



Published in final edited form as:

*Autism Res.* 2012 December ; 5(6): 407–418. doi:10.1002/aur.1252.

## Deficits in Mental State Attributions in Individuals with 22q11.2 Deletion Syndrome (Velo-Cardio-Facial Syndrome)

Jennifer S. Ho<sup>1,\*</sup>, Petya D. Radoeva<sup>2,\*</sup>, Maria Jalbrzikowski<sup>3</sup>, Carolyn Chow<sup>1</sup>, Jessica Hopkins<sup>1</sup>, Wen-Ching Tran<sup>1</sup>, Ami Mehta<sup>1</sup>, Nicole Enrique<sup>1</sup>, Chelsea Gilbert<sup>1</sup>, Kevin M. Antshel<sup>4</sup>, Wanda Fremont<sup>4</sup>, Wendy R. Kates<sup>4,\*</sup>, and Carrie E. Bearden<sup>1,3,\*</sup>

<sup>1</sup>Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles: 760 Westwood Plaza Los Angeles, CA 90095

<sup>2</sup>Department of Neuroscience and Physiology, State University of New York Upstate Medical University, Syracuse, New York: 750 East Adams Street Syracuse, NY 13210-2375

<sup>3</sup>Department of Psychology, University of California, Los Angeles: 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563

<sup>4</sup>Department of Psychiatry and Behavioral Sciences, State University of New York Upstate Medical University, Syracuse, New York : 750 East Adams Street Syracuse, NY 13210-2375

### Abstract

Velo-cardio-facial syndrome (VCFS; 22q11.2 deletion syndrome) results from a genetic mutation that increases risk for Autism Spectrum Disorder (ASD). We compared Theory of Mind (ToM) skills in 63 individuals with VCFS (25% with an ASD diagnosis) and 43 typically-developing controls, and investigated the relationship of ToM to reciprocal social behavior. We administered a video-based task to assess mentalizing at two sites (UCLA and SUNY Upstate Medical University). The videos depicted interactions representing complex mental states (ToM condition), or simple movements (Random condition). Verbal descriptions of the videos were rated for Intentionality (i.e., mentalizing) and Appropriateness. Using Repeated Measures ANOVA, we assessed the effects of VCFS and ASD on Intentionality and Appropriateness, and the relationship of mentalizing to Social Responsiveness Scale (SRS) scores. Results indicated that individuals with VCFS overall had lower Intentionality and Appropriateness scores than controls for ToM, but not for Random scenes. In the SUNY sample, individuals with VCFS, both with and without ASD, performed more poorly than controls on the ToM condition; however, in the UCLA sample, only individuals with VCFS without ASD performed significantly worse than controls on the ToM condition. Controlling for site and age, performance on the ToM condition was significantly correlated with SRS scores. Individuals with VCFS, regardless of an ASD diagnosis, showed impairments in the spontaneous attribution of mental states to abstract visual stimuli, which may underlie real-life problems with social interactions. A better understanding of the social deficits in VCFS is essential for the development of targeted behavioral interventions.

### Keywords

22q11.2 deletion syndrome; Velo-cardio-facial syndrome; Theory of Mind; Autism Spectrum Disorders; reciprocal social behavior; social cognition

---

Correspondence should be addressed to: Carrie E. Bearden, Ph.D., cbearden@mednet.ucla.edu, UCLA Psychiatry & Biobehavioral Sciences, BOX 956968, 300 Medical Plaza, Rm 2267, Los Angeles, CA 90095-6968, Phone: 310-206-2983, Fax: 310-794-9517.  
\*equal authorship contribution of the first two authors and the two senior authors

Social cognition, the ability to understand others' intentions, emotions, and perspectives, is considered a fundamental aspect of human interaction. Deficits in this ability are a hallmark feature of several psychological disorders, including Autism Spectrum Disorders (ASD) and schizophrenia. Velo-cardio-facial syndrome (VCFS), also known as 22q11.2 deletion syndrome or DiGeorge syndrome, is a neurogenetic syndrome that is one of the highest known risk factors for both of these complex neuropsychiatric illnesses (Green, et al., 2009). Approximately 30% of individuals with VCFS develop schizophrenia in adolescence or young adulthood (Green et al., 2009; Murphy, 2002) and Autism Spectrum Disorder (ASD) is diagnosed in 14–50% of children with the deletion (Antshel et al., 2007; Fine et al., 2005; Kates et al., 2007; Vorstman et al., 2006).

VCFS results from a hemizygous deletion at the 22q11.2 locus, and is the most commonly known, contiguous gene deletion syndrome occurring in approximately 1 out of every 2,000–7,000 births (Shprintzen, 2005). Most cases of VCFS occur as *de novo* mutations, although about 10% of cases are familial (McDonald-McGinn et al., 2001). The phenotype is highly variable, but common characteristics include cardiac defects, immune deficiency, palatal anomalies, and neurocognitive deficits (Bearden et al., 2001; McDonald-McGinn et al., 2001; Swillen et al., 2005).

The phenotype of VCFS has considerable overlap with the phenotype of behaviorally defined ASD (Kates et al., 2007). Specifically, significant social impairment characterizes both VCFS (Kiley-Brabeck & Sobin, 2006; Kobrynski & Sullivan, 2007) and idiopathic ASD (Landa, Holman, & Garrett-Mayer, 2007). It has long been hypothesized that deficits in theory of mind (ToM), or the ability to attribute mental states to the self and others, may underlie social impairments observed in ASD (Baron-Cohen, Leslie, & Frith, 1985). Individuals with ASD consistently demonstrate impairments in ToM as compared to healthy controls (Castelli, Frith, Happé, & Frith, 2002; Premack & Woodruff, 1978). However, the remarkable heterogeneity of ASD – at both the phenotypic and genetic level – has posed a challenge for identifying consistent, biologically informative endophenotypes associated with the illness. As such, the study of a genetic subtype, such as VCFS, can serve as an ideal model for elucidating gene-brain-behavior relationships.

Previous studies have reported ToM deficits in individuals with this disorder (Campbell et al., 2009; Chow, Watson, Young, & Bassett, 2006; Niklasson, Rasmussen, Óskarsdóttir, & Gillberg, 2002). However, these studies have either lacked typically developing comparison groups (Campbell et al., 2009; Campbell et al., 2002), or focused on differences between adults with VCFS, with and without schizophrenia (Chow et al., 2006). Recently, Campbell et al. (2011) expanded upon these findings by examining both social-perceptual and social-cognitive aspects of ToM. They found that children and adolescents with VCFS showed impairments in social-perceptual aspects of ToM, as measured by face processing, relative to sibling controls. However, using false belief paradigms, they found that social-cognitive ToM deficits were present only in the younger participants with VCFS, and only for more advanced, second-order false belief tasks. Given that older children with VCFS did not show impairments on these tasks, the authors suggested that these findings could reflect delayed development of the social-cognitive aspects of ToM in youth with VCFS, rather than a deficit. However, interpretation of this pattern of findings is challenging, given that false-belief tasks are grammatically complex, and rely heavily on both working memory and language competence. Additionally, this study did not examine how ToM deficits related to phenotypic variability (i.e. ASD diagnosis) within the syndrome. Thus, examining these issues using a task that assesses implicit aspects of mentalizing may help to better characterize ToM deficits in individuals with VCFS.

Here, we employed the Animations Task (Abell, Happé, & Frith, 2000; Castelli, Happé, Frith, & Frith, 2000) as a measure of ToM in children and adolescents with VCFS, their unaffected siblings, and healthy controls, in order to examine the relationship between ToM skills, social functioning, and ASD diagnosis in VCFS. The Animations task consists of silent video clips depicting triangles that simulate human interactions (“ToM” condition) or triangles that simply move around the screen and serve as a control (“Random”) condition (Abell et al., 2000; Castelli et al., 2000). The Animations Task is a unique measure of ToM that is ideally suited for study in neurodevelopmental disorders because: 1) the stimuli are devoid of social meaning, and thus may provide a more accurate measure of spontaneous mental state attributions; and 2) the task does not rely heavily on verbal comprehension skills or working memory, and thus is less dependent on general cognitive abilities.

