



Brain responses to biological motion predict treatment outcome in young adults with autism receiving Virtual Reality Social Cognition Training: Preliminary findings

Y.J. Daniel Yang^{a, b, *}, Tandra Allen^c, Sebiha M. Abdullahi^b, Kevin A. Pelphrey^a, Fred R. Volkmar^b, Sandra B. Chapman^c

^a Autism and Neurodevelopmental Disorders Institute, The George Washington University and Children's National Health System, Washington, DC 20052, USA

^b Child Study Center, Yale University School of Medicine, New Haven, CT 06520, USA

^c Center for BrainHealth, The University of Texas at Dallas, Dallas, TX 75235, USA

ARTICLE INFO

Article history:

Received 10 November 2016

Received in revised form

2 March 2017

Accepted 27 March 2017

Available online 29 March 2017

Keywords:

Virtual reality

Emotion recognition

Theory of mind

Autism

Predictive biomarker

Biological motion

fMRI

Intervention

ABSTRACT

Autism Spectrum Disorder (ASD) is characterized by remarkable heterogeneity in social, communication, and behavioral deficits, creating a major barrier in identifying effective treatments for a given individual with ASD. To facilitate precision medicine in ASD, we utilized a well-validated biological motion neuroimaging task to identify pretreatment biomarkers that can accurately forecast the response to an evidence-based behavioral treatment, Virtual Reality-Social Cognition Training (VR-SCT). In a preliminary sample of 17 young adults with high-functioning ASD, we identified neural predictors of change in emotion recognition after VR-SCT. The predictors were characterized by the pretreatment brain activations to biological vs. scrambled motion in the neural circuits that support (a) language comprehension and interpretation of incongruent auditory emotions and prosody, and (b) processing socio-emotional experience and interpersonal affective information, as well as emotional regulation. The predictive value of the findings for individual adults with ASD was supported by regression-based multivariate pattern analyses with cross validation. To our knowledge, this is the first pilot study that shows neuroimaging-based predictive biomarkers for treatment effectiveness in adults with ASD. The findings have potentially far-reaching implications for developing more precise and effective treatments for ASD.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

One of the major barriers in therapy for individuals with Autism Spectrum Disorder (ASD) (APA, 2013) is the difficulty to identify appropriate and effective treatments for a given individual with ASD. On the one hand, there is a remarkable variation and heterogeneity within the spectrum (Hahamy, Behrmann, & Malach, 2015; Lombardo et al., 2016), which makes it difficult for a single treatment to fit all individuals with ASD. On the other hand, people often need to spend considerable amount of resources (e.g., time, money) in trying out various treatment protocols before they are able to identify the most appropriate intervention. This problem is

particularly severe in adults with ASD, where intervention research has been very limited. To facilitate the fitting process and reduce potential waste of resources, it is crucial to develop objective predictors for treatment outcome in ASD, especially for adults with ASD, which would directly accelerate the long-term goal of precision medicine (Insel, 2014) in ASD.

In this research, we used a well-validated biological motion functional magnetic resonance imaging (fMRI) paradigm (Kaiser et al., 2010), which robustly engages the neural circuits supporting both socio-emotional and socio-cognitive components of social information processing, to identify pretreatment predictive biomarkers that can accurately forecast the response to an evidence-based behavioral intervention in young adults with ASD. The biological motion videos feature an adult engaging in children's games and social actions (e.g., waving, pat-a-cake, and peek-a-boo). Prior research has shown that social orienting to biological motion is evolutionarily well-conserved and fundamental to adaptive social

* Corresponding author. Autism and Neurodevelopmental Disorders Institute, The George Washington University and Children's National Health System, 2300 I St. NW, Washington, DC 20052, USA.

E-mail address: danielyang@gwu.edu (Y.J.D. Yang).

Abbreviations

ACS-SP	Advanced Clinical Solutions-Social Perception
ADOS	Autism Diagnostic Observation Schedule
ASD	Autism Spectrum Disorder
BIO	Biological Motion
BOLD	Blood Oxygen Level Dependent
CDT	Cluster-Defining Threshold
GLM	General Linear Model
LOSO	Leave-One-Subject-Out
MRI	Magnetic Resonance Imaging
MVPA	Multivariate Pattern Analyses
SCR	Scrambled Motion
SRS	Social Responsiveness Scale
VR-SCT	Virtual Reality-Social Cognition Training

engagement (Heberlein & Adolphs, 2004; Johnson, 2006; Simion, Regolin, & Bulf, 2008; Vallortigara, Regolin, & Marconato, 2005; Yang, Rosenblau, Keifer, & Pelphrey, 2015). The biological motion fMRI task has revealed key brain regions implicated in core ASD deficits (Allison, Puce, & McCarthy, 2000; Kaiser et al., 2010; McKay et al., 2012; Yang et al., 2015), including the ventrolateral prefrontal cortex (vlPFC), ventromedial prefrontal cortex (vmPFC), posterior superior temporal sulcus (pSTS), amygdala, and fusiform gyrus (FFG). These regions are implicated in various functions. Generally speaking, vlPFC, vmPFC, and amygdala are more closely related to socio-emotional processing and emotion regulation (Etkin, Buchel, & Gross, 2015; Kanske, Heissler, Schonfelder, Bongers, & Wessa, 2011; Phelps & LeDoux, 2005), while pSTS and FFG are more closely related to socio-cognitive processing and social information integration (Deen, Koldewyn, Kanwisher, & Saxe, 2015; Saggat, Shelly, Lepage, Hoeft, & Reiss, 2014; Yang et al., 2015). Recently, research has successfully applied the biological motion fMRI task in identifying predictive biomarkers for treatment outcome in young children with ASD receiving Pivotal Response Treatment (Yang et al., 2016), which marks the first evidence that neuroimaging-based task can effectively predict treatment effectiveness in ASD. This adds to the likelihood that the biological motion fMRI task may be used to identify neuropsychological biomarkers in young adults with ASD.

The treatment approach investigated in this research is Virtual Reality-Social Cognition Training (VR-SCT) (Kandalaf, Didehbani, Krawczyk, Allen, & Chapman, 2013), which is a short-term trial and consisted of 5-week treatment: two 1-h sessions per week with a total of 10 sessions. Recent research highlights the potential benefits of using Virtual Reality (VR) as an effective tool in training social skills for individuals with ASD (Bellani, Fornasari, Chittaro, & Brambilla, 2011; Kandalaf et al., 2013; Maskey, Lowry, Rodgers, McConachie, & Parr, 2014; Parsons & Mitchell, 2002; Wainer & Ingersoll, 2011). Reviews on VR studies suggest that there are several advantages of using VR environments to train social skills (Bellani et al., 2011; Parsons & Mitchell, 2002). Specifically, VR can simulate real-world contexts in a safe, non-threatening setting in which participants can practice commonly encountered social interactions (Bellani et al., 2011; Kandalaf et al., 2013; Parsons, Mitchell, & Leonard, 2005). VR also affords the user the opportunity to be immersed into the training by promoting engagement and a sense of presence within the simulated experience (Wallace et al., 2010). As presented in the previous studies (Didehbani, Allen, Kandalaf, Krawczyk, & Chapman, 2016; Kandalaf et al., 2013), the format of immersive role-play in VR-SCT can afford the participant a variety of opportunities to become engaged in

training, while reducing social anxiety and allowing for a dynamic practice experience without negative real-world social consequences. It is safe to try and fail, because scenarios are controlled by a clinician and allow for repeated practice using targeted social strategies. As well, the technology itself supports the *realism* of immersive role-play conversation by allowing a clinician to change his/her appearance, voice or even the physical setting of the conversation, which are key elements that in-person treatment can hardly provide.

Furthermore, VR and computer technologies as a training method are highly motivating platforms for many individuals with ASD (Chen & Bernard-Opitz, 1993; Moore & Calvert, 2000; Parsons & Mitchell, 2002). The ability to customize and practice dynamic social scenarios across multiple training sessions is also a strength of VR, which facilitates opportunities for generalization of social skills learned in VR to everyday life interactions (Bellani et al., 2011; Didehbani et al., 2016; Parsons & Cobb, 2011; Tzanavari, Charalambous-Darden, Herakleous, & Poullis, 2015). Generalization of trained skills to the real-world was also reported in a previous study (Maskey et al., 2014), whereby gradual exposure to a specific anxiety was presented in a visual context through VR, which combined with cognitive behavioral therapy resulted in participants reporting reduced anxiety in their everyday life. Similarly, in the previous VR-SCT study (Kandalaf et al., 2013), participants reported that their social functioning has continued to improve several months after the end of training. For adolescents and young adults with ASD, previous research has also shown that VR offers the flexibility to target social skills in isolation, such as social appropriateness with spatial proximity and knowing what to say in a job interview (Cheng, Moore, McGrath, & Fan, 2005; Parsons, Mitchell, & Leonard, 2004; Smith et al., 2014; Trepagnier, Olsen, Boteler, & Bell, 2011), or to target multiple skills in one platform (Kandalaf et al., 2013), including emotion recognition, theory of mind, and social functioning collectively.

The main principles of the VR-SCT intervention utilized in this study were based on prior VR-SCT studies (Didehbani et al., 2016; Kandalaf et al., 2013). In these studies, VR-SCT involved a semi-manualized structured prompt used by the clinicians for all participants. Clinicians used a scripted personality and response style for each character they played that further promoted standardization across participants. Even though the prompts were structured and repeated across scenarios, participants' individual responses allowed for flexible real-time responses by both the clinician and the participant. Similarly, in the current study, a manualized approach was used to engage participants in a conversation and the participants partially customized their experience through their own responses. As shown in these prior studies (Didehbani et al., 2016; Kandalaf et al., 2013), the role-play method has been utilized in both young adult and pediatric populations with similar improvements, regardless of the age group, in emotion recognition and theory of mind. Overall, VR-SCT offers an engaging, interactive, and individualized platform for training and improving socio-emotional and socio-cognitive abilities for individuals with ASD.

