaluators will learn how to query appropriate follow-up questions/examples while remaining on-topic.

Format: The SMI is a 2-part semi-structured interview. The interview protocol contains 23 questions, and each one should be administered to derive total scores. The SMI interview protocol is used by the clinician to record and score youth and caregiver responses. The goal of the semi-structured interview is to elicit as much information as possible about the youth's level of social motivation. To achieve this, questions will be asked verbally by the clinician to youth and caregiver dyads. No more than 2 caregivers/parents are encouraged to participate in the interview to ensure efficiency in administration.

The interview should be conducted in a setting which allows for minimal external distractions. The clinician should sit opposite the youth and caregiver dyads to facilitate reciprocal conversation. The semi-structured format of the interview allows for the theme of social motivation to be explored through initial specific prompts, which can be read verbatim. However, the clinician should probe for specific examples or ask follow-up questions if the participant offers minimal information. The clinician should occasionally verbally summarize the clinician's interpretation of the content given by the youth and caregiver dyads in order to make sure that the assumptions made are validated by the individuals present in the room.

Instructions: The youth and one caregiver/parent are encouraged to be present for the duration of the interview. Clinicians will verbally ask youth and caregiver dyads the interview prompts by reading the numbered queries in the left-hand side of the box. The interview prompts should be administered in the order presented in the protocol. The parent's role within the interview is as a co-informant. The child should be encouraged to actively participate in the interview through verbal or non-verbal communication. Responses from the youth and parent dyads may be verbal or non-verbal. For example, a youth participant may use assistive communication devices, write, or use other means to provide answers. However, if the youth is not able to verbally or non-verbally respond to the assessor, ratings will rely more heavily on the parent's report. If disruptive behaviors are observed either prior to the interview or during the interview, the interviewer may request that the interview be administered to the parent only in order to ensure efficient administration, in which distractions are minimized.

The Social Motivation Interview seeks to assess **current** thoughts, behaviors, and emotions. As such, item responses and corresponding scoring should be limited to the past 7 days. To best understand social motivation temporally, the interview may be repeated, at minimum, every 2 weeks.

The wording of the questions and prompting of specific examples from the parent/youth will help to make internal events (thoughts and emotions) more directly observable. Participants should be encouraged to provide detailed descriptions as opposed to more general summaries. Additionally, the SMI might include hypothetical scenarios and scales in which the youth can indicate the degree to which he/she might feel/do something.

Scoring:

The assessor will record examples of behavior and cognitions as reported by the participants in the space listed as ***Clinician Notes***. Clinician notes may consist of verbatim responses from the youth and caregiver, a summary of content described by those present in the room, and/or behavioral observations.

The clinician will also circle a quantitative 0-3 rating in the right side of the query box. Rating assignments will be made by the clinician, based on all available information (participant communication, behavioral observations, and clinician judgement). All ratings follow the same Likert headers (0 = *None at all*, 1 = *A Little*, 2 = *A Moderate*

*Amount, 3 = A Lot*). Qualitative text under the Likert headers may aid in choosing the appropriate rating. Additionally, the following rating examples for a variety of contextual factors below should serve as exemplars. That is, these domains (as appropriate to each individual question) should be referenced and/or considered prior to assigning a rating for any specific item.

During the interview, it is important to assess for contextual factors. The clinician should keep the following queries in mind and use them as necessary in order to obtain as much contextual information as possible. It is essential to be fully familiar with these exemplars, because they inform the scoring for each item and ultimately may inform case conceptualization/treatment.

