**Supplemental Materials**

**Methods**

***Participant*s**

Exclusion criteria included: (1) diagnosis or physical signs of known genetic conditions, (2) significant medical or neurological conditions or sensory impairments, (3) birth weight < 2000 g and/or gestational age < 36 weeks, (4) significant perinatal adversity and/or in-utero exposure to neurotoxins, (5) contraindication for MRI, (6) predominant home language other than English, and (7) 1st degree relative with psychosis, schizophrenia, bipolar disorder(Estes et al., 2015). ASD diagnosis in an older sibling was confirmed using the Autism Diagnostic Interview-Revised(Rutter, Le Couteur, & Lord, 2003) to establish an infant was at elevated familial ASD risk. Approximately 20% of high-risk infant siblings in IBIS meet criteria for ASD(Estes et al., 2015), consistent with prior infant sibling populations(Ozonoff et al., 2011). The clinical best estimate diagnosis was determined at age 24 months given this an age has strong evidence of diagnostic stability(Lord et al., 2006) and is the standard used by the Infant Brain Imaging Study during its longitudinal infant data collection, which launched in 2007 (Wolff et al., 2012; Estes et al., 2015).

***Measures***

The following instruments contained items incorporated in the Social Motivation Index (SMI). The 6-month social motivation index (SMI) had 22 items drawn from the Vineland Adaptive Behavior Scales (VABS) (Sparrow, Balla, & Cicchetti, 2005) and the Infant Behavior Questionnaire-Revised (IBQ-R)(Gartstein & Rothbart, 2003). The 12-month SMI had 40 items from the VABS; IBQ-R; Macarthur-Bates Communicative Development Inventory (M-CDI)(Fenson, Marchman, Thal, Dale, Reznick, & Bates, 2006); and the First Year Inventory (FYI) (Watson, Baranek, Crais, Reznick, Dykstra, & Perryman, 2007). The 24-month SMI had 16 items from the VABS and M-CDI.

**Vineland Adaptive Behavior Scale (VABS).** The VABS is a parent-report measure of adaptive function normed from 3 months of age through adulthood(Sparrow, Balla, & Cicchetti, 2005). The parent interview form of the VABS was used. SMI items were selected for ages 6, 12, and 24 months from expressive and receptive language items in the communication domain as well as the interpersonal relationships and play and leisure sections of the socialization domain. Recommended starting item of the scale advances with age; therefore, item-level data were included only in cases for which a specific item was assessed. Vineland items, noted by subscale, that were included on the SMI included: SMI6- Receptive: 2, 3; Expressive: 2, 5; Interpersonal Relationships: 4, 5, 6, 8, 9, 10; Play and Leisure: 3; SMI12 Receptive: 2, 3; Expressive: 2, 5; Interpersonal Relationships: 4, 5, 6, 8, 9, 10, 11, 12, 14; Play and Leisure: 3, 5; SMI24: Interpersonal Relationships: 11, 12, 14, 15, 17, 23; Play and Leisure: 5, 6, 7, 9, 14, 15, 17.

**Infant Behavioral Questionnaire-Revised (IBQ-R).** This widely used parent-report index of temperament, which detects constitutionally-based individual differences in reactivity and self-regulation(Gartstein & Rothbart, 2003), is validated for infants ages three to 12 months of age. SMI items for ages 6 and 12 months were selected from the cuddliness, smiling and laughter, high pleasure, and approach subscales. Items included on the SMI at 6 and 12 months were as follows: 5, 6, 67, 79, 80, 103, 126, 148, 149, 165, 172, 173).

**First Year Inventory (FYI).** This 63-item parent-report questionnaire, which probes social communication and sensory regulatory functions, is designed as a general population screening tool to identify 12-month old infants at risk for ASD or related developmental disorders(Watson et al., 2007). Social items included in the 12-month SMI were as follows: 1, 6, 12, 18, 19, 20, 29, 31, 52, 53, 58.

**Mac-Arthur Bates Communicative Development Inventory (M-CDI).** The M-CDI Words and Gestures Long Form, a parent-report measure indexing receptive and expressive language development, was completed for infants ages 12 and 24 months(Fenson et al., 2006). Scores from a subset of items on gestures commonly used in affiliative contexts (M-CDI, Section IIA, items 1, 2, and 4) were included in the 12- and 24-month SMI.

Instruments below were used in testing convergent and/or divergent validity of the Social Motivation Index.

