Measuring changes in autism: Focus on social communication

Catherine Lord, Ph.D.
George Tarjan Distinguished Professor
UCLA David Geffen School of Medicine
Semel Institute of Neuroscience and Human Behavior

Conflicts of Interest

- I receive royalties from Western Psychological Services for diagnostic and screening instruments. I am on scientific advisory boards of the Autism Science Foundation, Autism Speaks, the Child Mind Institute, Tilray, Springtide and Gateway Learning Groups. I am working with Roche and GW Pharmaceuticals/Greenwich Biosciences/Signant on particular projects. I have research funding from NICHD, NIMH, NIDCD, DoD and the Simons Foundation.
- The BOSCC is currently available without charge with permission of WPS, which owns the copyright and Making Better Measures, my little business that sends out kits and offers training (not for profit).
- The BOSA is available for free via an online webinair with permission from WPS but requires a competent ADOS user to code it (not to administer it).

UCLA

Goals for this talk

- Learn and think about formal goals (standardized measures) and your goals as a clinician for treatment, particularly of autism

- Think about how we can measure if we're meeting these goals?
  - Talk about what studies have shown change: how much change, what change and what proportion of patients change?

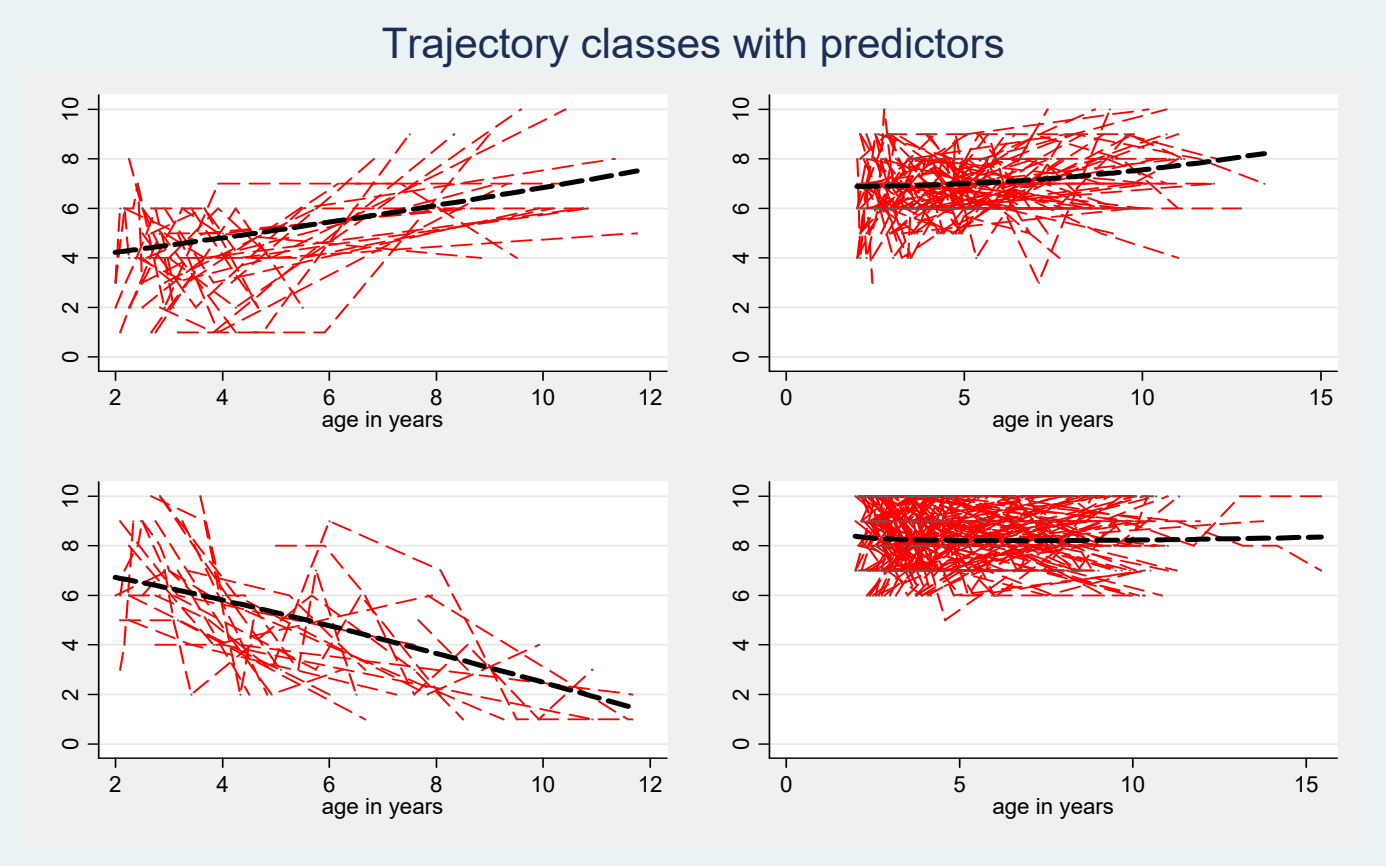
- Tell you a little about what we're trying to do with new measures (both the ups and downs)

What are our treatment goals in autism?

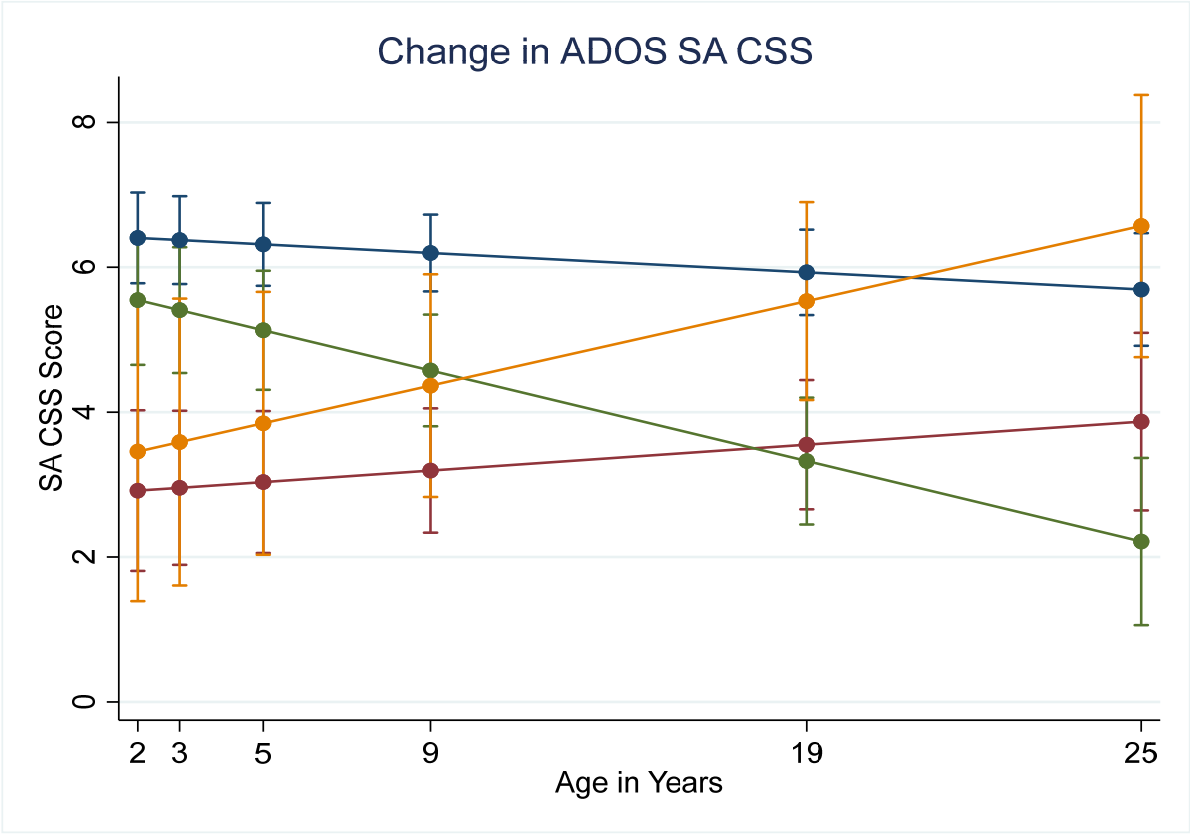
- **No autism? Language? Social interaction? Quality of life? Less pain? More joy? Access to adequate services**
- Avoiding or diminishing secondary problems? Comorbidity with depression, ADHD, OCD; anxiety disorders
- Building on strengths? Finding own resources? Executive functioning, motivation, initiative
- **Importance of a social network from preschool through adulthood and the importance of families**

Trajectory Classes from a longitudinal study (n=253)

Modeled with Risk Variables (VIQ, NVIQ, Gender, Race) (Gotham et al, 2015)



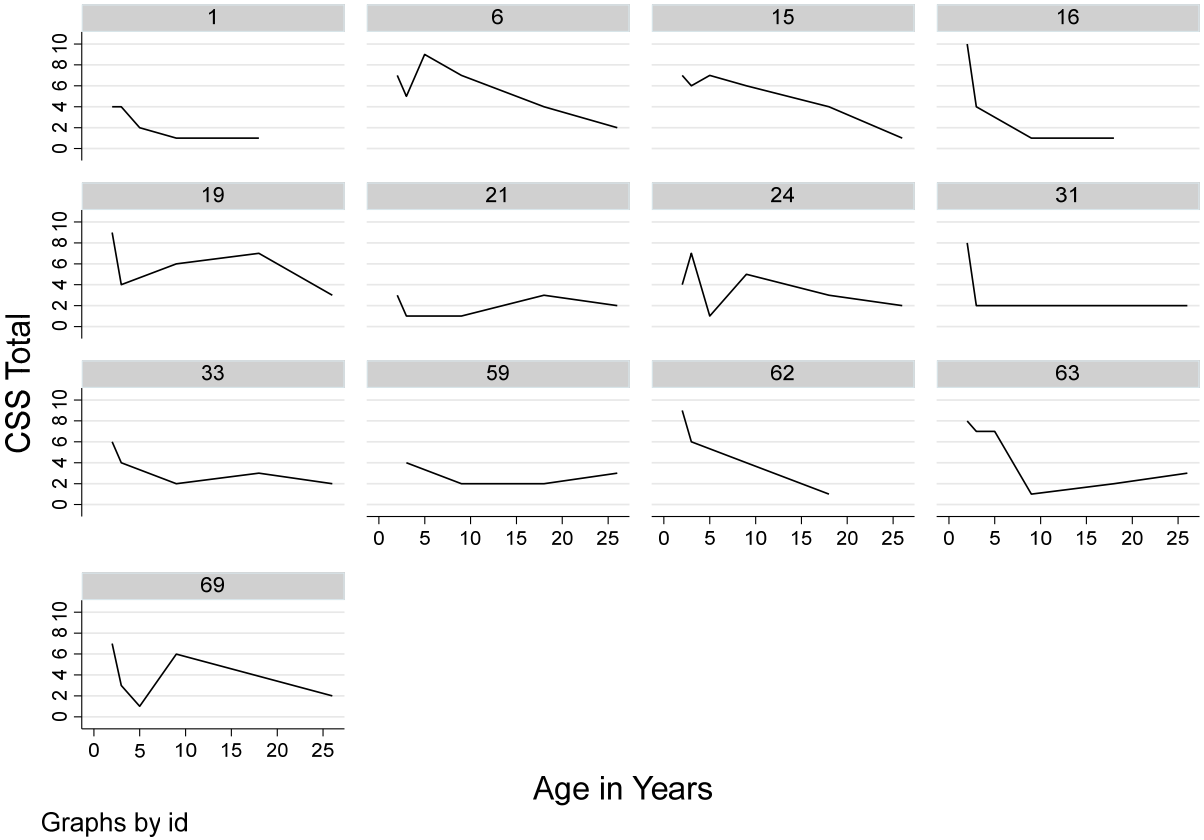
Trajectories by Diagnosis Group (MA; *n*=69)



Elias & Lord, In Preparation

Individual Trajectories by Diagnosis Group

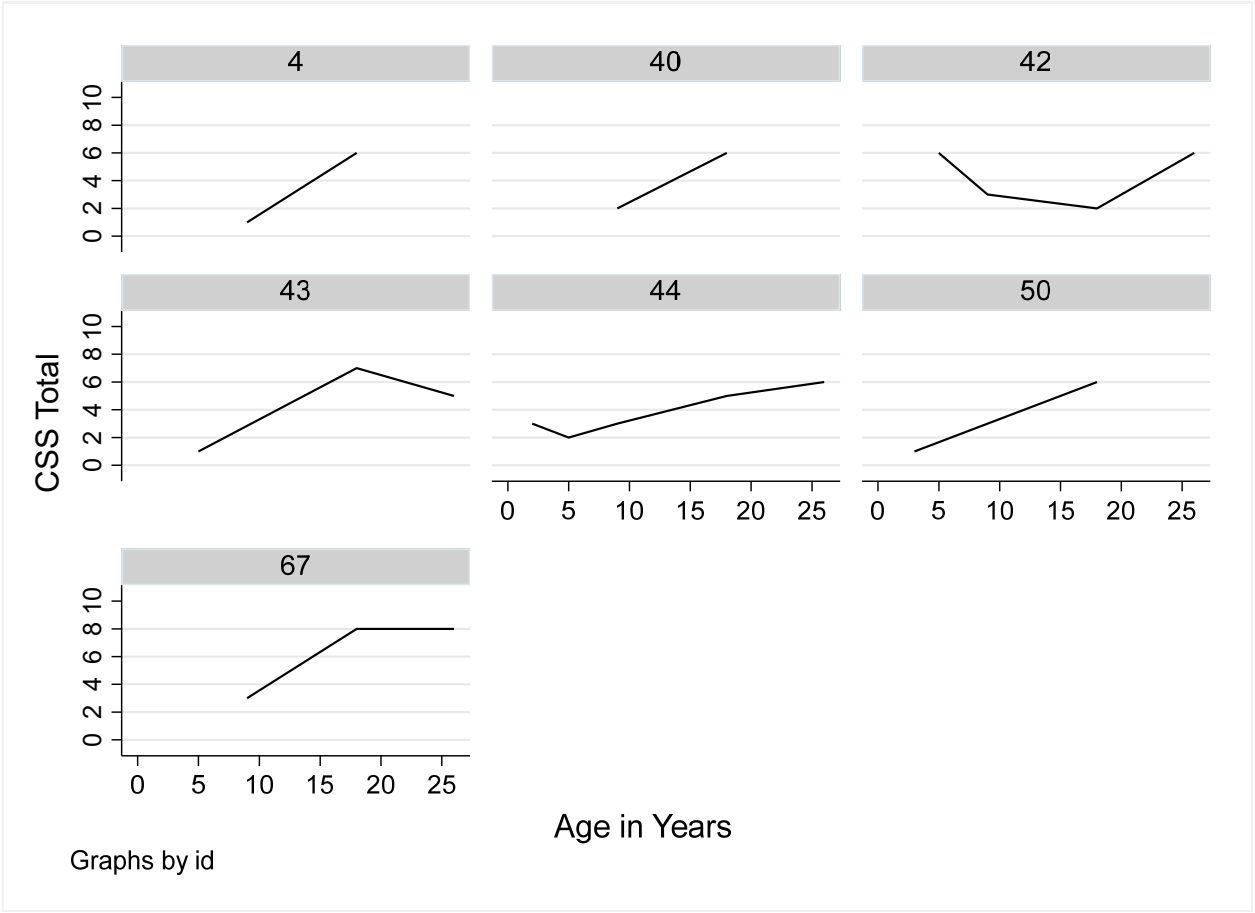
Lost Diagnosis-MA (n = 13)



