The development of co-speech gesture and its semantic integration with speech in 6- to 12-year-old children with autism spectrum disorders

Wing-Chee So1, Miranda Kit-Yi Wong1, Ming Lui2 and Virginia Yip3,4

Abstract
Previous work leaves open the question of whether children with autism spectrum disorders aged 6–12 years have delay in producing gestures compared to their typically developing peers. This study examined gestural production among school-aged children in a naturalistic context and how their gestures are semantically related to the accompanying speech. Delay in gestural production was found in children with autism spectrum disorders through their middle to late childhood. Compared to their typically developing counterparts, children with autism spectrum disorders gestured less often and used fewer types of gestures, in particular markers, which carry culture-specific meaning. Typically developing children’s gestural production was related to language and cognitive skills, but among children with autism spectrum disorders, gestural production was more strongly related to the severity of socio-communicative impairment. Gesture impairment also included the failure to integrate speech with gesture: in particular, supplementary gestures are absent in children with autism spectrum disorders. The findings extend our understanding of gestural production in school-aged children with autism spectrum disorders during spontaneous interaction. The results can help guide new therapies for gestural production for children with autism spectrum disorders in middle and late childhood.

Keywords
autism spectrum disorders, communication and language, gesture, school-aged children, semantic integration

Introduction

Gesture development in typically developing children
Children gesture when they talk. Co-speech gestures are spontaneous hand movements accompanying speech (McNeill, 1992, 2005). Typically developing (TD) children gesture well before they use speech (i.e. nouns, pronouns) to label objects (e.g. Bates, 1976). For example, 10-month-old infants point to objects to share their interest with adults or to make requests. Iconic gestures (gestures that represent actions or the shapes of objects, such as flapping one’s hands to indicate “bird”) and beat gestures (movements reflecting the prosody, rhythm, and structure of speech without conveying semantic information such as right hand flipping outward) are markedly absent or rare during early periods, but begin to emerge around 18–24 months of age (Özçalışkan et al., 2013; Özçalışkan and Goldin-Meadow, 2011). In later language development (from 36 months of age onward), children begin to show a marked increase in their use of iconic and beat gestures (McNeill, 1992; Nicoladis et al., 1999; Özçalışkan and Goldin-Meadow, 2011).
Gestures have a close relationship with speech. Initially, co-speech gestures reinforce the semantic information conveyed in speech (e.g. saying, “cookie” while pointing to a cookie). A few months later, children produce co-speech gestures to disambiguate speech (e.g. saying “I like this” while pointing to a cookie). Then, they produce gestures to supplement speech (e.g. saying, “I eat” while pointing to a cookie). During this period, disambiguating and supplementary gestures are produced more often than reinforcing gestures (McEachern and Haynes, 2004; Pizzuto and Capobianco, 2005). Recent research has shown that 3- to 5-year-old children use their gestures to clarify a referent that is ambiguous in speech but ought to be specified (on discourse grounds; So et al., 2010, 2014). For the example above, the child points to the cookie to specify the food being eaten. Previous research has also shown that production of supplementary gestures predicts the emergence of multi-word speech (Rowe and Goldin-Meadow, 2009).

**Gesture development in children with autism spectrum disorders**

Children with autism spectrum disorders (ASDs) have significant impairments in communication and social interaction (Diagnostic and Statistical Manual of Mental Disorders (5th ed.; DSM-V; American Psychiatric Association (APA), 2013)). In particular, they have slow understanding and development of gestures (e.g. Asperger, 1944; Bartak et al., 1975; Wetherby and Prutting, 1984). Abundant research has shown that young children with ASD gesture less often than both TD and developmentally delayed (DD) children (e.g. Bono et al., 2004; Camaioni et al., 1997; Mastrogiuseppe et al., 2014; Medeiros and Winsler, 2014; Stone et al., 1997; but see Attwood et al., 1988; Capps et al., 1998). In addition to the delay in producing certain types of gestures, children with ASDs are found to have difficulty in producing certain types of gestures. It is a robust finding that young children with ASDs are particularly impaired in “proto-declarative” pointing gestures, a type of gesture used to draw attention to an object and share interest in it (e.g. a child points to a dog in order to direct his mother’s attention to it) (Baron-Cohen, 1989; Carpenter et al., 2002; Wetherby and Prizant, 2002), although their ability to generate proto-imperative pointing, a type of gesture used in making requests is relatively spared (e.g. a child points to a cookie to request his mother to give him one) (Baron-Cohen, 1989; Wetherby and Prizant, 2002). The delay in use of gesture to communicate among children with ASDs also includes the use of markers, which carry culturally specific meaning, such as the raised thumb for hitchhiking. But ASD children show relatively intact use of transitive or pantomime gestures, which describe actual objects or object uses (e.g. using a finger to represent a toothbrush) (Ham et al., 2010). Besides proto-declarative gestures and markers, studies have also shown that young children with ASDs may have delayed production of other types of gestures, such as iconic gestures and speech beats (Charman et al., 2003; Luyster et al., 2007; Wetherby et al., 2004). However, most autism research on gestural production has been conducted among young children. Very little is known about the ability to produce different types of gestures among children in the grade school age range. It has yet to be established whether children with ASDs aged 6–12 years still lag behind their TD peers in terms of gestural production.