The present study, which was conducted at two academic medical center sites, addresses the following primary research questions: 1) Do individuals with VCFS show impairments in ToM when compared with healthy controls?; 2) Is there differential impairment in individuals with VCFS who have an ASD diagnosis?; and 3) Is ToM related to real-world social functioning? Given previous literature indicating social and theory of mind deficits associated with idiopathic schizophrenia and ASD, and the elevated risk for these disorders associated with VCFS, we predicted that: 1) the VCFS group – particularly those with an ASD diagnosis – would show selective impairment on the ToM task relative to healthy controls (but would not differ from controls for the Random condition); and 2) within children and adolescents with VCFS, ToM abilities would be associated with real-world social functioning.

## Methods

### Participants

**UCLA**—Individuals (age range: 6–25 years, Table 2) with a confirmed microdeletion at the 22q11.2 locus were recruited through various cardiology, pediatrics, and genetics clinics in Southern California, and via the Velo-cardio-facial Syndrome Educational Foundation website. Age-matched healthy controls were recruited from local schools and screened for Axis I psychiatric disorders using the Structured Clinical Interview for DSM-IV (First, Spitzer, & Williams, 1997) and/or Computerized Diagnostic Interview Schedule for Children (CDISC; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000). Controls were excluded if they had a parent or sibling with a diagnosed psychotic disorder or if they met criteria for a major mental illness (i.e., psychotic disorder, current mood or anxiety disorder), neurological disorder, substance abuse/dependence, intellectual disability, head injury, and/or serious medical illness.

**SUNY**—Individuals with VCFS (age range: 14–22 years) had a confirmed microdeletion at the 22q11.2 locus, and were part of a longitudinal study of VCFS (Antshel et al., 2006; Antshel et al., 2010). Age-matched healthy siblings served as controls. Exclusion criteria (for both individuals with VCFS and controls) were: identifiable neurological conditions such as traumatic brain injury, seizure disorder, and/or birth weight below 2,500 gm (as reported by parent). Controls were also excluded if they met criteria for a major mental illness or intellectual disability. Initial screening for intellectual disability and/or mental illness was conducted over the phone by asking if the child was in regular classrooms at school, and whether parents had any concerns about emotional or behavioral problems. If the answer was negative to both questions, the controls were scheduled to participate in the study, and underwent the same IQ testing and structured psychiatric interviews (K-SADS-PL) that were conducted on participants with VCFS.

The studies at both sites were approved by the Institutional Review Boards at the University of California, Los Angeles (UCLA) and at SUNY Upstate Medical University (which will be referred to as “SUNY” throughout the manuscript), respectively, and all participants and/or their parents/legal guardians gave informed consent in accordance with the human subject research protocol.

## Measures

Given that this collaborative study between UCLA and SUNY Upstate was not initially planned as a multi-site study, there were minor differences between the two sites with regards to the assessment of autism spectrum disorders, cognitive assessment, and administration of the Animations Task, which we detail below. Importantly, it should be noted that the identical version of the Animations Task was administered at both sites, with slight variations in verbal instructions provided to participants. Due to these cross-site differences, we sought to directly investigate and statistically control for these differences in our analyses.

### Autism/ASD Measures

**UCLA:** Diagnoses of autism and/or ASD were determined using the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) administered to the child, and the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994), administered to the subject’s parent/primary caretaker. Participants were classified as having ASD, based on the ADI-R, if scores were above threshold for the Reciprocal Social Interaction domain, as well as either Communication Impairment or Repetitive Behaviors and Stereotyped Patterns (see Antshel et al., 2007; Antshel et al., 2006; Kaufmann et al., 2004). Scores from the ADOS and ADI-R were used to determine a consensus diagnosis of ASD. In a small number of cases (~10%), the ADI-R and ADOS results were not consistent. When these classifications disagreed, the ADI-R and ADOS assessors discussed the case and came up with a consensus diagnosis, based on all available information. Additionally, six subjects in the VCFS group were over the age of 18. Adult subjects participated in a partially overlapping clinical research protocol, which did not include an evaluation by the UCLA Autism Phenotyping Core. As a result, autism spectrum diagnoses for these individuals were made based on DSM-IV diagnostic criteria for autism spectrum disorder, as assessed by SCID interview with an additional developmental disorders module (First, Spitzer, & Williams 2009) with both the patient and their parent, and all available medical records. Individuals who did not meet the diagnostic threshold for Autistic Disorder but who met criteria for impairment in social interaction and restricted interests/stereotyped behavior without a history of significant language delay were given a diagnosis of Asperger’s syndrome; otherwise they received a diagnosis of PDD-NOS.

**SUNY:** Diagnoses of autism and/or ASD were made using the ADI-R, which was administered by a licensed clinician trained in the reliable administration of the instrument to the participant’s parent at SUNY Upstate Medical University, or over the phone. Following the criteria used in prior studies, a diagnosis of autism was assigned if a subject met the empirically established cut-off scores on all three domains of the ADI-R (Antshel et al., 2007; Kates et al., 2007; Kaufmann et al., 2004). A diagnosis of ASD was based on the same ADI-R scoring criteria used at UCLA, and in our prior studies (e.g., Antshel et al., 2007; Antshel et al., 2006; Kates et al., 2007).

**Cognitive Assessment**—At UCLA, Full Scale IQ (FSIQ) estimates were derived from the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) Vocabulary and Matrix Reasoning subtests, administered by a Master’s-level clinician (supervised by CEB). At SUNY, the Wechsler Intelligence Scale for Children— Third Edition (WISC-III;

Wechsler & Corporation, 1991) was administered by a licensed psychologist to participants under 17 years of age, and the Wechsler Adult Intelligence Scale (WAIS-III; Kaufman & Lichtenberger, 1999) to subjects 17 years of age or older.

**Reciprocal Social Behavior**—Both sites assessed real-world social behavior using the Social Responsiveness Scale (SRS; Constantino & Todd, 2005) a well-validated parent-report questionnaire designed to measure social impairments traditionally linked to ASD. A higher score indicates greater reciprocal social behavior impairment.

### **Animations Task**

**UCLA:** The Animations Task was administered using the guidelines from (Castelli et al., 2002). Using SuperLab stimulus presentation software, eight 34–45 second video clips were shown to participants, with each depicting two triangles moving around on a white screen. Subjects were told that they were going to watch a series of animations, after which they were to describe what happened in each clip. No time limit was imposed and administrators provided no feedback on responses. Two sample clips were shown at the beginning for practice.

Participants were then shown four ToM and four Random video clips in randomized order. The ToM condition entailed triangles performing a scripted interaction of: coaxing, seducing, mocking, or surprising. These scripts involved complex mental states that were not directly observable, as well as interactions at a mental level. The Random condition clips depicted the triangles moving about without any purpose or interaction: moving around as billiards, stars, tennis, or drifting.

**SUNY:** Participants were shown the same eight video clips used at UCLA as well as four additional videos representing “goal-directed” motion (data not included in the current analysis, as this condition was not administered in the UCLA cohort) in randomized order, and 3 practice clips (Abell et al., 2000; Castelli et al., 2002). Task instructions were based on the guidelines outlined by Abell et al. (2000). Participants were told that they would see triangles that move around the screen. Prior to each cartoon, the interviewer explained the roles that the triangles would play. For the ToM videos, these included “mother and child” (for the coaxing video), “grandmother and grandson” (surprising); “prisoner and guard” (seducing); and “teacher and boy” (mocking). For the Random videos, participants were told that the triangles will be “just triangles”. Participants were then asked the same question after every video clip: “What happened in the cartoon?” If the participant’s response was difficult to interpret based on scoring guidelines from Abell et al. (2000), the interviewer used probes to clarify the participant’s response. At both sites, responses were recorded with a digital recorder and later transcribed for coding.