Elias & Lord, In Preparation

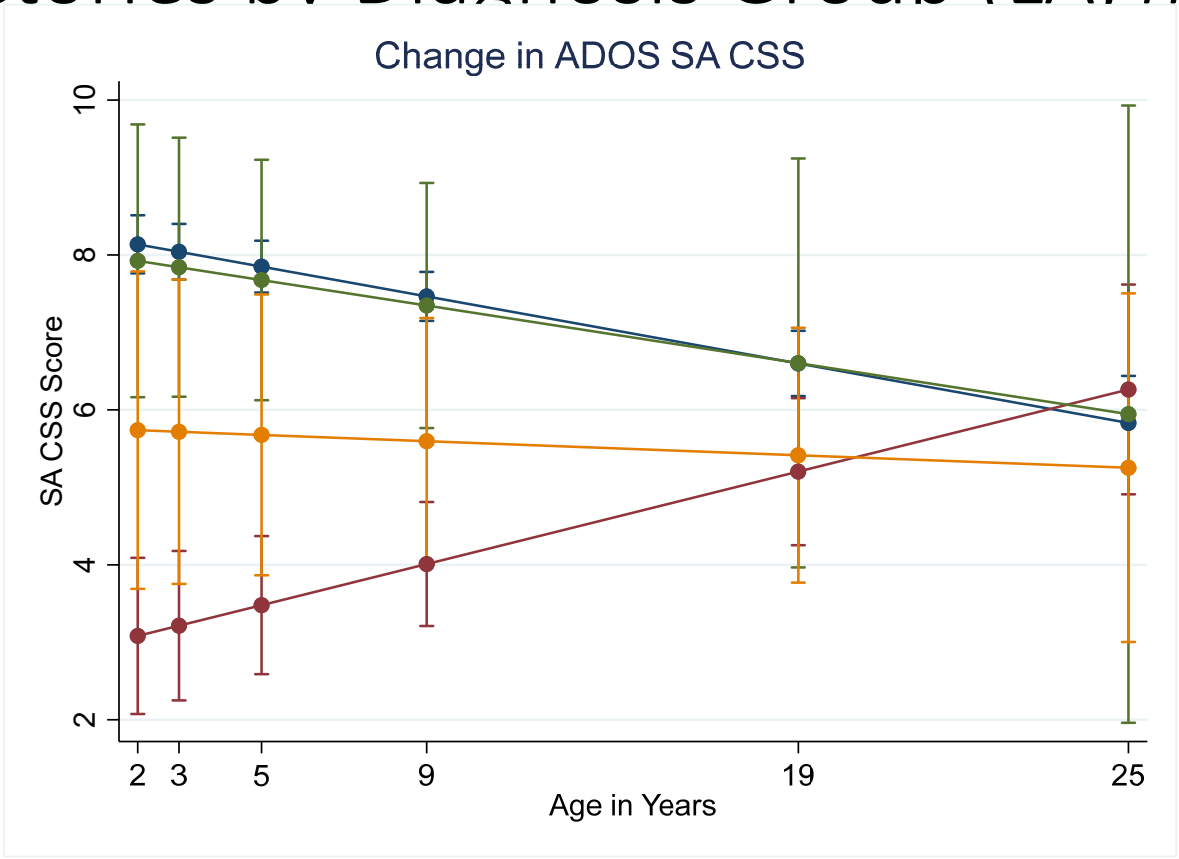
Individual Trajectories by Diagnosis Group

Gained Diagnosis-MA (n = 7)



Elias & Lord, In Preparation

Trajectories by Diagnosis Group (LA: *n*=86)



Elias & Lord, In Preparation

Longitudinal Trajectories of Co-Occurring Symptoms

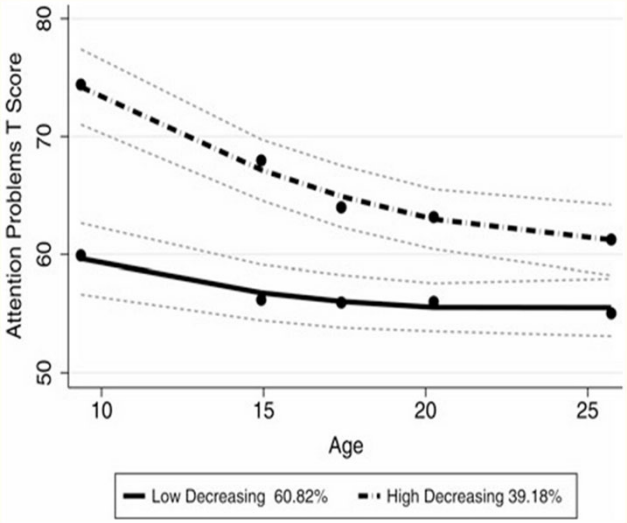


Figure 1.

Trajectory classes derived from the attention problems symptom T score. Both groups have significant quadratic effects in the model. Dots represent observed means, lines represent estimated trajectories, and dotted lines represent the 95th confidence interval for estimated trajectories.

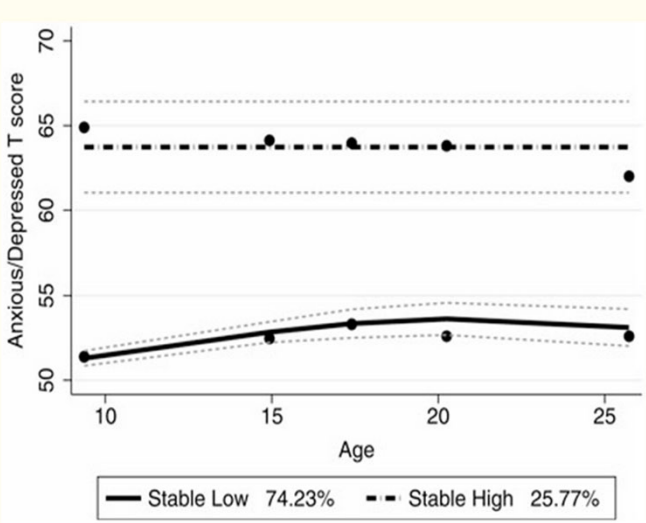


Figure 2.

Trajectory classes derived from the anxious/depressed symptom T score. Group 1 has a significant quadratic slope, and Group 2 is an intercept model. Dots represent observed means, lines represent estimated trajectories, and dotted lines represent the 95th confidence interval for estimated trajectories.

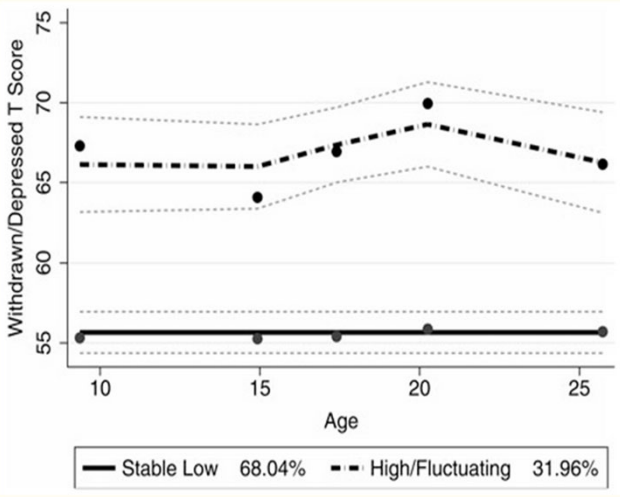
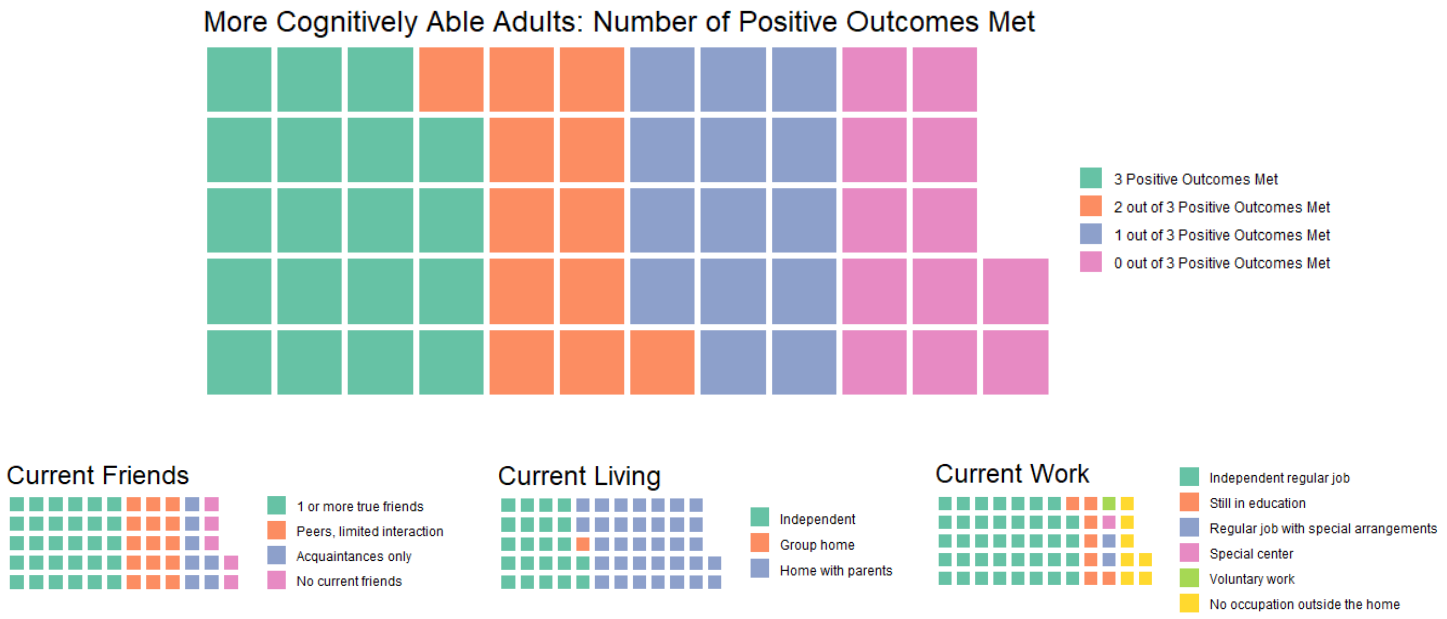


Figure 3.

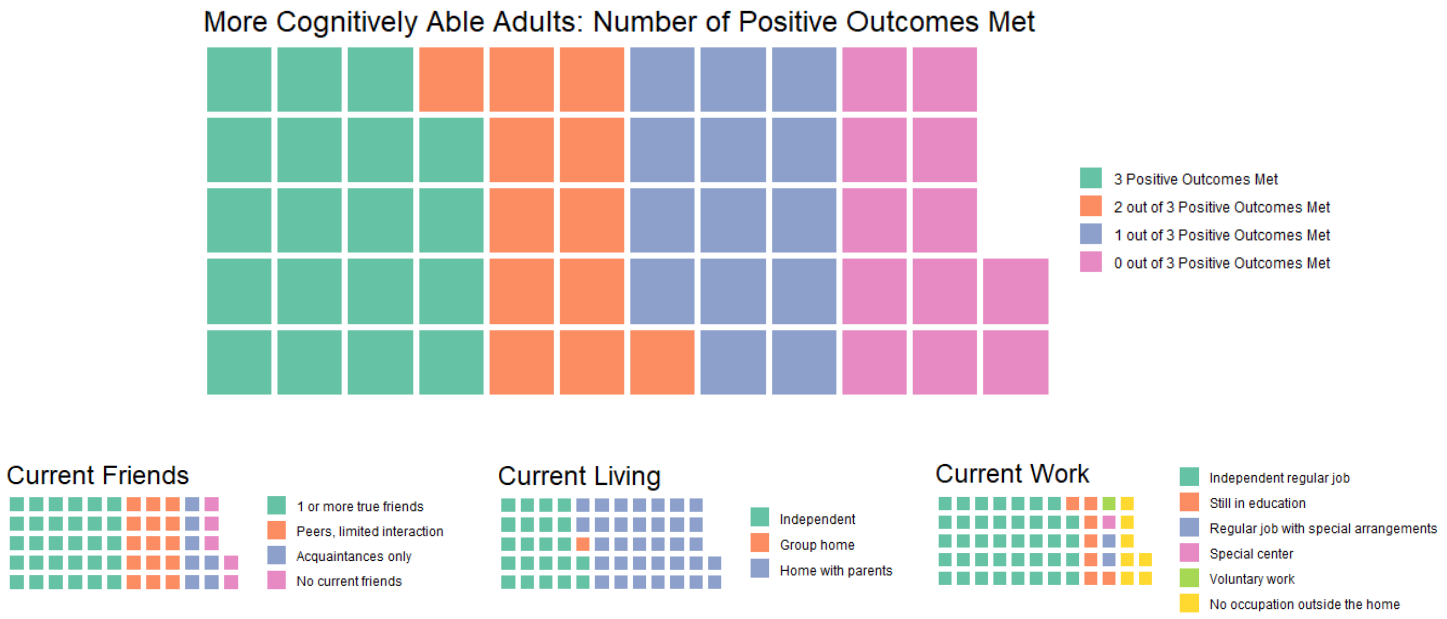
Trajectory classes derived from the withdrawn/depressed symptom T score. Group 1 is typified by a significant intercept model, Group 2 is typified by a significant cubic model. Dots represent observed means, lines represent estimated trajectories, and dotted lines represent the 95th confidence interval for estimated trajectories.

McCauley JB, Elias R, Lord C. Trajectories of co-occurring psychopathology symptoms in autism from late childhood to adulthood. Dev Psychopathol. 2020 Oct;32(4):1287-1302..