Similarly, little is known about the variables related to gestural production in school-aged children with ASDs. Previous studies have shown that cognitive and language abilities are strong predictors of language and communication skills in children with ASDs (e.g. Charman et al., 2003; Luyster et al., 2008). Three recent studies have examined whether and how cognitive and language abilities are correlated to gesture use in preschool children with ASDs. One study found that higher numbers of gesture types were positively associated with better language comprehension, language expression, and nonverbal thinking skills among 20- to 51-month-old children with ASDs (Braddock et al., 2013). Another study found that cognition and age explained only one-fourth of the variance in production of actions and gestures among 24- to 63-month-old children (Kjellmer et al., 2012). In contrast, the severity of autism symptoms explained an additional 11% of variance in production of actions and gestures. One shortcoming of these findings is that Kjellmer et al. (2012) did not look separately at gestures and actions. In addition, they did not have specific tasks examining children’s cognitive level; rather, they classified them into three cognitive groups (intellectual disability, developmental delay, and normal intelligence) based on psychiatric records.

While Braddock et al. (2013) and Kjellmer et al. (2012) interviewed parents of children with ASDs or asked them to complete questionnaires, Mastrogiuseppe et al. (2014) examined the relationship between language, cognitive abilities, and gestural production by collecting data through standardized assessments and naturalistic parent-child interactions. They found that the number of ideative gestures, which include iconic gestures and markers, was positively correlated with reasoning skills but negatively correlated with language skills among 30- to 66-month-old preschool children. To our knowledge, no previous studies have addressed this issue in school-aged children. More importantly, most of these studies did not administer cognitive tasks in order to examine children’s cognitive abilities (but see Mastrogiuseppe et al., 2014). Among various cognitive abilities, previous research has shown that verbal and spatial memories are associated with gesture frequency among normal individuals (Chu et al., 2014; Hostetter and Alibali, 2007; Sassenberg et al., 2011). For example, individuals with low verbal skills but high spatial
skills gesture more often than others (Hostetter and Alibali, 2007). Another study reported that individuals with poorer visual and spatial working memory gesture more often than those who have stronger visual and spatial working memory (Chu et al., 2014). Based on these findings, our study included assessments of verbal and spatial memory skills of children with ASDs and examined their correlation to gesture use.

Children with ASDs may also have problems synchronizing gestures with speech. However, very little is known about whether children with ASDs integrate speech and gesture in their speech production. Recently, Sowden et al. (2008, 2013) conducted two case studies on this question. One study examined the gestures produced by two 2-year-old children with ASDs who had attended an intensive intervention program on social and communicative skills (Sowden et al., 2008). They found that after training, only one child produced reinforcing gestures (but not supplementary or disambiguating gestures). They later reported findings from a longitudinal case study of four 2- to 3-year-old children with ASDs who had attended the same intervention program based on their spontaneous conversation (Sowden et al., 2013). As expected, there was individual variation in their ability to integrate speech and gestures. However, supplementary/disambiguating gestures were either absent or extremely rare in all children, whereas reinforcing gestures were common. Despite the small sample sizes (N≤5), Sowden et al.’s (2008, 2013) studies are pioneering in the study of the integration of speech and gestures in children with ASDs. This work suggests that young children with ASDs may show delay in supplementary/disambiguating gestures, which are a key milestone in the development of speech–gesture integration.

However, two key questions remain unaddressed in Sowden et al.’s (2008, 2013) studies. First, these studies did not separately analyze supplementary and disambiguating gestures (see Table 2 in Sowden et al., 2013, p. 929.). Unlike disambiguating gestures, supplementary gestures add semantic information to the message conveyed in speech (Özçalışkan and Goldin-Meadow, 2005) and help children to ease the cognitive burden they experience in the transition to further development (Goldin-Meadow, 2003). Second, it is not clear whether the impairment in speech–gesture coordination is also found in older children. Most autism research on the integration of speech and gesture has been conducted among children in early childhood. Very little is known about the development of the ability to integrate speech with gestures among children in the grade school age range.