In this study, we investigated whether a pre-treatment biological motion fMRI task could predict therapeutic response to VR-SCT in young adults with ASD. The biological motion fMRI task was chosen also because it measures key socio-emotional and socio-cognitive processing, which correspond to the treatment targets of VR-SCT. Linking to the biological motion fMRI task and VR-SCT, we utilized two separate behavioral tasks to measure behavioral changes in emotional and cognitive aspects of social information processing, respectively: for the emotional component, we measured behavioral changes in emotion-recognition ability, while for the cognitive component, we measured behavioral changes in theory-of-mind ability. As VR-SCT has been demonstrated to

improve emotion recognition and theory of mind in prior research (Didehbani et al., 2016; Kandalaf et al., 2013), we expected that there will be similar behavioral improvement on average (one-tailed). Moreover, central to the aim of the current study, we hypothesized that the biological motion fMRI task will be able to identify pretreatment neurobiological markers that can predict change in emotion recognition and theory of mind, respectively.

2. Methods

2.1. Participants

Study participants included 17 young adults with a primary diagnosis of ASD (M age = 22.50 years, SD = 3.89; 2 females, 15 males), recruited from two research sites: Yale Child Study Center (YCSC) for 7 participants, and Center for BrainHealth at The University of Texas at Dallas (CBH-UTD) for the other 10 participants. A recent neuroprediction study involving young children with autism receiving Pivotal Response Treatment (PRT) (Yang et al., 2016) shows that the effect size as measured by the linear bivariate correlation between pretreatment brain activities and behavioral changes induced by treatment is about $r = 0.54$ – 0.81 . Accordingly, for power = 0.80 and $\alpha = 0.05$, G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) suggests that it requires at least 17 participants to be sufficiently powered to detect similar neuropredictive effects, while assuming that the neuropredictive effect sizes are similar across different age groups (young children vs. young adults) and different treatment approaches (PRT vs. VR-SCT).

IQ was measured using the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999, 2011). In terms of full-scale IQ, all participants were relatively high-functioning ($IQ \geq 80$; M $IQ = 109.65$, $SD = 13.32$). All participants met DSM-V (APA, 2013) diagnostic criteria for ASD as determined by the results of a gold-standard diagnostic instrument, the Autism Diagnostic Observation Schedule (Gotham, Risi, Pickles, & Lord, 2007; Hus & Lord, 2014; Lord et al., 2000)—administered by research-reliable clinicians and licensed clinical psychologists. Unfortunately, the ADOS subdomain scores were not available for two participants because they were evaluated by psychologists involved in other projects, although their autism diagnosis was re-confirmed and current before they were included in this project. Pretreatment clinical behavioral characterization was based on the self-reported Social Responsiveness Scale, 2nd edition (SRS-2) (Constantino, 2012) (M total raw = 82.41, $SD = 33.43$), which assesses ASD symptom severity in five domains: social awareness, social cognition, social communication, social motivation, and restricted interests and repetitive behavior. Comprehensive demographics and characterization information are provided in Table 1. This is a pretest-posttest treatment-only study and all of the 17 participants were assigned to receive VR-SCT intervention. The study is registered at ClinicalTrials.gov (ID: NCT02139514; NCT02922400).

Inclusion criteria for all participants included being within the ages of 18–40 years, high-functioning ($IQ \geq 80$), having a diagnosis for ASD, and having a mean length of utterance greater than 5 words (as required by the intervention method). Exclusion criteria for all participants included not being fluent in English (the intervention has not been validated in other languages), significant hearing loss, and a history of significant head trauma or serious psychiatric illness (other than ASD). Ten adults (beyond the 17 participants) were excluded: five did not meet inclusion criteria (one was >40 years of age, one had $IQ < 80$, three were non-spectrum), three were unavailable (two lived too far away, one had schedule conflicts), and two were discontinued after treatment began due to either the therapist being unavailable or the participant being unable to make the treatment visits. All of the remaining

Table 1

Participants demographics and pretreatment autism symptom severity profile ($N = 17$).

Variable	Mean (S.D.)	Range
Pretreatment Age (years)	22.50 (3.89)	18.06–31.08
Gender, male (0 = f, 1 = m)	0.88 (0.33)	—
Full-scale IQ	109.65 (13.32)	88–131
Handedness (1 = right, 0 = ambi., –1 = left)	1.00 (0.00)	—
ADOS Module 4 ($n = 15$) ^a		
SA Domain	10.73 (3.63)	7–19
RRB Domain	0.93 (1.03)	0–3
Total	11.67 (3.85)	7–20
Pretreatment SRS-2 self-reported raw scores		
Social awareness	9.59 (3.22)	5–15
Social cognition	14.35 (5.79)	6–28
Social communication	25.41 (11.77)	6–47
Social motivation	15.18 (7.72)	5–27
Restricted interests and repetitive behavior	17.88 (7.78)	6–30
Total	82.41 (33.43)	33–142

Note. SA, Social Affect; RRB, Restricted and Repetitive Behaviors.

^a Unfortunately, the ADOS subdomain scores were not available for two participants because they were evaluated by psychologists involved in other projects, although their autism diagnosis was re-confirmed and current before they were included in this project.

17 participants passed the MRI (Magnetic Resonance Imaging) safety screening, including being free of any metal implants and having no evidence of claustrophobia. All of the 17 participants were right-handed. Written informed consent was obtained from each participant. The Human Investigations Committee (HIC) at Yale University and the Institutional Review Board (IRB) at the University of Texas at Dallas approved this study.

2.2. Treatment approach: Virtual Reality–Social Cognition Training (VR-SCT)

2.2.1. Technology

VR-SCT training technology, *Virtual Gemini*, was developed by programmers at Center for BrainHealth, using an Unreal® engine and transmitted over the Internet. The platform included facial emotion tracking using Faceshift Studio® software, which displayed the participant and clinician's facial movements in real time. Mumble® server was utilized to transmit audio as well as Morph-Vox (Screaming_Bee, 2005) was utilized for voice-modulation software. VR-SCT ran on a standard Windows computer (minimum specifications of a Mobile core i7, 750 m, 4 gigabytes of RAM), a web cam, and a headphone with built-in microphone.

2.2.2. Training

VR-SCT is a strategy-based immersive role-play intervention program designed to strengthen socio-emotional processing and socio-cognitive reasoning abilities in both children and young adults with ASD. As documented in previous research studies (Didehbani et al., 2016; Kandalaf et al., 2013), VR-SCT has been shown to specifically target and improve emotion recognition and theory of mind as well as executive function and daily social function. Based upon previous research findings (Didehbani et al., 2016; Kandalaf et al., 2013), we expected to see improvements in emotion recognition and theory of mind after VR-SCT in the current study. For the purposes of the study, sessions took place at both YCSC and CBH-UTD. The coach clinician conducted the training and hosted each session from CBH-UTD via the Internet. Across both sites, participants arrived for the session and were set up on the computer by the coach clinician or research staff. Once a participant logged into the virtual platform, he/she independently interacted with the coach clinician online. During the training session, both the clinician and participant interacted entirely through virtual

avatar characters (see Fig. 1 for sample)

For every participant, the training program lasts for five weeks, with two 1-h sessions per week and thus 10 h in total. VR-SCT presented hierarchical socio-emotional and socio-cognitive strategies that increased in complexity over the course of the 10 sessions. The first three sessions targeted learning three core social strategies (recognizing others, responding to others, self-assertion) and the remaining seven sessions focused on integrating all strategies together across varied and complex social situations (e.g., dealing with confrontation, job interview, blind date). Each session allowed for multiple conversations and practicing the same social objective for that day, so as to build a dynamic learning opportunity of the social strategies and to encourage generalization to real-life conversation.

At the beginning of each session, a social learning objective was presented and reviewed with the participant by a coach clinician. After being given a social prompt, (e.g., “You will be meeting a new neighbor at the apartment building”), the participant engaged in a semi-structured live conversation with a confederate clinician posing as a conversational partner. The confederate clinician could change avatar appearance and modify his/her voice to quickly change from one character to the next. Each character, played by the confederate, had a unique pre-determined conversational opening and emotional style (e.g., pleasant demeanor and easy-going, or a rush and acting rude style). However, the responses were dynamic and individualized as each participant partially determined the outcome based upon his/her response. Following each practice conversation, the coach and participant engaged in feedback discussion of the conversation that included self-ratings.

The participant was then given additional conversational opportunities to integrate the discussed feedback into subsequent scenarios.

2.2.3. Treatment targets of VR-SCT

The strategy-based program was presented in a top-down fashion, to strengthen socio-emotional and social-cognitive abilities of recognizing others, responding to others, and self-assertion. The first strategy of *recognizing others* targeted (a) filtering and blocking social distractions, (b) identifying key and relevant social cues, and (c) inferring social meaning in other's expressions. The next strategy of *responding to others* built on the social perceptions of the other person to generate meaningful connections by (a) considering one's own social emotional cues and how they are being conveyed, (b) formulating clear and direct responses that relate to the situation or the other person, and (c) building back-and-forth conversation to allow deeper conversation beyond the surface level. Finally, engagement of socio-emotional and social-cognitive control processes was further facilitated by the final *self-assertion* strategy of (a) considering multiple perspectives, (b) considering possible outcomes and reflecting on past mistakes, and (c) applying new knowledge to new situations.

2.3. Primary clinical outcome: changes in emotion recognition and theory of mind

Treatment effectiveness was measured by behavioral changes in two distinct domains of social abilities: emotion recognition (tapping change in socio-emotional processing abilities) and theory of mind (tapping change in socio-cognitive processing abilities), respectively.