- *Types of individuals (e.g., significant other, same aged peers, adults)*
  - 0 = The individual interacts socially only when something is needed. Socialization is purely for utilitarian or functional purposes.
  - 1 = The individual interacts socially only with individuals conveniently present (e.g., family members, teachers, classmates).
  - 2 = Insufficient evidence of age-typical social partners/playmates.
  - 3 = The individual displays normative social behavior/preference relative to the social role of the target individual. Displays obvious preference for individuals with similar interests and of similar age. Age typical.
- *Proximity to individual (e.g., close friend, stranger)*
  - 0 = Friend(s) are clearly absent or cited as strangers.
  - 1 = Friendships are notably superficial. Examples may include only interacting with others in the context of structured (e.g., club) activities; people who are only corresponded with on the internet and citing service-workers such as cashiers, clerks, and aides as friends.
  - 2 = At least 2 friends which are reciprocal in nature, and similar in age.
  - 3 = A best friend is present. Demonstrates at least 2 clear examples of the nature/purpose of a close friendship. Age-typical.
- *Quantity (e.g. many friends, none, or a few good friends)*
  - 0 = Friend(s) are clearly absent
  - 1 = Friend(s) in less than 2 settings, all of which are structured (e.g., school, group therapy). Difficulty providing clear examples
  - 2 = At least 1 friend present in across a variety of contexts (at least 1 example given). Friendship not reciprocal in nature.
  - 3 = At least 2 friends present across a variety of contexts (at least 2 examples given). Reciprocal in nature. Age-typical.
- *Past successes/failures (e.g., high peer rejection)*
  - 0 = The client does not seek out social interaction

- 1 = The client experiences past/present peer rejection, yet attempts social interaction with the encouragement/facilitation of an external agent (e.g., parent, teacher, behavior chart)
- 2 = The client occasionally (at least 1 example given) attempts social interaction despite past/present peer rejection
- 3 = The client demonstrates many attempts at social interaction, across a variety of contexts (at least 2 examples given) despite past/present peer rejection. Age-typical.

A participant’s cultural context and chronological age must be considered within the context of assigning an appropriate rating.

The rating scheme is created so that higher scores correspond to higher levels of intrinsic motivation or social behavior, regardless of social competence. Intrinsic motivation refers to spontaneous curiosity related to social situations, not driven by an external force. An example of a response which displays intrinsic motivation is “I ask my mom to help me set-up playdates because hanging with my friends makes me happy.” Extrinsic motivation refers to behavior or desire which is driven by external rewards. An example of a response which displays extrinsic motivation is, “I say hello to at least 2 people at school because I get a candy reward at the end of the day if I do.” Responses driven by extrinsic motivation receive lower scores than responses driven by intrinsic motivation.

**INSERT SCANNED EXAMPLE HERE:**

For both Part 1: Social Wanting and Part 2: Social Behavior, a sum [S] score will be assigned. Each individual item rating in the subsection will be added to acquire the subsection’s sum [S] score. Summed ratings for each subsection should be assigned at the termination of the interview.



**S**

The interview consists of two subsections: 1) Part 1: Social Wanting and 2) Part 2: Social Behavior. Instead of labeling the subsections by content domains, participants will be informed if they are participating in Part 1 or Part 2 of the interview.

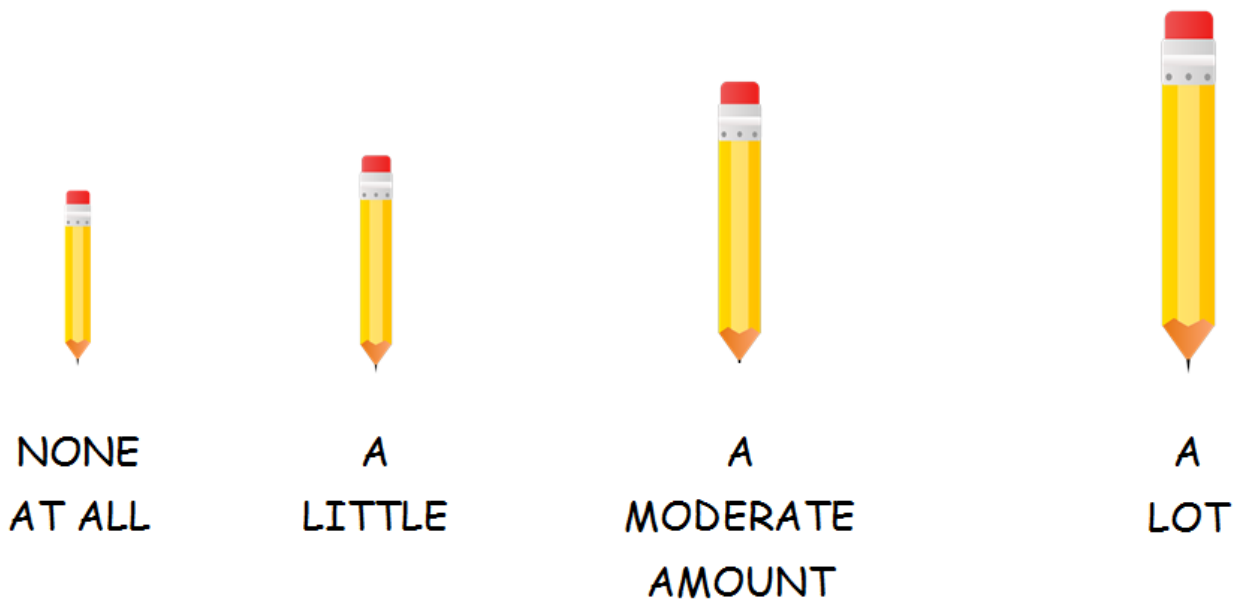
At the end of the interview, the clinician will provide several ratings. First, the clinician will rate how much the parent and youth each contributed to interview responses on a standardized scale (i.e., participation). The next standardized scale will assess parent and youth agreement. The Part 1: Social Wanting subsection and the Part 2: Social Behavior subsections are summed to form the total score.



