HHS Public Access

Author manuscript

J Autism Dev Disord. Author manuscript; available in PMC 2022 April 26.

Published in final edited form as:

J Autism Dev Disord. 2022 February ; 52(2): 569–583. doi:10.1007/s10803-021-04957-2.

Social Inferences in Agenesis of the Corpus Callosum and Autism: Semantic Analysis and Topic Modeling

Tiffany Renteria-Vazquez[1], **Warren S. Brown**[1,2], **Christine Kang**[1], **Mark Graves**[1], **Fulvia Castelli**[3], **Lynn K. Paul**[1,2,4]

[1]Fuller Graduate School of Psychology, Travis Research Institute, Pasadena, CA, USA

[2]International Research Consortium for the Corpus Callosum and Cerebral Connectivity (IRC5), Pasadena, CA, USA

[3]Brain and Behavioral Science Department, University of Pavia, Sezione di Psicologia Piazza Botta, Pavia, Italy

[4]Division of Humanities and Social Sciences, California Institute of Technology, Baxter MC 228-77, 1200 E California Blvd, Pasadena, CA 91125, USA

Abstract

Impoverished capacity for social inference is one of several symptoms that are common to both agenesis of the corpus callosum (AgCC) and Autism Spectrum Disorder (ASD). This research compared the ability of 14 adults with AgCC, 13 high-functioning adults with ASD and 14 neurotypical controls to accurately attribute social meaning to the interactions of animated triangles. Descriptions of the animations were analyzed in three ways: subjective ratings, Linguistic Inquiry and Word Count, and topic modeling (Latent Dirichlet Allocation). Although subjective ratings indicated that all groups made similar inferences from the animations, the index of perplexity (atypicality of topic) generated from topic modeling revealed that inferences from individuals with AgCC or ASD displayed significantly less social imagination than those of controls.

Keywords

Autism; Corpus callosum; Agenesis of the corpus callosum; Social inference; Mental attribution

Introduction

Individuals with agenesis of the corpus callosum (AgCC) without evidence of other neuropathology and with intelligence in the normal range (FSIQ > 80), a condition often referred to as Primary AgCC, are characterized by a consistent pattern of mild to

Lynn K. Paul, lkpaul@hss.caltech.edu.

Author Contributions LP and WB designed the study. LP and CK collected the data. CK and FC administered standard scoring, MG applied Topic Modeling, and TR applied LIWC. TR, MG, CK, LP and WB analyzed the data. All authors contributed to the manuscript.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10803-021-04957-2.

moderate cognitive deficiencies (Brown & Paul, 2019). Previous research (e.g., Young et al., 2019), as well as anecdotal reports from family members and friends of high-functioning individuals with AgCC, suggest that these cognitive deficits contribute to impairments in social functioning and increased risk of autism (Paul et al., 2014).

In the current study we used responses from the Animations Test (Abell, Happé, & Frith, 2000; Castelli et al., 2002) to study the capacity to semantically develop social situations (i.e., social imagination) and for mental state attribution (i.e. theory of mind, ToM) in individuals with AgCC compared to both neurotypical controls and high-functioning individuals with ASD. The Animations Test is typically scored in a manner that requires subjective decision-making regarding the presence of social and emotional inferences. However, we also employed analyses of topic models identified by Latent Dirichlet Allocation (LDA), and Linguistic Inquiry and Word Count (LIWC), to better characterize responses to these videos and explore potential differences between social processing deficits in AgCC and ASD.

Agenesis of the Corpus Callosum

AgCC is a congenital brain disorder involving complete or partial absence of the corpus callosum (Jinkins, Whitmore, & Bradley, 1989; Paul et al., 2007). Paul et al. (2006) defined Primary AgCC as “absence of the corpus callosum, with minimal additional neuropathology and general cognitive functioning in the normal range (i.e., FSIQ > 80)” (p. 47). While FSIQ is within the normal range (and occasionally above normal), individuals with Primary AgCC nevertheless tend to have mild to moderate cognitive deficits. It has recently been hypothesized that individuals with AgCC have a core syndrome characterized by reduced interhemispheric sensory-motor interactions, slowed cognitive processing speed, and difficulty in complex reasoning and novel problem-solving (Brown & Paul, 2000, 2019).

The latter two dimensions of this core syndrome give rise to a wide range of associated cognitive challenges that disrupt social functioning. For example, on a variety of tasks involving complex cognitive processes, performance in AgCC is typically characterized by reduced accuracy and increased response times (Brown et al., 1999; Brown, Thrasher, & Paul, 2001; Hines, Paul, & Brown, 2002; Marco et al., 2012). In the context of social processing, Brown and Paul (2019) posit that the core deficit in complex reasoning and problem solving contributes to difficulty imagining more complex possibilities not immediately obvious in the particular context (Symington et al., 2010; Young et al., 2019).

Social functioning is one of the most challenging areas for individuals with AgCC. Previous research exploring social processing in individuals with AgCC has described impoverished comprehension of non-literal language (Brown et al., 2005a, 2005b; Paul et al., 2003; Rehmel, Brown, & Paul, 2016), reduced capacity to deal with emotions in themselves and others (Brown, Anderson, & Paul, 2015) and ToM (Symington et al., 2010). These deficits were all evident in a study that required adults with AgCC to interpret social interactions presented in The Awareness of Social Inference Test (TASIT) video vignettes (Symington et al., 2010): deficits in understanding paradoxical sarcasm, recognizing others’ emotions, and integrating multiple sources of information in order to interpret social situations. This constellation of challenges not only interferes with the ability to accurately comprehend complex social situations, it also interferes with the ability to imagine the mental states of others (ToM) and contemplate alternate possibilities in social scenarios.

This difficulty imagining more complex possibilities not immediately obvious in the particular social context (i.e. social imagination) has been demonstrated in several studies of individuals with AgCC. For example, analysis of narratives generated from six emotionally charged pictures from the Thematic Apperception Test (TAT; Paul, Schieffer, & Brown, 2004) determined that narratives from a small sample of adult individuals with AgCC were less logical and socially relevant than narratives generated by neurotypical controls (Paul et al., 2004). A subsequent evaluation of linguistic content in narratives from a larger sample of children and adults with AgCC found that individuals with AgCC used fewer words related to cognitive processes, emotionality, social processes, and tended to use more present tense verbs and first-person pronouns (Turk et al., 2010). Likewise, when presented with social situations that led up to a decision-point with several potential alternative responses, individuals with AgCC had difficulty describing the consequences of various possible decisions (Young et al., 2019). Overall, their responses were adequate when scenarios had limited alternatives, but as complexity of interacting factors increased and implications of decisions became increasingly subtle, the responses of persons with AgCC became more deficient. In sum, individuals with AgCC have a deficit in social imagination characterized by difficulty imagining the wider-ranging consequences of decisions, inferring how decisions might impact others, and describing circumstances beyond their own experience, as well as difficulty making social inferences that involve attribution of mental and emotional states.

It is unclear if deficits in social imagination are, as proposed by Brown and Paul (2019), a secondary effect of the core deficit in complex reasoning and problem solving (e.g., a general deficit in imagining what is not currently the case as a possible solution to a problem), or if they are in fact a separate and additional domain of core deficiency in AgCC that is specific to imaging social scenarios. Moreover, it is unclear if deficits in ToM are distinct from, or a product of, deficient social imagination in AgCC. To explore these relationships, this study compares the capacity for social imagination and for ToM in persons with AgCC to individuals with a diagnosis that is defined primarily by social deficiencies, autism.

Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that is characterized by “deficits in social emotional reciprocity, in nonverbal communicative behavior used for social interactions, and in developing, maintaining, and understanding relationships” (American Psychiatric Association, 2013, p. 50). In some respects, individuals with AgCC share behavioral and social impairments with high-functioning individuals with ASD (Lau et al., 2013; Paul et al., 2014). For example, research in ASD suggests deficits in social inferences (Baron-Cohen, 1992; Baron-Cohen, Leslie, & Frith, 1985, 1986; Castelli et al., 2002) that are similar to those found in individuals with AgCC (Paul et al., 2014; Symington et al., 2010; Young et al., 2019).

A number of studies have shown that individuals with ASD have a diminished ability to effectively attribute mental states (Baron-Cohen, 1992; Baron-Cohen et al., 1985; Baron-Cohen, Leslie, & Frith, 1986; Hobson, 1993). For example, Thiébaut et al. (2016) and Zalla et al. (2009) examined ToM in high functioning adults with ASD using the advanced ToM Faux Pas test (Baron-Cohen et al., 1999), finding that individuals with ASD had difficulty detecting a faux pas, but engaged in “over detection” of embarrassment.