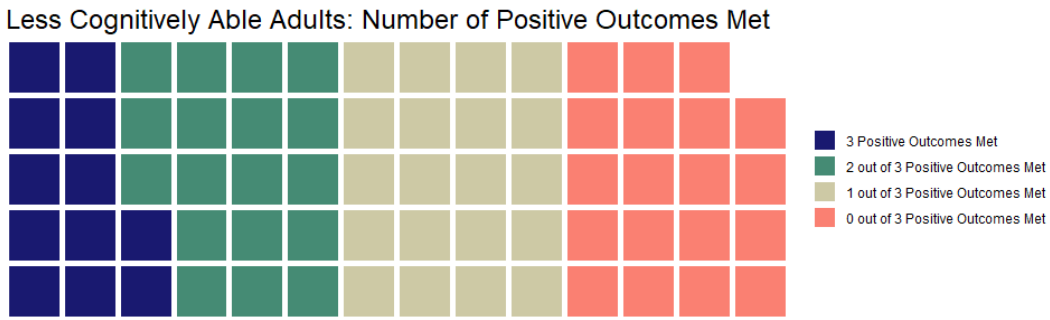
Adult outcomes for adults with autism without intellectual disability

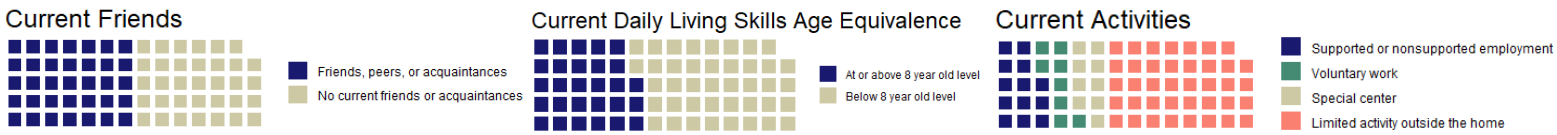


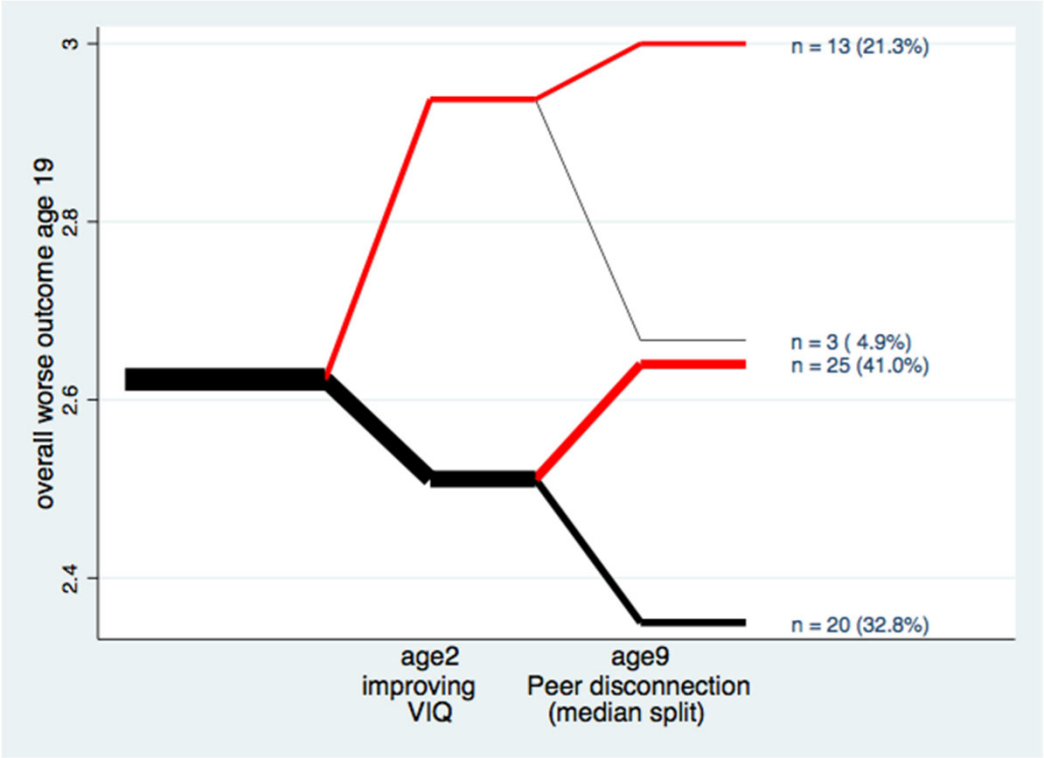
Adult outcomes for adults with autism without intellectual disability



Outcomes for autistic adults with limited language or intellectual abilty






overall worse outcome age 19
3
2.8
2.6
2.4
n = 13 (21.3%)
n = 3 (4.9%)
n = 25 (41.0%)
n = 20 (32.8%)
age2
improving
VIQ
age9
Peer disconnection
(median split)

Conclusions so far

- Changes are gradual
- Many start early but there are different patterns
- Some factors (e.g., the ability to speak well enough to talk to strangers; the ability to comply with requests and sit still long enough to participate in regular education or community activities; better academic skills on tests; sufficient social skills and executive functioning to have a paid part-time job) predict greater independence later
- Subjective measures (happiness, well-being, quality of life) are different from objective measures (work, residence, school, friends)

How much change should we expect? How fast? Some speculations:

- ASD involves difficulties in production of basic social behaviors and response to social and other information.
- ASD also involves seeking of repetition or predictability, unusual in intensity or type, often associated with various sensory differences.
- Also often associated with these difficulties are motor problems, language delay, anxiety, aggression, organization/executive functioning problems and attention to details, as well as intellectual disability.

- Our treatments can address any of these individual factors, as well as socio-cultural and familial issues that may make a difference.

Goals for this talk

- Learn and think about formal goals (standardized measures) and your goals as a clinician for treatment, particularly of autism

- Think about how we can measure if we're meeting these goals?
  - Talk about what studies have shown change: how much change, what change and what proportion of patients change?

- Tell you a little about what we're trying to do with a new measure (both the ups and downs)

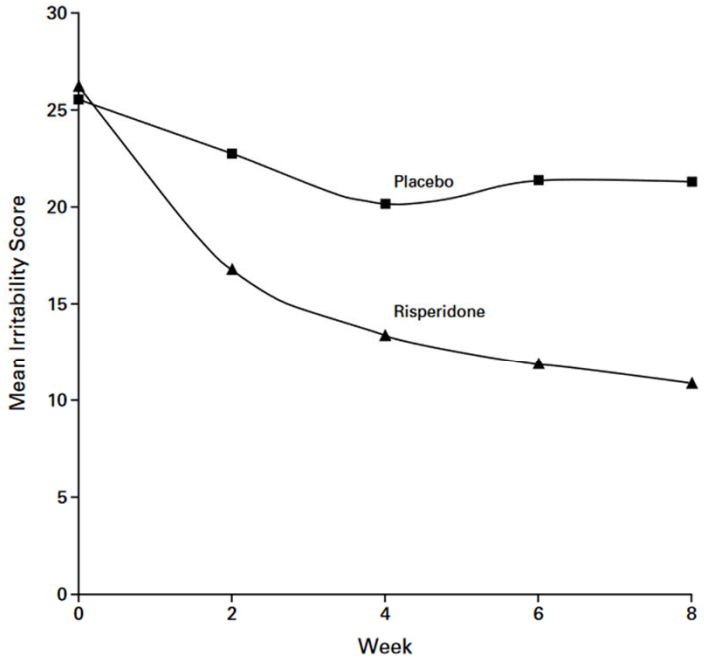
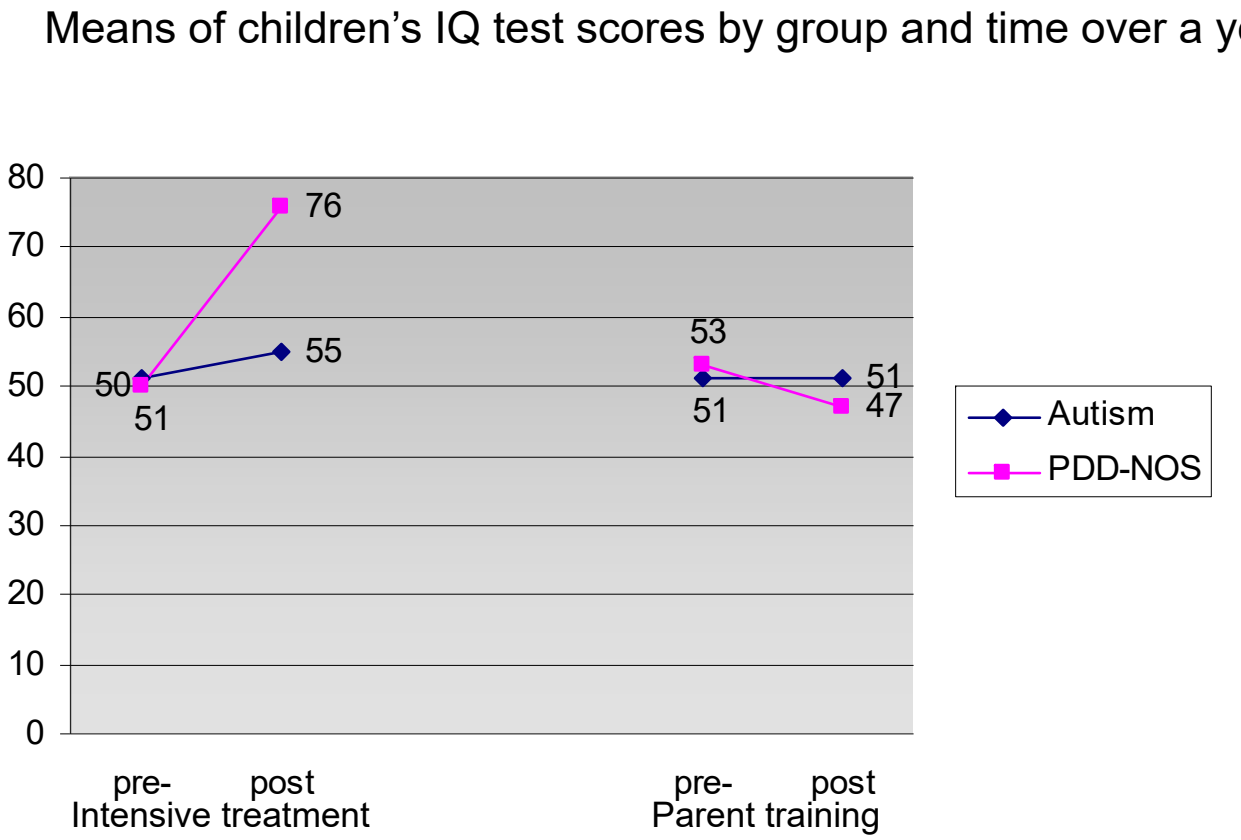


Figure 1. Mean Scores for Irritability in the Risperidone and Placebo Groups during the Eight-Week Trial. Data are for all 101 children (49 assigned to the risperidone group and 52 assigned to the placebo group). Higher scores indicate greater irritability.

McCracken JT, McGough J, Shah B, Cronin P, Hong D, Aman MG, Arnold LE, Lindsay R, Nash P, Hollway J, McDougle CJ, Posey D, Swiezy N, Kohn A, Scahill L, Martin A, Koenig K, Volkmar F, Carroll D, Lancor A, Tierney E, Ghuman J, Gonzalez NM, Grados M, Vitiello B, Ritz L, Davies M, Robinson J, McMahon D; Research Units on Pediatric Psychopharmacology Autism Network. Risperidone in children with autism and serious behavioral problems. N Engl J Med. 2002

Smith's 30 hours a week ABA trial



Adapted from: Smith, Groen & Wynn (2000)

Early Start Denver Model (first RCT: Dawson et al, 2009)

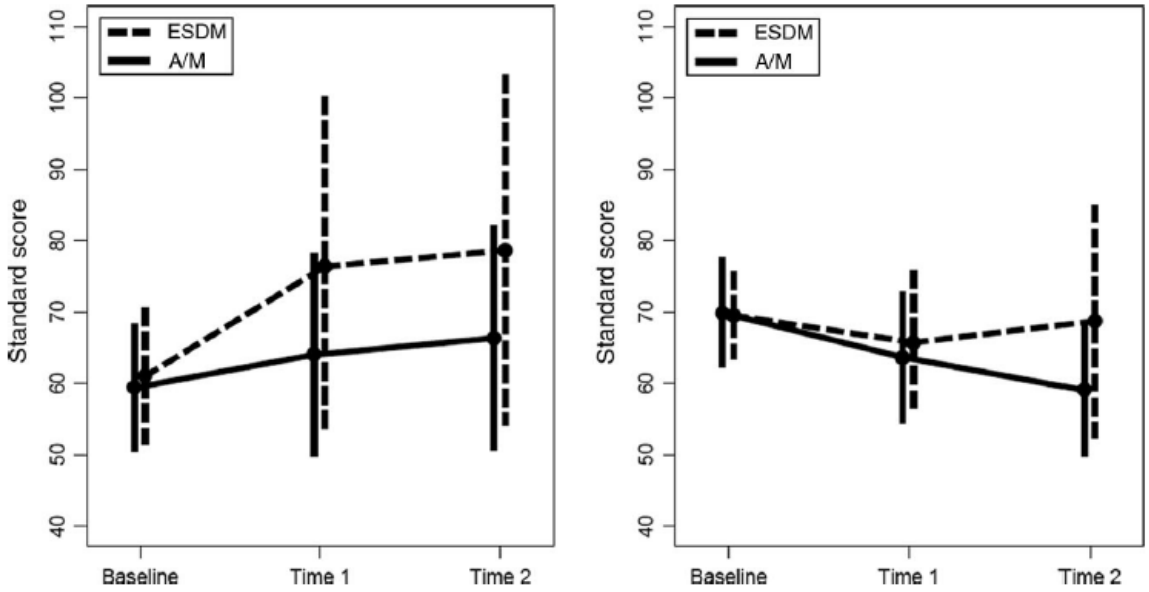
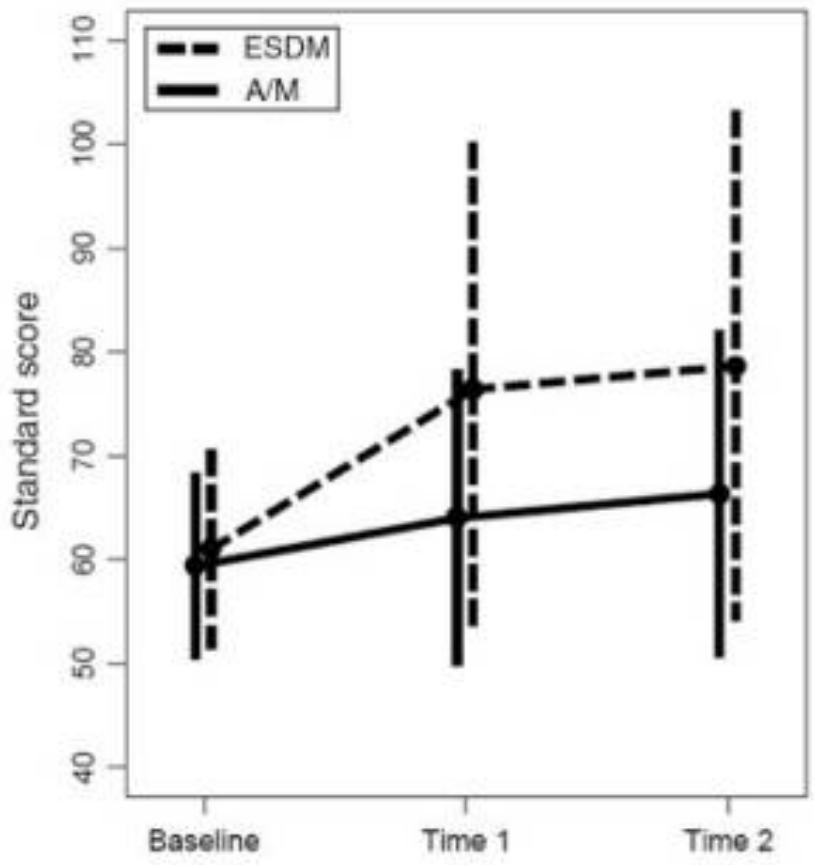


FIGURE 2

Mean scores on the MSEL (left) and the VABS composite (right) for children in the ESDM and A/M groups 1 and 2 years after entering study. Error bars indicate ±1 SD.

ESDM

- Change seems to be driven by language gains on the MSEL @ 2 years
  - ESDM
    - Receptive: Increased 18.9 points
    - Expressive: Increased 12.1 points
  - A/M
    - Receptive: 10.2 points
    - Expressive: 4.0 points



Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., ... & Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics*, *125*(1), e17-e23.