**Coding:** Verbal responses from both sites were coded by two independent raters from UCLA, who were blind to diagnostic status, following the scoring system developed by Castelli et al. (2000). Responses received a score for Appropriateness and Intentionality. The Appropriateness score reflects how accurately the responses described the events in the animation; this score ranged from 0–3, with a 0 indicating poorer performance. Additionally, each response received a score for Intentionality, which reflects the degree to which purposeful movements, interactions, and mental states are described. This was primarily based on scoring verbs. Verbs that conveyed mental states and interactions received higher scores than verbs that did not. This score ranged from 0–5, with a 5 indicating the highest level of Intentionality (Table 1). Intentionality scores were independent of whether or not the verbs accurately described the events in the animation (e.g., a transcript could potentially receive a high Intentionality score, but a low Appropriateness score). Inter-rater reliability



was calculated using intra-class correlations on a random subset of participants (25%; UCLA:  $n=15$ ; SUNY:  $n=12$ ), revealing very high consistency between raters ( $ICC = 0.83-0.98$ ).

## Statistical Analyses

All statistical analyses were performed using SPSS software (SPSS Inc. Chicago, Illinois). We compared baseline demographic characteristics between the two groups using independent samples t-tests for continuous variables and chi-squared tests for categorical variables. For the primary analyses, group differences on Animations Task performance were evaluated using  $2$  (Group: VCFS vs. Controls)  $\times$   $2$  (Condition: ToM vs. Random)  $\times$   $2$  (Site: UCLA vs. SUNY) repeated measures ANOVAs (RMANOVAs) to assess the two dependent variables, i.e., Intentionality and Appropriateness. Because the age distributions differed across sites (Table 1), we conducted secondary analyses in which residualized scores were obtained for Intentionality and Appropriateness scores by removing the effects of age, and the RMANOVA models described above were re-run with the age-adjusted dependent variables. IQ was not included in the model because, as a variable that differed by 2 standard deviations or more between the groups, it is considered a group-defining variable (Campbell et al., 2011); thus, we chose to examine the effects of IQ within each group using correlational analyses. Moreover, IQ did not differ between sites for individuals with VCFS or for controls.

To identify whether the performance of individuals with VCFS and a diagnosis of ASD (VCFS+ASD) significantly differed from individuals with VCFS without an ASD diagnosis (VCFS-ASD), we also employed a  $2$  (Condition: ToM vs. Random)  $\times$   $3$  (Group: VCFS+ASD vs. VCFS-ASD vs. controls) RMANOVA. The VCFS+ASD group included participants with VCFS and a diagnosis of either Autistic Disorder, or ASD (based on ADI-R criteria described above), which included individuals with Asperger's Syndrome and Pervasive Developmental Disorder-NOS (PDD-NOS). Because of site differences in how the ASD diagnosis was made, separate RMANOVAs were conducted for each site. Accordingly, we did not use age-adjusted scores in these analyses, because age did not differ between groups *within* each site. The effects of age on task performance within each site were also evaluated using correlational analyses.

For all RMANOVAs, interactions were followed up with independent samples t-tests. The p-values from these analyses were corrected for multiple comparisons, using false discovery rate (FDR) correction (Benjamini & Hochberg, 1995) in the program R (<http://www.r-project.org/>).

To assess the relationship between performance on the Animations Task and real-life social behavior in individuals with VCFS, we conducted partial Pearson correlations between the Intentionality and Appropriateness scores (for both conditions) and SRS T-scores, controlling for site and age.

## Results

### Participant Characteristics

As shown in Table 2, individuals with VCFS and controls at both sites did not significantly differ in gender distribution or age ( $p>0.05$ ); although, as expected, IQ was significantly higher in controls ( $p<0.001$ ). Additionally, there were no differences in gender or IQ between individuals with VCFS across the two sites, but the UCLA sample was significantly younger, on average, than the SUNY sample ( $p<0.01$ ). At UCLA, of the 10 participants meeting criteria for an autism spectrum disorder diagnosis, 3 of these subjects (30%) met criteria for Autistic Disorder, and 7 (70%) met criteria for a broader autism spectrum

diagnosis (Asperger's or PDD-NOS). At SUNY, of the 6 participants meeting criteria for an autism spectrum disorder diagnosis, 2 (33%) met criteria for Autistic Disorder, and 4 (67%) met criteria for an autism spectrum diagnosis. ADI-R profile and ADOS scores are shown in Figure 1.

### Overall Group Differences (VCFS vs. Controls)

**Intentionality**—Results of main effects and interactions for all analyses are presented in Table 3, and descriptive statistics are shown in Figure 2. For the ToM condition, Intentionality scores were significantly higher in the SUNY sample relative to the UCLA sample (UCLA < SUNY: VCFS,  $p_{FDR} < 0.001$ ; controls,  $p_{FDR} < 0.001$ ), but scores did not differ across site for the Random Condition. Within each site, individuals with VCFS significantly differed from controls on the ToM condition (VCFS < Controls: UCLA,  $p_{FDR} < 0.05$ ; SUNY,  $p_{FDR} < 0.001$ ), but did not differ on the Random Condition, as depicted in Figure 2. After adjusting for the effects of age, VCFS and control groups at the SUNY site continued to differ on the ToM, but not Random, Condition, but group differences no longer reached statistical significance at the UCLA site.

**Appropriateness**—Across groups, SUNY's scores on Appropriateness were significantly higher than UCLA's for the ToM (VCFS,  $p_{FDR} < 0.001$ ; controls,  $p_{FDR} < 0.001$ ), but not the Random, Condition (see Table 3 and Figure 2b). Moreover, within each site, individuals with VCFS had significantly lower Appropriateness scores than controls on the ToM condition (VCFS < Controls: UCLA,  $p_{FDR} < 0.05$ ; SUNY,  $p_{FDR} < 0.001$ ), but did not differ on the Random Condition, again resulting in a significant Condition  $\times$  Group interaction (Table 3). After adjusting for the effects of age, differences in Appropriateness between the VCFS and control group remained significant for the ToM Condition at both sites.

### Effect of ASD Diagnosis on Animations Task Performance (VCFS+ASD vs. VCFS-ASD vs. Controls)

**Intentionality**—The main effects, interactions, and descriptives are shown in Table 4 and Figure 3. Notably, we found a significant Condition  $\times$  Group interaction for Intentionality scores for both sites (Table 4). Post-hoc analyses revealed that at the SUNY site, controls had significantly higher Intentionality ratings in the ToM condition than individuals with VCFS both with and without ASD (VCFS-ASD < Controls,  $p_{FDR} < 0.001$ ; VCFS +ASD < Controls,  $p_{FDR} < 0.001$ ), but the groups did not significantly differ from each other in the Random condition. No significant differences were found in ToM Intentionality scores between VCFS+ASD vs. VCFS-ASD in the SUNY sample. Similarly, for the UCLA site, post-hoc tests revealed that controls performed better than the VCFS-ASD group for the ToM condition ( $p_{FDR} < 0.05$ ), but there were no significant between-group differences for the Random condition. There were no significant differences in ToM Intentionality scores between the VCFS+ASD group vs. the other two groups (VCFS-ASD, controls).

**Appropriateness**—As shown in Table 4 and Figure 3, the SUNY sample demonstrated a significant Condition  $\times$  Group interaction effect, such that controls had higher Appropriateness scores than both individuals with VCFS, with and without ASD, on the ToM, but not Random, condition (VCFS+ASD < Controls:  $p_{FDR} < 0.001$ ; VCFS -ASD < Controls:  $p_{FDR} = 0.001$ ). In the UCLA sample, controls significantly outperformed only VCFS-ASD and not VCFS+ASD in the ToM condition (VCFS+ASD < Controls:  $p_{FDR} = 0.26$ ; VCFS-ASD < Controls:  $p_{FDR} = 0.015$ ). No significant differences were found in Appropriateness scores for ToM condition between VCFS+ASD vs. VCFS-ASD at both sites. Finally, there were no group differences in the UCLA cohort for the Random condition.

### **Relationship between Task Performance and Reciprocal Social Behavior—**

After adjusting for age and site effects, participants with VCFS showed significant negative correlations between SRS scores and Intentionality ( $r=-0.283$ ,  $p<0.05$ ) and Appropriateness ( $r=-0.314$ ,  $p<0.05$ ) ratings for the ToM, but not Random, condition. In other words, lower mentalizing ability and less appropriate responses on ToM scenes in the Animations Task were associated with higher (i.e., more impaired) SRS scores, but task performance on Random scenes was not associated with SRS score.