Individuals with ASD also have difficulty in recognition of emotions, but this depends on the age of the individual and the method of assessment of the ability. For example, individuals with ASD were found to be unimpaired in recognition of emotions compared to typically developing controls when the method of testing included using perceptually oriented tasks, using unambiguous stimuli, and providing adequate processing time (Gepner, Deruelle, & Grynfeltt, 2001; Hobson, Ouston, & Lee, 1988; Humphreys et al., 2007; Ozonoff, Pennington, & Rogers, 1990; Piggot et al., 2004; Rump et al., 2009). However, other studies found impairments in individuals with ASD when stimuli were more complex (Adolphs, Sears, & Piven, 2001; Heerey, Keltner, & Capps, 2003; Humphreys et al., 2007) or were presented briefly (Critchley et al., 2000; Hobson et al., 1988; Mazefsky & Oswald, 2007; Pelphrey et al., 2002). Additional social cognitive deficits in ASD have been found in perspective taking (Mizuno et al., 2011), appraisals of social context (Wang et al., 2006), and regulation of emotion (Samson, Huber, & Gross, 2012).

In addition to similar sorts of deficiencies in social information processing, recent reports indicate elevated rates of diagnosable ASD in the AgCC population. One study of adults with AgCC reported the frequency of an autistic spectrum behavioral profile in their sample was 18% (Lau et al., 2013) and another study reported that ~ 30% of their sample met criteria for an autistic spectrum behavioral profile (Paul et al., 2014). In children with AgCC, 45% exceeded the autism-screening cut-off score of 26 on the child version of the Autism Quotient (Lau et al., 2013). Finally, family members of older children with AgCC (ages 6–11) and of age and IQ matched individuals with ASD both reported that these children had significant problems in social, emotional, and behavioral functioning, although the group with AgCC were significantly less impaired than individuals with ASD (Badaruddin et al., 2007).

The Animations Test

The Animations Test (Abell et al., 2000; Castelli et al., 2000, 2002) assesses the capacity to make appropriate social and mental state inferences based on imagination of the meaning of interactions between geometric shapes. In 1944, Heider and Simmel made a short film of interacting geometric shapes and found that neurotypical adults described the geometric figures as having complex mental states and intentional action. In 2000, Klin introduced the Social Attribution Task, a system for scoring narrative descriptions of the Heider and Simmel (1944) animated figures and reported that adolescents with ASD had impaired capacity to attribute social meaning to the animated shapes (Burger-Caplan et al., 2016; Klin, 2000; Klin & Jones, 2006).

The Animations Test, also based on the idea of Heider and Simmel (1944), uses a set of 12 brief animations to assess an individual’s ability to accurately make social inferences through attribution of mental states and emotions to the interactions of animated triangles (Abell et al., 2000; Castelli et al., 2000, 2002). Compared to the Heider and Simmel animations, the Animations Test stimuli offer greater diversity (12 animations, in color) and greater simplicity (each animation contains only 2 geometric figures). In eight animations, movement patterns of the two triangles (a big red triangle, and a small blue triangle) are designed to evoke the attribution of social interaction: goal-directed interactions (GD: *chase, dance, fight, lead*), and theory of mind interactions indicating mental states and emotions (ToM: *coax, mock, seduce, surprise*). In addition, there are four animations where movement patterns are random and do not display intentionality. Following each animation, the observer verbally describes what was seen and responses are rated for Appropriateness (accuracy of described activity or social script, i.e., dancing, mocking, etc.), and Intentionality (accuracy of mental state attribution to the triangles).

Multiple studies have demonstrated ToM deficits in individuals with ASD using the Animations Test. Using the free-response format, Abell, Happé, and Frith (2000) found responses to animations in the ToM category were significantly poorer in high functioning children with autism than children without autism (including children with other developmental delays). Likewise, when responding using a multiple-choice format, individuals with ASD were less accurate in selecting the emotions typically attributed to the animated objects in these videos, and were less accurate in identifying the presence of mental and physical interactions between agents (White et al., 2011). Finally, Castelli et al. (2002) found that adults with high functioning autism and Asperger's Syndrome made fewer mental state attributions and were less accurate compared to healthy controls for ToM animations, but performed similarly to healthy controls when interpreting goal directed interactions. Functional brain imaging during the test was consistent with the behavioral findings, with autism participants exhibiting less activation than neurotypical controls in the medial frontal cortex, amygdala, temporal pole and the superior temporal sulcus—brain areas associated with ToM and interpretation of biological motion (Castelli et al., 2002).

The Animations Test has also been used to examine ToM and ability to imagine and infer social intent in a small sample of individuals with AgCC (Kang, 2008), compared to high functioning individuals with autism and neurotypical controls. Although individuals with AgCC or ASD were able to identify simple interactions present in the random and goal directed categories, they had significant difficulty in describing the more complex mental states and behavioral intentions evidenced in the ToM animations (i.e. lower scores for subjectively judged Appropriateness on more complex items). While both AgCC and ASD groups exhibited poorer attribution of mental states than controls (as evidenced by lower Intentionality Scores), this ability was somewhat better in individuals with AgCC than in those with ASD.

The present study extends beyond prior studies of social imagination and mental state inference in AgCC and autism by comparing Animations Test performance in larger samples of adults with these diagnoses to performance of neurotypical controls.

Computational Linguistic Analyses of Animation Responses

The current study also introduces significant methodological advancements for the Animations Test: objective scoring using linguistic analyses. Previously, White et al. (2011) implemented a multiple-choice response system to circumvent subjective scoring. However, this approach negated a primary strength of the Animations Test—spontaneous imagination and generation of social inferences (Castelli et al., 2002). To overcome limitations of subjective scoring and more fully explore spontaneously generated social inferences, this study applied two forms of computational linguistic analysis: topic modeling and Linguistic Inquiry Word Count (LIWC).

Topic modeling is an unsupervised machine-learning technique that identifies dominant semantic content in a text or document (Atkins et al., 2012; Blei, 2012; Liu et al., 2016). Latent Dirichlet Allocation (LDA) is one example of topic modeling that is useful for searching the data to produce a summary of the semantic content while reducing dimensionality by producing word clusters or lists representing the topics within each document (Blei, 2012; Blei & Lafferty, 2006). Each topic is a probability distribution over a fixed vocabulary (Blei, 2012). Like factor analysis, topic models include an estimate of how much each word is associated with the given topic (Atkins et al., 2012), but in topic modeling each word may belong to multiple topics, and each document can consist of multiple topics.

In the present study, an LDA model derived from the conglomeration of all control participants' responses on the Animations Test represented the typical core semantic content and served as the referent model for characterizing each individual participant's responses. The response from each individual participant with AgCC or ASD was compared to the referent model based on responses from all control participants, and the response from each control participant was compared to a referent model based on responses of all *other* controls. Comparison between the topic model of an individual participant's response and the referent model provided an index of *perplexity.*

The *perplexity* index describes the correspondence between models from different texts. The theoretical concept of perplexity is a "canonical measure of goodness [of fit] that is used in language modeling to measure the likelihood of held-out data to be generated from the underlying (learned) distributions of the model" (AlSumait, Barbará, & Domeniconi, 2008, p. 6). Lower perplexity indicates a greater likelihood that the test text was generated from the same distribution as the model derived from the referent text (AlSumait et al., 2008). Thus, in this study *lower* perplexity values indicate greater similarity between the topic model from an individual participant and the model from the referent control group, and *higher* perplexity values indicate greater deviance of an individual participant's response compared to the semantic core of the referent control group response.

Linguistic Inquiry Word Count 2015 (LIWC2015; Pennebaker et al., 2015) was used in to quantify how frequently participants used words from various semantic and syntactic categories in their descriptions of the animations. LIWC was originally developed to study language and discourse in expressive writing (Francis & Pennebaker, 1993). It has been used to identify linguistic features of narratives regarding traumatic experiences and to predict improvements in health (Pennebaker, Francis, & Booth, 2001). This program analyzes text by classifying words into 90 semantic and syntactic categories, and stylistic elements of writing (Pennebaker et al., 2015). Of the 41 semantic categories tapping psychological concepts, the current research focused on three psychological processes relevant to social inference: Affective, Social, and Cognitive. These LIWC semantic categories were chosen based on previous research which found that individuals with AgCC used fewer of these words in responses to stimuli from the Thematic Apperception Test (TAT; Turk et al., 2010).

Hypotheses

The current study aimed to clarify the nature of impairments in social imagination and attribution of mental states in adults with AgCC or ASD as demonstrated on the Animations Test and glean greater insight regarding the similarities and differences of social processing deficits in these populations. As in previous studies, we hypothesized that individuals with AgCC or ASD would exhibit limitations in social imagination and impoverished mental state attributions in response to the ToM scenarios of the Animations Test, as evident in lower scores on conventional, subjective ratings of Intentionality and Appropriateness and in a significant group difference in perplexity measures from topic modeling. The direction of group differences in perplexity was expected to provide new insights regarding the social imagination and mental state attribution deficits in AgCC and ASD. Specifically, *lower* perplexity scores in AgCC and ASD than in the control participants (i.e. AgCC and ASD responses more similar to the referent) would indicate more conventional responses and restricted social imagination, and *higher* perplexity scores in AgCC and ASD participants than controls would indicate greater deviation from the thematic/semantic core and more imaginative responses. Nevertheless, with respect to use of words in semantic categories, based on previous findings, we expected that individuals with AgCC and ASD would use proportionally fewer words expressing social interactions, cognitive states, and emotions.