Part 1: Social Wanting Score\_\_\_\_\_ + Part 2: Social Behavior Score\_\_\_\_\_ = Total  
Score\_\_\_\_\_

The interview may be re-administered no more than once per month to determine temporal changes.

A visual rating scale is used (typically placed on the table for interviewee to see) to supplement verbal queries and facilitate youth participation. Use of the visual rating scale is not mandatory and is to be used at the clinician's discretion in order to get more information when needed.



## SMI Interview Protocol

**Introduction:** *I am going to ask you some questions about hanging out with other kids. I will ask how you feel when you're around other people, your reasons for acting in a certain way when others are around, or what you want your social life to be like in the future. There are no right or wrong answers on this interview. I just want to find out a little bit more about you. Your parent will also be answering questions about your social likes and dislikes. Sometimes you might both agree on an answer and sometimes you might not. It is okay to disagree; we know that having different opinions from a parent or your child is normal. Answer these questions based on your thoughts, behaviors, and emotions over the past 7 days. I will be taking notes so I can remember what you tell me.*

Who is the caregiver/informant for this interview? *Please indicate all that apply:*

**Caregiver 1**

Relationship to Youth: \_\_\_\_\_

Gender: \_\_\_\_\_

**Caregiver 2 [If Present]**

Relationship to Youth: \_\_\_\_\_

Gender: \_\_\_\_\_

**PART 1:**

1. Do you tend to notice what kids your age are doing?

[Optional Probes: In room with same age kids, such as a classroom or birthday party]

*Clinician Notes:*

0	•None at all
1	•A little •Shows little effort/desire to understand other people
2	•A moderate amount •Shows some effort/desire to understand other people
3	•A lot •Shows a lot of effort/desire to understand other people

2. Do you ever try to figure out what someone is thinking or feeling? Can you give me an example?  
[Optional Probes: Who/Family/To What Means]  
*Clinician Notes:*

- 0 •None at all
- 1 •A little  
•Shows little effort/desire to understand other people
- 2 •A moderate amount  
•Shows some effort/desire to understand other people
- 3 •A lot  
•Shows a lot of effort/desire to understand other people

3. What annoys you about hanging with peers? Why?  
[Optional Probes: Other people/groups/clubs]  
*Clinician Notes:*

- 0 •None at all
- 1 •A little  
•Shows little tolerance for differences in social interaction
- 2 •A moderate amount  
•Shows some tolerance for differences in social interaction
- 3 •A lot  
•Shows a lot of tolerance for differences in social interaction

4. Do you want other kids your age to be your friends?  
Tell me more about that.  
*Clinician Notes:*

- 0 •None at all
- 1 •A little  
•Shows little preference for social interaction
- 2 •A moderate amount  
•Shows some preference for social interaction
- 3 •A lot  
•Shows clear preference for social interaction

5. If you took a trip to the park/arcade/swimming pool (clinician: choose age appropriate activity), what types of things would you be interested in doing? Would you like to do them alone or with others your age? Tell me more about that.