**Relationship between Task Performance, Age and IQ—**In the UCLA sample, healthy controls showed significant correlations between age and ToM Appropriateness ( $r=0.533$ ;  $p=0.002$ ); Random Appropriateness ( $r=0.371$ ;  $p=0.043$ ); and Random Intentionality ( $r=-0.436$ ;  $p=0.016$ ) scores, which were not observed in the VCFS group (see Figure 4). In the SUNY sample, which had a narrower age range, no significant correlations between age and task performance were observed in either the VCFS or control group.

Within the SUNY VCFS group, there were significant correlations between FSIQ and Intentionality scores on the ToM condition ( $r=0.511$ ,  $p<0.01$ ) and Appropriateness scores on both conditions (ToM:  $r=0.463$ ,  $p<0.01$ ; Random:  $r=0.453$ ,  $p<0.01$ ), but there were no significant correlations between FSIQ and any of the Animations Task measures for the UCLA VCFS sample (all  $p$ -values  $>0.25$ ).

## **Discussion**

To our knowledge, this is the first study to use ambiguous, abstract visual stimuli to assess ToM abilities in individuals with VCFS. We administered the Animations Task, which consists of a series of silent videos depicting the interactions of triangles (including vignettes of purposeful, mentalizing interactions and ‘random’ scenes), to individuals with VCFS and age-matched controls at two different academic medical centers. Consistent with our hypotheses, we found that, in comparison to controls, individuals with VCFS showed significant impairments in the ToM (mentalizing) condition, in both the ability to explain purposeful behavior (Intentionality) and to accurately describe the events going on in the scene (Appropriateness). We also predicted that individuals with VCFS who have a diagnosis of ASD would perform worse than individuals with VCFS who do not have ASD; however, we found that regardless of ASD diagnosis, participants with VCFS were significantly impaired in ToM performance. Additionally, when controlling for age and site effects, we found that lower scores on Intentionality and Appropriateness in the ToM condition were related to greater real-world social impairment (as measured by the SRS) in individuals with VCFS.

These findings extend the burgeoning literature on deficits in ToM and other aspects of social cognition in VCFS. Consistent with results reported by both Chow et al. (2006) and Campbell et al. (2011) using other paradigms, we found that individuals with VCFS displayed significant impairments in ToM compared with healthy controls. Unlike many ToM tasks that require subjects to make logical inferences about the mental states of others (i.e., ‘explicit’ mentalizing), performance on the Animations task used in the current study depends on spontaneous mental state attribution, or ‘implicit’ mentalizing (Frith, 2004). Our findings add to the existing literature as they indicate that: 1) implicit mentalizing is also deficient in VCFS; and 2) that such deficits are significantly associated with real-world social behavior.

Deficits in ToM skills may be attributable, in part, to neuroanatomic anomalies that have been reported in both VCFS and idiopathic ASD that are critically involved in ToM functions. The role of the temporo-parietal junction, basal temporal region (including the



fusiform gyrus and temporal poles), and the medial prefrontal cortex has been described in ToM (for review, see Adolphs, 2009), including a study of brain activation while viewing Animations videos (Castelli et al., 2000). Castelli et al. (2002) also reported decreased activations for ToM videos in individuals with idiopathic ASD (as compared to controls) in these three brain regions, along with decreased functional connectivity between the extrastriate cortex and the temporo-parietal junction. Notably, several of these brain regions have been previously shown to have structural and functional deficits in VCFS (Bearden et al., 2009; Eliez, Schmitt, White, & Reiss, 2000; Kates et al., 2001). Thus, it is possible that deficits in these regions, and/or disruptions in connectivity, may underlie the ToM deficits that we observed in individuals with VCFS. Accordingly, future studies should investigate the underlying neural correlates of ToM deficits in individuals with VCFS.

Contrary to our hypotheses, individuals with VCFS and a diagnosis of ASD did not show differential ToM impairment. This pattern of findings suggests that ToM deficits are part of the VCFS phenotype, and that persons with VCFS will show deficits in ToM regardless of whether or not they meet categorical diagnostic criteria for ASD. This interpretation is consistent with our findings of a linear relationship between Intentionality and Appropriateness ratings for the ToM condition and SRS scores within the VCFS group; moreover, it supports the notion of a dimensional relationship between deficits in ToM, as measured by the Animations task, and social behavior. Thus, ASD traits may be more broadly characteristic of individuals with VCFS (Kates et al., 2007), and better assessed by a dimensional scale such as the SRS. We cannot rule out that the current study was underpowered to detect a categorical effect of ASD diagnosis; thus, future studies with larger sample sizes are important for further characterizing ToM deficits in individuals with VCFS who have an ASD diagnosis vs. those without ASD.

Although previous studies of ToM performance using the Animations Task in healthy individuals and youths with VCFS have found that older individuals exhibit greater mentalizing skills than younger participants (Campbell et al., 2006; Castelli et al., 2002), the effect of age on task performance in our study is less consistent. Campbell and colleagues (2011), who examined a sample of youth with VCFS ranging in age from 6 to 16.75 years, observed ToM differences between subjects with VCFS and controls in younger children only, leading them to conclude that ToM was delayed, rather than deficient, in youth with VCFS. However, in contrast to our Animations task, the false-belief tasks used in that study were heavily dependent on both working memory and language competence, which likely played a role in the observed age relationships. Intriguingly, we observed that age was significantly correlated with task performance in typically developing controls at the UCLA site, but not in individuals with VCFS at either site. While prospective longitudinal studies are needed to validate this cross-sectional finding, these results suggest the possibility of an abnormal trajectory of ToM development in children and adolescents with VCFS.

We further observed that both Intentionality and Appropriateness scores were significantly higher across both conditions for participants at the SUNY site relative to the UCLA site. Interestingly, the two sites had separate task administration instructions, which may have, in part, accounted for these performance differences. At UCLA, examiners asked the participants to describe what had just happened in the animation and did not provide additional probes or feedback, according to the procedures described by Castelli et al. (2002), whereas at SUNY, examiners both disclosed in advance the roles of the triangles, and probed for clarifications once the participants responded, according to Abell et al. (2000). Due to the age differences between sites, we performed additional, exploratory analyses to confirm that site differences were in fact due to differences in verbal instructions associated with task administration, independent of age. Using a subgroup of age-matched subjects from the two sites (age range 14 to 22), we conducted a 2×2×2 repeated measures

ANOVA for Intentionality and Appropriateness scores, with site, condition, and group as independent variables. The results of these analyses were comparable to those of our original analyses in which we statistically covaried for age, indicating that site differences persist even after accounting for age differences. Specifically, they indicated that the Upstate sample (both VCFS and controls) had significantly higher Intentionality and Appropriateness scores on the ToM condition relative to the UCLA sample; but there was no main effect of site for the Random condition. As seen in Figure 2, performance of individuals with VCFS at SUNY was similar to performance of healthy control participants at UCLA. Prior role designation may have, in part, increased both Intentionality and Appropriateness scores by setting up expectations for the context and content of the interaction. Abell et al. (2000) suggested that assigning roles in the ToM condition may prime participants to use more mental state attributions. Moreover, clinical populations with more variable ToM skills may be more prone to effects of specific task instructions. Although speculative, the higher scores achieved by individuals with VCFS with more structured task instructions suggest some possibilities for cognitive remediation; in particular, ToM abilities may be improved when additional context is provided.

In addition, we did not observe a consistent relationship between task performance and IQ across sites. It is possible that prior role assignment may account for the correlation that we observed between FSIQ and ToM scores in the SUNY but not UCLA site. In particular, participants with higher IQ might have had better understanding of the cues, and/or may have been able to utilize the cues more effectively than participants with lower IQ. Further studies are needed to examine more directly the effects of instructions on ToM performance (particularly in populations with deficits in reciprocal social interaction) and to assess whether providing contextual cues may help remediate performance in individuals with ToM deficits.