Although AgCC and ASD groups did not differ from controls on conventional, subjective scoring of responses to goal directed animations in previous studies, we hypothesized that social imagination deficits in AgCC and ASD would be evident in group differences on perplexity measures from topic modeling. However, group differences would be greater for ToM than goal-directed animations due to the combined impact of mental state attribution deficits overlaid on social imagination deficits.

Based on previous research, it was expected that individuals with AgCC and ASD would show similar patterns of performance on the Animations Test, but individuals with ASD would generally show a greater discrepancy from neurotypical controls. However, variations of group differences in perplexity and semantic usage across the two types of animations (those depicting simple goal directed social interactions and those depicting interactions) was expected to further clarify commonalities and differences in the mechanisms underlying social deficits in ASD and AgCC.

Methods

Participants

Participants included 14 individuals with AgCC; 13 with ASD; and 14 controls (see Table 1 for a summary of group characteristics). In the AgCC group, 10 participants had complete AgCC and 4 had partial AgCC. For all participants with AgCC, diagnosis was confirmed through review of MRI scans. To avoid confounding effects due to limited general intellectual function, full scale intelligence quotient (FSIQ) ≥ 80 was required, as well as 12 or more years of education. The exclusionary criteria for all participants included intractable epilepsy, history of moderate-to-severe head injury, and drug abuse as assessed by clinical interview.

Group comparisons for demographic variables are shown on Table 1. Groups did not differ in age. There was a significant group difference in FSIQ, $\eta^2_p = 0.153$. The ASD group had a significantly higher FSIQ compared to the AgCC group, $t = -2.43$, $p < 0.02$, $d = 0.93$, but the AgCC and ASD groups did not differ significantly in VCI, $d = 0.69$, or POI, $d = 0.58$. There were no other significant differences between groups for FSIQ: AgCC vs controls, $d = 0.72$; control vs ASD, $d = 0.30$; VCI: AgCC vs controls, $d = 0.81$; control vs ASD, $d = 0.10$; or POI: AgCC vs controls, $d = 0.56$; control vs ASD, $d = 0.01$ (VCI and POI scores were missing for 2 controls). Given the significant differences in FSIQ between the AgCC and ASD groups, all analyses were run covarying FSIQ. A χ^2 test of independence demonstrated that there was not a significant association between gender and group membership, $\chi^2 = 5.01$, $p = 0.08$.

Measures

Animations Test—The Animations Test consists of 12 video clips, ranging in length from 34 to 45 s, depicting 2 triangles moving against a framed white background. Half of the animations also involve a rectangular enclosure with an opening (see Supplemental Materials for examples). Three different types of interactions of the moving triangles are presented: Random (RD), Goal-Directed (GD) and Theory of Mind (ToM). The four ToM animations depict one triangle pretending, persuading, seducing, or surprising the other. The four GD animations show the two triangles dancing, chasing, fighting, or leading. The four RD animations depict the two triangles randomly bouncing, drifting, spinning, or floating. After watching each video, participants are asked to describe what was happening in the animation, with no hint given from the examiner regarding the video's intended content. Responses to the Animations Test were scored using three separate methods: standard subject scoring (Abell et al., 2000; Castelli et al., 2002), topic modeling (Atkins et al., 2012; Blei, 2012; Liu et al., 2016) and Linguistic Inquiry Word Count 2015 (LIWC2015; Pennebaker et al., 2015).

Standard Scoring: The *standard scoring* of the Animations Test is based on three major categories, with each category having its own criteria and point system (Abell et al., 2000; Castelli et al., 2002). The Intentionality index assesses the ability and quality of attribution of mental states to the triangles. Scores range from 0 (descriptions of purposeless movement) to 5 (attributions of mental states to characters). The Appropriateness index assesses whether or not the individual was able to accurately capture the script (i.e., dancing, mocking, etc.), with scores ranging from 0 to 3. Finally, the Length index assessed the length of each response by counting the number of clauses, with shortest to longest scores ranging from 0 to 4, respectively (Castelli et al., 2000).

Topic Modeling: *Topic modeling* was used in order to escape the subjectivity of the typical scoring of the Animations Test. For each GD or ToM response we characterized departures from common themes using a perplexity measure that compared topic models. Topic modeling involves Latent Dirichlet Allocation (LDA), a Bayesian graphical model implemented for text documents which were represented as "bags-of-words" (Blei, Ng, & Jordan, 2003; Griffiths & Steyvers, 2004). In topic modeling, each document (in this case, participant response) is modeled as a multinomial distribution over some number of topics, where each topic is a multinomial distribution of a subset of words. Typically, only a small number of words are important (have high likelihood) in each topic, and only a small number of topics are present in each document (Lau et al., 2010).

For this study, LDA models were trained on referent documents (Animations Test responses from control participants) and compared to a particular document of interest (the testing document involving the response of one participant), deriving a measure of perplexity as an index of document similarity (Lau et al., 2010). Separate referent models involving responses from all control participants (a semantic core) were created for each of the four GD and four ToM animations. In addition, referent models were computed for the across-animation combined responses of controls, again separately for GD and ToM. Perplexity measures the likelihood of a test document having been generated by the training model. In statistical comparisons, we used the inverse log of perplexity, called "logword bound", as a proxy for perplexity, as the range and distribution of values for logword bound are more convenient and better fit the assumptions of statistical analysis—specifically, Perplexity = 2**(−LogWordBound). The logword bound is an upper bound on the number of bits-per-word needed to communicate the text efficiently (i.e., compress the text) given the model. The perplexity value reflected how "perplexed" the LDA model of a particular individuals response was when compared to the GD or ToM referent (semantic core) model. A lower perplexity value reflects greater thematic/semantic similarity between the testing document (in this case, an individual response) and the trained LDA model (the semantic common core).

LIWC2015: *LIWC2015* was also used to analyze Animations Test responses with respect to the proportion of words that fall into various syntactic and semantic categories. For each word, LIWC searches for a dictionary match. If there is a dictionary match, then the word is assigned to the corresponding variable count. There are 90 output variables available including total word count, 4 summary language variables, 3 general descriptor categories, 21 linguistic dimensions, 41 categories of psychological constructs, 6 personal concern categories, 5 informal language markers, and 12 punctuation categories. This study was particularly concerned about the proportions of words in participant responses that involve the psychological categories of social interaction, cognition, and emotion as indications of social imagination and inferences within responses to the GD and ToM animations.

Intelligence Measures—The Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997) was administered to calculate the FSIQ, VCI, and POI for the participants with AgCC and six individuals in the control group. The remaining eight participants in the control group and all participants with ASD were given the Wechsler Abbreviated Scales of Intelligence (WASI; Wechsler, 1999).

Procedure

The Animations Test was uploaded onto a website for administration via internet. At a prescheduled administration time, the participant received the website address via email, and the examiner phoned the participant to confirm that the participant was in a room with no distractions and in front of a computer with internet access. Instructions were then read aloud by the examiner. The presentation order of the twelve animations was counterbalanced between different participants. The participant and examiner remained on the telephone throughout the entire administration. After viewing each animation, participants orally responded to the question: "What was happening in the animation?" Responses were audio recorded, transcribed, and compiled into subject-specific (de-identified) files for scoring.

All responses to GD and ToM animations were scored by one of the original authors of the test for Intentionality (degree of mental state attribution, 0–5 points) and Appropriateness (degree of correctness of descriptions, 0–3 points). Length was not scored for this study.

Separate LDA models were calculated for reference texts based on the combined responses of all controls to each of the four GD animations (i.e., chase, dance, fight, lead), and each of the ToM animations (i.e., coax, mock, seduce, surprise). In addition, models were computed for the combination of all controls over all GD responses, and separately over all ToM responses. Each participant's response was compared to the model resulting from the reference text and a perplexity measure generated for each comparison. Generally, high perplexity values represent reduced fit of the testing document model within the training model.

Since the perplexity values of control participants would be affected by the presence of their own responses within the reference text, a leave-one-out procedure was used. LDA models were calculated based on reference documents that included the responses from all of the control participants except that from the control participant to be tested (i.e., systematically removing one control participant's animation response from referent documents). The testing document (participant response) was then compared to the leave-one-out referent document LDA model to calculate a perplexity measure. Analyses were conducted for texts combined across all ToM animations, and for texts for each animation separately.

In addition, the GD and ToM responses from participants were analyzed using LIWC to identify the percentage of target words for three semantic categories: Affect, Social, and Cognitive. The proportion of words in each of these categories was separately analyzed for responses combined across GD animations and across ToM animations.

Results

Subjective Scoring of Animations

Results of the standard scoring (subjective ratings) of the animations test for each group and type of animation (GD vs ToM) are presented in Table 2. Scores for Appropriateness and Intentionality were analyzed separately.