2.3.1. Emotion recognition

The *Advanced Clinical Solutions for WAIS-IV and WMS-IV Social Perception Subtest (ACS-SP)* (Kandalaf et al., 2012; Pearson, 2009), administered by trained research staff in our research centers, was utilized to measure emotion recognition abilities. Three subscales are generated from the subtest tasks: (a) *SP-Affect Naming*, a measure of face emotion recognition; (b) *SP-Prosody*, a measure of vocal affect recognition; (c) *SP-Pairs*, a measure of non-literal language interpretation. Across ACS-SP scores, average internal consistency has been reported as $r = 0.69–0.81$, test–retest stability coefficient as corrected $r = 0.60–0.70$, and inter-scorer agreement from 0.98 to 0.99. Normative scaled scores are available for all ACS-SP subtests. Treatment effectiveness on emotion recognition is modeled as the Δ change scores of the ACS-SP scaled scores, that is, post minus pre, such that positive (or negative) delta change scores indicate increase (or decrease) in emotion recognition abilities. In the prior pilot study involving VR-SCT and young adults with autism (Kandalaf et al., 2013), emotion recognition changed significantly from pretreatment ($M = 7.63$, $SD = 3.42$) to posttreatment ($M = 9.63$, $SD = 3.78$), $t(7) = 2.83$, $p = 0.03$, Cohen's d_{rm} (Lakens, 2013) = 0.55.

2.3.2. Theory of mind

The Social Attribution Task, also known as the triangles task (Abell, Happe, & Frith, 2000) was administered to measure a person's abilities of theory of mind. Videos were adapted from a previous study (Heider & Simmel, 1944), in which participants were asked to narrate the movements of triangles presented in six separate brief videos. Narratives were recorded, transcribed, and double-scored by two blind raters. Based upon the scoring criteria established in previous research (Abell et al., 2000), participants' narratives were first scored on accuracy and attribution, respectively, and the two scores were then summed up to derive a total



Fig. 1. The computer set-up and example screenshots of a virtual reality training session.

score. The accuracy and attribution aspects were scored separately using a 4-point Likert scale (0–3 point scale) for each video, with 18 as the maximum possible score across all 6 videos for both accuracy and attribution, respectively. More points were awarded when the participant stated descriptions that were accurate to the nature of the video (for the accuracy aspect), or when more mentalizing or emotional words were utilized to describe the movement of the triangles (for the attribution aspect). The triangles task has been shown to have a high test-retest reliability of $r = 0.76$ to 0.88 and concurrent validity $r = 0.78$ to 0.93 (Hu, Chan, & McAlonan, 2010). The order of the videos was randomized and participants were presented with different sets of videos at pre- and post-intervention testing. In the prior pilot study involving VR-SCT and young adults with autism (Kandalaf et al., 2013), theory of mind changed significantly from pretreatment ($M = 12.63$, $SD = 4.93$) to posttreatment ($M = 15.38$, $SD = 4.81$), $t(7) = 3.45$, $p = 0.01$, Cohen's d_{rm} (Lakens, 2013) = 0.56 .

2.4. fMRI imaging task

We measured the pretreatment BOLD (blood oxygen level dependent) responses using a well-established biological motion fMRI task (Bjornsdotter, Wang, Pelphrey, & Kaiser, 2016; Kaiser et al., 2010; Ventola et al., 2015; Yang et al., 2016; Yang et al., 2017). We selected this paradigm because it measures the neural activities of two key components of social information processing (Kaiser et al., 2010), namely, socio-emotional and socio-cognitive processing, which correspond to the treatment targets of VR-SCT and the two behavioral measures of emotion recognition and theory of mind, respectively. Before the treatment, the participants were scanned while viewing coherent and scrambled point-light displays of biological motion created from motion capture data. The coherent biological motion displays featured an adult male actor performing movements relevant to early childhood experiences, such as playing pat-a-cake (Klin, Lin, Gorrindo, Ramsay, & Jones, 2009), and contained 16 points corresponding to major joints. The scrambled motion animations were created by selecting all the 16 points from the biological motion displays and randomly plotting their trajectories on a black background (see [Supplementary Material 1](#) for sample fMRI stimuli). Thus, the coherent and scrambled displays contained the same local motion information, but only the coherent displays contained the configuration of a person (Johansson, 1973). During the MRI scan, stimuli were presented using E-Prime 2.0 software (Psychological Software Tools, Pittsburgh, PA, USA). Six coherent biological motion clips (BIO) and six scrambled (SCR) motion clips were presented once each in an alternating-block design (time per block, ~24 s). The experiment began with a 20-s fixation period and ended with a 16-s fixation period. The total duration was 328 s. The movies were presented without audio. The participants were asked to watch the videos and reminded to remain still and alert. The imaging task and stimuli are available from the authors upon reasonable request.

2.5. Imaging acquisition and processing

Scanning was performed on a Siemens MAGNETOM 3 T Tim Trio scanner at the Yale Magnetic Resonance Research Center (for YCSC participants) or a Philips 3 T MR system (for CBH-UTD participants) within one week before the treatment began. For each YCSC participant, a structural MRI image series was acquired with a 12-channel head coil, a high-resolution T1-weighted MPRAGE sequence, and the following parameters: 176 slices; TR = 2530 ms; TE = 3.31 ms; flip angle = 7 deg; slice thickness = 1.00 mm; voxel size = $1 \times 1 \text{ mm}^2$; matrix = 256×256 . Afterwards, BOLD T2*-weighted functional MRI images were acquired using the

following parameters: 164 vol; TR = 2000 ms; TE = 30 ms; flip angle = 90°; slice thickness = 4.00 mm; voxel size = $3 \times 3 \text{ mm}^2$; matrix = 64×64 ; number of slices per volume = 34. For each CBH-UTD participant, a structural MRI image series was acquired with an 8-channel head coil, a high-resolution T1-weighted MPRAGE sequence, and the following parameters: 176 slices; TR = 7.730 ms; TE = 3.53 ms; flip angle = 7 deg; slice thickness = 1.00 mm; voxel size = $1 \times 1 \text{ mm}^2$; matrix = 256×256 . Afterwards, BOLD T2*-weighted functional MRI images were acquired using the following parameters: 164 vol; TR = 2000 ms; TE = 30 ms; flip angle = 90°; slice thickness = 4.00 mm; voxel size = $3 \times 3 \text{ mm}^2$; matrix = 64×64 ; number of slices per volume = 34. The site variable (YCSC vs. CBH-UTD; dummy-coded and mean-centered) was included as a covariate of no interest across all neuroimaging analyses. Although each site used a different scanner, controlling for site has statistically controlled for the differences due to the two sites, including using two different scanners, head coils, etc.

The T1-weighted MPRAGE structural scan was segmented by SPM12 into gray matter, white matter (WM), and cerebrospinal fluid (CSF) images. This method is highly accurate and has reduced bias, comparable to manual measurement (Malone et al., 2015).

The fMRI data were processed using FSL (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012) v5.0.9 and the participant-level preprocessing steps followed a standardized processing stream—ICA-AROMA (ICA-based strategy for Automatic Removal of Motion Artifacts) (Pruim et al., 2015). This consisted of the following sequence: (a) motion correction using MCFLIRT; (b) interleaved slice timing correction; (c) BET brain extraction; (d) grand mean intensity normalization for the whole 4D data set; (e) spatial smoothing with 5 mm FWHM; (f) data denoising with ICA-AROMA (Pruim et al., 2015), which uses a robust set of theoretically motivated temporal and spatial features to remove motion-related spurious noise; (g) nuisance regression using time-series for WM and CSF signal to remove residual, physiological noise, and finally, (h) high-pass temporal filtering (100 s). The first 4 s were discarded to establish T1 equilibrium. Registration of the fMRI data was performed using both the subject's structural scan and then the Montreal Neurological Institute (MNI152) standard brain. Preprocessed data were then pre-whitened using FSL's FILM to remove time series autocorrelation.

To model the BIO and SCR conditions, the timing of the corresponding blocks was convolved with the default gamma function (phase = 0 s, standard deviation = 3 s, mean lag = 6 s) with temporal derivatives. The participant-level contrast of interest is BIO > SCR, which served as inputs for the subsequent mass univariate, whole-brain, group-level GLM analyses and multivariate pattern analyses.

2.6. Mass univariate group-level GLM analyses

We conducted mass univariate voxel-wise GLM analyses across the whole brain to identify clusters where pretreatment BOLD activation to the contrast of BIO > SCR predicted treatment effectiveness. The analyses were conducted using mixed-effects modeling with FSL's FLAME (FMRIB's Local Analysis of Mixed Effects) 1 + 2 inference algorithm, which provides highly accurate estimation of group-level results that are generalizable to the population. Consistent with our prior neuropredictive research (Yang et al., 2016), we employed a stringent cluster-defining threshold (CDT) of $Z > 2.33$, $p < 0.01$, while correcting for multiple comparisons at a cluster-level threshold of $p < 0.05$. Information about the surviving clusters was reported, including number of voxels in the cluster, the anatomical regions covered by the clusters based on the Desikan-Killiany atlas (Desikan et al., 2006), the coordinates of the peak voxels within each of the anatomical regions,

and the Z-statistics associated with the peak voxels. Site, age, IQ, sex, and pretreatment autism symptom severity using the SRS total raw scores were mean-centered and controlled for as covariates of no interest in all group-level univariate GLM analyses. This was to ensure that the results could be generalized to different sites, ages, IQ levels, sexes, and levels of pretreatment autism symptom severity. The pretreatment autism symptom severity was not significantly correlated with either pretreatment emotion recognition, $r = -0.39$, $p = 0.13$, or pretreatment theory of mind, $r = 0.17$, $p = 0.51$, which could be due to a number of possible reasons (e.g., the generality-specificity differences, autism symptom severity encompassing some different domains, small sample size, difference in responders/coders, etc.).