As a result, previous work leaves open the questions of whether children with ASDs aged 6–12 years have delay in producing various types of gestures compared to their TD peers and whether their gesture frequency is related to their language and cognitive abilities, specifically verbal and spatial memory skills. We also do not know whether children with ASDs in this age range produce supplementary gestures. Therefore, we designed this study to examine gesture in relation to language in school-aged children with ASDs compared to age- and IQ-matched TD children and to investigate whether their gesture frequency is related to their language and cognitive skills.

We hypothesized that (1) children with ASDs would produce fewer types of gestures and gesture less often than TD children, (2) gesture frequency in children with ASDs would be related to the severity of their autism symptoms rather than their language and cognitive skills (Kjellmer et al., 2012), and (3) children with ASDs would be less likely than TD children to produce supplementary co-speech gestures, based on what Sowden et al. (2008, 2013) found in toddlers. We hypothesized that the difficulty in producing supplementing co-speech gesture would carry on until later childhood among children with ASDs.

### Method

#### Participants

A total of 30 Cantonese-speaking children aged 6–12 years participated in this study. Of them, 16 had been diagnosed with an ASD or autistic disorder (two females; aged 8.68 ± 1.29 years, mean ± standard deviation (SD), ranging from 6.93 to 12.15 years), and 14 were age- and IQ-matched TD children (eight females; aged 9.04 ± 1.77 years, ranging from 6.38 to 11.58 years). ASD and TD children did not differ significantly in age, t(28) = 0.68, p = not significant (ns). Neither TD nor children with ASDs had any history of traumatic brain injury, birth-related injury, or seizure disorder, and no TD children had a family history of ASD or other diagnosed developmental disorders or impairments. All procedures were approved by The Chinese University of Hong Kong’s institutional review board in compliance with the Declaration of Helsinki. We obtained the parents’ informed consent prior to the study.

IQs were assessed with the Wechsler Intelligence Scale for Children®, Fourth Edition (Hong Kong; WISC-IV HK) by a qualified clinical psychologist. The participants with ASDs had IQs ranging from 72 to 122 (mean = 91.12), and the TD children had IQs ranging from 77 to 112 (mean = 95.38), t(28) = 0.91, p = ns.

Autism or autism spectrum diagnoses were confirmed by a licensed clinical psychologist with the Autism Diagnostic Observation Schedule™, Second Edition (ADOS™-2; Lord et al., 2012). A total score ≥7 confirmed the presence of an ASD (n = 11), and a total score ≥9 confirmed the presence of an autistic disorder (n = 5).

#### Procedures

Children with ASDs were recruited through two primary schools and three family organizations. TD children were recruited through three primary schools. Children were
tested individually. In order to avoid the stress of overloading the children and parents with experimental tasks, the assessments were separated into three sessions in different venues. Children first took the ADOS-2 (exclusively for children with ASDs) and WISC-IV HK in a clinical psychologist’s clinic. They then took the Hong Kong Cantonese Oral Language Assessment Scale (HKCOLAS; T’sou et al., 2006) either in a speech therapist’s clinic or in their primary schools. In the last session, they played with farm blocks with their caregivers and completed the Hong Kong List Learning Test–Form One (HKLLT-Form One; Chan et al., 1998, 2000, 2002; Cheung et al., 2000) and the Rey Complex Figure Test and Recognition Trial (RCFT; Meyers and Meyers, 1995) in a university laboratory under instructions from research assistants. Parents filled out the Chinese version of the Social Communication Questionnaire (SCQ; Rutter et al., 2003), while their children had their language and cognitive skills assessed in the laboratory. The order of the tasks was identical for all participants.

Tasks

Spontaneous interaction with caregivers. Caregivers were instructed to interact naturally with the children. A farm blocks play set of 36 brightly painted wooden blocks was provided to facilitate communication. This set allowed children to build a barn and play with animals together with their caregivers. Caregivers and children were not instructed or prompted to gesture. The session lasted for approximately 20 min for each child and was videotaped.

Language assessment. A speech therapist administered the narrative test (a subtest in HKCOLAS) which measured children’s ability to retell the content of the story, construct sentences, introduce and switch references, and use connectives to join sentences in Cantonese. We focused on the narrative test because previous research has shown that narratives are segmented into two utterances.4

Assessment of cognitive abilities. We assessed children’s verbal and spatial memory abilities.