Appropriateness

In an ANCOVA comparing 3 groups by 2 animation categories (GD vs. ToM) controlling for FSIQ, the groups did not differ overall, $F(2,37) = 0.773$, $p = 0.469$, $\eta_p^2 = 0.040$, but there was a trend toward a significant group-by-animation interaction, with a medium effect size, $F(2,37) = 2.753$, $p = 0.077$, $\eta_p^2 = 0.130$.

Further analyses were conducted to examine the trend toward a significant interaction. Three-group univariate comparisons were conducted separately for each animation type (GD and ToM), with no significant group effect for GD, $\eta_p^2 = 0.029$, nor for ToM, although the ToM comparison had a moderate effect size, $\eta_p^2 = 0.086$. Although no group-wise comparisons were significant, difference from the control group was greater for the ASD group, $\eta_p^2 = 0.067$, than the AgCC group, $\eta_p^2 = 0.000$.

Intentionality

Neither group, $F(2,37) = 1.352$, $p = 0.271$, $\eta_p^2 = 0.068$, nor interaction effects, $F(2,37) = 1.095$, $p = 0.345$, $\eta_p^2 = 0.056$, were significant in the 3-group by 2-animation category comparison of Intentionality ratings. Although no group-wise comparisons were significant, difference from the control group was greater in the ASD group, $\eta_p^2 = 0.104$, than the AgCC group, $\eta_p^2 = 0.052$.

Overall, the results from the subjective ratings of the Animation Test did not yield any significant differences among the groups for the Intentionality index. For Appropriateness ratings, there was a trend toward a significant interaction of the 3-groups by animation category, which was driven primarily by an interaction effect in the comparison of ASD and control group involving lower appropriateness ratings in the ASD on the ToM items.

Topic Model Perplexity of Responses to GD and ToM Animations

Summary statistics and results of group comparisons for perplexity scores for both ToM and GD animations are in Table 3. Responses combined across animations within each category (GD and ToM) were first compared in a group-by-animation type ANCOVA, covarying FSIQ. There was a significant group effect, $F(2,37) = 31.43$, $p < 0.001$, $\eta_p^2 = 0.629$, but the effect of animation was not significant despite overall higher mean perplexity for GD animations, $F(1,37) = 0.41$, *ns*, $\eta_p^2 = 0.011$. The group-by-animation interaction was significant, $F(2,37) = 10.02$, $p < 0.001$, $\eta_p^2 = 0.351$, with group differences being somewhat greater for GD than for ToM for responses combined across specific animations.

Goal Directed—Perplexity values for combined responses to all GD animations were compared using a 3-group ANCOVA controlling for FISQ. There was a significant overall group difference in perplexity values, $F(2,37) = 27.86$, $p < 0.001$, $\eta_p^2 = 0.601$. Follow-up comparisons between each pair of groups revealed significantly higher perplexity values in the control group than both the AgCC, $F(1,25) = 40.85$, $p < 0.001$, $\eta_p^2 = 0.620$, and ASD groups, $F(1,24) = 19.21$, $p = 0.001$, $\eta_p^2 = 0.445$, suggesting greater differences from the core topic model among controls. The AgCC and ASD groups were also significantly different from each other $F(1,24) = 12.12$, $p = 0.002$, $\eta_p^2 = 0.336$, with the mean for ASD between that of AgCC and controls.

Perplexity scores for each of the four GD animations were analyzed with a 3-group by 4-animation MANCOVA. There was a significant multivariate main effect of group, $F(8, 70) = 4.07$, $p < 0.001$, $\eta_p^2 = 0.317$. The three groups differed significantly for three of the four animations: *chase,* $\eta_p^2 = 0.459$, *dance,* $\eta_p^2 = 0.222$, and *lead,* $\eta_p^2 = 0.376$, but were not different for *fight,* $\eta_p^2 = 0.105$. In pairwise group comparisons, perplexity values for all four animations were significantly higher for the control group than the AgCC group (η_p^2 between 0.183 and 0.527). The control and ASD groups were significantly different for *chase* and *lead* animations ($\eta_p^2 = 0.259$ and 0.241, respectively), but did not differ for *dance* or *fight* ($\eta_p^2 = 0.000$ and 0.084, respectively). The AgCC and ASD participants differed significantly for *chase*, *dance*, and *lead* ($\eta_p^2 = 0.220$, 0.224, and 0.211, respectively) with values higher for ASD than AgCC, but they were not significantly different in the perplexity of their responses to the *fight* animation ($\eta_p^2 = 0.043$).

Theory of Mind—Perplexity values for each participant were also calculated for combined responses to all ToM animations (All ToM) and compared using a 3-group ANCOVA controlling for FISQ. There was a significant overall group difference between perplexity values, $F(2,38) = 22.46$, $p < 0.001$, $\eta_p^2 = 0.550$. Follow-up comparisons between each pair of groups revealed significantly higher perplexity values in the control group than both the AgCC, $F(1,25) = 33.55$, $p = 0.000$, $\eta_p^2 = 0.573$, and ASD groups, $F(1,24) = 20.63$, $p = 0.000$, $\eta_p^2 = 0.462$, again suggesting greater differences from the core topic model for controls. The AgCC and ASD groups were not significantly different from each other $F(1,24) = 1.61$, $p = 0.216$, $\eta_p^2 = 0.063$.

Perplexity scores for each of the four ToM animations were analyzed with a 3-group by 4-animation MANCOVA covarying FISQ. A significant multivariate main effect of group was found, $F(8, 72) = 5.52$, $p < 0.001$, $\eta_p^2 = 0.510$. The groups also differed significantly for each of the four animations: *coax,* $\eta_p^2 = 0.577$, *mock,* $\eta_p^2 = 0.591$, *seduce,* $\eta_p^2 = 0.717$, and *surprise,* $\eta_p^2 = 0.609$. For all four animations, the control group's perplexity values were significantly higher than both the AgCC (η_p^2 between 0.538 and 0.719) and ASD (η_p^2 between 0.474 and 0.592) groups. Comparisons of the AgCC and ASD groups yielded significantly larger perplexity scores for ASD for *seduce* ($\eta_p^2 = 0.239$) and similar trends for *surprise* ($\eta_p^2 = 0.111$) and *mock* ($\eta_p^2 = 0.100$), but the groups did not differ for *coax.*

Semantic Category Word Counts to GD and ToM Animations

LIWC was used to investigate the nature of perplexity differences in the topic models of controls and the AgCC and ASD groups. Summary statistics for percentage of words in the semantic categories of Affect, Social and Cognitive for animation responses combined within each category (GD and ToM), as well as the results of group comparisons for each semantic category, are presented in Table 4. LIWC was not used to compute proportions of words in semantic categories for each individual animation because there were too few words per individual response for stable LIWC results.

Goal Directed—For the combined responses to all GD animations, a MANCOVA comparing 3 groups by 3 semantic categories, controlling for FSIQ, revealed a significant multivariate group difference, $F(6, 72) = 3.36$, $p = 0.006$, $\eta_p^2 = 0.219$. The control group used a significantly higher percentage of words from these categories than the ASD group, $F(3,22) = 5.10$, $p = 0.008$, $\eta_p^2 = 0.410$, with a trend toward significantly higher percentage than the AgCC group, $F(3,23) = 2.72$, $p = 0.068$, $\eta_p^2 = 0.262$, but the AgCC and ASD groups did not differ from one another, $\eta_p^2 = 0.202$. With respect to the three semantic categories, only the percentage of Cognitive words differed between groups, $F(2,41) = 10.74$, $p < 0.001$, $\eta_p^2 = 0.367$, with controls using a significantly greater proportion of Cognitive words than the AgCC group, $\eta_p^2 = 0.235$, and the ASD group, $\eta_p^2 = 0.394$.

Theory of Mind—For the combined responses to all ToM animations, a MANCOVA comparing 3 groups by 3 semantic categories (Affective, Social, and Cognitive) and controlling for FSIQ did not find a significant multivariate group difference, $F(6,72) = 1.03$, $p = 0.411$, $\eta_p^2 = 0.079$.

Discussion

Linguistic analysis (topic modeling) of responses to the GD and ToM scenarios of the Animations Test revealed impoverished social imagination and attribution of mental states by individuals with AgCC and ASD compared to a matched control group. Semantic analysis (LIWC) revealed overall lesser use of psychological descriptions by the individuals with AgCC and ASD compared to controls on GD animations, but no group differences on ToM. Surprisingly, group differences were not evident in the standard subjective scoring of either GD or ToM responses.

This combination of results from three different methods of analyzing of Animations Test responses provides important insights about imagination and mental attribution in individuals with AgCC and ASD compared to neurotypical controls, as well as commonalities and differences in the nature of deficits in AgCC and ASD. To best understand these outcomes, it is important to first probe more deeply into the nature of the differences in perplexity scores from LDA topic modeling, which will then provide a better understanding of analyses of the outcomes of subjective ratings and LIWC analyses.

Perplexity Scores from LDA Topic Models

As explained previously, perplexity scores reflect the degree to which an individual participant’s response deviates from the semantic core contained in the conglomeration of all control responses. This approach, generating the referent topic model from a conglomeration of all control participant responses (using a leave-one-out method to generate a unique referent model for each control participant), effectively removed the individual variability in controls’ responses and preserved in the topic model what was most common among controls. Thus, a low perplexity score indicates that an individual's response was very similar to the semantic core shared among the responses of controls, and a high perplexity score indicates that a particular response was notably different from the shared semantic core.