2.7. Meta-analytical reverse inference

To understand the functional relevance of the surviving clusters, we performed a quantitative reverse inference using NeuroSynth (<http://www.neurosynth.org/>). The NeuroSynth dataset v0.6 (July 2015 release) contains activation data for over 11,406 studies and feature information for over 3,300 term-based features. The term-based features were derived from the abstracts of articles in the NeuroSynth database. For each feature, the database stores the whole-brain, reverse inference, meta-analysis map, $P(\text{Term} | \text{Activation})$, that is, the likelihood that a feature term is used in a study given the presence of reported activation (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011). Each surviving cluster was decoded with NeuroSynth, which computed the voxel-wise Pearson correlation between the cluster image file and the meta-analytical image file associated with each of the 3,300 feature terms. The top 10 psychological functional terms with the highest positive correlation were retained and reported, while we omitted non-functional terms, such as (but not limited to) those describing an anatomical region (e.g., frontal operculum, vlpfc), a technique/method/task (e.g., signal task, nogo), a population (e.g., speakers), or being relatively generic (e.g., reference, difficulty, conveyed, etc.).

2.8. Multivariate pattern analyses (MVPA)

To guard against data over-fitting and to gain understanding of how different voxels in the neuropredictive network derived from the mass univariate GLM analyses worked together in predicting change in emotion recognition, we utilized regression-based Multivariate Pattern Analyses (MVPA) (Haxby, Connolly, & Guntupalli, 2014). In MVPA, the samples were divided into training and testing data sets, which constitute a *cross validation framework* in which the predictive model is first trained with the training set and then used to predict the regression labels of the sample in the testing set. This type of cross validation provides approximately unbiased estimates of effects, generalizable to new samples, helping to minimize the likelihood that the results over-fit the data. Moreover, in contrast to the mass univariate voxel-wise GLM analyses, MVPA draws on the multivariate information across many voxels comprising neural networks, which may capture how the voxels or regions work together to achieve complex functions. All these characteristics render MVPA well suited for establishing robust predictive biomarkers. MVPA has been applied to fMRI data to establish predictive biomarkers for treatment response or long-term outcome in a number of neuropsychiatric or neurocognitive disorders, such as depression (van Waarde, Scholte, van Oudheusden, Verwey, Denys, & van Wingen, 2015), dyslexia (Hoeft et al., 2011), social anxiety disorder (Mansson et al., 2015; Whitfield-Gabrieli et al., 2015), panic disorder (Hahn et al., 2015), and more recently autism spectrum disorder (Yang et al., 2016).

MVPA were performed using the Pattern Recognition for

Neuroimaging Toolbox (Schrouff et al., 2013) (PRoNT) v2.0 in Matlab and followed several steps. First, each participant's pretreatment Z-statistic BIO > SCR contrast image (up-sampled to the standard MNI152 space using trilinear interpolation) was inputted into the MVPA. The surviving clusters derived from the univariate analysis as a network was used as an analytical mask, while treatment effectiveness was entered as the regression target. Second, PRoNT computed a linear kernel (that is, dot product) between the voxel intensities within the mask for each pair of the input images, thereby generating a 17×17 similarity matrix, which served as the input feature set for the subsequent machine learning algorithm. Third, we used kernel ridge regression (KRR) (Chu, Ni, Tan, Saunders, & Ashburner, 2011) as the multivariate regression method. This is the dual-form formulation of ridge regression and solves regression problems with high dimensional data in a computationally efficient way. Cross validation was based on a leave-one-subject-out (LOSO) framework with mean-centered features across training images. We selected LOSO (which is equal to 17-fold cross-validation with our sample) because a larger number of folds may reduce bias of the estimates, even at the cost of increasing variance of the estimates, and should provide more accurate estimates of neural predictability, especially when sample sizes are small. For each fold, one input image was left out and served as the testing set. The KRR machines were trained to associate treatment effectiveness with the multivariate information in the remaining sample of 16 participants. The trained KRR machines were then used to predict treatment effectiveness in the left-out image. This step was repeated for each of the 17 folds. Across all folds, predictive accuracy was calculated as the Pearson's correlation coefficient (r), coefficient of determination (R^2), and normalized mean squared error ($nMSE$) between predicted and actual treatment effectiveness. Fourth, the significance of the predictive accuracy statistics was evaluated using a permutation test, consisting of 50,000 iterations. In each iteration, the regression targets were randomly permuted across all participants and the cross-validation procedure was repeated. The p -values of r , R^2 , and $nMSE$ were then calculated as the proportion of all permutations where r , R^2 , and $nMSE$ were greater than (or less than, in the case of $nMSE$) or equal to the obtained r , R^2 , and $nMSE$, respectively.

3. Results

3.1. Primary clinical outcome

As illustrated in Fig. 2(A) and as hypothesized, VR-SCT significantly improved emotion recognition in terms of the ACS-SP scaled scores from pretreatment ($M = 11.41$, $SD = 4.42$) to posttreatment ($M = 12.94$, $SD = 3.51$), $\Delta = 1.53$, $S.D. \text{ of } \Delta = 2.72$, $t(16) = 2.32$, $p = 0.03$ (two-tailed), Cohen's d_{rm} (Lakens, 2013) = 0.37. In addition, as shown in Fig. 2(B), VR-SCT marginally improved theory of mind in terms of the total scores from the triangles task from pretreatment ($M = 19.41$, $SD = 3.89$) to posttreatment ($M = 20.35$, $SD = 3.84$), $\Delta = 0.94$, $S.D. \text{ of } \Delta = 1.95$, $t(16) = 1.99$, $p = 0.06$ (two-tailed), Cohen's d_{rm} (Lakens, 2013) = 0.24. This effect was significant one-tailed, $p = 0.03$, and therefore also as hypothesized. The change scores of emotion recognition were not significantly correlated with those of theory of mind, $r(15) = -0.22$, $p = 0.40$, suggesting that they were tapping improvement in distinctively different domains. The raw data for individual participants were reported in Supplementary Material 2.

3.2. Mass univariate GLM analyses

For change in emotion recognition, as illustrated in Fig. 3 and as hypothesized, the whole-brain mass univariate GLM analyses of the

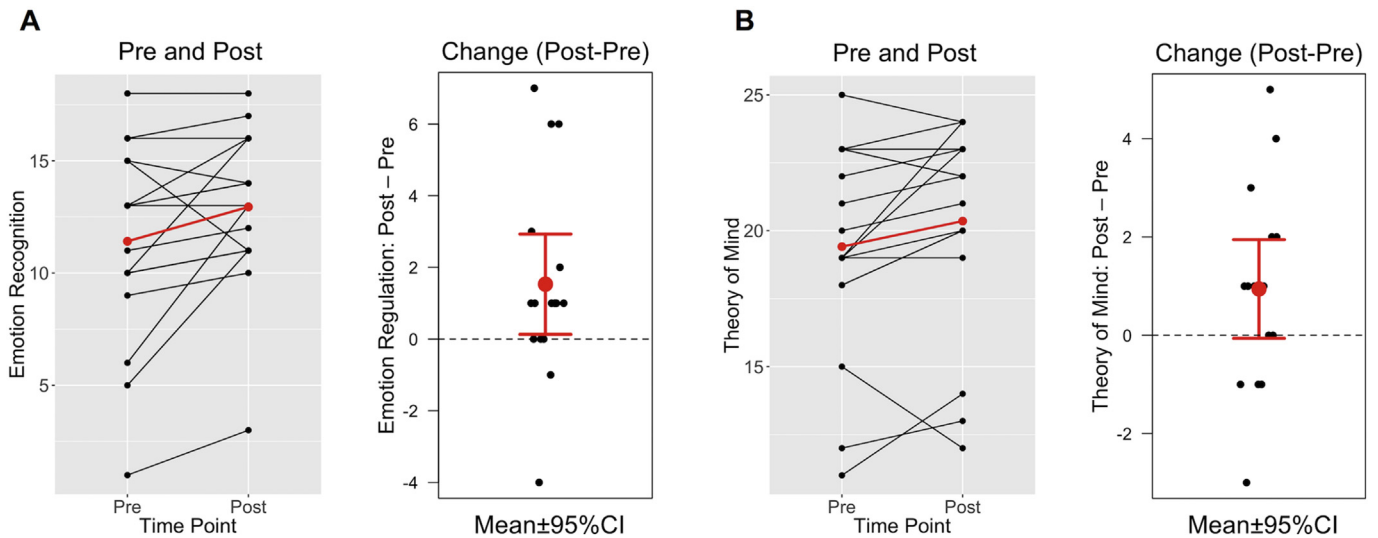


Fig. 2. Treatment effectiveness quantified as the changes in (A) emotion regulation and (B) theory of mind. In each panel, left: The black lines indicate each patient's change in the behavioral measure from pretreatment to posttreatment, while the red line denotes the group mean and the 95% confidence interval (CI) of Δ , the change score (that is, post minus pre). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

pre-treatment brain BOLD responses to BIO vs. SCR on the change in emotion recognition from baseline to treatment endpoint revealed a network (758 voxels) of two distinct clusters of neuropredictive activities. Cluster#1 has 319 voxels and included primarily the left posterior superior temporal sulcus, left superior temporal gyrus, and left middle temporal gyrus. Cluster#2 has 439 voxels and included primarily the right insula, right orbitofrontal cortex, and right inferior frontal gyrus. The scatterplot in Fig. 3 also illustrates the form of the neuropredictive relationship between the change in emotion recognition (y-axis) vs. pretreatment BIO > SCR activation (x-axis) for each network and the estimated effect size in terms of Pearson's $r = 0.71$. It should be noted that the data points in the scatterplot were retrieved from the surviving neuropredictive network; they serve the purpose of illustration only and were not the results of a separate analysis. As can be seen, greater levels of pretreatment activation in the network were positively correlated with improvement of emotion recognition brought about by VR-SCT. There was no region that showed negative correlation

between pretreatment activation and change in emotion recognition. Table 2 lists the peak significance, peak coordinates, spatial extent, and anatomical locations encompassed by each predictive cluster within the network.