**Autism Observation Schedule in Infants (AOSI).** This semi-structured play assessment is designed to detect and monitor early signs of ASD infants ages 6-18 months(Bryson, Zwaigenbaum, McDermott, Rombough, & Brian, 2008; Zwaigenbaum et al., 2021). Higher scores indicate lower functional performance in aspects of behavior frequently disrupted in ASD, including visual tracking, coordinated attention, social-communicative behaviors, reactivity, and sensorimotor development.

**Autism Diagnostic Observation Schedule (ADOS).** The ADOS is a semi-structured play assessment of characteristic features of ASD in the domains of communication, social interaction, play skills and restricted interests/repetitive behavior(Lord, Rutter, DiLavore, & Risi, 2000). ADOS Module 1 or 2 was administered to all subjects at age 24 months by certified, research-reliable evaluators at all sites. Calibrated severity scores (Gotham, Risi, Pickles, & Lord, 2007; Hus, Gotham, & Lord, 2014)were used to test relationships between level of core ASD features and social motivation. Higher scores corresponded to more ASD-related behaviors.

**Mullen Scales of Early Learning (MSEL).** The Mullen Scales of Early Learning was incorporated as a covariate in models testing relationships between social motivation and ASD outcome. The MSEL is a standardized direct assessment of cognitive development normed for ages from birth to 68 months(Mullen, 1995), which includes subscales measuring major early abilities, including expressive and receptive language, gross and fine motor function, and visual reception. The MSEL also includes an Early Learning Composite (ELC), a standardized index of overall cognitive function.

***Statistical Analyses for Calculation of Cohen’s d in Hierarchical Linear Models***

To follow Cohen’s procedure for calculating effect size (Cohen’s d) from simpler statistical models(Cohen, 1988), we utilized the EMATools package in R(Kleiman, 2019) that calculates d according to the following equation: d = estimate for fixed effect / (square root of sum of variances of random effects).

***Multiple Imputation***

To assess for potential biases from missing data, primary statistical models testing relationships between social motivation and ASD diagnosis were re-run following multiple imputation (Rubin, 1996; Enders, 2007), a robust approach even when a large proportion of data is missing (Madley-Dowd, et al., 2019). Multiple imputation was performed on SPSS version 28 via fully conditional specification (Lee and Carlin, 2010; Van Buuren et al., 2006) at the default setting of 10 iterations. This method fits a univariate model using all other available variables in the model as predictors and replaces missing values following the creation of multiple complete data sets using a Markov Chain Monte Carlo and expectation maximization method. Social motivation scores were estimated using predictive mean matching, and auxiliary variables of sex, diagnosis, and cognitive development, as indexed by the ELC on the Mullen Scales of Early Learning, were included as predictors in the imputation model (Yoo, 2009). Data were assumed to be missing at random, meaning missingness could be fully explained by other observed variables. The percentage of missing SMI scores among participants with a score for ≥ 1 time point were as follows: 6 months, 41%; 12 months, 44%; 24 months, 25%; and 24-month change scores: 46%. Each of the 10 imputed datasets were analyzed separately, followed by pooling of estimates and adjustments of confidence intervals (Barnard & Rubin, 1999).

***Exploratory Factor Analysis (EFA)***

To test the hypothesis that social motivation items across 6, 12, and 24 months loaded onto a single factor, we conducted an EFA using principal axis factoring. Included participants (n=114) had complete item level data for the 78 items queried at all three time points. Because our sample size did not meet a commonly recommended ratio for number of participants to items for EFA, previously noted as 10:1 (Kunce, Cook, & Miller, 1975), we used a parcellation strategy, an established approach in measurement modeling (Stucky, Gottfredson, & Panter, 2012), to consolidate the number of items into 10 composite item indicators. While parcellation approaches do not quantify loadings for each individual item, they have been shown to increase the ratio of common to unique variance across items and to reduce random measurement error (Matsunaga, 2008).

Using an *a priori* distributive item selection strategy, ten parcels were generated to include a balanced proportion of instrument type and assessment time point. Items queried at multiple time points were also distributed across parcels. An oblimin rotation was run in SPSS to allow exploration of the possibility of more than one correlated factor, which has previously been observed across social behaviors related to ASD (Frazier et al., 2014).

The planned approach for evaluating candidate factor solutions was to first prioritize factors occurring above inflection(s) in the scree plot with eigenvalues ≥ 1, as per the Kaiser criterion of factor retention. Factors were then to be reviewed for interpretability; in the case of factor solutions with poor interpretability, a third step planned step was to constrain the solution to a single factor. Item loadings ≥ 0.3 were considered to make substantive contributions to factor variance.