JASPER

<u>Total socially communicative utterances at Week 12 (midpoint)</u>:

JASP+EMT+SGD: 54.4 utterances

JASP+EMT: 35.3 utterances

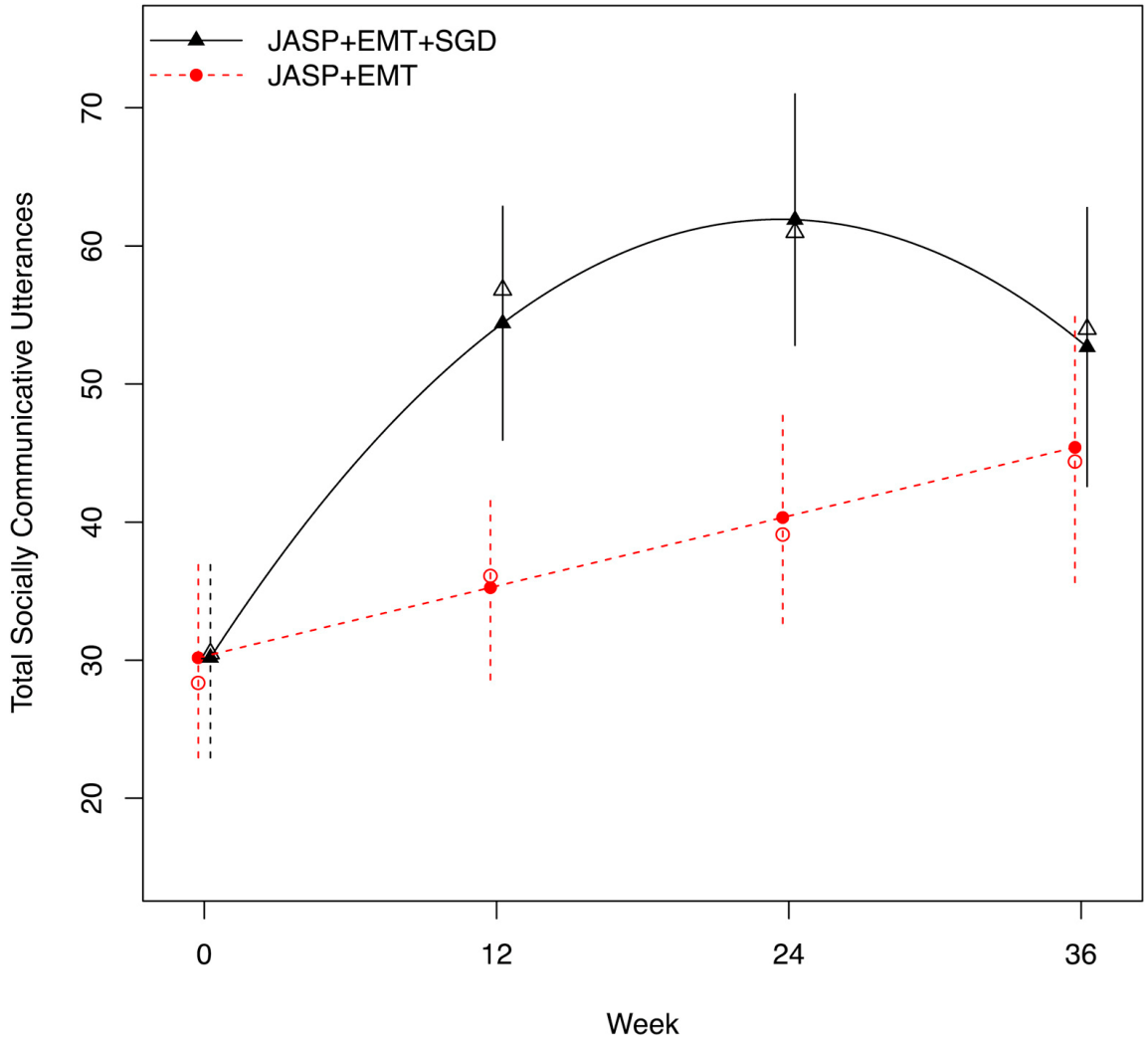
Treatment effect size: .57

<u>Total socially communicative utterances at Week 24 (end of treatment)</u>:

JASP+EMT+SGD: 61.9 utterances

JASP+EMT: 40.3 utterances

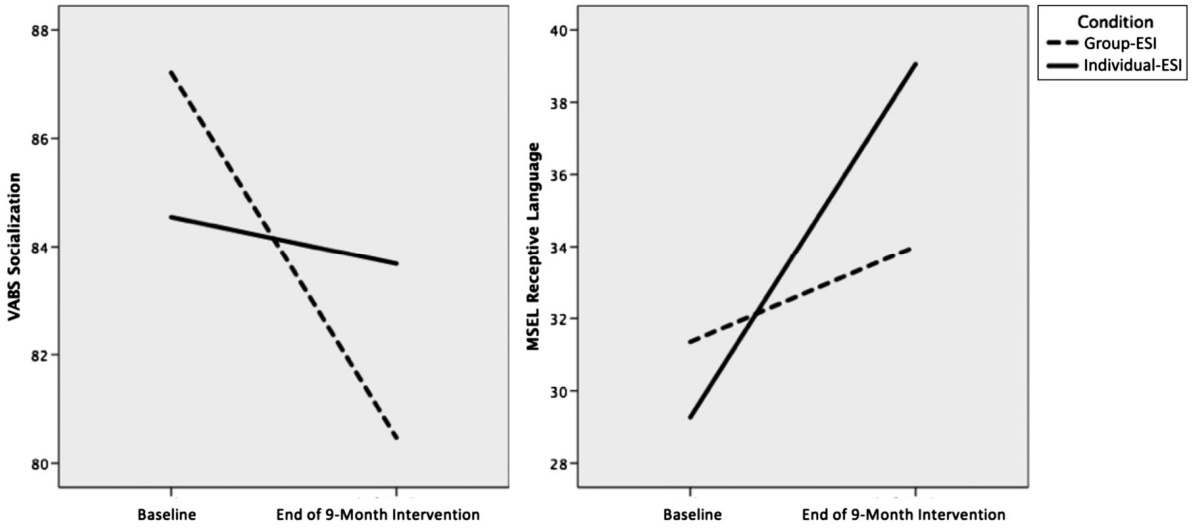
Treatment effect size: .62



Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., ... & Almirall, D. (2014). Communication interventions for minimally verbal children with autism: A sequential multiple assignment randomized trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(6), 635-646.

Early Social Interaction (ESI)

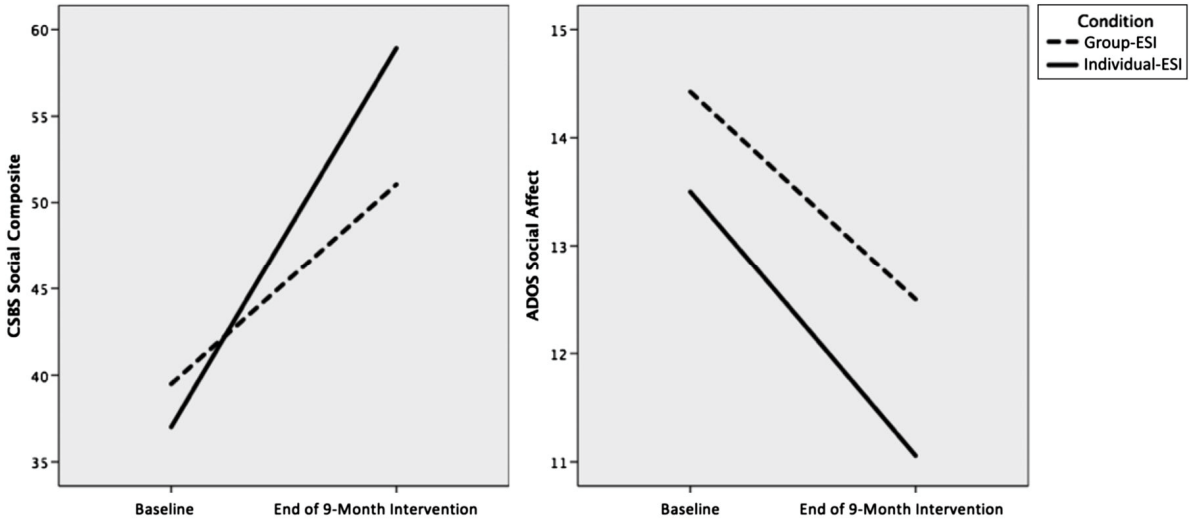
- VABS Communication
  - Individual ESI: Δ11.8
  - Group ESI: Δ3.1
- VABS Socialization
  - Individual ESI: Δ-.9
  - Group ESI: Δ-5.6
- MSEL receptive language
  - Individual ESI: Δ10.1
  - Group ESI: Δ2.8



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

Early Social Interaction (ESI)

- CBCS Social Composite
  - Individual ESI: Δ22.3
  - Group ESI: Δ12.4
- ADOS Social Affect
  - Individual ESI: Δ-.2.6
  - Group ESI: Δ-2.1



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

PACT: Long-term follow-up

Pickles, A., Le Couteur, A., Leadbitter, K., Salomone, E., Cole-Fletcher, R., Tobin, H., ... & Green, J. (2016). Parent-mediated social communication therapy for young children with autism (PACT): long-term follow-up of a randomised controlled trial. *The Lancet*, *388*(10059), 2501-2509.

	PACT intervention	Treatment as usual
Autism symptoms (ADOS CSS*)		
Baseline	8·0 (1·4)	7·9 (1·4)
Post-treatment	6·7 (1·7)	7·3 (1·6)
Follow-up	7·3 (2·0)	7·8 (1·8)
Child initiations (dyadic)†		
Baseline	22·7% (18·8)	26·1% (18·7)
Mid-treatment	41·0% (21·8)	30·1% (18·8)
Post-treatment	34·0% (18·7)	27·2% (17·6)
Follow-up	30·1% (17·5)	26·7% (17·0)
Language composite‡		
Follow-up	84·8 (38·6)	80·0 (40·0)
Parent synchrony (dyadic)§		
Baseline	30·7% (14·2)	31·1% (16·0)
Mid-treatment	53·8% (20·2)	33·8% (14·5)
Post-treatment	53·0% (20·9)	33·4% (14·4)
Follow-up	44·4% (16·1)	43·1% (15·7)
Conversation turns (dyadic)¶		
Follow-up	28·3 (24·4)	26·2 (19·4)
Teacher-rated adaptive behaviour‖		
Follow-up	66·3 (21·3)	60·4 (16·6)
Parent-rated adaptive behaviour‖		
Baseline	65·3 (8·1)	65·5 (9·0)
Post-treatment	67·5 (13·0)	65·2 (12·2)
Follow-up	63·2 (14·8)	60·7 (11·3)
Social communication (SCQ score)**		
Follow-up total	27·4 (5·8)	29·0 (5·1)
Repetitive behaviour (RBQ score††)		
Sensory-motor (follow-up)	4·8 (3·4)	8·3 (4·2)
Sameness (follow-up)	7·1 (4·3)	11·6 (6·3)

(Table 2 continues in next column)

	PACT intervention	Treatment as usual
(Continued from previous column)		
Strength and difficulties (SDQ score‡‡)		
Peer problems (follow-up)	5·0 (2·4)	5·7 (1·7)
Prosocial (follow-up)	4·7 (2·9)	3·7 (2·7)
Co-occuring disorder at follow-up (DAWBA§§)		
Depression	2/50 (4%)	3/44 (7%)
Conduct/oppositional disorder	17/50 (34%)	17/44 (39%)
Hyperkinesis	6/50 (12%)	7/44 (16%)
Anxiety/OCD	12/50 (24%)	16/46 (35%)

Social ABCs

Table 3. Performance on Video-Coded Outcome Measures by Group at Baseline, Post-Training, and Follow-up

	Social ABCs			Control				
Variable Mean (SD)	Baseline[a]	Post-Training[b]	Follow-Up[c]	Baseline[d]	Post-Training[e]	Follow-Up[f]	R^2 (*P value*)	Significant* Follow-up Contrasts
Responsiveness to caregiver prompts (%)	21.64 (18.99)	66.50 (19.98)	59.88 (20.85)	25.30 (19.4)	30.99 (21.52)	40.21 (22.04)	0.42 (<.001)	b>e, c>f b>a
Initiations (rate/minute)	.36 (.33)	1.06 (.92)	.93 (.70)	.46 (.47)	.36 (.44)	.43 (.42)	0.28 (<.001)	b>e, c>f b>a
Shared smiling (% intervals)	4.26 (7.06)	8.32 (9.59)	8.29 (9.18)	6.66 (9.46)	7.71 (8.17)	5.74 (6.43)	0.04 (.143)	b>a
Parent smiling (% intervals)	26.63 (19.99)	38.35 (20.38)	34.76 (22.38)	31.12 (19.62)	31.29 (19.24)	27.72 (17.45)	0.09 (.017)	b>a
Social orienting (% intervals)	25.31 (16.92)	37.68 (19.74)	37.44 (19.54)	29.67 (19.71)	33.17 (18.81)	30.52 (17.05)	0.06 (.054)	b>a
Parent fidelity (% correct)	48.07 (12.31)	86.97 (6.60)	82.63 (11.52)	51.31 (9.43)	55.28 (8.74)	55.62 (8.99)	0.71 (<.001)	b>e, c>f b>a

Note. For all behaviours ($n = 29–30$ Treatment, 31–32 Controls).

*Significance at corrected alpha = 0.025 (.05/2) for between-group contrasts (Mann–Whitney test) and alpha = 0.0125 (0.05/4) for within-group contrasts (Wilcoxon Signed Rank test).

Brian, J. A., Smith, I. M., Zwaigenbaum, L., & Bryson, S. E. (2017). Cross-site randomized control trial of the Social ABCs caregiver-mediated intervention for toddlers with autism spectrum disorder. *Autism Research*, *10*(10), 1700-1711.

Background: Caregiver report is the most common measure of change in pediatric psychiatry. Yet, placebo rates can be up to 50%, posing significant challenges to interpreting results of clinical trials.