To further interpret the possible functions of each distinct cluster in the neuropredictive networks for predicting change in emotion recognition, we conducted a NeuroSynth-based (<http://neurosynth.org>) reverse inference analysis. As seen in Table 3, cluster#1 correlates with language processing, meaning comprehension and integration, semantic inference, and resolution of meaning conflicts, such as interpretation of incongruent auditory emotions and prosody (Mitchell, 2006). Cluster#2 correlates with socio-emotional experience, interpersonal affective information, emotional regulation (Nummenmaa, Hirvonen, Parkkola, & Hietanen, 2008; Walter et al., 2009), and social information processing (Yang et al., 2015). The image files from this analysis are available at <http://neurovault.org/collections/1924> so that interested readers may independently decode the image files with

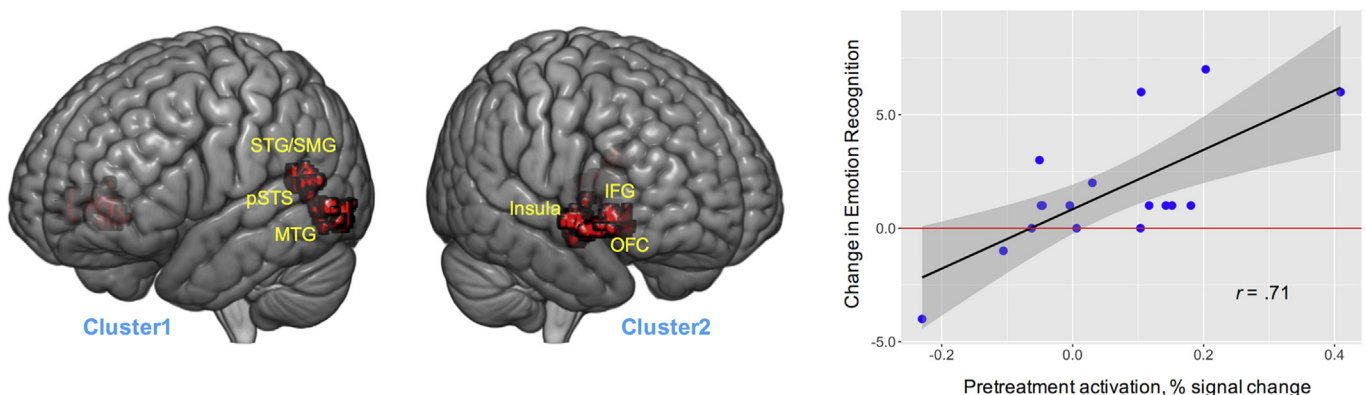


Fig. 3. Neural prediction of change in emotion recognition induced by VR-SCT using biological motion fMRI and univariate GLM. Greater pretreatment BOLD activations to biological vs. scrambled motion in these regions were associated with more improvement in emotion recognition due to VR-SCT. The scatterplot illustrated the relationship between pretreatment BOLD activations and actual changes in emotion recognition (post minus pre), with a horizontal reference line at $y = 0$ indicating no change from pretreatment to posttreatment (that is, post = pre). The regression lines and the 95% confidence intervals were plotted. The results were based on cluster-defining threshold of $Z > 2.33$, $p < 0.01$ and cluster-level threshold of $p < 0.05$. Site, IQ, age, sex, and pretreatment autism symptoms severity were included as covariates of no interest. pSTS, posterior superior temporal sulcus; MTG, middle temporal gyrus; STG, superior temporal gyrus; SMG, supramarginal gyrus; IFG, inferior frontal gyrus; OFC, orbitofrontal cortex; VR-SCT, Virtual Reality-Social Cognition Training.

Table 2

Neural prediction of change in emotion recognition due to VR-SCT using biological motion fMRI and mass univariate GLM.

Cluster	N_{voxels}	Local Maxima					
		Anatomical Region		x	y	z	Z_{peak}
#1	319	L	Banks of the STS	−58	−50	14	3.58
		L	Inferior temporal gyrus	−54	−56	−2	3.64
		L	Middle temporal gyrus	−54	−56	−2	3.64
		L	Superior temporal gyrus	−58	−50	14	3.58
		L	Supramarginal gyrus	−56	−50	16	2.77
#2	439	R	Insula	42	0	−14	3.25
		R	Lateral orbitofrontal cortex	42	28	−8	3.60
		R	Pars opercularis	52	10	−2	3.54
		R	Pars orbitalis	44	28	−8	3.72
		R	Precentral gyrus	54	8	0	3.07
		R	Superior temporal gyrus	52	8	−6	4.30

Note. The coordinates are in MNI152 space, mm. L, Left; R, Right; STS, Superior temporal sulcus; VR-SCT, Virtual Reality-Social Cognition Training. The analysis was corrected, with voxel-level threshold $Z > 2.33$, $p < 0.01$, and cluster-level threshold $p < 0.05$. Greater pretreatment activation to the social perception contrast: biological motion vs. scrambled motion in these regions was associated with greater improvement in emotion recognition brought out by VR-SCT. Site, IQ, age, sex, and pretreatment autism symptoms severity were included as covariates of no interest.

NeuroSynth through links within the NeuroVault website.

For change in theory of mind, contrary to our hypothesis, the neuropsychological analysis revealed no surviving clusters. However, to avoid possible Type II errors and the premature conclusion that biological motion fMRI cannot predict treatment outcome in terms of change in theory of mind, we followed a procedure recommended in the literature (Lieberman & Cunningham, 2009) by analyzing the uncorrected results with voxel-level threshold $Z > 3.29$, $p < 0.001$, and minimal extent $k = 20$ voxels. The analysis revealed three small distinct clusters of neuropsychological activities, which were reported in [Supplementary Material 3](#) for readers' reference.

Because the neuropsychological network for change in emotion recognition is implicated in verbal and non-verbal socio-emotional processing, the finding suggests that individuals who demonstrate greater activity in brain regions implicated in socio-emotional processing at baseline show greater treatment response. Here, there is a question whether the pretreatment activities in these regions are specific to predicting change in emotion recognition, or also correlated with pretreatment abilities of emotion recognition. To address this question, we conducted additional analyses to examine the neural correlates of pretreatment levels of emotion recognition. As shown in [Supplementary Material 4](#), the pretreatment levels of emotion recognition were correlated with brain regions that are distinct from the neuropsychological network for change in emotion recognition. This suggests that the neuropsychological regions are specific to treatment outcome and separable from the regions that are associated with variance in emotion recognition capacities at the pretreatment time point. The corresponding neural correlates for pretreatment levels of theory of mind were reported in [Supplementary Material 5](#). Compared to [Supplementary Material 3](#), the finding also suggests that there exist distinct pretreatment neural correlates for the longitudinal change

in theory of mind brought out by VR-SCT versus the cross-sectional pretreatment variance in theory of mind.

3.3. Multivariate pattern analyses with cross validation

To guard against the possibility of data over-fitting in the mass univariate analyses and to gain an understanding of how the voxels comprising the univariate clusters worked together in predicting change in emotion recognition, we applied regression-based Multivariate Pattern Analyses (MVPA) of pretreatment BOLD responses to the contrast of BIO > SCR with leave-one-subject-out (LOSO) cross-validation in the voxels comprising the univariate clusters. As shown in [Table 4](#), the neuropsychological network consisting of the two clusters survived cross validation—the multivariate pattern information from this brain network significantly predicted change in emotion recognition ($r = 0.79$, $p = 0.001$; $R^2 = 0.62$, $p = 0.002$; $nMSE = 0.24$, $p = 0.001$). [Fig. 4\(A\)](#) shows the weight map (i.e. model parameters) in the representative slices of this network derived from the multivariate modeling of pretreatment images predicting change in emotion recognition. [Fig. 4\(B\)](#) shows the scatterplot of actual versus predicted treatment response on emotion recognition. As estimated via the LOSO framework, each of the points in this plot was derived from a separate training set, and for a new unseen patient (testing set), the remaining 16 participants' data were used as the training set. Thus, the correlation is not a standard correlation derived from a single set of participants. Rather, each point reflects different combinations of training and testing sets. Finally, [Fig. 4\(C\)](#) shows the line plot of actual versus predicted treatment response on emotion recognition for each of the 17 folds.

In addition, we conducted MVPA analyses with comparison/control ROIs that we did not expect to be predictive of treatment outcome in ASD. The inferior occipital gyrus and middle occipital gyrus were selected because: (a) they respond strongly to a range of visual stimuli including the SCR and BIO stimuli used here, and (b) the occipital gyri were shown to be neuropsychological of treatment effectiveness in a markedly different neuropsychiatric condition, social anxiety disorder (Doehrmann et al., 2013). Furthermore, we conducted MVPA with the whole brain (including the neuropsychological network) to evaluate the specificity of our findings to the network of these univariate clusters. As shown in [Table 4](#), neither the comparison ROIs nor the whole brain analysis were predictive of treatment outcome ($ps > 0.05$).

3.4. Demographic and behavioral findings

To evaluate whether fMRI provides unique information concerning the prediction of response to VR-SCT on emotion recognition and theory of mind, we examined how a host of demographic and pretreatment ASD symptom severity measures predict changes in emotion recognition and theory of mind, respectively. We ran correlation analyses between each of the measures listed in [Table 1](#) and the delta change in emotion recognition and that in theory of mind. As seen in [Supplementary Material 6](#), no measure showed a significant correlation, $ps > 0.11$. This suggests that the fMRI

Table 3

Reverse inference analysis of the neuropsychological clusters for change in emotion recognition using NeuroSynth.