In this study, neurotypical controls had significantly higher average perplexity scores than either the AgCC or ASD groups for both GD and ToM animations overall, and for 7 out of 8 animations when analyzed separately (GD: *chase, dance,* and *lead*; ToM: *coax, mock, seduce,* and *surprise*). Higher perplexity scores reflect greater thematic diversity in the responses given by control participants and indicate a greater range of socially imaginative elaboration with respect to inferring and describing the most salient features of each animation (i.e. goal-directed actions and mental state attributions). In contrast, lower perplexity scores of responses from participants with AgCC or ASD reflect less diverse, more conventional semantic content, indicating that they were less elaborative and imaginative in their social inferences.

This interpretation—That lower perplexity scores in persons with AgCC indicates conventional responses lack imaginative elaboration—is consistent with previous findings in this population. For example, on the Social Norms Questionnaire (Kramer et al., 2013) adults with AgCC demonstrated deficient understanding of the application of social norms in the form of over-adherence to social norms (Brown et al., 2021). Individuals with AgCC also have lower scores on the index of Creative Strengths on the Torrance Tests of Creative Thinking, primarily due to lower scores in abstractness and elaboration (Garrels, 2004). Similarly, persons with AgCC responded to the pictures from the Thematic Apperception Test (TAT) using fewer social, emotional, and cognitive words, suggesting diminished ability to imagine and elaborate these dimensions when telling stories related to the TAT pictures (Turk et al., 2010). Finally, individuals with AgCC were found to have deficits in learning and memory for word-lists and rote word-pairs, but no deficits in learning and memory of verbal information presented in a narrative-context (Erickson, Paul, & Brown, 2014; Paul et al., 2016). This suggests that memory is restricted by a lack of elaborative encoding in the learning phase, which may be overcome with externally-provided elaboration (as in the story narrative). Thus, multiple lines of research suggest that diminished capacity for imaginative elaboration of current stimulus information is a consistent difficulty in individuals with AgCC.

Deficits in imagination have long been considered a core symptom of autism (Craig & Baron-Cohen, 2000; Lord, Leventhal, & Cook Jr. , 2001). In the current study, impoverished imagination and creativity are evident in low perplexity scores (relative to controls) among individuals with ASD as well. Crespi et al. (2016) reviewed results that suggested a deficient social imagination in persons with ASD. They defined as imagination the ability to form "new ideas, mental images, and concepts," considering imagination to be implicated in a matrix of other abilities such as narrative production, pretend play, generativity, and Theory of Mind. For example, children with ASD show reduced capacity to produce pretend play, but engage in pretend play more normally if structure is provided by others (reviewed by Jarrod, 2003). Similarly, children with ASD are less likely than neurotypical children to have an imaginary companion (e.g., Davis et al., 2018). Children with ASD also show difficulties in imaginative drawings (e.g., Ten Eycke & Müller, 2018; Craig, Baron-Cohen, & Scott, 2001) and generating an imaginative narrative (e.g., Ferretti et al., 2018).

Although deficiencies in elaboration of social imagination were evident in interpretation of both GD and ToM animations, in both AgCC and ASD groups the deficits were more pronounced on more complex social interactions for which inferential interpretation requires ToM (i.e. larger effect sizes for ToM animations). Each GD animation presented a single social interaction (chasing, dancing, etc.), whereas each ToM animation portrayed a developing narrative of relational interaction. Across all groups, perplexity scores were generally higher for GD than ToM animations, suggesting that GD animations were likely less constraining of socially imaginative elaborations and thus allowed greater range of elaboration, particularly for controls. In contrast, the greater narrative complexity of each ToM animation, while somewhat more constraining of response content for all participants, nevertheless revealed a more marked elaboration deficit in the AgCC and ASD groups presumably resulting from ToM-specific deficits overlaid on general deficiencies in elaboration of social imagination.

Subjective Scoring

Perplexity results from topic modeling may suggest why neither the AgCC nor the ASD group differed substantially from controls on standard subjective ratings of appropriateness and intentionality. These ratings assess whether responses include the social inferences and apparent intentions that are commonly inferred by neurotypical individuals, i.e. do they include basic and conventional social understandings and judgments. While a highly conventional response gets a positive rating on subjective scoring of the animations, it would get a low perplexity score on topic modeling—as was the case for the AgCC and ASD groups. On the other hand, higher perplexity scores in controls reflected a wider range of elaborations around this core, but such elaborative variety would not impact subjective scoring as long as the core meaning was also included in the responses. The information provided by topic modeling of the Animation Test responses is substantially different than the original scoring and captures important aspects of the deficits in AgCC and ASD.

The current results from subjective scoring fail to replicate previous findings in persons with ASD using the same animations (Abell et al., 2000; Castelli et al., 2002). Particularly noteworthy is the failure to replicate the previous results in adults with ASD reported by Castelli et al. (2002). Using the same animations and same rating system, adults with ASD in the current study were rated higher in intentionality (3.4 ± 0.93) than found by Castelli et al. (2.9 ± 0.6; $t = 1.54$, $p = 0.14$) and significantly higher in appropriateness (1.9 ± 0.7 versus 0.5 ± 0.2; $t = 6.14$, $p < 0.001$). It is possible that conventional scoring captured the impact of somewhat higher verbal general intelligence in the current sample of adults with ASD (111 or 77 percentile) compared cohort studied by Castelli et al. (61 percentile). This subtle difference in verbal ability may have been sufficient to facilitate more conventional mental state attributions in ASD, as would be captured by the subjective scoring system.

Semantic Category Word Counts

Similar arguments can be made with respect to the failure to find robust differences between groups in the LIWC analyses of the proportions of Cognitive, Emotion, and Social words within ToM responses. Although control participants used a greater proportion of psychological words in their responses to the GD animations than either AgCC or ASD participants, the groups did not differ on ToM responses. Given the rudimentary nature of the social interactions that can be inferred from the animations, the imaginative elaborations that resulted in more diverse topic models from control participants' responses likely involved a different topic-related utilization of words that are outside of these three LIWC semantic categories.

The LIWC results of this study are also discrepant from a previous finding that persons with AgCC used fewer words pertaining to emotion, cognitive processes, and social interactions than neurotypical controls in their responses to TAT pictures (Turk et al., 2010). This prior study used an earlier version of LIWC, with the categories examined here refined and expanded in the newer LIWC version. If the LIWC update had any influence on this analysis, one would expect the newer version would be more sensitive to finding group differences. Thus, it is unlikely that discrepant findings can be attributed to the different LIWC versions.

It is more likely that this discrepancy reflects differences in the nature of the Animations Test and TAT stimuli. The Animations Test depicts rather simple interactions intended to elicit basic social inferences (e.g., *chasing, dancing, coaxing, mocking,* etc.). In contrast, TAT stimuli are more complex, depicting persons in social and environmental contexts that are richly suggestive of emotion, cognitive processes and social interactions, and contain information about facial expressions, body posture, and social/environmental context. Stories prompted by TAT pictures would normally be richer in emotional, social, and cognitive content, but would also be more thematically constrained toward semantic content to these categories—that is, imaginative elaborations in TAT responses are more likely to be primarily about what the characters in the pictures are thinking, feeling, and how they interact socially. Thus, the imaginative elaborations of controls would involve more words in these categories than expressed in the presumably less well elaborated (or less appropriate) stories of persons with AgCC.

Finally, in addition to differences in stimuli, the Animations Test and TAT use different prompts to elicit responses. The Animations Test asks participants to merely "tell what you saw," while respondents to the TAT are explicitly asked to tell a story about the picture with a beginning, middle, and end and to tell what the characters are thinking, feeling, and doing. Consequently, responses to the animations are more constrained by the events actually observed (what was seen), while responses to TAT pictures depend much more heavily on a participant's ability to elaborate and expand a narrative beyond what is seen.

AgCC Versus ASD

Previous research has suggested that there are similarities between high functioning individuals with AgCC and high functioning persons with ASD (Paul et al., 2014). As hypothesized for all 3 methods of analysis in this study, comparison with the control group indicated a generally similar pattern of deficits in the AgCC and ASD groups. However, the pattern of relative deviation from the control group varied across animation type (GD vs. ToM) and scoring method. While effect sizes from group comparisons using standard subjective scores (GD and ToM) and LIWC (ToM only) were consistent with our expectation that performance in ASD would be more discrepant from controls, the opposite pattern emerged for perplexity. Perplexity scores in both ASD and AgCC were significantly below the control group, but the discrepancy from controls was *smaller* for ASD, particularly on GD animations. In fact, perplexity scores on GD animations were significantly lower in AgCC than in ASD. Thus, while both groups exhibited greater deficits relative to controls on ToM than on GD animations, the type of animation had less impact on performance of AgCC group.