Table 3. Outcome measures for ASD Clinical Trials ranked by Sensitivity to Change

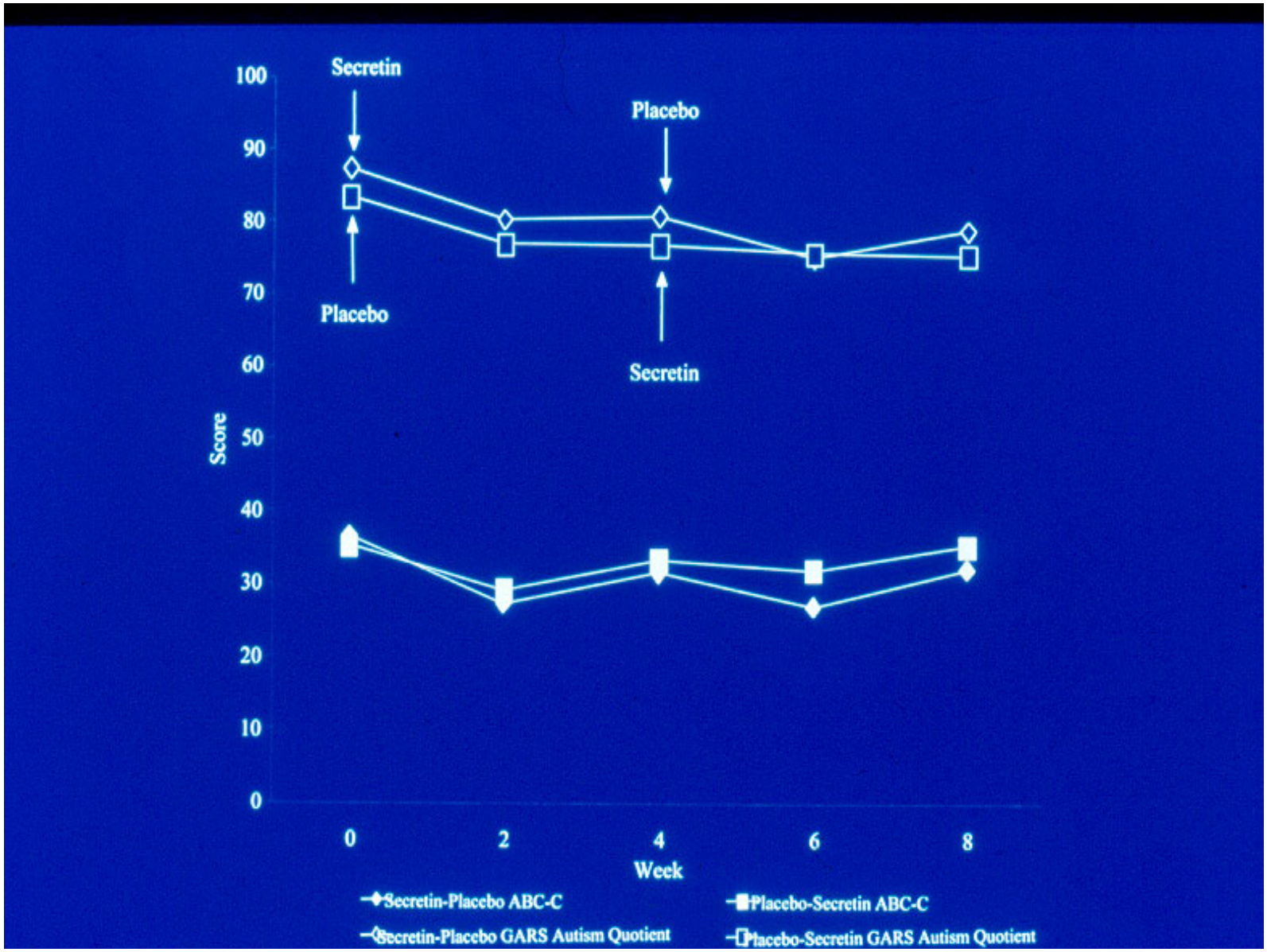
Domain	Instrument	Sensitivity to Change	Informant	Burden
ASD Severity				
	Social Responsiveness Scale - 2 (SRS-2) Total	++	S, P, T	Low
	Brief Observation of Social Communication Change (BOSCC)	+	Ex	Medium
(With Conditions)				
	Aberrant Behavior Checklist (ABC) - Total	+	P, T	Low
	Parent Target Problems	+	Ex	Low
(New/Needs More Data)				
	Autism Behavior Inventory (ABI)	N/A	P, Ex	Low
	SRS-2 Item Response Theory 16-item Short Form	N/A	S, P, T	Low
	SRS-2 11-item Factor Analyzed Short Form	N/A	S, P. T	Low
	Autism Impact Measure (AIM)	N/A	P	Low
	Autism Diagnostic Observation Schedule-2 Calibrated Severity Metric (CSS) (ADOS-2 CSS)	N/A	Ex	High
Social/ Communication				
	Vineland Adaptive Behavior Scales (VABS Communication) II	+++	P, Ex	Medium
	Social Skills Improvement System (SSIS)	++	S, P, T	Low
	Natural Language Sample	++	Ex	Medium
(With Conditions)				
	ABC Lethargy- Social Withdrawal subscale	++	P, T	Low
	Behavior Assessment System for Children- 2 (BASC-2)	++	P, T	Low
	Early Social Communication Scale (ESCS-JAMES)	+	Ex	Medium
	Communication and Symbolic Behavior Scales (CSBS)	+	Ex	High
New/Needs More Data				
	Autism Impact Measure (AIM)	N/A	P	Low
	Pervasive Developmental Disorder-Behavior Inventory (PDD-BI)	N/A	P	Low
	Social Communication Interaction Test (SCIT)	N/A	Ex	Medium
Repetitive Behaviors/Restricted Interests				
	Children's Yale-Brown Obsessive Compulsive Scale-Modified for ASD	+++	Ex	Medium
	Aberrant Behavior Checklist-Stereotypy subscale	+++	P, T	Low
	Repetitive Behavior Scale-Revised	++	P	Low
(With conditions)				
	Stereotyped Behavior Checklist	N/A	P	Low
	Repetitive Behavior Questionnaire	N/A	P	Low
New/Needs More Data				
	Autism Behavior Inventory	N/A	P, Ex	Low
Anxiety				
	Anxiety Disorders Interview Schedule--Clinician Severity Rating (ADIS-CSR)	+++	Ex	High
	Pediatric Anxiety Rating Scale (PARS)	+++	Ex	Medium
(With conditions)				
	Multi-dimensional Anxiety Scale for Children - Parent (MASC-P)	+	S, P	Low
	Child and Adolescent Symptom Inventory -4 21-item Anxiety Scale (CASI-4-ANX	N/A	Ex	Medium
(New/Needs more data)				
	Parent-Rated Anxiety Scale for ASD (PRAS-ASD)	N/A	P	Low
	Spence Children's Anxiety Scale Parent (SCAS-P)	N/A	P	Low

Legend:

+ - one positive comparative study or mixed **++ - >1 study showing change to active treatment** **+++ - multiple studies showing robust, consistent changes**

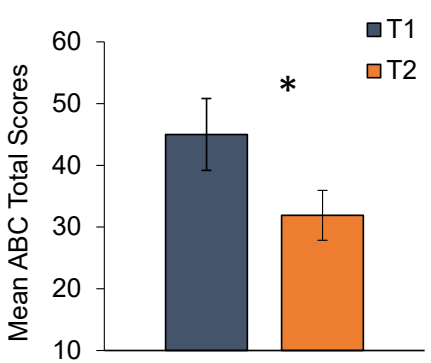
S - Self **P - Parent/ Caregiver** **EX- examiner** **T - Teacher**

European J of Neuropsychopharmacology, McCracken, et al


Secretin
Placebo
Placebo
Secretin
100
90
80
70
60
50
40
30
20
10
0
Score
0
2
4
6
8
Week
Secretin-Placebo ABC-C
Placebo-Secretin ABC-C
Secretin-Placebo GARS Autism Quotient
Placebo-Secretin GARS Autism Quotient

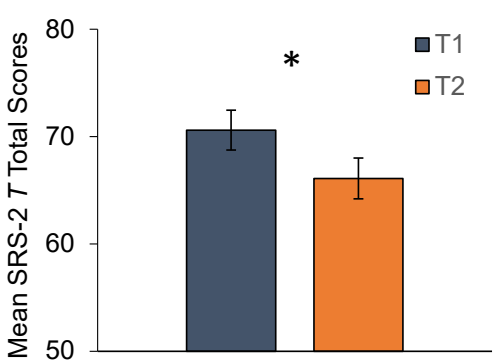
Placebo-like response without treatment

Aberrant Behavior Checklist (ABC)



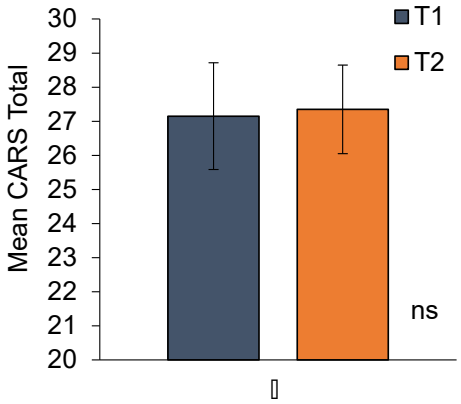
29% Decrease

Social Responsiveness Scale (SRS)



7% Decrease

Clinical Ratings (CARS)



80 – 85% of caregivers reported an improvement in their child's behavior in the absence of clear intervention or treatment

Jones et al. *2017*

What are our treatment goals?

- No autism? Language? Quality of life? Less pain? More joy? Access to adequate services
- Avoiding secondary problems? Comorbidity with depression, ADHD, OCD; anxiety disorders
- Building on strengths? Finding own resources? Executive functioning, motivation, initiative
- Importance of a social network from preschool through adulthood and the importance of families

Goals for this talk

- Learn and think about formal goals (standardized measures) and your goals as a clinician for treatment, particularly of autism

- Think about how we can measure if we're meeting these goals?
  - Talk about what studies have shown change: how much change, what change and what proportion of patients change?
  - Could we do a better job of measuring change? Or do we even need to know if small changes are happening?

- Tell you a little about what we're trying to do with a new measure (both the ups and downs)

Objectives of the BOSCC: Brief Observation of Social Communication - Change

To create an easy to use and code observational instrument that is sensitive to change in social communication behaviors over short periods of time.

To have versions of the instrument for children and adults ranging from toddlers to adults and ranging in language skill from minimally verbal to verbally fluent.

To develop an online coding system to make scoring more efficient.

To use engineering strategies and machine learning to develop automatic codes (e.g., of eye contact, of vocalizations) so that eventually coding could be automatic.

What can the BOSCC potentially do?

- Truly blinded or masked assessment of social communication
- With an instrument that is easy to use and mobile (record on phone)
- That can be used to look at social communication for people from toddlers to adults
- And allow study-outcomes to be compared
- Not intended to replace parent or patient report, clinical impressions or finer grained instruments related to theory behind interventions
- But could allow probes into individual differences in trajectories and responses over time as well as to different treatments

- BOSCC
  - 12- 15 minute videotaped play or informal interaction between child/adult and adult (research assistant, parent, peer)
  - Standardized set of toys or materials
  - MV version: 4 min play, 2 min bubbles, 4 min play, 2 min bubbles
  - V versions: 4 minutes of an activity plus questions, 2 min conversation (repeated) and sometimes bubbles!
  - For Kids Connect: Completed at entry and exit (10 weeks of program enrollment)
  - Can be carried out by parents, naïve research assistants, teachers or therapists and maybe some day by peers (it has easy instructions)

Based on Confirmatory Factor Analysis (CFA) and theory.

ITEM	DOMAIN		
Eye Contact	Basic Social	Social Communication	Overall
Facial Expressions			
Gestures			
Integration of Vocal and Non-Vocal			
Frequency of Social Overtures			
Quality of Social Overtures	Interaction Quality		
Consistency of Social Responses			
Quality of Social Responses			
Conversation			
Offering Information			
Stereotyped and Echoed Speech	RRBs		
Unusual Sensory Interests/Behavior(s)			
Hand and Finger or Complex Body Mannerisms			
Unusual, Repetitive, Stereotyped Interest/Behavior(s), and Compulsions			
Activity Level	Other		
Disruptive or Oppositional Behavior			
Anxious Behavior			

RMSEA= 0.05, CFI = 0.81, TLI = 0.76

Method

- BOSCC Coding
  - 15 items coded on a 6-point scale
  - Items categorized into two domains: social communication (SC) and restricted and repetitive behaviors (RRBs), as well as a core total
  - * Coding training took place during COVID and were conducted remotely
    - % agreement ±1 point=.904
    - ICC_{SC}=.975
    - ICC_{RRB}=.982
    - ICC_{Total}=.985
  - Similar ICCs for inter-rater reliability across studies
  - Coders watch the video twice in 6 minute segments and mark codes
  - The participant cannot be masked but we have tried this with masked examiners

Median BOSCC Scores from Entry to Exit (n = 25)

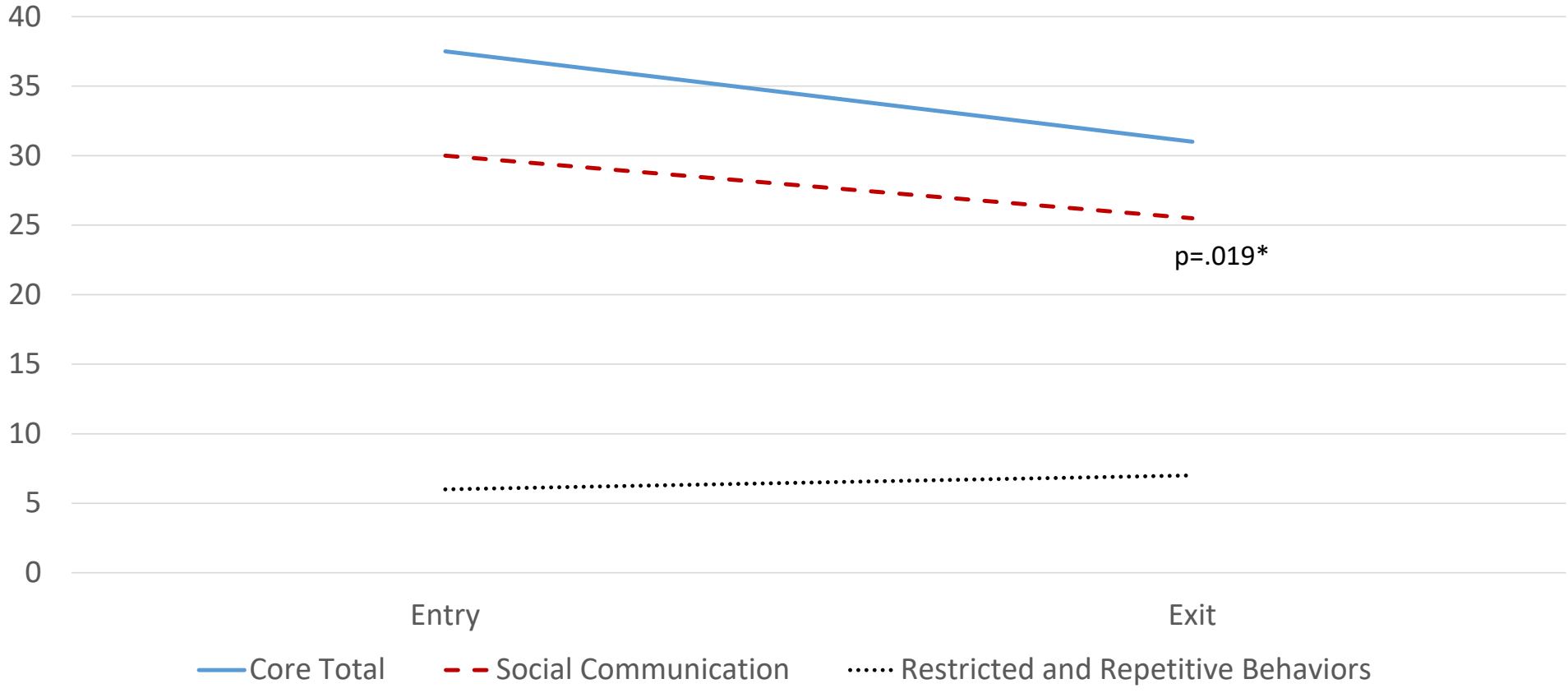
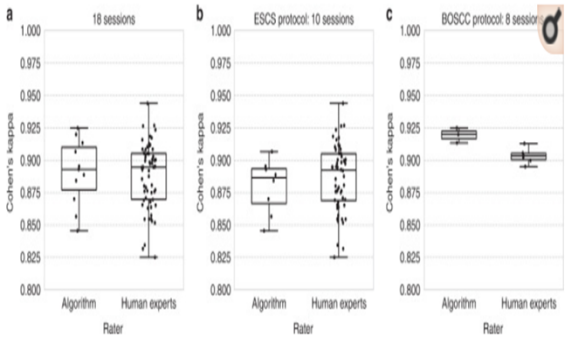


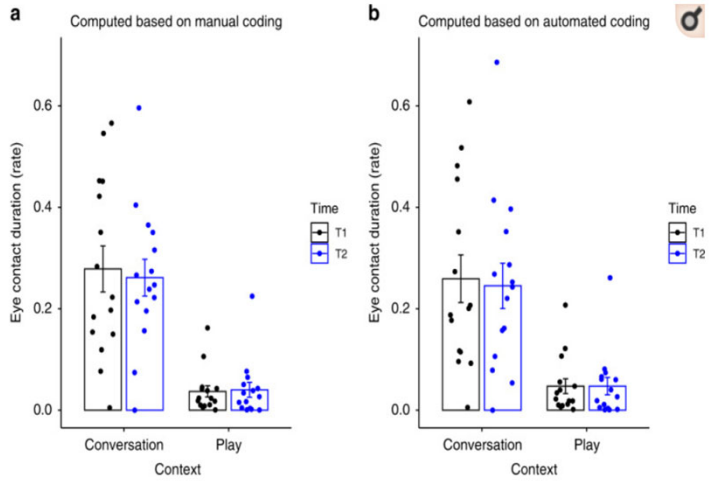
Figure 3 illustrates the Cohen's κ statistics for the human–human and human–detector comparison

Fig. 3



Pairwise Cohen's kappa distributions among all human pairs and human–algorithm pairs, represented as box plot. **a** 18 validation sessions, **b** ESCS, **c** BOSCC. Generally, kappa scores above 0.8 are considered an almost perfect agreement. On all sessions annotated by ten human experts, agreements among humans and agreements between each human and algorithm are similar in terms of kappa values.