Verbal memory tests. The HKLLT-Form One consists of 16 two-character Chinese words, which were presented to each participant three times in three learning trials. We calculated the total learning score by adding the number of words recalled in the first three learning trials (Ho et al., 2003) and reported the standardized score. Scores indicate children’s ability to acquire and retain verbal information.

Visual spatial memory test. In the RCFT task, children are asked to reproduce a complicated line drawing, first by copying it, followed by recalling it 3 min later (immediate recall). We scored participants’ performance on immediate recall, which is indicative of their spatial memory, based on an established manual (Meyers and Meyers, 1995). Standardized scores were reported.

Severity of autism symptoms. We used SCQ (lifetime) to measure the severity of autism symptoms in communication and social functioning. It consists of 40 items (e.g. “Is he/she able to produce short phrases or sentences?”). We scored the SCQ based on the established manual. Higher scores represent more severe communication and social skill impairment.

Coding

We first transcribed the participants’ speech and then coded their gestures.

Speech transcription. All conversations between children and caregivers were transcribed by research assistants who were native Cantonese speakers. All transcripts were then checked by a second coder who was also a native Cantonese speaker.

The stream of speech was segmented into utterances in which each consists of a clause that expresses a proposition1 and includes a predicate2 (Crystal, 1980; Hartmann and Stork, 1972; Pei and Gaynor, 1954). Utterances containing more than one clause connected by a conjunction—for example, and (Cantonese ŋung6 jau5) or but (daam6 hai6)—were segmented into two utterances.4

Gesture coding. We followed Goldin-Meadow and Mylander (1984) (see also Iverson and Goldin-Meadow, 2005; Özçalışkan and Goldin-Meadow, 2005) in excluding hand movements that involved direct manipulation of an object (e.g. placing a block on the table) or were part of a ritualized game (e.g. putting a puzzle in a puzzle slot). Gestures were of four types. Definitions and examples of each type of gesture are given below:

1. Iconic gestures bear a resemblance to the objects they represent or the actions associated with those objects (e.g. index finger and thumb forming a rectangle, classified as a reference to small block; index finger and thumb turned counterclockwise, classified as an action of rotation).
2. Deictic gestures serve to pick out objects (e.g. index finger pointing to a block or holding it up, classified as a reference to block).
3. Markers express culturally specific meaning (e.g. head nod signifying agreement).
4. Speech beats do not carry semantic content but follow the rhythm of the accompanying speech (e.g. index finger flips outward).
We counted the total number of gestures of all types produced by the two groups of children. We also calculated the proportion of each type of gestures children with ASDs and TD children produced when interacting with their caregivers, which was calculated as the total number of gestures in each of these categories divided by the total number of gestures produced during the interaction. Each gesture was then assigned a meaning. The meaning of a gesture was determined by its form in conjunction with the speech in the utterance with which it occurred.

**Utterances with spatial information.** We classified each utterance into two categories: utterances with and without spatial information. Spatial information, including size, shape, spatial dimension, location, direction, and orientation of blocks, could be conveyed in speech and/or gestures.

In terms of speech, examples of utterances with spatial information were, “This block is very small (nei1 go3 zik1muk6 hou2 sai3)”; “It is a triangle (nei1 go3 hai6 saam1gok3jying4)”; “Put this block on top of this (zoeng! nei1 go3 zik1muk6 fong3 hai2 soeng6min6)”; “Let’s turn it around (bat1jyu4 dou3 zyun3 keoi5)”; “There is a cow over there (go2 dou6 jau5 zek3 ngau4).” Some utterances did not contain spatial information, such as “This block is pretty (nei1 go3 zik1muk6 hou2 leng3).”

Gestures are inherently spatial because they are produced in space (McNeill, 1992), and abundant research has shown that gestures convey spatial information (Emmorey et al., 2000; Sauter et al., 2012). Examples of gestures conveying spatial information were index finger pointing to a block (a pointing gesture), which was assumed to refer to a particular block in a specific location, and two hands with open palms moving toward each other (an iconic gesture), which was assumed to refer to putting the blocks next to each other. Children could convey spatial information in gestures alone, even if their accompanying speech expressed nonspatial content. For example, a child pointed to a block while saying, “I like it.” In this example, the pointing gesture indicated the location of the block. However, markers and speech beats did not contain spatial information. If the speech in the corresponding utterances also did not contain spatial information, these utterances would be excluded from analyses.