Overall, these results demonstrate that with respect to controls, deficiencies in imaginative elaboration and social inference are markedly similar in AgCC and ASD and are not isolated to ToM processes for either group. Nevertheless, significantly lower perplexity values in AgCC than ASD on GD animations suggest that while both conditions impair social imagination and mental attribution, this deficit is more directly relevant for ToM processing in ASD and in more broadly distributed across social processing skills in AgCC.

It is worth noting that the presence of ASD symptoms in some individuals with AgCC confounds attempts to distinguish between these conditions. In the current study, 11 participants with AgCC were assessed for ASD as part of a prior study (Paul et al., 2014) study and 2 met criteria for an ASD diagnosis based on current behaviors. For one of the participants with AgCC who met criteria for ASD, perplexity scores were within 1 standard deviation *below* the AgCC group mean (All GD = 9.92, All ToM = 9.27). Perplexity scores for the other participant with AgCC and ASD were somewhat more similar to the pattern in the ASD group, with perplexity for ToM animations within 1.5 standard deviations *above* the AgCC group mean (All ToM = 13.22) and the highest perplexity score in the AgCC group for GD animations (All GD = 18.17, above the ASD group mean). Thus, inclusion of these individuals does not account for the low perplexity outcomes in the AgCC group, nor does it diminish the finding of a shared pattern of mental attribution deficits in AgCC and ASD relative to controls, but their divergent outcomes do suggest individual variability in the relative impact of AgCC and ASD when they co-occur.

Limitations and Future Research

The present study focused on individuals with primary AgCC and high functioning ASD, and it is unclear how these findings might generalize to individuals with complex AgCC (i.e. AgCC with associated anomalies or syndromes) or to lower functioning individuals with ASD.

Generalizability of the study may also be limited by small group size. However, the group size was quite sufficient to detect large differences (large effect sizes) in perplexity scores. It is more likely that the small groups may have obscured group differences in the subjective ratings and LIWC word counts, which might emerge in larger samples. Nevertheless, the pattern of results in this study—robust differences in perplexity contrasted with very much smaller and insignificant differences in ratings and word counts—provided important information about AgCC and ASD. In further research it would be valuable to characterize in greater detail the forms of elaboration present in the responses of controls, but absent in persons with AgCC or ASD. In addition, further research comparing AgCC and ASD with respect to social inferences of this kind using other measures of social inference and accounting for potential moderating factors (e.g. related cognitive skills) is warranted.

Continued use of LDA topic modeling and the perplexity measures is strongly suggested by this research. For example, analysis of topic-model perplexity scores in the TAT stories of persons with AgCC and neurotypical controls (i.e., the responses analyzed by Turk et al., 2010) would be informative and perhaps supportive of the outcome and conclusions of the current study. Most importantly, the current research illustrates possibilities for analyzing free verbal response data in ways that go beyond the limitations inherent in subjective ratings or word classifications most often used with these sorts of data.

Conclusions

Based on the ratings of appropriateness and intentionality, as well as the semantic word counts from LIWC, the outcome of this study suggests that high functioning persons with either AgCC or ASD can make conventional social inferences about the animations in a manner similar to neurotypical controls. What was different between individuals with AgCC or ASD and neurotypical controls appears to be the greater degree of variability in imaginative elaboration in the responses of controls as indexed by their significantly higher perplexity values. For goal-directed scenarios, the ASD group provided more imaginative responses than the AgCC group. This outcome is consistent with the claim made by Brown and Paul (2019) that the core deficit of diminished capacity for complex novel problem-solving in high functioning individuals with AgCC may secondarily limit the generation of elaborative and imaginative solutions in understanding the nature of novel and complex social interactions, as well as inferring mental states of others.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dr. Ralph Adolphs and Brian Cheng for assisting with recruitment and testing of ASD participants. This research was funded in part by a Grant from Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under Award Number R01HD092430 (LP). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Portions of this manuscript were included in the doctoral dissertations of CK and TR.

References

Abell FF, Happé FF, & Frith UU (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. Cognitive Development, 15(1), 1–16. 10.1016/S0885-2014(00)00014-9

Adolphs R, Sears L, & Piven J (2001). Abnormal processing of social information from faces in autism. Journal of Cognitive Neuroscience, 13, 232–240. 10.1162/089892901564289 [PubMed: 11244548]

AlSumait L, Barbará D, & Domeniconi C (2008). On-line LDA: Adaptive topic models for mining text streams with applications to topic detection and tracking. In Proceedings—IEEE International Conference on Data Mining, ICDM (pp. 3–12). s10.1109/ICDM.2008.140.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders: DSM-5 (5th ed.). American Psychiatric Association.

Atkins DC, Rubin TN, Steyvers M, Doeden MA, Baucom BR, & Christensen A (2012). Topic models: A novel method for modeling couple and family text data. Journal of Family Psychology, 26(5), 816–827. 10.1037/a0029607 [PubMed: 22888778]

Badaruddin D, Andrews GL, Bolte S, Schilmoeller KJ, Schilmoeller G, Paul LK, & Brown WS (2007). Social and behavioral problems of children with agenesis of the corpus callosum. Child Psychiatry and Human Development, 38, 287–302. 10.1007/s10578-007-0065-6 [PubMed: 17564831]

Baron-Cohen S (1992). Out of sight or out of mind? A naturalistic study of deception in autism. Journal of Child Psychology and Psychiatry, 33, 1141–1155. 10.1111/j.1469-7610.1992.tb00934.x [PubMed: 1400697]

Baron-Cohen S, Leslie AM, & Frith U (1985). Does the autistic child have a "theory of mind"? Cognition, 21(1), 37–46. 10.1016/0010-0277(85)90022-8 [PubMed: 2934210]

Baron-Cohen S, Leslie A, & Frith U (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children. British Journal of Developmental Psychology, 4, 113–125. 10.1111/j.2044-835X.1986.tb01003.x

Baron-Cohen S, O'Riordan M, Stone V, Jones R, & Plaisted K (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism (1999). Journal of Autism and Developmental Disorders, 29(5), 407–418. 10.1023/a:1023035012436 [PubMed: 10587887]

Blei DM (2012). Probabilistic topic models. Communications of the ACM, 55(4), 77–84. 10.1145/2133806.2133826.

Blei DM, & Lafferty JD (2006). A correlated topic model of science. The Annals of Applied Science, 1(1), 17–35. 10.1214/07-aoas114

Blei DM, Ng AY, & Jordan MI (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3(4–5), 993–1022. 10.1162/jmlr.2003.3.4-5.993

Brown WS, Anderson L, & Paul L (2015). Emotional intelligence in agenesis of the corpus callosum. Archives of Clinical Neuropsychology, 32(3), 267–279. 10.1093/arclin/acx001

Brown WS, Burnett K, Vaillancourt A, & Paul LK (2021). Appreciation of social norms in agenesis of the corpus callosum. Archives of Clinical Neuropsychology. 10.1093/arclin/acab003

Brown WS, Jeeves MA, Dietrich R, & Burnison DS (1999). Bilateral field advantage and evoked potential interhemispheric transmission in commissurotomy and callosal agenesis. Neuropsychologia, 37(10), 1165–1180. 10.1016/S0028-3932(99)00011-1 [PubMed: 10509838]

Brown WS, & Paul LK (2000). Cognitive and psychosocial deficits in agenesis of the corpus callosum with normal intelligence. Cognitive Neuropsychiatry, 5(2), 135–157. 10.1080/135468000395781

Brown WS, & Paul LK (2019). The neuropsychological syndrome of agenesis of the corpus callosum. Journal of the International Neuropsychological Society, 25, 324–330. 10.1017/S135561771800111X [PubMed: 30691545]

Brown WS, Paul LK, Symington M, & Dietrich R (2005a). Comprehension of humor in primary agenesis of the corpus callosum. Neuropsychologia, 43(6), 906–916. 10.1016/j.neuropsychologia.2004.09.008 [PubMed: 15716161]

Brown WS, Symington M, Van Lancker-Sidtis D, Dietrich R, & Paul LK (2005b). Paralinguistic processing in children with callosal agenesis: Emergence of neurolinguistic deficits. Brain and Language, 93(2), 135–139. 10.1016/j.bandl.2004.09.003 [PubMed: 15781301]

Brown WS, Thrasher ED, & Paul LK (2001). Interhemispheric Stroop effects in partial and complete agenesis of the corpus callosum. Journal of the International Neuropsychological Society, 7, 302–311. 10.1017/S1355617701733048 [PubMed: 11311031]

Burger-Caplan R, Saulnier C, Jones W, & Klin A (2016). Predicting social and communicative ability in school-age children with autism spectrum disorder: A pilot study of the Social Attribution Task, Multiple Choice. Autism, 20(8), 952–962. 10.1177/1362361315617589 [PubMed: 27121244]

Castelli F, Frith C, Happé F, & Frith U (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. Brain, 125(8), 1839–1849. 10.1093/brain/awf189 [PubMed: 12135974]

Castelli F, Happé F, Frith U, & Frith C (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. Neuroimage, 12(3), 314–325. 10.1080/13607863.2010.513038 [PubMed: 10944414]

Craig J, & Baron-Cohen S (2000). Story-telling ability in children with autism or Asperger Syndrome: A window into the imagination. Israel Journal of Psychiatry and Related Sciences, 37(1), 64–70.