***Confirmatory Factor Analysis (CFA)***

A CFA using a maximum likelihood estimator was also performed using the cfa function in R within the lavaan package (Rosseel, 2012) to evaluate fit indices for a unitary factor structure. Full information maximum likelihood estimation was applied to allow inclusion of data from participants who had missing scores for some parcels but not others. Fit indices included the Comparative Fit Index (CFI), Tucker Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). Per common guidelines, thresholds for good fit were CFI and TLI ≥ 0.95 and RMSEA ≤ 0.05 and SRMR ≤ 0.08 (Hu and Bentler, 1999) and for acceptable fit, CFI and TLI ≥ 0.90 and RMSEA < 0.08 (Browne & Crudeck, 1993).

**Results**

***Participant Characteristics***

No significant differences in proportions of males and females (ꭓ2(2)=1.23, *p*=.54), ASD diagnosis (ꭓ2(2)=0.35, *p*=.84), and racial (ꭓ2(6)=5.32, *p*=.50) and ethnic composition (ꭓ2(2)=0.35, *p*=.84) were observed across age groups. Mean Mullen ELC scores, indicating overall cognitive development, were close to the normed mean of 100 and did not differ across age (FELC(2,1072)=1.64, *p*=.19).

***Measurement Properties of the Social Motivation Index***

An EFA was performed on 10 parcels derived for SMI items across all ages in 114 participants with full item level data at 6, 12, and 24 months. The Kaiser-Meyer Olkin Measure of Sampling Adequacy was equal to 0.88, consistent with good sampling adequacy. The scree plot revealed a single inflection point for a factor having an eigenvalue of 4.22 and accounting for 42.21% of the variance. The second potential factor had an eigenvalue < 1 (0.97); thus, both the inflection point and eigenvalue were consistent with a unitary factor solution. The ten item parcels showed moderate to strong loadings on the single factor (in descending order: Parcel 1: 0.76, Parcel 2: 0.70, Parcel 3: 0.69, Parcel 4: 0.69, Parcel 5: 0.68, Parcel 6: 0.64, Parcel 7: 0.61, Parcel 8: 0.59, Parcel 9: 0.56, Parcel 10: 0.55).

In a test for divergent validity, SMI scores at each age showed weaker correlations with concurrent Mullen fine motor scores in comparison to correlations for the AOSI total score and ADOS calibrated severity score, as determined by Zou’s method to compute confidence intervals (CIs) for differences in correlations. At all ages, differences comparing correlations of social motivation with the AOSI/ADOS scores versus Mullen fine motor scores produced confidence intervals that did not cross zero, indicating correlations were significantly different: 6 months, 95% CI [0.02, 0.28]; 12 months, 95% CI [0.07, 0.35]; and 24 months, 95% CI [0.14, 0.35].

Shape

Description automatically generated with medium confidence**Supplemental Tables and Figures**

Shape

Description automatically generated with medium confidence

Shape

Description automatically generated with medium confidenceShape

Description automatically generated with medium confidence

Shape

Description automatically generated with medium confidence

Shape

Description automatically generated with medium confidence

Shape

Description automatically generated with low confidence

A picture containing shape

Description automatically generated

Chart, histogram

Description automatically generated

Graphical user interface

Description automatically generated

**Supplemental References**

Barnard, J., & Rubin, D.B. (1999). Small-sample degrees of freedom with multiple imputation. *Biometrika, 86*(4), 948–955. doi.org/10.1093/biomet/86.4.948

Browne MW, Cudeck R. Alternative Ways of Assessing Model Fit. In: Bollen K, Long J, editors. *Testing Structural Equation Models.* Sage; Newbury Park, CA: 1993. pp. 136–162.

Bryson, S.E., Zwaigenbaum, L., McDermott, C., Rombough, V., & Brian, J. (2008). The Autism Observation Scale for Infants: scale development and reliability data. *Journal of Autism and Developmental Disorders*, *38*, 731-738.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E.S., & Schultz, R.T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*, 231-239.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York: Routledge.

Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour research and therapy*, 98, 4-18. Doi:10.1016/j.brat.2016.11.008

Estes, A., Zwaigenbaum, L., Gu, H., St John, T., Paterson, S., Elison, J. T., Hazlett, H., Botteron, K., Dager, S. R., Schultz, R. T., Kostopoulos, P., Evans, A., Dawson, G., Eliason, J., Alvarez, S., Piven, J., & IBIS network. (2015). Behavioral, cognitive, and adaptive development in infants with autism spectrum disorder in the first 2 years of life. *Journal of Neurodevelopmental Disorders*, *7*, 24.

Fenson, L., Marchman, V.A., Thal, D.J., Dale, P.S., Reznick, J.S., & Bates, E. (2006). *MacArthur-Bates Communicative Development Inventories: User’s Guide and Technical Manual* (2nd ed.). Baltimore: Brookes Publishing Company.