We counted the proportion of utterances that contained spatial information (in speech only, in both speech and gestures, and in gestures only) produced by children with ASDs and TD children when interacting with their caregivers. This was calculated by dividing the total number of utterances in each of these categories by the total number of utterances with spatial information produced during the interaction.

We further classified utterances containing spatial information conveyed in both speech and gestures into three categories, depending on the semantic relationship between language and gesture (Özçalışkan and Goldin-Meadow, 2005). A reinforcing relation was coded when gesture emphasized or conveyed the same spatial information as the co-occurring speech: for example, a child said “the block is so small” while forming a rectangle with his index finger and thumb. A disambiguating relation was coded when the gesture clarified an underspecified referent: for example, a child said “move this” while pointing to a block. A supplementary relation was coded when gesture added spatial information that was not explicitly conveyed in the co-occurring speech: for example, a child said “rotate” while pointing to the blocks.

We computed the proportions of utterances with reinforcing, disambiguating, and supplementary relations by totaling the sum of utterances in each of these categories divided by the total number of utterances with both spatial language and gestures. We performed an arcsine transformation on all proportions prior to further analyses in order to meet the criteria for parametric analyses.

**Reliability**

To assess inter-coder reliability, we randomly selected 6 of the 30 videos (3 per group) for independent transcription and coding by a second trained coder, who was a native Cantonese speaker. Regarding the number of utterances containing spatial information, we examined the number of utterances with spatial language agreed upon by both coders and divided it by the total number of utterances identified by both coders. The inter-rater agreement was 0.94 (N=656; Cohen’s kappa=0.91, p<0.001). We also examined the number of utterances in which the number of gestures, types of gestures, gesture meanings, and semantic relation between speech and its accompanying gestures agreed upon by both coders. The inter-rater agreement was 0.85 for identifying the same number of gestures (N=144; Cohen’s kappa=0.81, p<0.001), 0.95 for determining types of gestures (N=122; Cohen’s kappa=0.92, p<0.001), 0.91 for identifying the meaning of gestures (N=116; Cohen’s kappa=0.88, p<0.001), and 0.88 for determining the semantic relation between speech and co-occurring gestures (N=617; Cohen’s kappa=0.83, p<0.001).

**Results**

On average, children with ASDs produced 43.00 utterances (SD=27.69, ranging from 14 to 130), and TD children produced 44.43 utterances (SD=27.20, ranging from 12 to 101). Using a Mann–Whitney test, there was no significant difference between groups, U=110.5, p=ns. Of all utterances, spatial information was contained in 49.51% (SD=14.75%, ranging from 0.25 to 0.84) of utterances produced by children with ASDs and in 54.45% (SD=15.36%, ranging from 0.17 to 0.71) produced by TD...
children, U=79, p=ns. In the following analyses, we only included the utterances containing spatial information in speech and/or gesture.

**Gesture frequency and its association with language and cognitive variables**

On average, children with ASDs produced 5.13 gestures of all types (SD=4.24, ranging from 0 to 12), and TD children produced 17.50 gestures of all types (SD=10.29, ranging from 0 to 36, U=29.50, p<0.001. Four children with ASDs and one TD child did not produce any gestures. Table 1 shows the median numbers, ranges, median proportions, and average deviations of different types of gestures in both groups of children.

More than half of the gestures produced by children with ASDs and TD children were deictics. This is not surprising given that the children were playing with blocks, which were easily referred to by pointing gestures. A repeated-measures analysis of variance (ANOVA) was conducted with group (children with ASDs, TD children) as the between-subject independent variable, gesture type (deictic, iconic, marker, beats) as the within-subject independent variable, and the proportion of gestures as the dependent variable. We found significant effects for gesture type, F(3, 84)=22.21, p<0.001, η²=0.44, and interaction, F(3, 84)=3.78, p<0.01, η²=0.12, and a non-significant effect for group, F(1, 28)=1.44, p=ns. Mann–Whitney tests were conducted to examine the differences in the proportions of different types of gestures between both groups of children. We found that TD children produced a significantly higher proportion of markers than children with ASDs. Other pairs did not differ significantly.

We next examined whether language and cognitive skills were correlated with gesture frequency in each group of children. Table 2 shows descriptive statistics for the language and cognitive measures. Using Mann–Whitney tests, we found significant differences in the SCQ, HKCOLAS connectivity, and immediate spatial recall in RCFT between groups. Specifically, children with ASDs scored higher on the SCQ than TD children, U=23.5, p<0.001. TD children scored higher on the narrative measure of connectivity, U=65.50, p<0.05, and on the spatial recall task, U=56, p<0.02, than children with ASDs.