Craig J, Baron-Cohen S, & Scott F (2001). Drawing ability in autism: A window into the imagination. Israel Journal of Psychiatry and Related Sciences, 38(3–4), 242–253.

Crespi B, Leach E, Dinsdale N, Mokkonen M, & Hurd P (2016). Imagination in human social cognition, autism, and psychotic-affective conditions. Cognition, 150, 181–199. 10.1016/j.cognition.2016.02.001 [PubMed: 26896903]

Critchley HD, Daly EM, Bullmore ET, Williams SCR, Van Amelsvoort T, Robertson DM, Rowe A, Phillips M, McAlonan G, Howlin P, & Murphy DG (2000). The functional neuroanatomy of social behaviour: Changes in cerebral blood flow when people with autistic disorder process facial expressions. Brain: A Journal of Neurology, 123(11), 2203–2212. 10.1093/brain/123.11.2203 [PubMed: 11050021]

Davis PE, Simon H, Meins E, & Robins DL (2018). Imaginary companions in children with autism spectrum disorder. Journal of Autism and Developmental Disorders, 48, 2790–2799. 10.1007/s10803-018-3540-y [PubMed: 29564680]

Erickson RL, Paul LK, & Brown WS (2014). Verbal learning and memory in agenesis of the corpus callosum. Neuropsychologia, 60, 121–130. 10.1016/2014.06.003 [PubMed: 24933663]

Francis ME, & Pennebaker JW (1993). LIWC: Linguistic Inquiry and Word Count. Southern Methodist University.

Ferretti F, Adornetti I, Chiera A, Nicchiarelli S, Valeri G, Magni R, Vicari S, & Marini A (2018). Time and narrative: An investigation of storytelling abilities in children with autism spectrum disorder. Frontiers in Psychology, 9, 944. 10.3389/fpsyg.2018.00944 [PubMed: 29971024]

Garrels SR (2004). Divergent thinking and abstract problem solving in children and adults with agenesis of the corpus callosum. Doctoral Dissertation. Retrieved from Proquest Dissertations and Theses database. Order number 3124710.

Gepner B, Deruelle C, & Grynfeltt S (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. Journal of Autism Developmental Disorders, 31(1), 37–45. 10.1023/A:1005609629218 [PubMed: 11439752]

Griffiths TL, & Steyvers M (2004). Finding scientific topics. Proceedings of the National Academy of Science of USA, 101 (Supplement 1), 5228–5234.

Heerey EA, Keltner D, & Capps LM (2003). Making sense of self-conscious emotion: Linking theory of mind and emotion in children with autism. Emotion, 3(4), 394–400. [PubMed: 14674831]

Heider F, & Simmel M (1944). An experimental study of apparent behavior. American Journal of Psychology, 57, 243–259. 10.2307/1416950

Hines RJ, Paul LK, & Brown WS (2002). Spatial attention in agenesis of the corpus callosum: Shifting attention between visual fields. Neuropsychologia, 40, 1804–1814. 10.1016/s0028-3932(02)00032-5 [PubMed: 12062892]

Hobson R (1993). Autism and the development of mind. L. Erlbaum Associates.

Hobson R, Ouston J, & Lee A (1988). Emotion recognition in autism: Coordinating faces and voices. Psychological Medicine, 18(04), 911. 10.1017/s0033291700009843 [PubMed: 3270834]

Humphreys K, Minshew N, Leonard G, & Behrmann M (2007). A fine-grained analysis of facial expression processing in high-functioning adults with autism. Neuropsychologia, 45(4), 685–695. 10.1016/j.neuropsychologia.2006.08.003 [PubMed: 17010395]

Jinkins JR, Whitmore AR, & Bradley WG (1989). Magnetic resonance imaging of callosal and corticocallosal dysgenesis. American Journal of Neuroradiology, 10(2), 339–344. https://pdfs.semanticscholar.org/38f7/aa5adc4b2f8c4fc62b7c61308d0e9a8314b7.pdf. [PubMed: 2494854]

Kang CH (2008). Mental state attribution in agenesis of the corpus callosum versus high functioning autism. Doctoral Dissertation. Available from ProQuest Dissertations and Theses database 304412195.

Klin A (2000). Attributing social meaning to ambiguous visual stimuli in high-functioning autism and Asperger syndrome: The Social Attribution Task. Journal of Child Psychology and Psychiatry, 41, 831–846. [PubMed: 11079426]

Klin A, & Jones J (2006). Attributing social and physical meaning to ambiguous visual displays in individuals with higher-functioning autism spectrum disorders. Brain and Cognition, 61(1), 40–53. 10.1016/j.bandc.2005.12.016 [PubMed: 16497422]

Kramer JH, Mungas D, Possin KL, Rankin KP, Boxer AL, Rosen HJ, Bostrom A, Sinha L, Berhel A, & Widmeyer M (2013). NIH EXAMINER: Conceptualization and development of an executive function battery. Journal of the International Neuropsychological Society, 20(1), 11–19. 10.1017/81355617713001094 [PubMed: 24103232]

Lau JH, Grieser K, Newman D, & Baldwin T (2010). Automatic Evaluation of Topic Coherence. Paper presented at the meeting Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association Computational Linguistics. Los Angeles, CA.

Lau YC, Hinkley LB, Bukshpun P, et al. (2013). Autism traits in individuals with agenesis of the corpus callosum. Journal of Autism and Developmental Disorders, 43(5), 1106–1118. 10.1007/s10803-012-1653-2 [PubMed: 23054201]

Liu L, Tang L, Dong W, Yao S, & Zhou W (2016). An overview of topic modeling and its current applications in bioinformatics. SpringerPlus, 5(1), 1–22. 10.1186/s40064-016-3252-8 [PubMed: 26759740]

Lord C, Leventhal BL, & Cook EH Jr. (2001). Quantifying the phenotype in autism spectrum disorders. American Journal of Medical Genetics, 105(1), 36–38. [PubMed: 11424991]

Marco EJ, Harrell KM, Brown WS, Hill SS, Jeremy RJ, Kramer JH, Sherr EH, & Paul LK (2012). Processing speed delays contribute to executive function deficits in individuals with agenesis of the corpus callosum. Journal of the International Neuropsychological Society, 18(3), 521–529. 10.1017/S1355617712000045 [PubMed: 22390821]

Mazefsky CA, & Oswald DP (2007). Emotion perception in Asperger's syndrome and high-functioning autism: The importance of diagnostic criteria and cue intensity. Journal of Autism and Developmental Disorders, 37(6), 1086–1095. 10.1007/s10803-006-0251-6 [PubMed: 17180461]

Mizuno A, Liu Y, Williams DL, Keller TA, Minshew NJ, & Just MA (2011). The neural basis of deictic shifting in linguistic perspective-taking in high-functioning autism. Brain, 134(8), 2422–2435. 10.1093/brain/awr151 [PubMed: 21733887]

Ozonoff S, Pennington BF, & Rogers SJ (1990). Are there emotion perception deficits in young autistic children? Journal of Child Psychology and Psychiatry, 31(3), 343–361. 10.1111/j.1469-7610.1990.tb01574.x [PubMed: 2318918]

Paul LK, Brown WS, Adolphs R, Tyszka JM, Richards LJ, Mukherjee P, & Sherr EH (2007). Agenesis of the corpus callosum: Genetic, developmental and functional aspects of connectivity. Nature Reviews: Neuroscience, 8(4), 287–299. 10.1038/nm2017 [PubMed: 17375041]

Paul LK, Corsello C, Kennedy DP, & Adolphs R (2014). Agenesis of the corpus callosum and autism: A comprehensive comparison. Brain, 137, 1813–1829. 10.1093/brain/awu070 [PubMed: 24771497]

Paul LK, Erikson R, Hartman J, & Brown WS (2016). Memory functioning in individuals with agenesis of the corpus callosum. Neuropsychologia, 86, 183–192. [PubMed: 27091586]

Paul LK, Lautzenhiser A, Brown WS, Hart A, Neumann D, Spezio M, & Adolphs R (2006). Emotional arousals in agenesis of the corpus callosum. International Journal of Psychophysiology, 61(1), 47–56. 10.1016/j.ijpsycho.2005.10.017 [PubMed: 16759726]

Paul LK, Schieffer B, & Brown WS (2004). Social processing deficits in agenesis of the corpus callosum: Narratives from the Thematic Apperception Test. Archive of Clinical Neuropsychology, 19(2), 215–222. 10.1016/S0887-6177(03)00024-6

Paul LK, Van Lancker-Sidtis D, Schieffer B, Dietrich R, & Brown WS (2003). Communicative deficits in agenesis of the corpus callosum: Nonliteral language and affective prosody. Brain and Language, 85(2), 313–324. 10.1016/S0093-934X(03)00062-2 [PubMed: 12735947]

Pelphrey KA, Sasson NJ, Reznick J, Paul G, Goldman BD, & Piven J (2002). Visual scanning of faces in autism. Journal of Autism and Developmental Disorders, 32(4), 249–261. 10.1023/A:1016374617369 [PubMed: 12199131]

Pennebaker JW, Boyd RL, Jordan K, & Blackburn K (2015). The development and psychometric properties of LIWC2015. University of Texas at Austin.