We hypothesized that there would be a relationship between performance on the Animations Task and SRS scores. Interestingly, this has never been directly tested before, either in individuals with VCFS or, to our knowledge, in persons with idiopathic ASD. In the VCFS sample, the ability to spontaneously attribute mental states to ambiguous stimuli, as measured by performance on the ToM condition of the Animations task, was significantly associated with real-world social behavior, as measured by the SRS. More specifically, worse performance on the ToM tasks was significantly correlated with greater social impairment in the participants with VCFS. This suggests that implicit aspects of social cognition are relevant to the observed social deficits that are so functionally debilitating for individuals with VCFS (Feinstein & Singh, 2007). It is important to address the possibility that the significant relationship reported between the ToM condition of the Animations task and SRS may be due to overlapping items found on both measures. While the underlying constructs captured by the ToM condition of the Animations task and some SRS items may overlap, the measures used to obtain the data are quite distinct (e.g., blind rater judgments of Intentionality and Appropriateness, based on study participants' audio-taped responses to experimental probes, as compared to parent ratings of real-world social behavior). From a theoretical perspective, we would expect impaired parent reported social functioning to be consistent with poorer performance on the attribution of Intentionality and its level of Appropriateness. The key issue is whether or not the two measures are correlated simply because they are measuring the same construct. We can evaluate this using the magnitude of the reported correlations. The correlations that we observed between the SRS subscales and ToM condition Animations scores were in the range of  $r = .2 - .3$ . These correlations fell below 0.8, which has been established in the field as the approximate cutoff point at which measures are completely redundant. Thus, we can infer that the two measures are capturing different constructs and that their relationship is not simply due to the fact that they share ToM-related items. In fact, the correlations that we observed between the Animations Task

and SRS can be regarded as support for convergent validity of the Animations Task as a measure of social cognition. These findings highlight the importance of assessing mechanisms underlying social problems in children with VCFS, in order to develop targeted interventions.

Deficits in ToM are also a key component of the schizophrenia phenotype (Pinkham, Penn, Perkins, & Lieberman, 2003) and a similar pattern of deficit has previously been reported in patients with idiopathic schizophrenia, using this task (Horan et al., 2009). In light of the fact that individuals with VCFS exhibit a phenotype involving symptoms of both ASD (Kates et al., 2007) and psychosis (Vorstman et al., 2006; Baker & Skuse, 2005), future research is warranted to establish whether or not autistic features ascribed in VCFS are better characterized as symptoms relating to the premorbid or prodromal phase of schizophrenia. Given the age range of participants in our study, we were only able to examine this question in a subset of our sample by testing correlations between task performance and the Brief Psychiatric Rating Scale (BPRS; Overall & Donald, 1962) and Structured Interview for Prodromal Syndromes (SIPS/SOPS; McGlashan et al. 2001), as well as the SRS. Results indicated that Intentionality and Appropriateness ratings for the ToM condition of the Animations task were significantly correlated with measures of reciprocal social behavior (i.e., SRS), but were not significantly associated with positive and negative prodromal/psychotic symptoms, at least in this subset of our overall sample. It may be that early ToM deficits may predict subsequent development of psychotic symptoms in individuals with VCFS, a possibility we plan to explore in subsequent longitudinal studies.

In conclusion, our results indicate impairments in the spontaneous attribution of mental states to abstract visual stimuli in individuals with VCFS, which may underlie real-life problems with social interactions. Because the VCFS endophenotype shares deficits and problems observed in more genetically heterogeneous clinical groups, these findings suggest that the spontaneous attribution of mental states may also impact real-world social functioning in a broader range of psychiatric disorders, e.g., ASD and schizophrenia. Moreover, as our study suggested that mentalizing deficits associated with VCFS may be remediated with contextual verbal information, there may be implications for its usefulness in persons with ASD and schizophrenia. Accordingly, our results could inform the extant literature on social – cognitive remediation in ASD (Abdi & Sharma, 2004; Frankel et al., 2010) and schizophrenia (Horan et al., 2011; Wölwer & Frommann, 2011) as well as form a basis for investigating the effectiveness of social cognitive interventions in individuals with autism.

## Acknowledgments

The funding sources for the study included grants from the National Institute of Mental Health (R01 MH085953 to CEB; R01 MH64824 to WRK), NIH/NICHHD (P50-HD-055784 (CART Pilot Project Grant to CEB)) and the Dennis Weatherstone Pre-Doctoral Fellowship from Autism Speaks (#7076 to PDR). Special thanks to Dr. Robert J. Shprintzen, a co-investigator on the R01 MH64824 grant, Anne Marie Higgins and Joanna Botti for coordination of the longitudinal study at SUNY Upstate Medical University, Kelly Wallace and Joanna Botti for assistance with transcription of verbal responses of participants at the SUNY site, Lauren Sanderson for assistance with the overall project, and Christopher McCarthy for administering the ToM paradigm to some of the participants at the SUNY site.

## Literature Cited

- Abdi Z, Sharma T. Social cognition and its neural correlates in schizophrenia and autism. *Cns Spectrums*. 2004; 9:335–343. [PubMed: 15115945]
- Abell F, Happé F, Frith U. Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*. 2000; 15(1):1–16.