Table 3 shows the correlation between total number of gestures and age, IQ, language, and cognitive variables in both groups. All language and cognitive scores were standardized. The total number of gestures was negatively correlated with age, verbal learning score, and connectivity score in TD children. Age was positively correlated with verbal learning scores, r(14)=0.60, p<0.002, but not connectivity scores, r(14)=−0.10, p=ns. Verbal learning scores were not correlated with connectivity measures, r(14)=−0.12, p=ns. Our findings

---

**Table 1. Median numbers, ranges, median proportions, and average deviations of deictic gestures, iconic gestures, markers, and speech beat in children with ASDs and TD children.**

<table>
<thead>
<tr>
<th>Group</th>
<th>Deictic</th>
<th>Iconic</th>
<th>Marker</th>
<th>Beat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median number</td>
<td>7.50</td>
<td>0.32</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Range</td>
<td>0−53</td>
<td>0.00−0.72</td>
<td>0.00−0.3</td>
<td>0.00−0.2</td>
</tr>
<tr>
<td>Median proportion</td>
<td>0.93</td>
<td>0.31</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Average deviation of proportion</td>
<td>0.08</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Median number of proportion</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Range of proportion</td>
<td>0.00−0.3</td>
<td>0.00−0.3</td>
<td>0.00−0.2</td>
<td>0.00−0.2</td>
</tr>
<tr>
<td>Median deviation of proportion</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

ASD: autism spectrum disorder; TD: typically developing; ns: not significant.
suggest that TD children who were younger, had poorer verbal learning skill, and were less able to express connectivity tended to gesture more.

As opposed to TD children, age, verbal learning score, and connectivity score were not associated with total number of gestures in children with ASDs. Rather, as expected, only SCQ was significantly correlated with total number of gestures. The more severe their communication and social functioning symptoms, the fewer the gestures children with ASDs produced.

**Semantic integration of speech and gesture**

We examined the proportions of utterances containing spatial information expressed in speech only, speech with gesture, and gesture only. Figure 1 shows the distribution of these three types of utterances. In both groups of children, a majority of utterances containing spatial information were composed of speech only. However, a Mann–Whitney nonparametric test showed that TD children produced a significantly higher proportion of utterances containing spatial information expressed in speech with gesture than children with ASDs, U = 48, *p* < 0.05. This result suggests that TD children were more capable of integrating speech and gestures to express spatial information than ASD children in their spontaneous conversations.

We then examined the semantic relation between speech and gesture by looking at the utterances combining speech and gesture in both groups. Figure 2 shows the median proportions of utterances containing speech and gesture classified as reinforcing, disambiguating, and supplementary. A repeated-measures ANOVA was conducted with group (children with ASDs, TD children) as the between-subject independent variable, type of semantic relation (reinforcing, disambiguating, supplementary) as the within-subject independent variable, and the proportion of utterances as the dependent variable. We found a nonsignificant main effect of group, *F*(1, 28) = 0.35, *p* = ns, but significant effects for semantic relation, *F*(2, 56) = 31.91, *p* < 0.001, and interaction between group and semantic relation, *F*(2, 56) = 3.88, *p* < 0.03.

Mann–Whitney tests were conducted for each semantic relation. Children with ASDs produced a lower proportion of utterances containing supplementary semantic relations than TD children, U = 40, *p* < 0.002. Approximately 30% of

<table>
<thead>
<tr>
<th>SCQ</th>
<th>Story content (HKCOLAS)</th>
<th>Referencing (HKCOLAS)</th>
<th>Connectivity (HKCOLAS)</th>
<th>Syntactic complexity (HKCOLAS)</th>
<th>Verbal learning (HKLLT-Form one)</th>
<th>Spatial memory (RCFT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASD Median</td>
<td>19.00</td>
<td>−0.45</td>
<td>−0.75</td>
<td>−1.00</td>
<td>−0.30</td>
<td>−2.57</td>
</tr>
<tr>
<td>Range</td>
<td>12.00 to 30.00</td>
<td>−2.50 to 1.80</td>
<td>−3.80 to 1.10</td>
<td>−1.50 to 0.50</td>
<td>−2.30 to 1.20</td>
<td>−5.97 to 0.41</td>
</tr>
<tr>
<td>Average deviation</td>
<td>3.98</td>
<td>1.10</td>
<td>1.25</td>
<td>0.37</td>
<td>0.78</td>
<td>1.39</td>
</tr>
</tbody>
</table>