*Clinician Notes:*

6. Rank the following activities in order of preference: [based off of youth's interests/age/ability]

- A) An act which is solitary (e.g., reading, TV)
- B) An act which involves others indirectly (e.g., online video game partner or chat)
- C) An act with in-person interaction

How did you decide which activities seemed most interesting to you?

*Clinician Notes:*

7. Do you like chatting with other kids your age? Why? Do you like having a conversation with other kids your age, even if it is about things you're not too interested in?

*Clinician Notes:*

0 •None at all

1 •A little  
•Shows little preference for social interaction

2 •A moderate amount  
•Shows some preference for social interaction

3 •A lot  
•Shows clear preference for social interaction

0 •None at all

1 •A little  
•Shows little preference for social interaction

2 •A moderate amount  
•Shows some preference for social interaction

3 •A lot  
•Shows clear preference for social interaction

0 •None at all

1 •A little  
•Shows little preference for social interaction

2 •A moderate amount  
•Shows some preference for social interaction

3 •A lot  
•Shows clear preference for social interaction

8. Would you like to spend more time with other kids, if given the opportunity? Why?

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little preference for social interaction

2

•A moderate amount

•Shows some preference for social interaction

3

•A lot

•Shows clear preference for social interaction

9. Would you like more or new friends? Why?

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little preference for social interaction

2

•A moderate amount

•Shows some preference for social interaction

3

•A lot

•Shows clear preference for social interaction

10. Would you like to spend more time with your current friends or peers? Would you like to become closer to any kids you know?

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little preference for social interaction

2

•A moderate amount

•Shows some preference for social interaction

3

•A lot

•Shows clear preference for social interaction

11. Do you have someone who is not a family member that you can trust with your secrets? Do you have a friend that you are really close to?

[Optional Probes: What makes them different from your other friends?]

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little effort/desire to understand other people

2

•A moderate amount

•Shows some effort/desire to understand other people

3

•A lot

•Shows a lot of effort/desire to understand other people

12. What makes someone “cool”? Do you do anything to make yourself seem “cool” from time to time?

[Optional Probes: Stand out in a crowd, get noticed]

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little effort/desire to understand other people

2

•A moderate amount

•Shows some effort/desire to understand other people

3

•A lot

•Shows a lot of effort/desire to understand other people

13. If you saw someone who appeared friendly, would you like to get to know them? What makes someone appear friendly or approachable?

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little effort/desire to understand other people

2

•A moderate amount

•Shows some effort/desire to understand other people

3

•A lot

•Shows a lot of effort/desire to understand other people

Part 1: Social Wanting



S

**PART 2**

1. Do you ever share or compromise with people outside of your family when you really don't have to? Are there ever times when sharing is difficult? Tell me more about that.

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little social behavioral engagement

2

•A moderate amount

•Shows some social behavioral engagement

3

•A lot

•Shows a lot of social behavioral engagement

2. Do you spend time with other kids? Why? What do you like to do together?

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little social behavioral engagement

2

•A moderate amount

•Shows some social behavioral engagement

3

•A lot

•Shows a lot of social behavioral engagement

3. What do you do during free or unstructured times during the school-day (e.g., recess, waiting for the bell for start of classes)? Do you like doing \_\_\_\_\_ at [recess]? Have you ever tried doing something else at [recess]? What was that like?

[Optional Probes: Any unstructured social time not at home]

*Clinician Notes:*

4. Do you do things to fit in? Do you do things to get kids to like or notice you (*clinician inquire about types of clothes, social media use, slang, trends in media, afterschool activities*)?

[Optional Probes: Facebook requests, giving compliments, being friendly]

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little social behavioral engagement

2

•A moderate amount

•Shows some social behavioral engagement

3

•A lot

•Shows a lot of social behavioral engagement

0

•None at all

1

•A little

•Shows little effort/desire to conform to social norms

2

•A moderate amount

•Shows some effort/desire to conform to social norms

3

•A lot

•Shows a lot of effort/desire to conform to social norms



5. Do you sit with other kids at lunch at school?  
*Clinician: If child is homeschooled inquire about public cafeteria, family gathering with peers the same age, etc. Tell me more about that.*

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little social behavioral engagement

2

•A moderate amount

•Shows some social behavioral engagement

3

•A lot

•Shows a lot of social behavioral engagement

6. Tell me about your friends. What do you like to do with your friends?

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little social behavioral engagement

2

•A moderate amount

•Shows some social behavioral engagement

3

•A lot

•Shows a lot of social behavioral engagement

7. Do you ever adjust your behavior to fit in with the other kids? What types of things do you do to not stand out in a bad way?