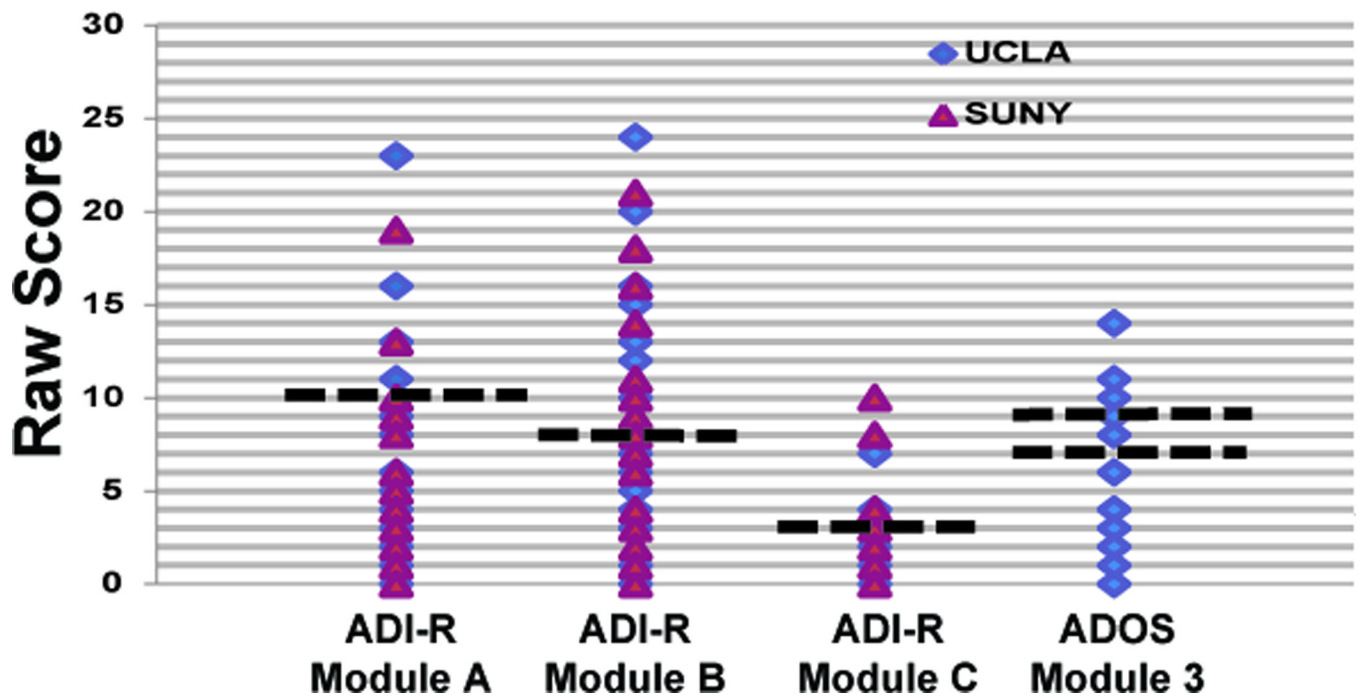
- Adolphs R. The social brain: neural basis of social knowledge. *Annual review of psychology*. 2009; 60:693.
- Antshel KM, Aneja A, Strunge L, Peebles J, Fremont WP, Stallone K, Kates WR. Autistic spectrum disorders in velo-cardio facial syndrome (22q11. 2 deletion). *Journal of Autism and Developmental Disorders*. 2007; 37(9):1776–1786. [PubMed: 17180713]
- Antshel KM, Fremont W, Roizen NJ, Shprintzen R, Higgins A, Dhamoon A, Kates WR. ADHD, major depressive disorder, and simple phobias are prevalent psychiatric conditions in youth with velocardiofacial syndrome. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2006; 45(5):596–603. [PubMed: 16670654]
- Baker KD, Skuse DH. Adolescents and young adults with 22q11 deletion syndrome: psychopathology in an at-risk group. *The British Journal of Psychiatry*. 2005; 186(2):115–120. [PubMed: 15684233]
- Baron-Cohen S, Leslie AM, Frith U. Does the autistic child have a "theory of mind"? *Cognition*. 1985; 21(1):37–46. [PubMed: 2934210]
- Bearden CE, Van Erp TGM, Dutton RA, Lee AD, Simon TJ, Cannon TD, Thompson PM. Alterations in midline cortical thickness and gyrification patterns mapped in children with 22q11. 2 deletions. *Cerebral cortex*. 2009; 19(1):115. [PubMed: 18483006]
- Bearden CE, Woodin MF, Wang PP, Moss E, McDonald-McGinn D, Zackai E, Cannon TD. The neurocognitive phenotype of the 22q11. 2 deletion syndrome: selective deficit in visual-spatial memory. *Journal of Clinical and Experimental Neuropsychology*. 2001; 23(4):447–464. [PubMed: 11780945]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1995:289–300.
- Campbell LE, Stevens A, Daly E, Toal F, Azuma R, Karmiloff-Smith A, Murphy KC. A comparative study of cognition and brain anatomy between two neurodevelopmental disorders: 22q11. 2 deletion syndrome and Williams syndrome. *Neuropsychologia*. 2009; 47(4):1034–1044. [PubMed: 19061904]
- Campbell LE, Stevens AF, McCabe K, Cruickshank L, Morris RG, Murphy DGM, Murphy KC. Is theory of mind related to social dysfunction and emotional problems in 22q11. 2 deletion syndrome (velo-cardio-facial syndrome)? *Journal of Neurodevelopmental Disorders*. 2011:1–10. [PubMed: 21484593]
- Campbell R, Lawrence K, Mandy W, Mitra C, Jeyakuma L, Skuse D. Meanings in motion and faces: Developmental associations between the processing of intention from geometrical animations and gaze detection accuracy. *Development and Psychopathology*. 2006; 18(01):99–118. [PubMed: 16478554]
- Castelli F, Frith C, Happé F, Frith U. Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*. 2002; 125(8):1839. [PubMed: 12135974]
- Castelli F, Happé F, Frith U, Frith C. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*. 2000; 12(3):314–325. [PubMed: 10944414]
- Chow EWC, Watson M, Young DA, Bassett AS. Neurocognitive profile in 22q11 deletion syndrome and schizophrenia. *Schizophrenia research*. 2006; 87(1–3):270–278. [PubMed: 16753283]
- Constantino JN, Todd RD. Intergenerational transmission of subthreshold autistic traits in the general population. *Biological Psychiatry*. 2005; 57(6):655–660. [PubMed: 15780853]
- Eliez S, Schmitt JE, White CD, Reiss AL. Children and adolescents with velocardiofacial syndrome: a volumetric MRI study. *American Journal of Psychiatry*. 2000; 157(3):409–415. [PubMed: 10698817]
- Feinstein C, Singh S. Social phenotypes in neurogenetic syndromes. *Child and Adolescent Psychiatric Clinics of North America*. 2007; 16(3):631–647. [PubMed: 17562583]
- Fine SE, Weissman A, Gerdes M, Pinto-Martin J, Zackai EH, McDonald-McGinn DM, Emanuel BS. Autism spectrum disorders and symptoms in children with molecularly confirmed 22q11. 2 deletion syndrome. *Journal of Autism and Developmental Disorders*. 2005; 35(4):461–470. [PubMed: 16134031]

- First, G.; Spitzer; Williams. Structured clinical Interview for DSM-IV axis I disorders clinician version SCID-I. Developmental disorder addition. American Psychiatric Association; 2009.
- First, MB.; Spitzer, RL.; Williams, JB. Structured clinical interview for DSM-IV axis I disorders SCID-I: clinician version, administration booklet. American Psychiatric Pub.; 1997.
- Frankel F, Myatt R, Sugar C, Whitham C, Gorospe CM, Laugeson E. A randomized controlled study of parent-assisted children's friendship training with children having autism spectrum disorders. *Journal of Autism and Developmental Disorders*. 2010; 40(7):827–842. [PubMed: 20058059]
- Frith CD. Schizophrenia and theory of mind. *Psychological Medicine*. 2004; 34(03):385–389. [PubMed: 15259823]
- Green T, Gothelf D, Glaser B, Debbane M, Frisch A, Kotler M, Eliez S. Psychiatric disorders and intellectual functioning throughout development in velocardiofacial (22q11.2 deletion) syndrome. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2009; 48(11):1060–1068. [PubMed: 19797984]
- Horan W, Nuechterlein K, Wynn J, Lee J, Castelli F, Green M. Disturbances in the spontaneous attribution of social meaning in schizophrenia. *Psychol Med*. 2009; 39:635–643. [PubMed: 18606048]
- Horan WP, Kern RS, Tripp C, Hellemann G, Wynn JK, Bell M, Green MF. Efficacy and specificity of Social Cognitive Skills Training for outpatients with psychotic disorders. *Journal of psychiatric research*. 2011
- Kates WR, Antshel KM, Fremont WP, Shprintzen RJ, Strunge LA, Burnette CP, Higgins AM. Comparing phenotypes in patients with idiopathic autism to patients with velocardiofacial syndrome (22q11 DS) with and without autism. *American Journal of Medical Genetics Part A*. 2007; 143(22):2642–2650. [PubMed: 17937445]
- Kates WR, Burnette CP, Jabs EW, Rutberg J, Murphy AM, Grados M, Pearlson GD. Regional cortical white matter reductions in velocardiofacial syndrome: a volumetric MRI analysis. *Biological Psychiatry*. 2001; 49(8):677–684. [PubMed: 11313035]
- Kaufman, AS.; Lichtenberger, EO. Essentials of WAIS-III assessment. John Wiley & Sons Inc.; 1999.
- Kaufmann WE, Cortell R, Kau ASM, Bukelis I, Tierney E, Gray RM, Stanard P. Autism spectrum disorder in fragile X syndrome: communication, social interaction, and specific behaviors. *American Journal of Medical Genetics Part A*. 2004; 129(3):225–234. [PubMed: 15326621]
- Kiley-Brabeck K, Sobin C. Social skills and executive function deficits in children with the 22q11 deletion syndrome. *Applied neuropsychology*. 2006; 13(4):258–268. [PubMed: 17362146]
- Kobrynski LJ, Sullivan KE. Velocardiofacial syndrome, DiGeorge syndrome: the chromosome 22q11.2 deletion syndromes. *The Lancet*. 2007; 370(9596):1443–1452.
- Landa RJ, Holman KC, Garrett-Mayer E. Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Archives of General Psychiatry*. 2007; 64(7):853. [PubMed: 17606819]
- Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, Rutter M. 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*. 2000; 30(3):205–223. [PubMed: 11055457]
- Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*. 1994; 24(5):659–685. [PubMed: 7814313]
- McDonald-McGinn DM, Tonnesen MK, Laufer-Cahana A, Finucane B, Driscoll DA, Emanuel BS, Zackai EH. Phenotype of the 22q11.2 deletion in individuals identified through an affected relative: Cast a wide FISH ing net! *Genetics in Medicine*. 2001; 3(1):23. [PubMed: 11339373]
- Murphy KC. Schizophrenia and velo-cardio-facial syndrome. *The Lancet*. 2002; 359(9304):426–430.
- Niklasson L, Rasmussen P, Óskarsdóttir S, Gillberg C. Chromosome 22q11 deletion syndrome (CATCH 22): neuropsychiatric and neuropsychological aspects. *Developmental Medicine & Child Neurology*. 2002; 44(1):44–50. [PubMed: 11811651]
- Overall JE, Donald RG. The brief psychiatric rating scale. *Psychological reports*. 1962; 10(3):799–812.