| TD Median | 11.00 | −0.20 | −0.20 | 1.30 | −1.25 | −2.16 | 0.34 |
| Range | 6.00 to 19.00 | −2.4 to 1.2 | −3.00 to 0.60 | 0.30 to 1.9 | −2.3 to 0.80 | −3.70 to 0.41 | −1.03 to 2.11 |
| Average deviation | 3.14 | 0.84 | 0.97 | 0.44 | 0.68 | 1.16 | 0.78 |

| U | 23.50 | 90.00 | 111.50 | 65.50 | 74.50 | 86.50 | 56.00 |
| p | <0.001 | ns | ns | 0.05 | ns | ns | 0.02 |

SCQ: Social Communication Questionnaire; HKCOLAS: Hong Kong Cantonese Oral Language Assessment Scale; HKLLT: Hong Kong List Learning Test; RCFT: Rey Complex Figure Test and Recognition Trial; ASD: autism spectrum disorder; TD: typically developing; ns: not significant.

**Table 3. Correlations between total number of gestures and age, IQ, SCQ, four narrative measures in HKCOLAS, verbal learning, and immediate spatial recall.**

<table>
<thead>
<tr>
<th>Age</th>
<th>IQ</th>
<th>SCQ</th>
<th>Story content (HKCOLAS)</th>
<th>Referencing (HKCOLAS)</th>
<th>Connectivity (HKCOLAS)</th>
<th>Syntactic complexity (HKCOLAS)</th>
<th>Verbal learning (HKLLT-Form one)</th>
<th>Immediate spatial recall (RCFT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD children</td>
<td>−0.57*</td>
<td>−0.38</td>
<td>−0.34</td>
<td>0.002</td>
<td>0.32</td>
<td>−0.54*</td>
<td>0.23</td>
<td>−0.68***</td>
</tr>
<tr>
<td>Children with ASDs</td>
<td>0.03</td>
<td>−0.28</td>
<td>−0.65***</td>
<td>−0.20</td>
<td>−0.07</td>
<td>−0.28</td>
<td>−0.27</td>
<td>−0.34</td>
</tr>
</tbody>
</table>

SCQ: Social Communication Questionnaire; HKCOLAS: Hong Kong Cantonese Oral Language Assessment Scale; HKLLT: Hong Kong List Learning Test; RCFT: Rey Complex Figure Test and Recognition Trial; TD: typically developing; ASD: autism spectrum disorder.

* *p* < 0.05; ** *p* < 0.001.
TD children’s speech–gesture utterances used gestures to supplement the accompanying speech, but this kind of utterance was absent in children with ASDs. Nine TD children produced supplementary gestures, but no children with ASDs did. There was no significant difference in the amount of reinforcing gestures, $U=97.5$, $p=\text{ns}$, or gestures disambiguating semantic relations, $U=94$, $p=\text{ns}$, produced by children with ASDs and TD children. One TD child and three children with ASDs produced reinforcing gestures. A total of 11 children in each group produced disambiguating gestures.

**Discussion**

This study showed that children with ASDs produced fewer gestures—especially markers—than TD children. Unlike TD children, whose gestural production was negatively related to age, gestural production was not related to age among ASD children. Instead, severity of autism symptoms was negatively correlated with one subscale of narrative measures, verbal learning skill, and gestural production. In terms of speech–gesture integration, we found that TD children produced a higher proportion of
co-speech gestures than children with ASDs. Among the three types of co-speech gestures, TD children produced a higher proportion of supplementary gestures than children with ASDs, while there was no difference for the other two types of co-speech gestures (reinforcing and disambiguating). Taken together, these findings suggest that children with ASDs in middle to late childhood demonstrate gestural delay.

**Gesture frequency and types**

Prior studies on ASDs have commonly reported communicative gestural delay (e.g. Bartak et al., 1975; Charman et al., 2003; Luyster et al., 2007; Medeiros and Winsler, 2014). However, most autism research on gestural production has been conducted among children in early childhood. To our knowledge, this study is the first in the literature that has analyzed gestural communication (both gesture frequency and types) in 6- to 12-year-old children with ASDs during spontaneous parent–child interaction. Our findings showed that children with ASDs and TD children produced a comparable number and proportion of utterances carrying spatial content. However, children with ASDs produced significantly fewer gestures than their TD peers. This finding suggests that gesture delay as shown in the quantity and diversity of gestures extends to middle and late childhood.