Fig. 4

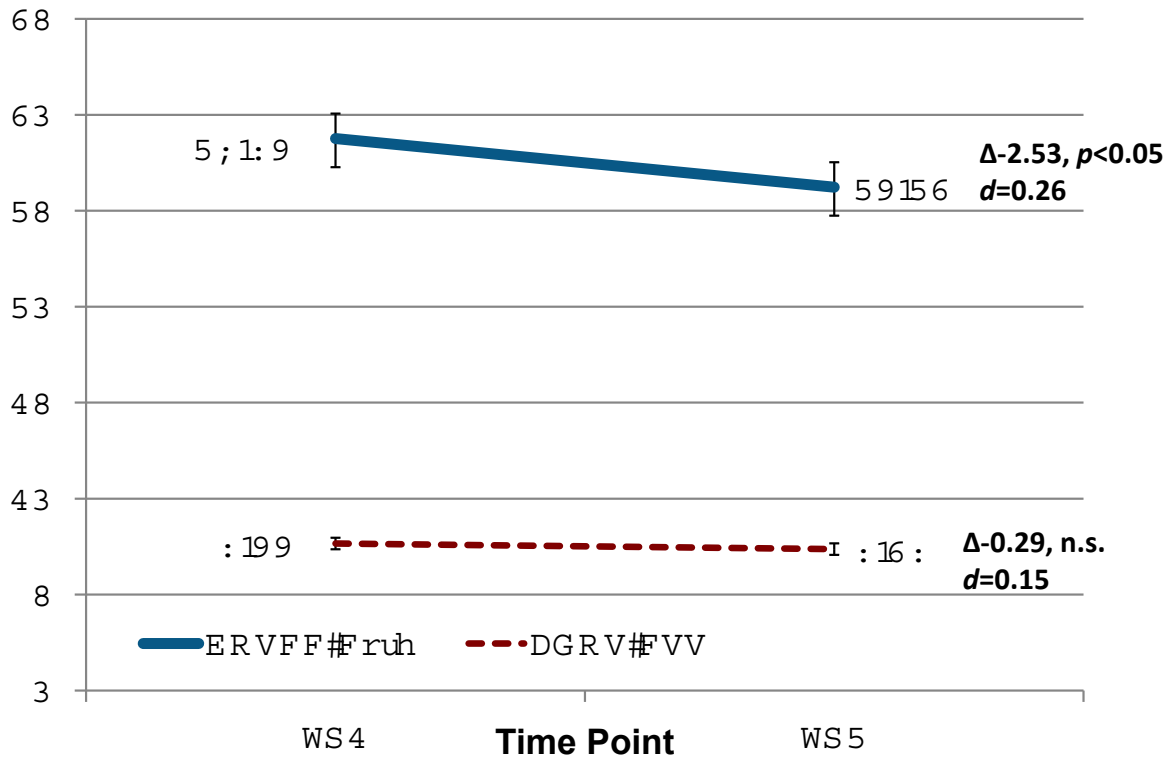


Average duration of eye contact during conversation and interactive play in child and adolescent samples ($n = 15$), measured at time 1 and time 2. **a** based on human coding, **b** based on automated coding. Data are presented as mean values ± SEM.

Chong E, Clark-Whitney E, Southerland A, Stubbs E, Miller C, Ajodan EL, Silverman MR, Lord C, Rozga A, Jones RM, Rehg JM. Detection of eye contact with deep neural networks is as accurate as human experts. Nat Commun. 2020

Results: Validity (JASPER, ESDM, ESI)

- Does the BOSCC measure change?



Est. over 6 mo:
Δ -4.20,
d= -0.37

Est. over 6 mo:
Δ 0.33,
d= 0.15

Grzadzinski et al., 2016

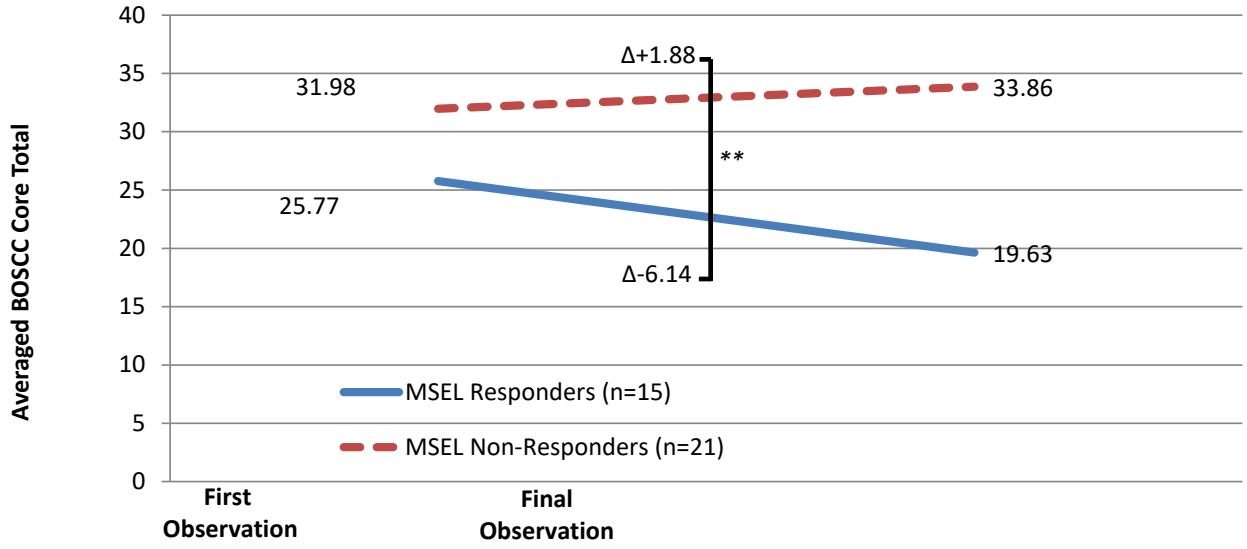
Results: Validity

- Do children that demonstrate change on other standard measures also show change on the BOSCC?

MSEL Receptive Language Groups:

Responders (n=15)= SS increase ≥5 (1/2 SD)

Non-Responders (n=21)



So where are we now?

- There are a number of other studies by us and independent groups showing similar results mostly significant, sometimes not quite.
- The BOSCC detects subtle changes in observed social communication associated with a brief (several months) intervention in MV children with ASD
- Provides different information than parent-reported change in social skills and behavioral problems
  - Potential to provide truly blinded data that can be applied across studies
- Limitations:
  - Lack of a control group in Kids Connect study (waiting to see if we can assess their wait list when covid is over) and original study (analyses under way)

New Missions…

- When we don't get a result, what can we ask?
- Is it that the context in which the BOSCC is applied is not capturing this change (e.g., generalization)?
    - Primary treatment location is 1:1 in school
- Applied the BOSCC coding scheme to 10 minutes of entry and exit treatment sessions
- Is it that the context in which the BOSCC is applied (interactions with mothers who received parent-mediated training) is too good? Do the changes generalize? Are the child's changes related to the parents? (which is good and not good if totally dependent on the immediate changes).

Grzadzinski et al., in prep

New Mission…

BOSCC Domain	TP1	TP2	Δ TP2-TP1	Paired t-test		
	Mean(SD)(SE)	Mean(SD)(SE)	Mean(SD)	t	df	p
SC	29.61 (6.09) (0.99)	27.34 (5.21) (1.15)	-2.27 (5.61)	-2.14	27	0.042
RRB	7.21 (3.27) (0.62)	7.51 (3.28) (0.62)	0.30 (2.90)	0.55	27	0.58
Core	36.82 (6.11) (1.15)	34.86 (5.52) (1.04)	-1.96 (6.16)	-1.69	27	0.10

N=28, preliminary data

6 kids (21%) decreased ≥ 7 (1SD) on **BOSCC SC** domain.

This proportion of improvers is larger than parent-child interaction results.

Grzadzinski et al., in prep

Figure 1. Raw and model-based carer MONSI and child BOSCC social-communication scores.

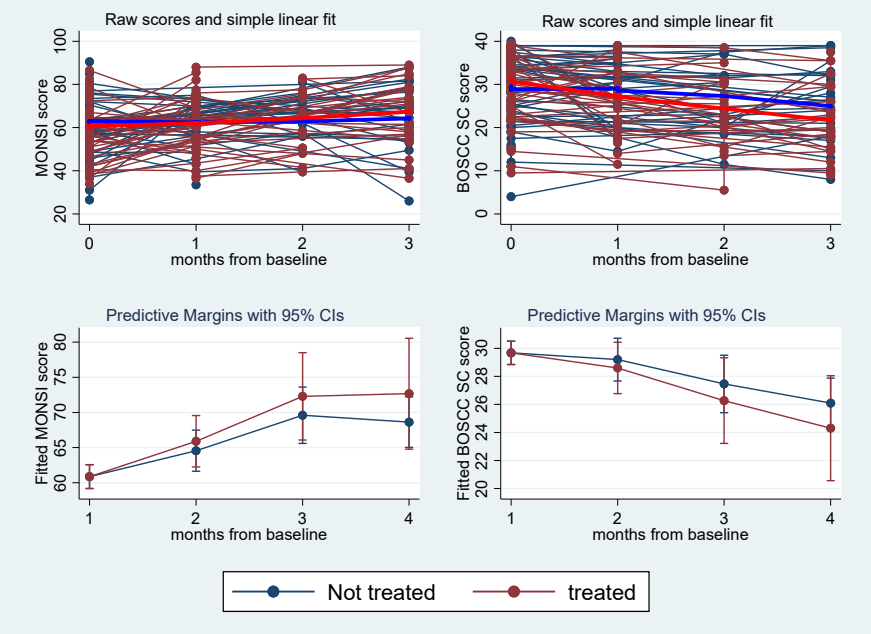
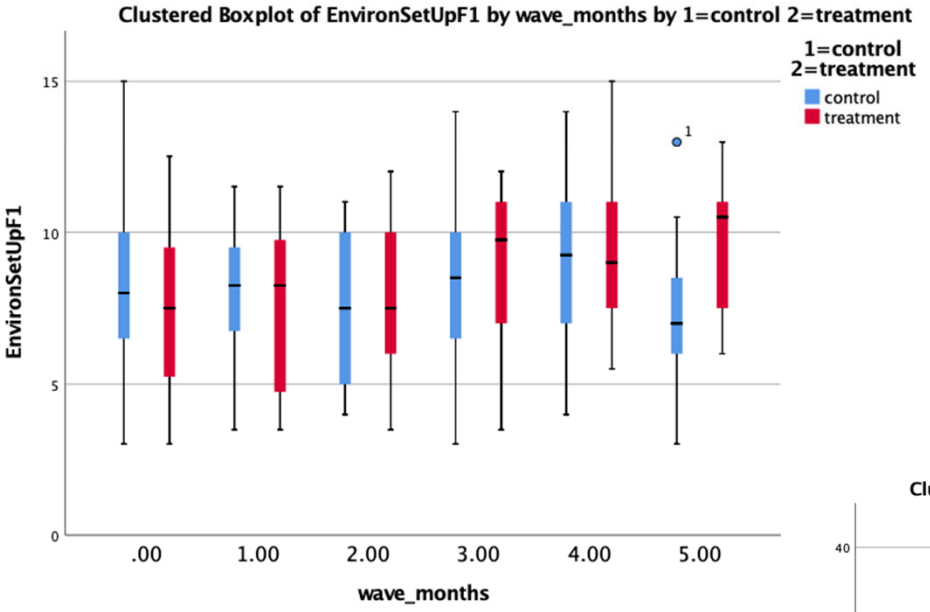
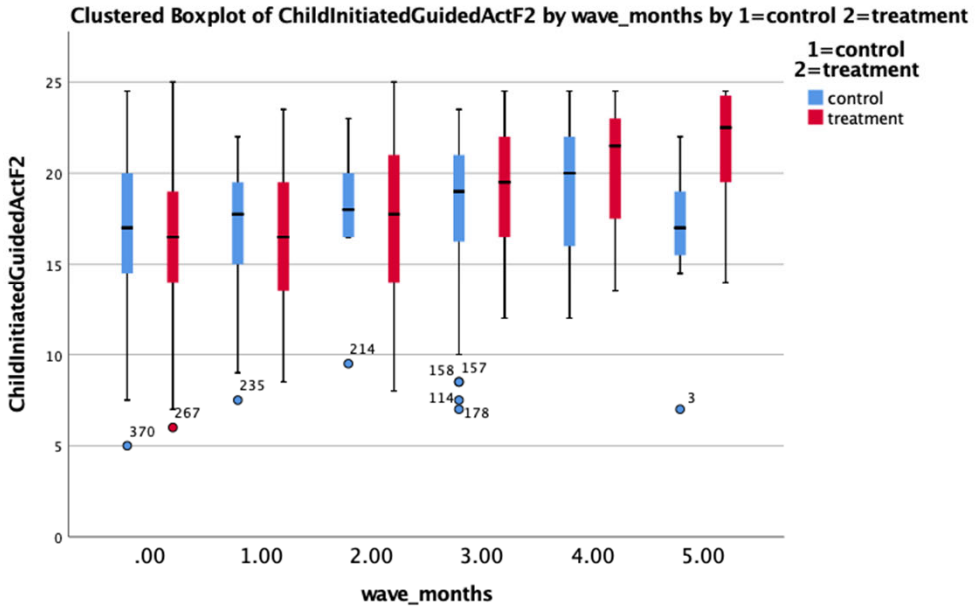