[Optional Probes: Do you ever stop doing something specifically because other kids did not like it?]

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little effort/desire to conform to social norms

2

•A moderate amount

•Shows some effort/desire to conform to social norms

3

•A lot

•Shows a lot of effort/desire to conform to social norms

8. Are you involved in any clubs, sports, or activities outside of school? Why do you participate in these activities? Are other people involved in them as well?

[Optional Probes: *If youth is involved in extracurricular activities:* Would you participate in these activities if your parent did not make you, or encourage you to participate?]

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little social behavioral engagement

2

•A moderate amount

•Shows some social behavioral engagement

3

•A lot

•Shows a lot of social behavioral engagement

9. Tell me about your use of social media.

[Optional Probes: Do you interact with people online (e.g., video game online multiplayer, chat rooms, instant messaging, Facebook, Twitter, and Instagram)?

Can you give me some examples of ways that you interact with people online? Have you met in-person some of the people you communicate with? If not, have you ever tried to set up a meeting when them? Do you have a preference for online or in-person social interactions?]

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little social behavioral engagement

2

•A moderate amount

•Shows some social behavioral engagement

3

•A lot

•Shows a lot of social behavioral engagement

10. If a group of kids are doing something you're interested in, would you want to be included? Why? Give me some reasons why you choose to approach or avoid groups of other kids.

*Clinician Notes:*

0

•None at all

1

•A little

•Shows little social behavioral engagement

2

•A moderate amount

•Shows some social behavioral engagement

3

•A lot

•Shows a lot of social behavioral engagement

**PART 2: Social Behaviors**



**S**

**PART 3: Acceptability of Interview Format**

1. To what degree was the format of the interview acceptable to you? This includes the time commitment, appropriateness of question content, and method of delivery?

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>not acceptable</b>		<b>moderately acceptable</b>		<b>very acceptable</b>

2. What are some things you liked and disliked about the interview?

---

---

---

**OVERALL:**

Physical and behavioral observations (e.g., appearance, attitude, mood, eye-contact, psychomotor activity) for both the child and caregiver participants:

---

---

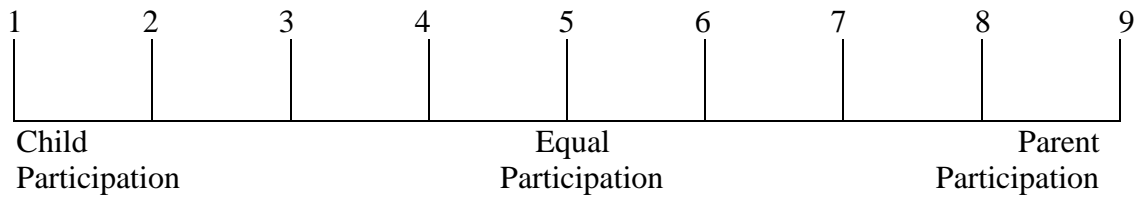
---

---

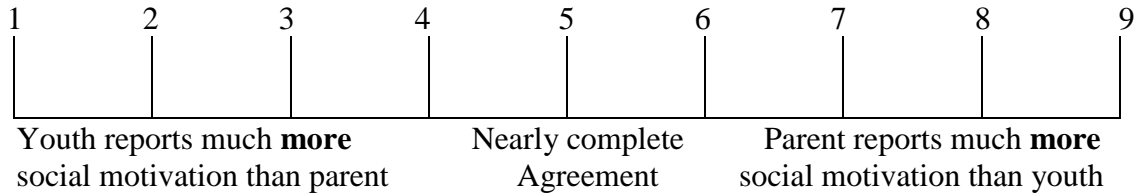
---

---

***Participation***



***Agreement***



***Total Score***

Part 1: Social Wanting Score\_\_\_\_\_ + Part 2: Social Behavior Score\_\_\_\_\_ =

Total Score\_\_\_\_