Pennebaker JW, Francis ME, & Booth RJ (2001). Linguistic Inquiry and Word Count (LIWC): LIWC2001. Lawrence Erlbaum Associates.

Piggot J, Kwon H, Mobbs D, Blasey C, Lotspeich L, Menon V, et al. (2004). Emotion attribution in high-functioning individuals with autistic spectrum disorder: A functional imaging study. Journal of the American Academy of Child and Adolescent Psychiatry, 43(4), 473–480. 10.1097/00004583-200404000-00014 [PubMed: 15187808]

Rehmel JL, Brown WS, & Paul LK (2016). Proverb comprehension in individuals with agenesis of corpus callosum. Brain and Language, 160, 21–29. 10.1016/j.bandl.2016.07.001 [PubMed: 27448531]

Rump KM, Giovannelli JL, Minshew NJ, & Strauss MS (2009). The development of emotion recognition in individuals with autism. Child Development, 80(5), 1434–1447. 10.1111/j.1467-8624.2009.01343.x [PubMed: 19765010]

Samson AC, Huber O, & Gross JJ (2012). Emotion regulation in Asperger's syndrome and high-functioning autism. Emotion, 12(4), 659–665. 10.1037/a0027975 [PubMed: 22642342]

Symington SH, Paul LK, Symington MF, Ono M, & Brown WS (2010). Social cognition in individuals with agenesis of the corpus callosum. Social Neuroscience, 5(3), 296–308. 10.1080/17470910903462419 [PubMed: 20162492]

Ten Eycke KD, & Müller U (2018). Drawing links between the autism cognitive profile and imagination: Executive function and processing bias in imaginative drawings by children with and without autism. Autism, 22(2), 149–160. 10.1177/136236131668293 [PubMed: 29490482]

Thiébaut F, White S, Walsh A, Klargaard S, Wu HC, Rees G, & Burgess P (2016). Does faux pas detention in adult autism reflect differences in social cognition or decision-making abilities? Journal of Autism and Developmental Disorder, 46(1), 103–112. 10.1007/s10803-015-2551-1

Turk A, Brown WS, Symington M, & Paul LK (2010). Social narratives in agenesis of the corpus callosum: Linguistic analysis of the Thematic Apperception Test. Neuropsychologia, 48(1), 43–50. 10.1016/j.neuropsycholgia.2009.08.009 [PubMed: 19686767]

Wang AT, Lee SS, Sigman M, & Dapretto M (2006). Developmental changes in the neural basis of interpreting communicative intent. Social Cognitive and Affective Neuroscience, 1(2), 107–121. 10.1093/scan/ns1018 [PubMed: 18985123]

Wechsler D (1997). WAIS-III: Wechsler Adult Intelligence Scale (3rd ed.). The Psychological Corporation.

Wechsler D (1999). Wechsler Abbreviated Scale of Intelligence. The Psychological Corporation.

White S, Coniston D, Rogers R, & Frith U (2011). Developing the Frith-Happé Animations: A quick and objective test of theory of mind for adults with autism. Autism Research, 4(2), 149–154. 10.1002/aur.17 [PubMed: 21480540]

Young CM, Folsom RC, Paul LK, Su J, Mangum R, & Brown WS (2019). Social cognition in agenesis of the corpus callosum: Computational linguistic analysis of the awareness of consequences scale. Neuropsychology, 33(2), 275–284. 10.1037/neu0000512 [PubMed: 30667251]

Zalla T, Sav AM, Stopin A, Ahade S, & Leboyer M (2009). Faux pas detection and intentional action in Asperger Syndrome. A replication on a French sample. Journal of Autism and Developmental Disorders, 9(2), 373–382. 10.1007/s10803-008-0634-y

Table 1

Summary statistics of participant demographic information

	AgCC, n = 14			Control, n = 14			ASD, n = 13			*F*	*p*	η_p^2
	M	***SD***	**Range**	***M***	***SD***	**Range**	***M***	***SD***	**Range**			
Age	32.00	09.81	21–51	35.79	15.89	18–61	28.85	12.17	18–55	0.98	0.38	.049
FSIQ[a]	99.14	11.71	84–129	107.29	10.82	88–121	110.92	13.51	93–133	3.43	0.04	.153
VCI[b]	101.64	13.28	83–131	110.00	5.99	101–122	111.08	14.20	90–135	2.43	0.10	.123
POI[b]	101.43	13.80	82–133	109.09	12.65	88–133	109.23	12.17	89–128	1.58	0.22	.083
M/F		6/8			11/3			10/3		$\chi^2 = 5.01$	0.08	

N = 41

FSIQ full-scale intelligence quotient, *VCI* verbal comprehension index, *POI* perceptual organization index

[a] Comparison of AgCC and ASD groups was significant at $p < .05$

[b] VCI and POI control n = 11

Table 2

Summary statistics for subjective scoring averaged across the 4 animations in the goal directed and theory of mind categories

Ratings	AgCC, *n* = 14			Control, *n* = 14			ASD, n = 13			*F*	*p*	η_p^2
	M	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range			
Appropriateness GD	2.64	0.32	1.75–3.00	2.68	0.21	2.25–3.00	2.81	0.26	2.50–3.00	0.55	0.58	.029
Appropriateness ToM	2.02	0.68	1.00–3.00	2.21	0.48	1.25–3.00	1.87	0.68	1.00–3.00	1.75	0.19	.086
Intentionality GD	2.00	0.24	1.50–2.50	2.23	0.37	1.75–3.00	2.15	0.30	2.00–3.00	1.25	0.30	.063
Intentionality ToM	3.38	0.77	2.00–4.25	3.77	0.70	2.75–4.50	3.42	0.93	1.75–4.75	1.25	0.30	.063

GD goal directed, *ToM* theory of mind

Table 3

Summary statistics for perplexity values with ANCOVA of group differences for combined ToM and combined gd and MANCOVA for each animation

Ratings	AgCC, *n* = 14			Control, *n* = 14			ASD, *n* = 13			*F*	*p*	η_p^2
	M	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range			
All GD	11.58	2.81	8.33–18.17	25.35	6.85	13.15–40.87	14.82	4.90	8.36–26.82	27.86	.000	.601
Chase	15.86	5.72	8.83–26.32	33.41	11.21	15.81–51.74	20.42	10.59	10.40–44.27	15.71	.000	.459
Dance	14.38	3.95	9.60–22.02	33.47	12.67	17.93–66.39	32.15	37.25	12.28–136.35	5.29	.010	.222
Fight	21.91	17.73	8.58–71.92	36.30	17.89	20.15–76.55	27.62	18.12	9.18–74.34	2.16	.129	.105
Lead	16.18	4.01	9.57–23.50	38.64	19.56	13.56–84.86	21.60	9.42	8.70–39.50	11.15	.000	.376
All ToM	10.38	2.20	7.90–14.67	17.26	4.29	9.73–24.36	10.82	2.44	7.93–15.52	22.46	.000	.548
Coax	11.49	5.77	6.65–27.43	25.87	9.11	15.16–45.56	10.81	2.81	7.18–14.90	25.01	.000	.576
Mock	8.95	2.60	6.58–14.94	23.63	9.48	12.89–52.29	10.24	2.97	6.84–15.67	25.90	.000	.583
Seduce	8.52	1.32	6.67–10.92	24.83	7.12	14.40–40.42	10.88	4.20	6.63–20.44	46.44	.000	.715
Surprise	9.52	2.84	6.62–16.47	24.16	9.54	12.66–53.79	10.44	2.57	6.86–16.12	28.12	.000	.600

Log word bound (function of the inverse log of the actual perplexity) was used as a proxy for perplexity

ToM theory of mind, *GD* goal directed

Table 4

Summary statistics for multivariate analysis of variance of LIWC semantic categories for goal directed and theory of mind animations

Animation	LIWC variable	AgCC $n = 14$		Control $n = 14$		ASD $n = 13$		F	p	η_p^2
		M	SD	M	SD	M	SD			
Combined GD	Affect	1.68	1.36	2.67	2.04	2.46	2.06	1.26	0.30	.064
	Social	7.21	3.97	8.49	3.95	6.83	4.08	0.63	0.54	.033
	Cognitive	7.66	2.81	11.69	3.14	7.41	3.35	10.74	0.00	.367
Combined ToM	Affect	2.04	1.45	2.25	1.60	2.38	1.61	0.58	0.86	.008
	Social	4.55	2.36	6.79	4.48	4.94	2.25	1.94	0.16	.093
	Cognitive	6.53	2.59	7.58	3.19	5.48	2.42	1.95	0.16	.093

ToM theory of mind, *GD* goal directed