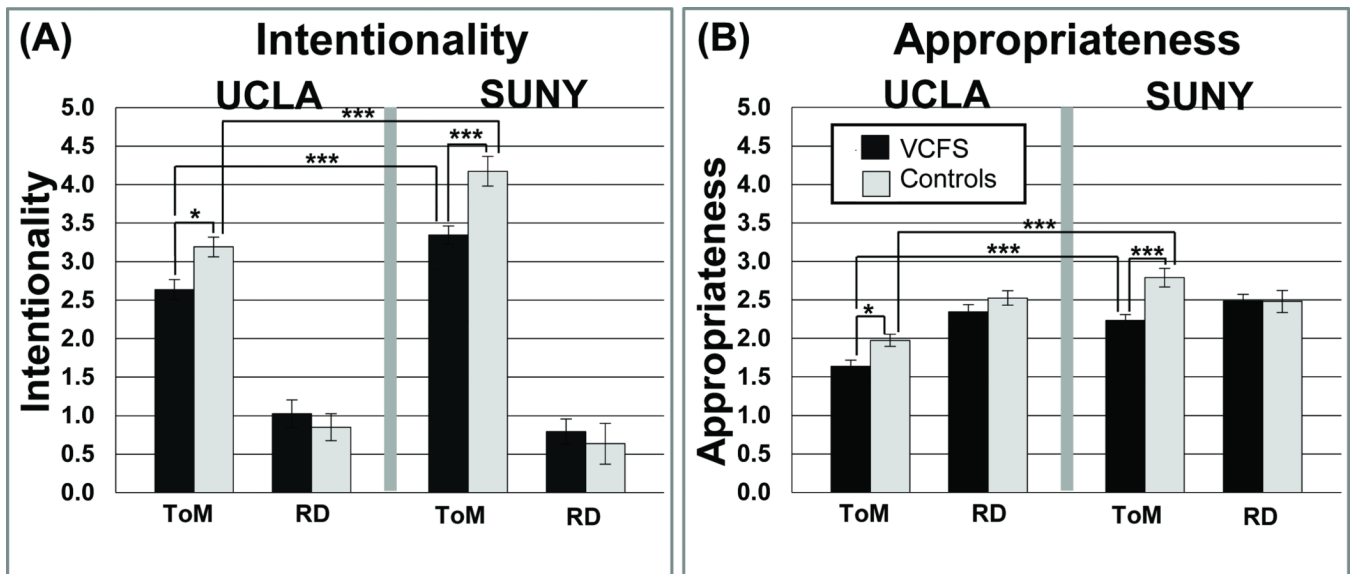


- Pinkham AE, Penn DL, Perkins DO, Lieberman J. Implications for the neural basis of social cognition for the study of schizophrenia. *American Journal of Psychiatry*. 2003; 160(5):815–824. [PubMed: 12727681]
- Premack D, Woodruff G. Does the chimpanzee have a theory of mind. *Behavioral and brain sciences*. 1978; 1(4):515–526.
- Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2000; 39(1):28–38. [PubMed: 10638065]
- Shprintzen RJ. Velo-cardio-facial syndrome. *Progress in Pediatric cardiology*. 2005; 20(2):187–193.
- Swillen A, Feys H, Adriaens T, Nelissen L, Mertens L, Gewillig M, Fryns JP. Early motor development in young children with 22q. 11 deletion syndrome and a conotruncal heart defect. *Developmental Medicine & Child Neurology*. 2005; 47(12):797–802. [PubMed: 16288668]
- Vorstman JAS, Morcus MEJ, Duijff SN, Klaassen PWJ, Boer H, Josien A, van Engeland H. The 22q11. 2 Deletion in Children: High Rate of Autistic Disorders and Early Onset of Psychotic Symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2006; 45(9): 1104–1113. [PubMed: 16926618]
- Wechsler, D. Wechsler abbreviated scale of intelligence. Psychological Corp.; 1999.
- Wechsler, D.; Corporation, P. WISC-III: Wechsler intelligence scale for children. The Psychological Corporation San Antonio; 1991.
- Wölwer W, Frommann N. Social-cognitive remediation in schizophrenia: generalization of effects of the training of affect recognition (tar). *Schizophrenia Bulletin*. 2011; 37(suppl 2):S63–S70. [PubMed: 21860049]

# VCFS ADI-R and ADOS Scores

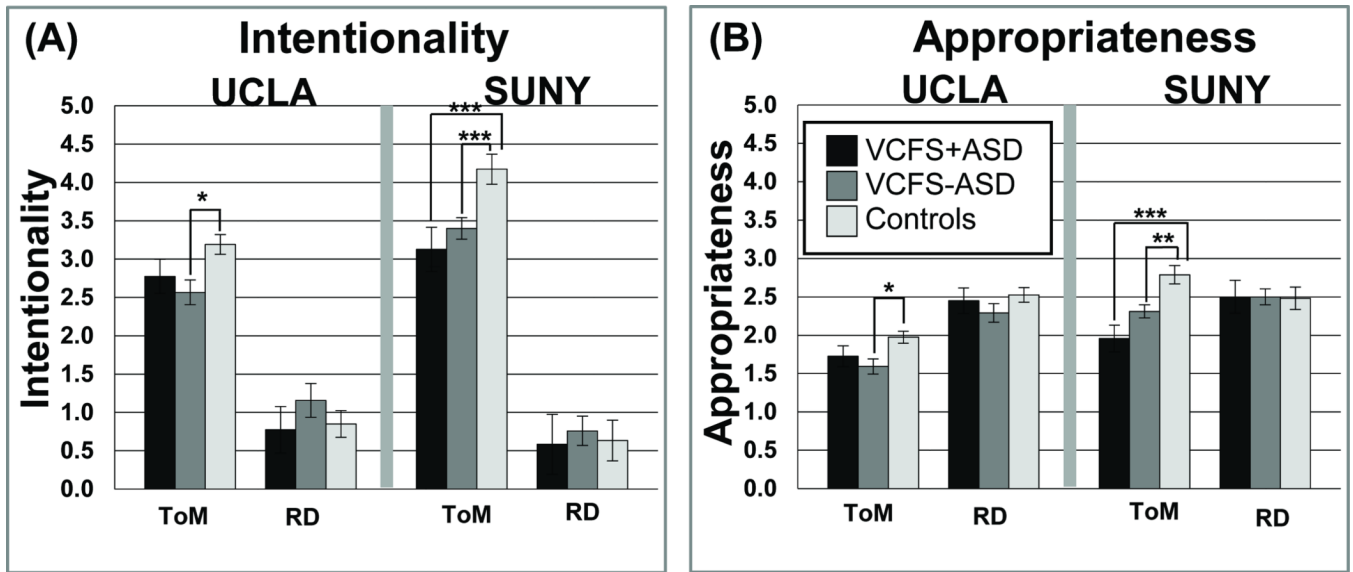


**Figure 1.** ADI-R and ADOS score profiles for VCFS patients at UCLA and SUNY. Blue represents UCLA and red represents SUNY. ADI-R Module A= Qualitative Abnormalities in Reciprocal Social Interaction; ADI-R Module B = Qualitative Abnormalities in Communication; ADI-R Module C = Restricted, Repetitive, and Stereotyped Patterns of Behavior. Dotted lines represent thresholds for autism criteria in each of the domains. For the ADI-R: Social Interaction = 10; Communication = 8; Stereotypes = 3. For the ADOS: ASD = 7; Autism = 9.