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2014). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the social responsiveness scale-2. *Autism: the international journal of research and practice*, *18*(1), 31–44. doi: 10.1177/1362361313500382

Gartstein, M.A. & Rothbart, M.K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development*, *26*, 64-86.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, *37*, 613-627.

Hus V., Gotham K., & Lord C. (2014). Standardizing ADOS domain scores: separating severity of social affect and restricted and repetitive behaviors. *Journal of Autism and Developmental Disorders*, *44*, 2400-2412.

Kleiman, E. (2019). EMAtools: Data Management Tools for Real-Time Monitoring/Ecological Momentary Assessment Data (Version 0.1.3) [Computer software].

Kunce, J. T., Cook, D. W., & Miller, D. E. (1975). Random variables and correlational overkill. Educational and Psychological Measurement, 35, 529-534.

Lee, K. J., & Carlin, J. B. (2010). Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *American journal of epidemiology*, 171(5), 624-632. doi:10.1093/aje/kwp425

Lord C., Risi S., DiLavore P.S., Shulman C., Thurm A., & Pickles A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry*, *63*, 694-701.

Lord, C., Rutter, M., DiLavore, P.C., & Risi, S. (2000). *Autism Diagnostic Observation Scale*. Los Angeles: Western Psychological Services.

Madley-Dowd, P., Hughes, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of clinical epidemiology*, 110, 63-73. doi:10.1016/j.jclinepi.2019.02.016

Matasunaga, M. Item Parceling in Structural Equation Modeling: A Primer. *Communication Methods and Measures*, *2*(4), 260–293, 2008 doi: 10.1080/19312450802458935

Mullen, E. (1995). *Mullen Scales of Early Learning*. Circle Pines: Guidance Service Publishing.

Nakagawa S., Johnson P.C.D., & Schielzeth H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, *14*, 20170213.

Ozonoff, S., Young, G.S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Bryson, S., Carver, L.J., Constantino, J.N., Dobkins, K., Hutman, T., Iverson, J.M., Landa, R., Rogers, S. J., Sigman, M., & Stone, W.L. (2011). Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics*, *128*, e488-e495.

Rosseel Y (2012). “lavaan: An R Package for Structural Equation Modeling.” Journal of Statistical Software, 48(2), 1–36. doi: 10.18637/jss.v048.i02

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American statistical Association*, 91(434), 473-489. doi:10.2307/2291635

Rutter M., Le Couteur A., & Lord C. (2003). *Autism Diagnostic Interview-Revised*. Los Angeles: Western Psychological Services.

Sparrow, S.S., Balla, D.A., & Cicchetti, D.V. (2005). *Vineland Adaptive Behavior Scales* (2nd ed.). Circle Pines: American Guidance Service.

Van Buuren, S., Brand, J. P., Groothuis-Oudshoorn, C. G., & Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. *Journal of statistical computation and simulation*, 76(12), 1049-1064. doi:10.1080/10629360600810434

Watson, L.R., Baranek, G.T., Crais, E.R., Reznick, J.S., Dykstra, J., & Perryman, T. (2007). The first year inventory: Retrospective parent responses to a questionnaire designed to identify one-year-olds at risk for autism. *Journal of Autism and Developmental Disorders*, *37*, 49-61.

Wolff, J. J., Gu, H., Gerig, G., Elison, J. T., Styner, M., Gouttard, S., Botteron, K. N., Dager, S. R., Dawson, G., Estes, A. M., Evans, A. C., Hazlett, H. C., Kostopoulos, P., McKinstry, R. C., Paterson, S. J., Schultz, R. T., Zwaigenbaum, L., Piven, J., & IBIS Network (2012). Differences in white matter fiber tract development present from 6 to 24 months in infants with autism. The *American journal of psychiatry, 169*(6), 589–600. doi.org/10.1176/appi.ajp.2011.11091447

Yoo, J. E. (2009). The effect of auxiliary variables and multiple imputation on parameter estimation in confirmatory factor analysis. *Educational and Psychological Measurement*, 69(6), 929-947. doi:10.1177/0013164409332225

Zwaigenbaum, L., Bryson, S.E., Brian, J., Smith, I.M., Sacrey, L., Armstrong, V., Roberts, W., Szatmari, P., Garon, N., Vaillancourt, T., & Roncadin, C. (2021). Assessment of autism symptoms from 6 to 18 months of age using the autism observation scale for infants in a prospective high‐risk cohort. *Child Development*, *92*, 1187-1198. doi.org/10.1111/cdev.13485