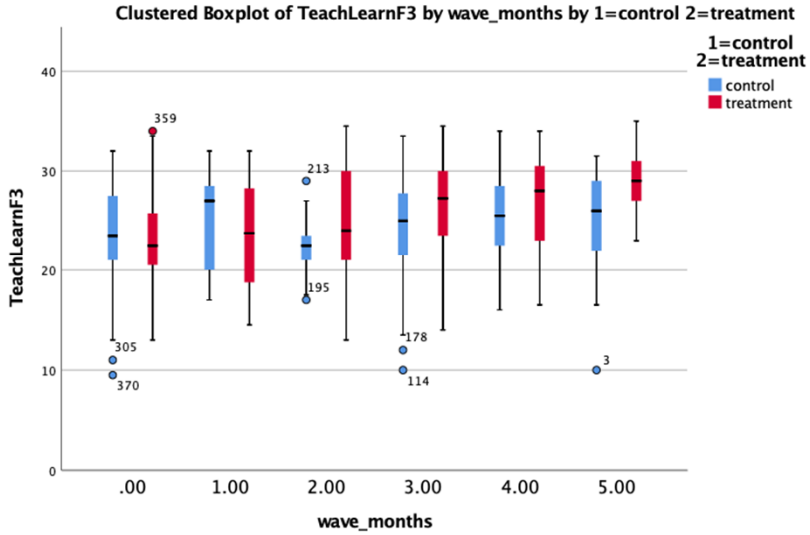
Table: Bivariate growth curve model with direct treatment effects on change from baseline (slope) for BOSCC and MONSI and indirect treatment effect on BOSCC via MONSI (MONSI-BOSCC Baseline/Intercept correlation -0.49)

	Carer (MONSI)			Child (BOSCC social communication)		
Coefficient	Estimate	95% CI	p-value	Estimate	95% CI	p-value
Intercept factor variance	1.0			1.0		
Intercept loading	5.67	4.18 to 7.16	<.001	3.41	2.71 to 4.12	<.001
Error variance	93.0			19.3		
Treatment effect per mth	1.99	0.97 to 3.01	<.001	-0.15	-0.66 to 0.36	0.561
Carer on child				-0.12	-0.18 to -0.05	<.001
Baseline/Intercept covariates						
CSS	-0.12	-0.25 to 0.00	.056	0.13	0.02 to 0.24	.021
Language level	0.18	-0.05 to 0.41	0.127	-0.66	-0.90 to -0.42	<.001
Nonverb IQ	0.27	-0.06 to 0.60	0.114	-0.50	-0.80 to -0.21	.001
Age	1.69	-2.33 to 5.71	0.410	-1.66	-5.09 to 1.77	0.343
Study 2	2.10	12.5 to 2.94	<.001	1.22	0.59 to 1.84	<.001
Study 3	1.38	0.36 to 2.40	.008	1.03	0.17 to 1.88	.019

MONSI Change by domain







What can we offer clinical researchers?

- Providing truly “blinded” or “masked” assessments of social communication (that don’t include general impressions of parents or clinicians for example)
- This is not to say that parent reports, clinical impressions and patient-reported outcomes aren’t incredibly important – but they are more liable to bias.
- Providing access to assessments for labs that don’t have a large behavioral “coding” staff

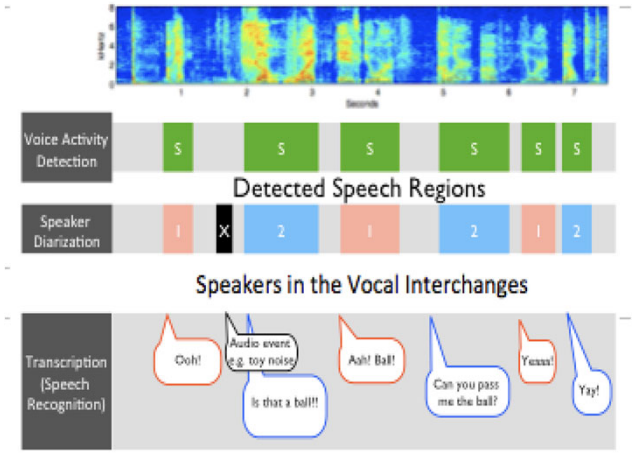
Unpacking Treatment Mechanisms: Combining Evidence from Three Early Intervention Models for ASD

5R01MH114925-02

National Institute of Mental Health

Treatment month	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Videos for ESI	T1		T2		T3		T4		T5		T6		T7		T8		T9		T10						
ESI intervention	Starting with individual ESI (Treatment) n=42									Followed by group ESI (Control)															
	Strating with group ESI (Control) n=40									Followed by individual ESI (Treatment)															
Video for ESDM	T1		*T2*		*T3*		*T4*		*T5*		*T6*		T7		*T8*		*T9*		*T10*		*T11*		*T12*		T13
ESDM Intervention	Treatment n=49																								
	Control n=48																								
Videos for JSAPER	T1			T2			T3																		
JASPER Treatment	Treatment n=59																								
	Control n=48																								

Italicized time points are only available for michigan sample (n=24, 10 tx cases)



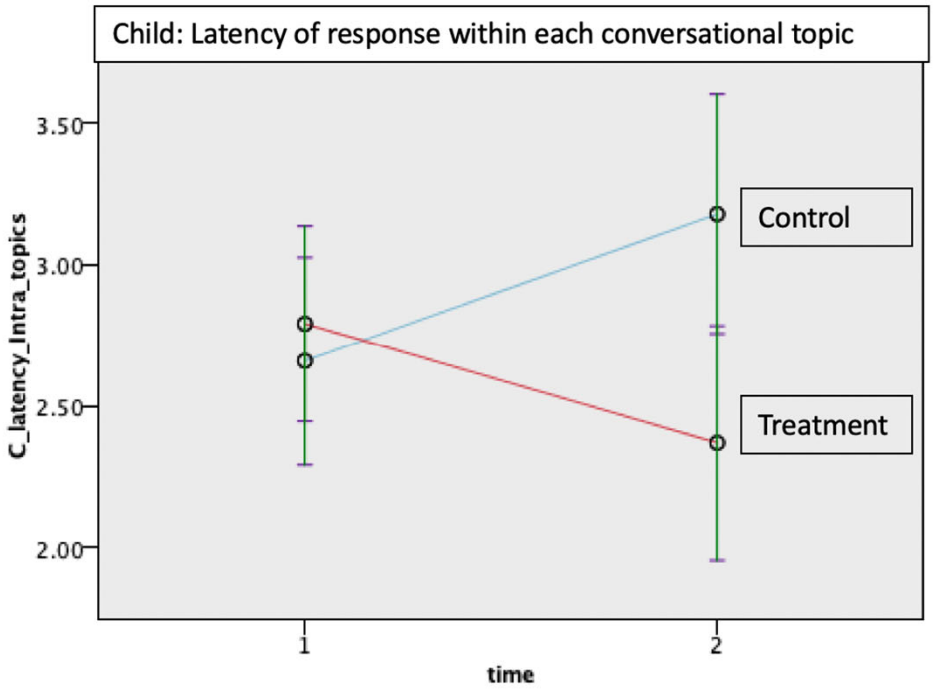
Changes in language features using machine learning algorithms

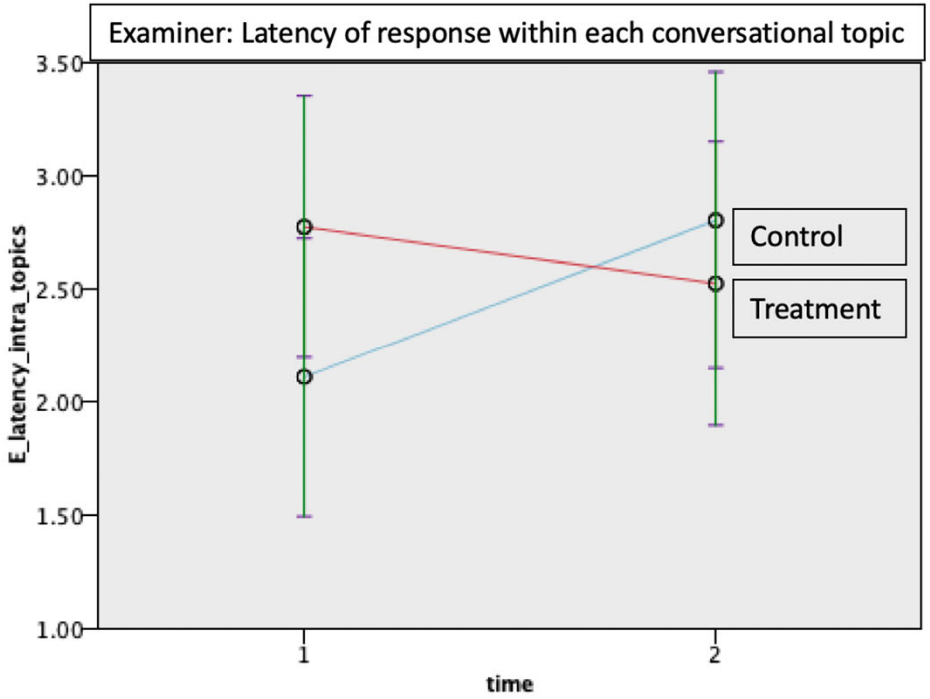
- N = 191 (99 treatment cases vs. 92 control cases) across 4 sites in NY
- Treatment varied from social skills training to oxytocin trials
- Treatment ranged from 2-8 months (M=3.8 mos, SD=1.5)

- <7 years: n=56 (26 tx cases)
- >=7 years: n=134 (72 tx cases)

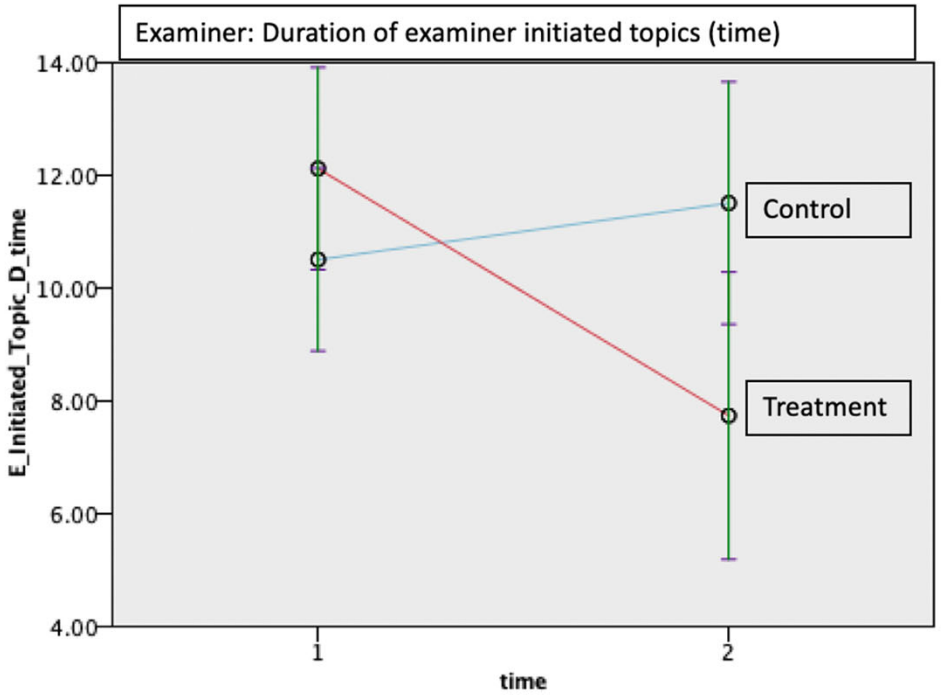
- Analyses: Generalized Linear Mixed Models (GLMM) to examine effects of time, treatment condition, and time*treatment condition on automatically extracted features from 12-min BOSCC sessions

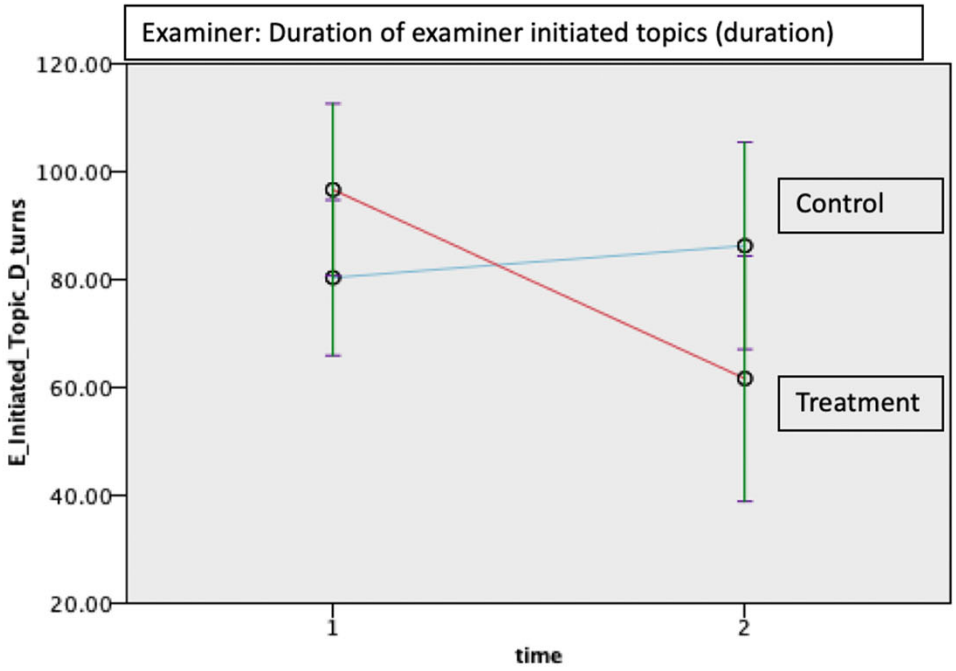
Significant treatment by time interaction for >=8 years