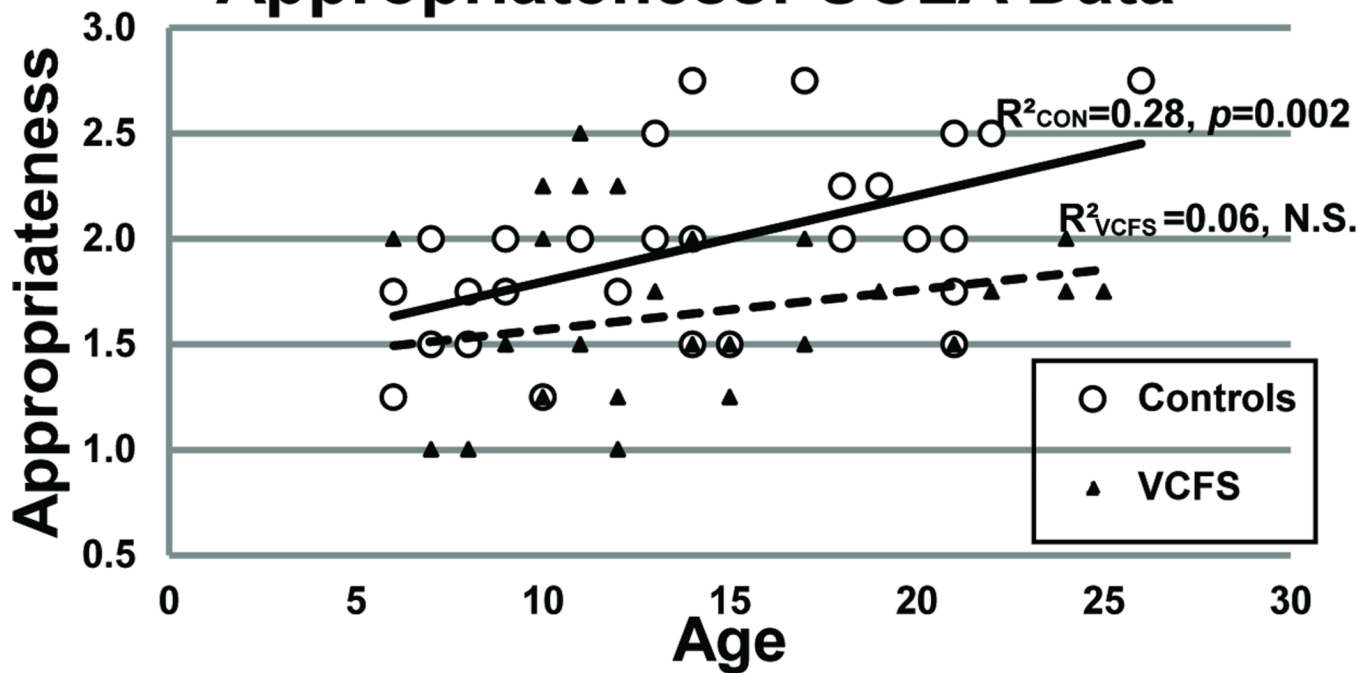
**Figure 2.**

(A): Intentionality scores on ToM and Random conditions, for UCLA and SUNY. Black represents individuals with VCFS and light grey represents controls. Across groups, performance in the ToM condition was significantly better at SUNY in comparison to UCLA. At both sites, in comparison to individuals with VCFS, controls had significantly higher performance in the ToM condition, but did not differ on the Random Condition. Significance levels are indicated as follows: \*\*\*  $p_{FDR} < 0.001$ ; \*\*  $p_{FDR} < 0.01$ ; \*  $p_{FDR} < 0.05$ . Raw scores, rather than age adjusted scores, are presented for visualization purposes. (B) Appropriateness scores on ToM and Random conditions, for UCLA and SUNY. A similar pattern of results was observed for both the Appropriateness and Intentionality ratings (Figure 2A).



**Figure 3.** (A) Intentionality and (B) Appropriateness for UCLA (left panel) and SUNY (right panel). The shading represents the groups as follows: VCFS+ASD=black; VCFS-ASD=dark grey; controls=light grey. Significance levels are marked as follows: \*\*\*  $p_{FDR}<0.001$ ; \*\*  $p_{FDR}<0.01$ ; \*  $p_{FDR}<0.05$ .

# Relationship between Age and Appropriateness: UCLA Data



**Figure 4.** Relationship between Age and ToM Appropriateness for UCLA. Note that there is a significant correlation between age and ToM appropriateness scores for the controls but *NOT* for the individuals with VCFS.



**Table 1**

## Explanation of Intentionality Scores

Score	Explanation of Score	Example
0	Verbs that convey no purposeful action or mental state attribution	“bouncing around”
1	Verbs that convey purposeful action but no interaction between agents and no mental states	“walking”
2	Verbs that convey purposeful action with interaction between agents, but no mental states	“fighting”
3	Verbs that convey goal-directed intentions without reciprocal interaction between agents	“trying”
4	Verbs that attribute mental states involving reciprocal interactions between agents at a mental state level	“mad at each other”
5	Verbs that convey that one agent is intentionally affecting or manipulating the other agent’s mental state	“tricking the other triangle”

**Table 2**

Demographics

Variable	UCLA		SUNY	UCLA vs. SUNY	
	VCFS	Controls		P-value	P-value
Female Gender (N, %)	11 (38%)	18 (60%)	6 (46%)	0.44	0.098
Age, in years (SD)	13.7 (5.5)	14.4 (5.6)	17.7 (1.7)	0.33	<0.01
FSIQ (SD)	80.5 (13.7)	113.3 (21.3)	109.8 (16.9)	<0.001	0.083
SRS Total T-Score (SD)	66.2 (14.9)	52.0 (14.7)	42.8 (6.1)	<0.001	0.117
					<0.05

**Table 3**

Repeated-Measures ANOVA results for Intentionality and Appropriateness ratings (Raw Scores and Age-adjusted residuals). Factors are Condition (ToM vs. Random), Group (VCFS vs. Controls), and Site (UCLA vs. SUNY).

Dependent Variable	Raw Scores			Age-Adjusted Residuals		
	F (df)	p	$\eta^2$	F (df)	p	$\eta^2$
<b>Intentionality</b>						
Condition	450.71 (1,102)	<0.001	0.815	1.04 (1,102)	0.310	0.010
Group	4.14 (1,102)	<0.05	0.039	2.98 (1,102)	0.087	0.028
Site	5.84 (1,102)	<0.05	0.054	0.81 (1,102)	0.371	0.008
Site × Condition	20.39 (1,102)	<0.001	0.167	0.98 (1,102)	0.324	0.010
Group × Condition	13.17 (1,102)	<0.001	0.114	5.67 (1,102)	<0.05	0.053
Group × Site	0.32 (1,102)	0.574	0.003	5.81 (1,102)	<0.05	0.054
Group × Site × Condition	0.30 (1,102)	0.588	0.003	4.49 (1,102)	<0.05	0.042
<b>Appropriateness</b>						
Condition	21.28 (1,102)	<0.001	0.173	0.79 (1,102)	0.377	0.008
Group	12.70 (1,102)	<0.01	0.111	13.42(1,102)	<0.001	0.116
Site	25.39 (1,102)	<0.001	0.199	0.71 (1,102)	0.401	0.007
Site × Condition	25.57 (1,102)	<0.001	0.200	0.83 (1,102)	0.364	0.008
Group × Condition	7.56 (1,102)	<0.01	0.069	7.68 (1,102)	<0.01	0.070
Group × Site	0.01 (1,102)	0.917	0.000	0.001 (1,102)	0.979	0.000
Group × Site × Condition	2.38 (1,102)	0.126	0.023	1.58 (1,102)	0.212	0.015

**Table 4**

Repeated-Measures ANOVA results for Intentionality and Appropriateness ratings, analyzed separately by site. Factors are Condition (ToM vs. Random) and Group (VCFS+ASD vs. VCFS-ASD vs. Controls).

Site		F	P	$\eta^2$		
Intentionality	UCLA	Condition	126.47 (1,56)	<0.001	0.693	
		Group	0.54 (2,56)	0.585	0.019	
		Group × Condition	3.61 (2,56)	<0.05	0.114	
	SUNY	Condition	246.81 (1,41)	<0.001	0.858	
		Group	3.78 (2,41)	<0.05	0.156	
		Group × Condition	3.57 (2,41)	<0.05	0.148	
	Appropriateness	UCLA	Condition	50.00 (1,56)	<0.001	0.472
			Group	4.78 (2,56)	<0.05	0.146
			Group × Condition	0.44 (2,56)	0.646	0.015
SUNY		Condition	1.71 (1,41)	0.199	0.040	
		Group	2.71 (2,41)	0.078	0.117	
		Group × Condition	4.75 (2,41)	<0.05	0.188	