Cluster	Top 10 NeuroSynth-decoded feature terms
#1	Conflicting (0.120), integration (0.107), lexical (0.104), sentence (0.100), verbs (0.085), linguistic (0.081), meaning (0.077), audiovisual (0.075), comprehension (0.072), reading (0.072)
#2	Musical (0.079), interpersonal (0.046), painful (0.046), monitor (0.041), regulating (0.037), nociceptive (0.034), social (0.034), affective (0.033), emotions (0.031), sad (0.030)

Note. Numbers within the parentheses are correlation coefficients between the surviving clusters and the meta-analysis maps of the feature terms in NeuroSynth.

Table 4

Predictive accuracy of the univariate neuropsychological network for change in emotion recognition, as estimated by MVPA with cross validation.

Mask	N_{voxels}	r	$p(r)$	R^2	$p^2(R)$	$nMSE$	$p(nMSE)$
Neuropsychological network	758	0.79	0.001	0.62	0.002	0.24	0.001
Inferior occipital gyrus	1930	-0.24	0.67	0.06	0.50	1.04	0.58
Middle occipital gyrus	5368	-0.08	0.46	0.01	0.83	0.88	0.44
Whole brain	228,453	-0.65	0.92	0.43	0.08	0.90	0.88

Note. Predictive accuracy was indicated by Pearson's correlation coefficient (r), coefficient of determination (R^2), and normalized mean squared error ($nMSE$) between predicted and actual change in emotion recognition. Significance (p -value) was determined with a random permutation test (50,000 iterations). Significant regions and statistics were displayed in bold. Cross validation was based on a 17-fold leave-one-subject-out (LOSO) framework.

measure provided unique advantage over behavioral measures on predicting behavioral responses to VR-SCT.

4. Discussion

As hypothesized, the biological motion fMRI task successfully identified brain regions in which pretreatment brain activations during passively viewing biological motion versus scrambled motion predicted change in verbal and non-verbal emotion recognition in a study of young adults with high-functioning autism, who received an evidence-based VR-SCT behavioral intervention.

Specifically, the key brain regions are implicated in functions supporting (a) language-related comprehension and meaning integration, as well as interpretation of incongruent auditory emotions and prosody, and (b) socio-emotional experience processing, interpersonal affective information processing, and emotional regulation. These two groups of functions are closely related to verbal and non-verbal emotion recognition, respectively. As such, it is possible that the regions may be involved in processing the socio-emotional aspect of the biological motion stimuli, that is, how the social emotions associated with adult's actions in children's games (e.g., waving, pat-a-cake, peek-a-boo) are decoded and interpreted.

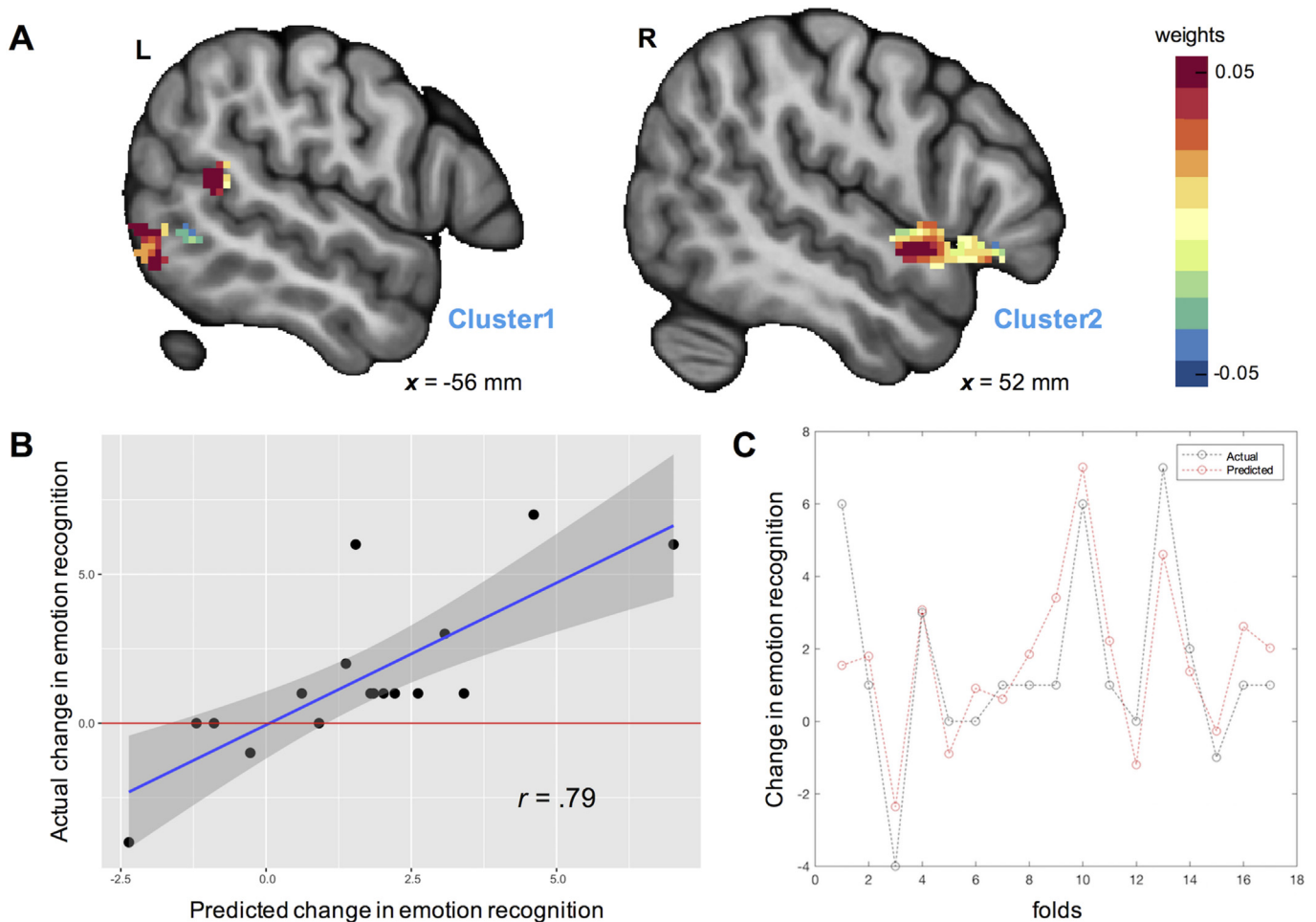


Fig. 4. Predictive accuracy of the univariate neuropsychological clusters, as estimated by MVPA with cross validation, for predicting change in emotion recognition due to VR-SCT. (A): Weight map showing the relative weights derived from the multivariate modeling of pretreatment response to biological motion that contributed to the prediction of change in emotion recognition (that is, post minus pre) at representative slices (MNI152 mm space). **(B):** Scatterplot illustrating actual and predicted changes in emotion recognition, with a horizontal reference line at $y = 0$ indicating no change from pretreatment to posttreatment (that is, post = pre). The regression lines and the 95% confidence intervals were plotted. Cross validation was based on a 17-fold leave-one-subject-out (LOSO) framework. **(C):** Line plot illustrating actual and predicted changes in emotion recognition for each of the 17 folds used in the cross-validation framework. The results were based on the neuropsychological network estimated with cluster-defining threshold of $Z > 2.33$, $p < 0.01$ and cluster-level threshold of $p < 0.05$. VR-SCT, Virtual Reality-Social Cognition Training.

Importantly, the results were supported by regression-based MVPA with a standard LOSO cross-validation framework, which suggests that the brain activities within the neuropsychological networks may serve as robust predictive biomarkers, generalizable to new, unseen participants. Moreover, we found that there existed distinctively different regions that are associated with variance in emotion recognition capacities at the pretreatment time point, which suggests that the neuropsychological regions may be specific to predicting treatment outcome. To our knowledge, the current findings provide the first evidence of neuroimaging-derived predictive biomarkers in young adults with ASD. The predictive biomarkers identified in this research may be interpreted as the pretreatment neurobiological readiness to respond to a specific treatment, VR-SCT, in the domain of emotion recognition.

For change in theory of mind, although the corrected results were not as hypothesized and revealed no surviving regions, the exploratory uncorrected analyses identified several neuropsychological regions for change in theory of mind. These regions need to be interpreted with caution. First, they were uncorrected in nature and due to skipping the procedure of correcting for multiple comparisons, there is a greater chance of Type I error here. Second, they were based on a very high voxel-level threshold and had relatively small extent. Thus, they may be ill-suited to be analyzed by MVPA because the small extent would render the results unreliable. Nonetheless, the exploratory analyses here raised the possibility that fMRI can predict change in not only emotion recognition but also theory of mind, and there likely exist separable neuropsychological regions for change in emotion recognition and theory of mind, respectively.

Compared to the previous neuropsychological study involving a 16-week Pivotal Response Treatment and young children with autism, and using the change in SRS total raw scores as the measure of treatment effectiveness (Yang et al., 2016), the current study has a number of similarities and differences. First, in terms of similarities, both studies used the same biological motion fMRI task (Kaiser et al., 2010) as the pretreatment neural predictor. This supports the usability of the task across a wide range of age, treatment modality, and outcome measures. Both studies also recruited high-functioning individuals with autism as participants, included both male and female participants, and employed the same data analytical pipeline and cluster thresholding. Second, in terms of differences, the current study involved a 5-week VR-SCT intervention and young adults with autism, and used changes in emotion regulation and theory of mind, respectively, as the treatment outcome variables, while the SRS total raw scores measured trait-like (6-month tendency) autism symptom severity and was not suitable for a 5-week intervention.