Significant treatment by time interaction for < 8 years





Speech and gaze integration during BOSCC Home session

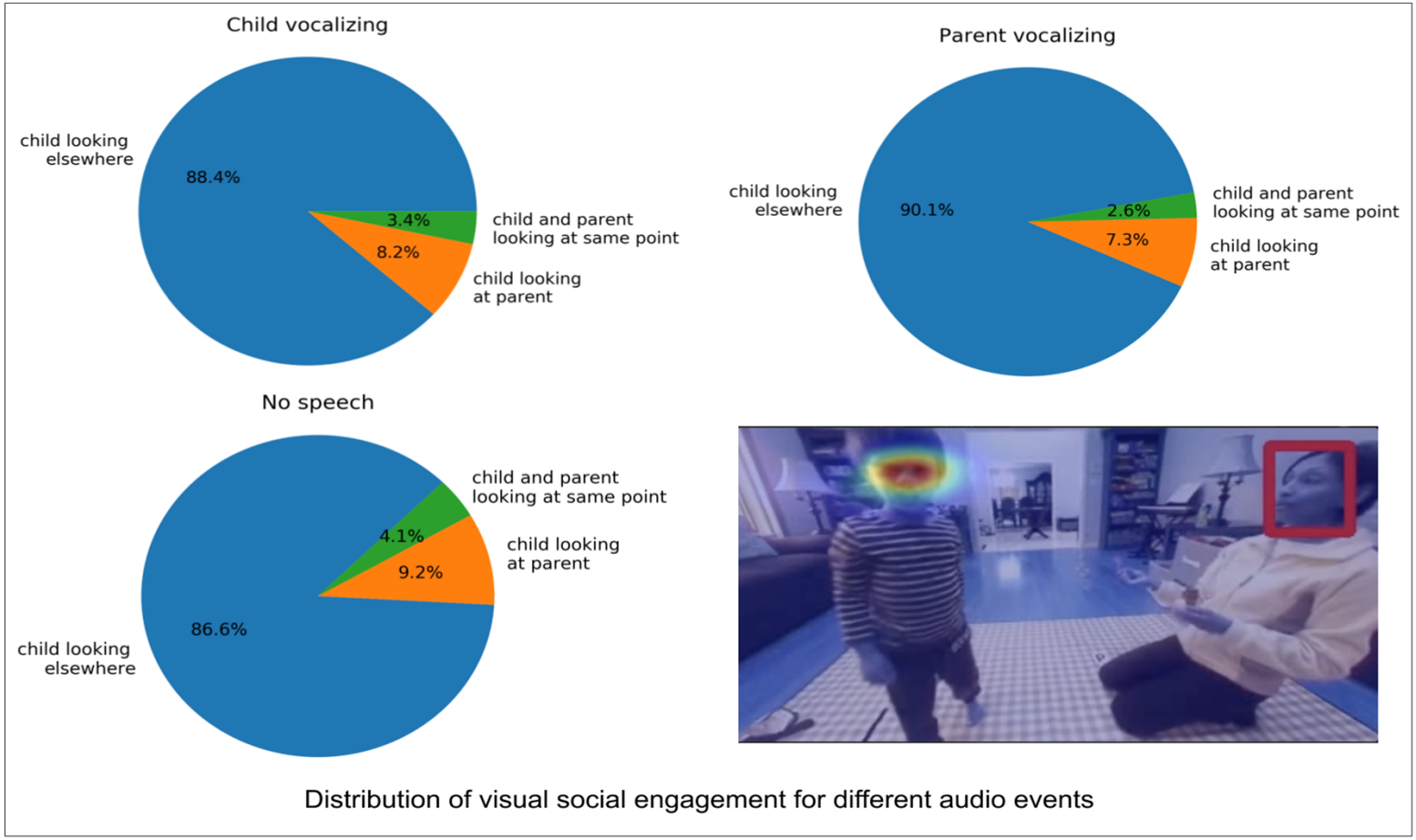
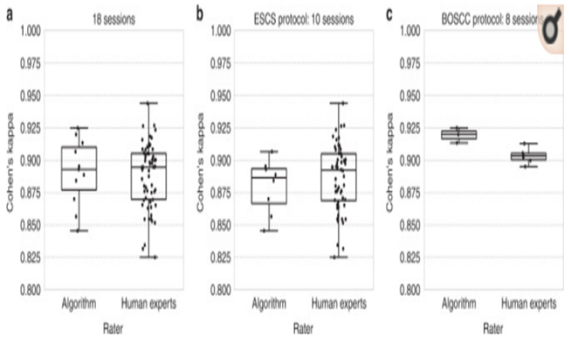


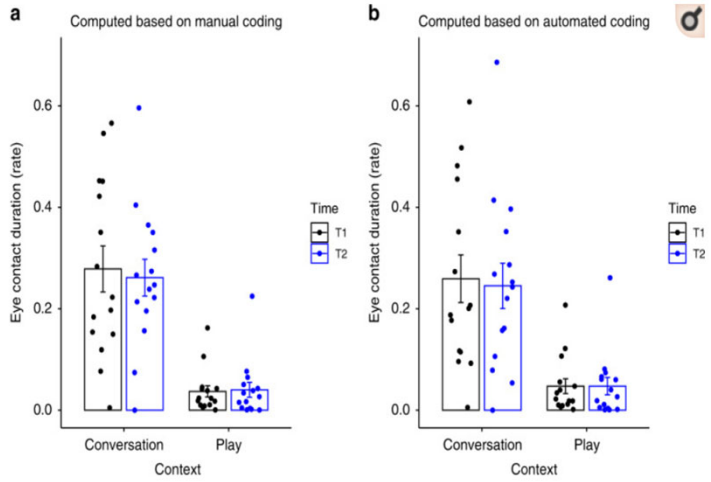
Figure 3 illustrates the Cohen's κ statistics for the human–human and human–detector comparison

Fig. 3



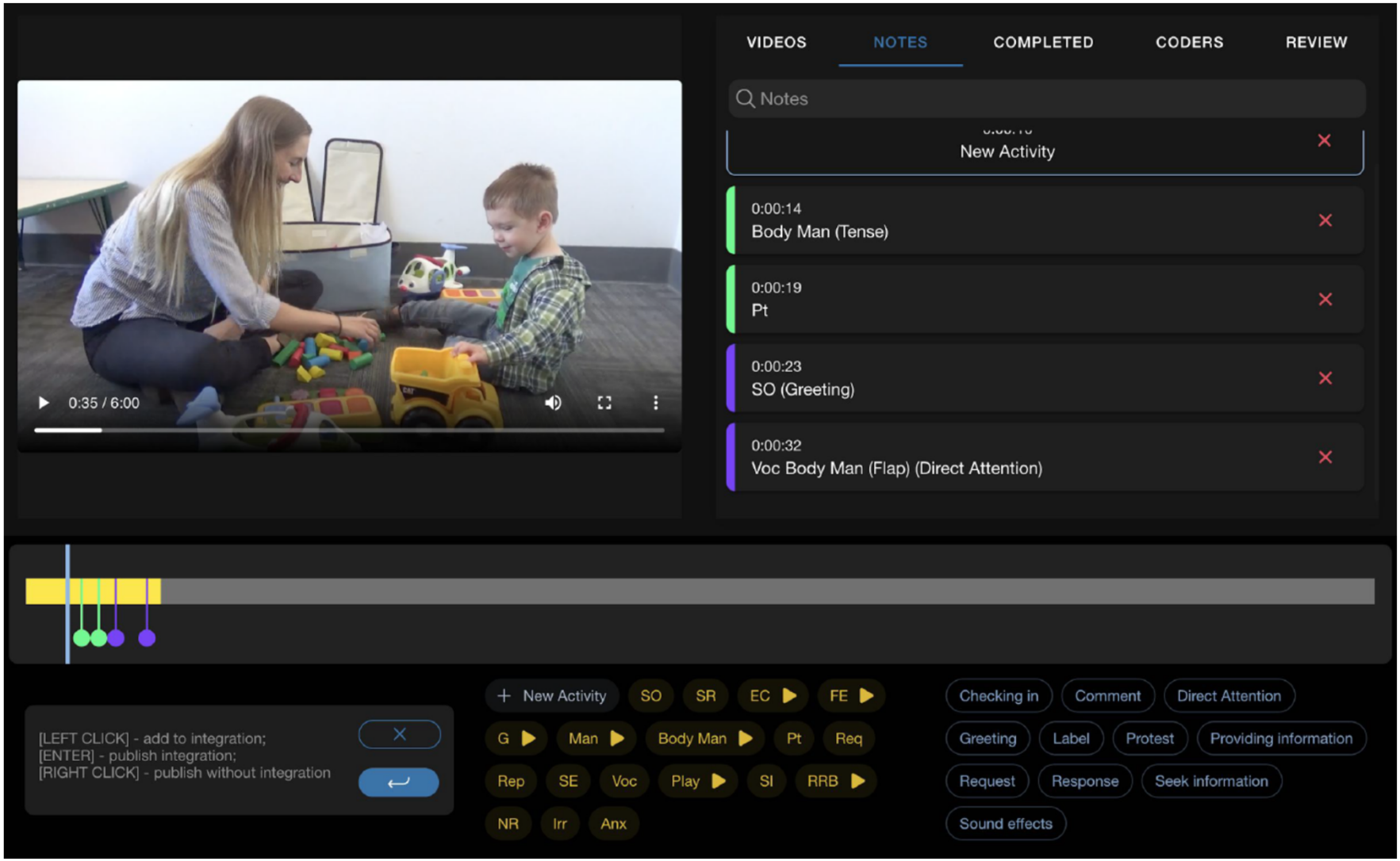
Pairwise Cohen's kappa distributions among all human pairs and human–algorithm pairs, represented as box plot. **a** 18 validation sessions, **b** ESCS, **c** BOSCC. Generally, kappa scores above 0.8 are considered an almost perfect agreement. On all sessions annotated by ten human experts, agreements among humans and agreements between each human and algorithm are similar in terms of kappa values.

Fig. 4

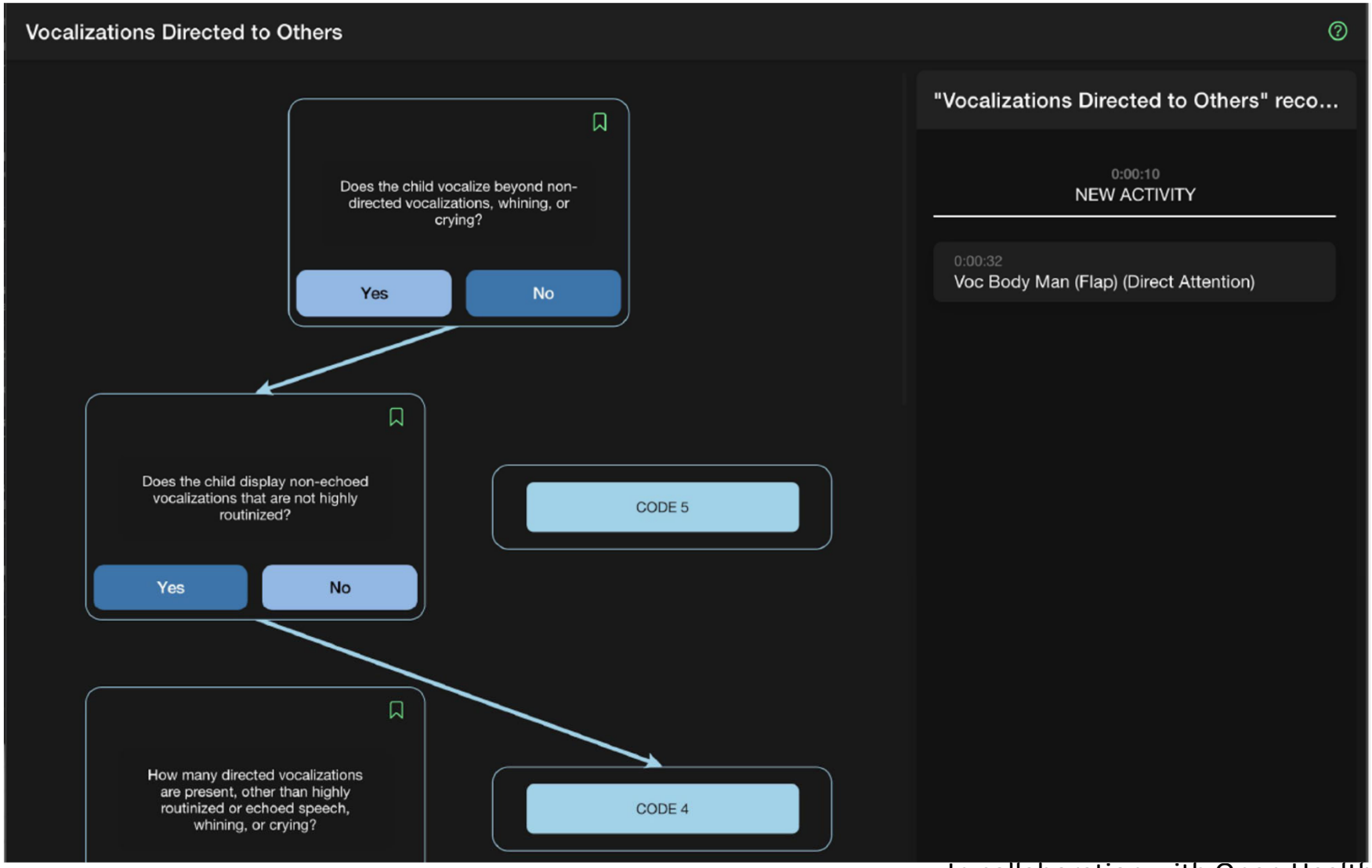


Average duration of eye contact during conversation and interactive play in child and adolescent samples ($n = 15$), measured at time 1 and time 2. **a** based on human coding, **b** based on automated coding. Data are presented as mean values ± SEM.

Chong E, Clark-Whitney E, Southerland A, Stubbs E, Miller C, Ajodan EL, Silverman MR, Lord C, Rozga A, Jones RM, Rehg JM. Detection of eye contact with deep neural networks is as accurate as human experts. Nat Commun. 2020



In collaboration with Open Health Network



In collaboration with Open Health Network

Gaps in the field

- Standard ways of measuring outcome
  - Each of the major studies, besides IQ and Vineland, has had their own way of measuring social and communicative behavior that is often directly tied to their intervention (which is fine, but harder to interpret)
- More information about older children, adolescents and adults
- More information about longer term, low intensity interventions and follow-up for older children, adolescents and adults
- Shorter term, multi-tier studies that allow for individual differences
- Building on SMART models, championed by Connie Kasari, to ask not just is one intervention better than another but for who?

Where are we?

- In part, we're still figuring out how to do this but we're making progress. communication that
- We can offer a standard not-project-specific measure of change in social moves beyond the 100 or so different measures used in different psychosocial trials, eliminates possible effects of bias either from parents or clinicians and gives us insight into differences that might help us understand what is actually happening (e.g, less able children who change when with their therapists but not with their parents or vice versa).
- One step closer to automating this (though it's been a long time coming)
- Still use parent reports, patient reports, clinician impressions and project-specific measures if they add important information!
- Remember mostly we're looking for incremental changes, kick-offs for cascades and keeping the trajectory going in the right direction!

Every day clinical implications

- Changes are usually incremental and relatively slow
  - Counting (for example, out of how many trials? Is seldom useful
    - Because it's tied to context and will vary from day to day
- Spot-checks and probes may make more sense
  - Less time consuming; more informative
  - Can be done across contexts (these findings may be depressing but are important)
- Many, though not all communicative behaviors, cluster – don't occur individually
  - We used to treat "eye contact;" now we work on looking related to different functioning
  - Treatment can make a difference across contexts

What are our treatment goals in autism?

- **Remember mostly we're looking for incremental changes, kick-offs for cascades and keeping the trajectory going in the right direction!**
- **No autism? Language? Social interaction? Quality of life? Less pain? More joy? Access to adequate services**
- Avoiding or diminishing secondary problems? Comorbidity with depression, ADHD, OCD; anxiety disorders
- Building on strengths? Finding own resources? Executive functioning, motivation, initiative
- **Importance of a social network from preschool through adulthood and the importance of families**

Thank you and thanks to all the participants in these studies and all the research partners

- Funders were NICHD, NIDCD, NIMH, Simons Foundation

- Co PIs or collaborators: So Hyun Kim (Cornell), Rebecca Grzadzinski (UNC), Alison Holbrook, Shri Naryanan (USC), James Rehg (Georgia Tech), Andrew Pickles (KCL), Eva Petkova (NYU), Connie Kasari, Sally Rogers (UCDavis), Amy Wetherby (FSU)
- LordLAB: Chrissie Toolan, Rebecca Elias, Sandra Gaspar, Niki Bahri, Nina Masjedi, Hannah Singer, Katherine Byrne, Kourtney Christopher, Elaine Clarke, Nicole Rosen