Given the differences, particularly on the variable of treatment effectiveness, we did not expect to identify the same neural predictive biomarkers in the current study. Specifically, the cluster#1 in the current study, predicting change in emotion recognition, was around the left pSTS region. In contrast, the cluster#1 in the previous study (Yang et al., 2016), predicting change in autism symptom severity, was around the right pSTS region. The left pSTS region was known for its role in language processing and resolution of conflicting meaning, which is arguably closely relevant to verbal emotion recognition, while the verbal behavioral measure of emotion recognition in this project consisted of prosody understanding (SP-Prosody) and conflicts between literal and non-literal meanings (SP-Pairs). In contrast, the right pSTS region was implicated in temporal and spatial sensory information integration, which is arguably closely related to the core symptoms of autism (Yang et al., 2015). Moreover, the cluster#2 in the current study was near the right insula but more dorsal and extended to the pars opercularis of the inferior frontal gyrus, while the cluster#3 in the

previous study was also near the right insula but more ventral and extended to the temporal pole. These two clusters are spatially close and located to the same hemisphere. This region at large is generally known for socio-emotional processing and emotion recognition, which is arguably closely aligned with non-verbal behavioral measure of emotion recognition in the current project (SP-Affect Naming), in which the participants were asked to identify the expressed emotion from a given list when presented various faces. Relatedly, difficulty in interpretation of facial expressions of emotion is also frequently cited as one of the main characteristics associated with the core symptoms of autism (Eack, Mazefsky, & Minshew, 2015).

There has been a significant barrier in identifying and predicting which treatments might be beneficial for a given individual with ASD, before the treatment is prescribed and delivered. Joining a recent discovery of predictive biomarkers in young children with ASD (Yang et al., 2016), our findings advanced the field one more step forward toward the goal of targeted, personalized treatment for individuals with ASD. Although more research is needed, it is promising that fMRI techniques may provide guidelines and suggestions for possible treatment(s) for those who are most likely to immediately benefit from the treatment(s). For those who would be otherwise unlikely to immediately benefit from the treatment, the current findings also raise the question as to whether increasing pretreatment activations through other types of pretreatment interventions, may theoretically lead to better response to treatment. Accordingly, more research is needed to investigate methods that would increase the pretreatment activation in these individuals and test the hypothesis as to whether they could become more ready to respond to treatment. For example, in adults with ASD, oxytocin has been shown to increase brain activation, particularly in regions similar to the predictive biomarkers identified in this research such as inferior frontal gyrus (Domes, Kumbier, Heinrichs, & Herpertz, 2014). It is thus possible that administration of intranasal oxytocin at pretreatment may increase brain readiness to respond to treatment in ASD (e.g., increasing learning rate of socio-emotional and cognitive training) for those who would not be able to immediately benefit from the behavioral intervention alone. This possibility may be tested in future research.

4.1. Limitations

Several limitations should be considered regarding this research. First, the neuropsychological findings were limited to one single treatment-only group in a pretest-posttest design. Future work should conduct randomized controlled trials to further establish these findings. A waitlist control group will generate additional insight into the robustness and specificity of the neural predictive biomarkers (they should have no predictive abilities in the absence of an ongoing treatment) as well as further evaluation of the test-retest reliability of the behavioral measures within the 5-week span of the current study. Second, the size of our preliminary sample ($N = 17$) is relatively small. The neuropsychological findings need to be further tested for reproducibility in an independent, larger sample. A larger sample would also increase the statistical power, which is needed to detect smaller sizes of effect. On the contrary, the non-significant neuropsychological results on theory of mind (after controlling for multiple comparisons) should not be interpreted as no predictive potential in this ability. The uncorrected exploratory analyses still revealed several regions of neuropsychological activities for change in theory of mind.

Third, the primary clinical outcomes are limited to changes in emotion recognition and theory of mind, respectively. Although these two abilities are among the most basic abilities in socio-emotional and socio-cognitive processing (Baron-Cohen, Leslie, &

Frith, 1985; Gallese, Keysers, & Rizzolatti, 2004) that underlie a wide range of social skills and are also central to our understanding of ASD deficits (Uljarevic & Hamilton, 2013), there is a need for future research to include other ASD-related measures, such as interaction behaviors (Rice & Redcay, 2016) and conversation skills (Scattone, 2008). Future studies are also needed to establish the minimal clinically important difference (MCID) in the change scores in these behavioral measures, and include naturalistic measures to test whether the behavioral effects can generalize to real-life daily functioning at the 5-week posttreatment endpoint. Fourth, all participants were high-functioning and it remains unclear whether the findings may apply to all individuals with autism. The VR-SCT has an inclusion criterion of $IQ \geq 80$ and future research needs to investigate other treatment approaches for those who may be cognitively impaired. Fifth, while the findings suggest that clinicians may potentially one day use brain activity patterns to identify individuals with autism who would benefit from a given treatment, wide use of such an approach might depend on clinicians' ability to gather the data using imaging methods robust to motion or relatively inexpensive, such as functional near-infrared spectroscopy (fNIRS).

Finally, while the current paper focused exclusively on testing the use of fMRI as a forecasting tool to facilitate subject selection to inform future treatment design, it did not address the question of why VR-SCT works on the brain level. The latter question concerns the neural mechanisms of change and can be answered by comparing brain activations before and after VR-SCT, or analyzing what brain changes track with behavioral changes induced by VR-SCT. This question requires a full consideration well beyond the scope of the current paper and should be further pursued in future research and analyses.

5. Conclusions

Despite the limitations, for the first time in the field of ASD, we provide evidence that treatment effectiveness at the individual level in adults with ASD in the domain of verbal and non-verbal emotion recognition can be accurately predicted by the pretreatment fMRI activations using biological motion task in brain circuits implicated in (a) language comprehension and integration, and processing incongruent auditory emotions and prosody, and (b) socio-emotional experience processing, interpersonal affective information processing, and emotional regulation. Relative to children with ASD, intervention research for adults with ASD is very limited. This study offers a key direction toward increasing the effectiveness of intervention for adults with ASD and extends the findings that the biological motion fMRI task can be used to predict treatment outcome from young children with ASD to young adults with ASD. Our results thus open a new avenue for important future research and should potentially accelerate progress toward developing more precise and effective treatments for individuals with ASD across the lifespan.

Conflict of interest

The authors declare no competing financial interests.

Acknowledgments

We thank the participants and their families included in this study for their time and participation and the research assistants in our research centers, making this research possible. This work was supported by the Harris Professorship at Yale Child Study Center to KAP, Autism Speaks Meixner Postdoctoral Fellowship in Translational Research (#9284) to DY, a gift from the Autism

Society—Northwestern Pennsylvania to DY, and the Yale University Biomedical High Performance Computing Center (NIH grants RR19895 and RR029676-01). Additionally, we thank the Rees-Jones Foundation, Vin and Caren Prothro Foundation, and the Crystal Charity Ball for their generous support of Center for BrainHealth's research. We also thank Jeffrey Spence for feedback with statistical analyses.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.brat.2017.03.014>.

References

- Abell, F., Happe, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*, 15, 1–16.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, 4, 267–278.
- APA. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5* (5th ed.). Washington, D.C: American Psychiatric Publishing.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21, 37–46.
- Bellani, M., Fornasari, L., Chittaro, L., & Brambilla, P. (2011). Virtual reality in autism: State of the art. *Epidemiology and Psychiatric Sciences*, 20, 235–238.
- Bjornsdotter, M., Wang, N., Pelphrey, K., & Kaiser, M. D. (2016). Evaluation of quantified social perception circuit activity as a neurobiological marker of autism spectrum disorder. *JAMA Psychiatry*, 73, 614–621.
- Chen, S. H., & Bernard-Opitz, V. (1993). Comparison of personal and computer-assisted instruction for children with autism. *Mental Retardation*, 31, 368–376.
- Cheng, Y. F., Moore, D., McGrath, P., & Fan, Y. L. (2005). Collaborative virtual environment technology for people with autism. 5th IEEE International Conference on Advanced Learning Technologies, Proceedings, 247–248.
- Chu, C., Ni, Y., Tan, G., Saunders, C. J., & Ashburner, J. (2011). Kernel regression for fMRI pattern prediction. *Neuroimage*, 56, 662–673.
- Constantino, J. N. (2012). *Social responsiveness scale* (2nd ed.). Torrance, CA: Western Psychological Services.
- Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015). Functional organization of social perception and cognition in the superior temporal sulcus. *Cerebral Cortex*, 25, 4596–4609.
- Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, 31, 968–980.
- Didehban, N., Allen, T., Kandalaf, M., Krawczyk, D., & Chapman, S. (2016). Virtual reality social cognition training for children with high functioning autism. *Computers in Human Behavior*, 62, 703–711.
- Doehrmann, O., Ghosh, S. S., Polli, F. E., Reynolds, G. O., Horn, F., Keshavan, A., et al. (2013). Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA Psychiatry*, 70, 87–97.
- Domes, G., Kumbier, E., Heinrichs, M., & Herpertz, S. C. (2014). Oxytocin promotes facial emotion recognition and amygdala reactivity in adults with asperger syndrome. *Neuropsychopharmacology*, 39, 698–706.
- Eack, S. M., Mazefsky, C. A., & Minshew, N. J. (2015). Misinterpretation of facial expressions of emotion in verbal adults with autism spectrum disorder. *Autism*, 19, 308–315.
- Etkin, A., Buchel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews Neuroscience*, 16, 693–700.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160.
- Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, 8, 396–403.
- 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.
- Hahamy, A., Behrmann, M., & Malach, R. (2015). The idiosyncratic brain: Distortion of spontaneous connectivity patterns in autism spectrum disorder. *Nature Neuroscience*, 18, 302–309.
- Hahn, T., Kircher, T., Straube, B., Wittchen, H. U., Konrad, C., Strohle, A., et al. (2015). Predicting treatment response to cognitive behavioral therapy in panic disorder with agoraphobia by integrating local neural information. *JAMA Psychiatry*, 72, 68–74.
- Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annual Review of Neuroscience*, 37, 435–456.
- Heberlein, A. S., & Adolphs, R. (2004). Impaired spontaneous anthropomorphizing despite intact perception and social knowledge. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 7487–7491.

- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243–259.
- Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., et al. (2011). Neural systems predicting long-term outcome in dyslexia. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 361–366.
- Hu, Z., Chan, R. C., & McAlonan, G. M. (2010). Maturation of social attribution skills in typically developing children: An investigation using the social attribution task. *Behavioral and Brain Functions*, 6, 10.
- Hus, V., & Lord, C. (2014). The autism diagnostic observation schedule, module 4: Revised algorithm and standardized severity scores. *Journal of Autism and Developmental Disorders*, 44, 1996–2012.
- Insel, T. R. (2014). The NIMH research domain criteria (RDoC) project: Precision medicine for psychiatry. *The American Journal of Psychiatry*, 171, 395–397.
- Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *Neuroimage*, 62, 782–790.
- Johansson, G. (1973). Visual-perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 201–211.
- Johnson, M. H. (2006). Biological motion: A perceptual life detector? *Current Biology*, 16, R376–R377.
- Kaiser, M. D., Hudac, C. M., Shultz, S., Lee, S. M., Cheung, C., Berken, A. M., et al. (2010). Neural signatures of autism. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 21223–21228.
- Kandalaf, M. R., Didehban, N., Cullum, C. M., Krawczyk, D. C., Allen, T. T., Tamminga, C. A., et al. (2012). The wechsler ACS social perception subtest: A preliminary comparison with other measures of social cognition. *Journal of Psychoeducational Assessment*, 30, 455–465.
- Kandalaf, M. R., Didehban, N., Krawczyk, D. C., Allen, T. T., & Chapman, S. B. (2013). Virtual reality social cognition training for young adults with high-functioning autism. *Journal of Autism and Developmental Disorders*, 43, 34–44.
- Kanske, P., Heissler, J., Schonfelder, S., Bongers, A., & Wessa, M. (2011). How to regulate emotion? Neural networks for reappraisal and distraction. *Cerebral Cortex*, 21, 1379–1388.
- Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, 459, U142–U157.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863.
- Lieberman, M. D., & Cunningham, W. A. (2009). Type I and type II error concerns in fMRI research: Re-balancing the scale. *Social Cognitive and Affective Neuroscience*, 4, 423–428.
- Lombardo, M. V., Lai, M. C., Auyeung, B., Holt, R. J., Allison, C., Smith, P., et al. (2016). Unsupervised data-driven stratification of mentalizing heterogeneity in autism. *Scientific Reports*, 6, 35333.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30, 205–223.
- Malone, I. B., Leung, K. K., Clegg, S., Barnes, J., Whitwell, J. L., Ashburner, J., et al. (2015). Accurate automatic estimation of total intracranial volume: A nuisance variable with less nuisance. *Neuroimage*, 104, 366–372.
- Mansson, K. N., Frick, A., Boraxbekk, C. J., Marquand, A. F., Williams, S. C., Carlbring, P., et al. (2015). Predicting long-term outcome of Internet-delivered cognitive behavior therapy for social anxiety disorder using fMRI and support vector machine learning. *Translational Psychiatry*, 5, e530.
- Maskey, M., Lowry, J., Rodgers, J., McConachie, H., & Parr, J. R. (2014). Reducing specific phobia/fear in young people with autism spectrum disorders (ASDs) through a virtual reality environment intervention. *PLoS One*, 9, e100374.
- McKay, L. S., Simmons, D. R., McAleer, P., Marjoram, D., Piggot, J., & Pollick, F. E. (2012). Do distinct atypical cortical networks process biological motion information in adults with Autism Spectrum Disorders? *Neuroimage*, 59, 1524–1533.
- Mitchell, R. L. C. (2006). How does the brain mediate interpretation of incongruent auditory emotions? The neural response to prosody in the presence of conflicting lexico-semantic cues. *European Journal of Neuroscience*, 24, 3611–3618.
- Moore, M., & Calvert, S. (2000). Brief report: Vocabulary acquisition for children with autism: Teacher or computer instruction. *Journal of Autism and Developmental Disorders*, 30, 359–362.
- Nummenmaa, L., Hirvonen, J., Parkkola, R., & Hietanen, J. K. (2008). Is emotional contagion special? An fMRI study on neural systems for affective and cognitive empathy. *Neuroimage*, 43, 571–580.
- Parsons, S., & Cobb, S. (2011). State-of-the-art of virtual reality technologies for children on the autism spectrum. *European Journal of Special Needs Education*, 26, 355–366.
- Parsons, S., & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research*, 46, 430–443.
- Parsons, S., Mitchell, P., & Leonard, A. (2004). The use and understanding of virtual environments by adolescents with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, 34, 449–466.
- Parsons, S., Mitchell, P., & Leonard, A. (2005). Do adolescents with autistic spectrum disorders adhere to social conventions in virtual environments? *Autism*, 9, 95–117.
- Pearson. (2009). *Advanced clinical solutions (ACS) for the WAIS-IV and WMS-IV*. San Antonio, TX: Pearson.
- Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron*, 48, 175–187.
- Pruim, R. H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-aroma: A robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage*, 112, 267–277.
- Rice, K., & Redcay, E. (2016). Interaction matters: A perceived social partner alters the neural processing of human speech. *Neuroimage*, 129, 480–488.
- Saggar, M., Shelly, E. W., Lepage, J. F., Hoeft, F., & Reiss, A. L. (2014). Revealing the neural networks associated with processing of natural social interaction and the related effects of actor-orientation and face-visibility. *Neuroimage*, 84, 648–656.
- Scattone, D. (2008). Enhancing the conversation skills of a boy with Asperger's Disorder through Social Stories and video modeling. *Journal of Autism and Developmental Disorders*, 38, 395–400.
- Schrouff, J., Rosa, M. J., Rondina, J. M., Marquand, A. F., Chu, C., Ashburner, J., et al. (2013). PRoNTo: Pattern recognition for neuroimaging toolbox. *Neuroinformatics*, 11, 319–337.
- Screaming_Bee. (2005). *MorphVox* (3.4.18 ed.). (In).
- Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 809–813.
- Smith, M. J., Ginger, E. J., Wright, K., Wright, M. A., Taylor, J. L., Humm, L. B., et al. (2014). Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44, 2450–2463.
- Trepagnier, C. Y., Olsen, D. E., Boteler, L., & Bell, C. A. (2011). Virtual conversation partner for adults with autism. *Cyberpsychology Behavior and Social Networking*, 14, 21–27.
- Tzanavari, A., Charalambous-Darden, N., Herakleous, K., & Poullis, C. (2015). Effectiveness of an Immersive Virtual Environment (CAVE) for Teaching Pedestrian Crossing to Children with PDD-NOS. 15th IEEE International Conference on Advanced Learning Technologies (ICALT 2015), 423–427.
- Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, 43, 1517–1526.
- Vallortigara, G., Regolin, L., & Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biology*, 3, e208.
- Ventola, P., Yang, D. Y., Friedman, H. E., Oosting, D., Wolf, J., Sukhodolsky, D. G., et al. (2015). Heterogeneity of neural mechanisms of response to pivotal response treatment. *Brain Imaging and Behavior*, 9, 74–88.
- van Waarde, J. A., Scholte, H. S., van Oudheusden, L. J., Verwey, B., et al. (2015). A functional MRI marker may predict the outcome of electroconvulsive therapy in severe and treatment-resistant depression. *Molecular Psychiatry*, 20, 609–614.
- Wainer, A. L., & Ingersoll, B. R. (2011). The use of innovative computer technology for teaching social communication to individuals with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 5, 96–107.
- Wallace, S., Parsons, S., Westbury, A., White, K., White, K., & Bailey, A. (2010). Sense of presence and atypical social judgments in immersive virtual environments. Responses of adolescents with Autism Spectrum Disorders. *Autism*, 14, 199–213.
- Walter, H., von Kalckreuth, A., Schardt, D., Stephan, A., Goschke, T., & Erk, S. (2009). The temporal dynamics of voluntary emotion regulation. *PLoS One*, 4, e6726.
- Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence (WASI)*. San Antonio, TX: Pearson.
- Wechsler, D. (2011). *Wechsler abbreviated scale of intelligence (WASI-II)* (2nd ed.). San Antonio, TX: Pearson.
- Whitfield-Gabrieli, S., Ghosh, S. S., Nieto-Castanon, A., Saygin, Z., Doehrmann, O., Chai, X. J., et al. (2016). Brain connectomics predict response to treatment in social anxiety disorder. *Molecular Psychiatry*, 21, 680–685. <http://dx.doi.org/10.1038/mp.2015.109>. published online 11 August 2015.
- Yang, D., Pelphrey, K. A., Sukhodolsky, D. G., Crowley, M. J., Dayan, E., Dvornek, N. C., et al. (2016). Brain responses to biological motion predict treatment outcome in young children with autism. *Translational Psychiatry*, 6, e948.
- Yang, D. Y., Rosenblau, G., Keifer, C., & Pelphrey, K. A. (2015). An integrative neural model of social perception, action observation, and theory of mind. *Neuroscience and Biobehavioral Reviews*, 51, 263–275.
- Yang, Y. J., Sukhodolsky, D. G., Lei, J., Dayan, E., Pelphrey, K. A., & Ventola, P. (2017). Distinct neural bases of disruptive behavior and autism symptom severity in boys with autism spectrum disorder. *Journal of Neurodevelopmental Disorders*, 9, 1.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8, 665–670.