Although children with ASDs produced comparable numbers of iconic and deictic gestures to TD children, they produced markedly fewer markers than TD children. Even those children with ASDs who did produce markers produced fewer than TD children. Markers are conventionally or culturally defined and are used for regulating interaction (McNeill, 1992, 2000). This result is in line with Mastrogiuseppe et al.’s (2014) study, which showed that preschool children with ASDs produced fewer markers. Our study adds to the findings that delay in producing markers does carry over into middle and late childhood. Such delay is not found in other special populations: individuals with Down’s Syndrome (DS), for example, do not have trouble producing markers in order to compensate for linguistic limitations (Abbeduto et al., 2007; Stefanini et al., 2007).

While most studies have used parental reports of children’s gesture use in communication and observation with experimental procedures, this study examined naturalistic spontaneous interaction between children and caregivers in which children with ASDs underperformed compared to TD children in terms of gestural production after controlling the content of speech. Our results are consistent with prior findings of studies with preschool-aged children (e.g. Mastrogiuseppe et al., 2014; Medeiros and Winsler, 2014; Stone et al., 1997; but see Attwood et al., 1988; Capps et al., 1998).

**Relationship of gestural production with language and cognitive skills**

Few studies have examined the relationships between gestural production and specific developmental characteristics and cognitive abilities, especially in the context of spontaneous communication and among school-aged children. In this study, younger TD children gestured more often than their older peers. In addition, TD children with lower ability in learning and retaining verbal information displayed more gestures. TD children who had poorer ability to connect sentences in narrative also tended to gesture more. It is possible that, as TD children grow, they mature linguistically and develop a more complex communication profile in which speech, but not gesture, is the preferred communication strategy. As a result, older TD children, who rely more heavily on speech to communicate, are less likely to gesture than younger TD children. In fact, recent studies on preschool-aged children have shown this developmental trajectory in both TD children and their peers with ASDs (Capirci et al., 2005; Sowden et al., 2013).

Our findings showed that children with poorer verbal learning skills gestured more frequently. Verbal learning skills involve the ability to acquire and retain verbal information, which is crucial in learning languages and other verbal concepts. It has been proposed that gestures, especially iconic and deictic gestures, can help speakers maintain mental representations (De Ruiter, 1998, 2000), lighten speakers’ working memory load (Ping and Goldin-Meadow, 2008; Wagner et al., 2004), and package spatial–motoric information for speaking (Chu and Kita, 2011; Kita and Davies, 2009). Therefore, it is possible that TD children who have poorer verbal learning produce gestures more frequently in order to facilitate mental operations during discourse and to lessen the mental load during communication.

Besides age and verbal learning skill, we found a negative correlation between gesture frequency and a measure of cohesion. Using connectives accurately enhances narrative skill by making the narration of a story cohesive. Narrative ability reflects communicative competence, especially pragmatic skills (Botting, 2002). The negative correlation between narrative ability and gesture frequency indicates that TD children who have not yet mastered pragmatic skills tend to rely on gestures to communicate (So et al., 2010, 2014).

As opposed to TD children, age and narrative and verbal learning skills were not significantly correlated with gesture frequency in children with ASDs. Our finding is in line with Kjellmer et al.’s (2012) findings on preschool children in suggesting that cognitive level and age have low explanatory value in nonverbal communication (production of gestures and actions) in children with ASDs. This finding implies that children with ASDs may rely less heavily on gestures to help conceptualize
and communicate information. Such an interpretation is consistent with the literature, which has shown that unlike TD children, young children with ASDs whose language development is significantly delayed rarely use gestures to aid their communication (e.g. Charman et al., 2003; Luyster et al., 2007).

Rather than cognition and age, severity of delay in communication skills and social functioning (as measured by the SCQ) was found to be related to gesture frequency in children with ASDs in our study. Specifically, children with ASDs who had more severe autism symptoms related to communication and social function tended to produce fewer gestures than those with less severe symptoms. Because gestures are a nonverbal communication skill, it is reasonable to assume that children with ASDs who have more severe impairments in communication and social functioning are less capable of gesturing to communicate. Another explanation is that motor deficits are positively associated with severity of autism symptoms (Dziuk et al., 2007; Mostofsky et al., 2006). Therefore, children with ASDs who display severe autism symptoms are less likely to gesture because of their motor constraints.

**Semantic integration of speech and gestures**

Finally, we found that children with ASDs overall were less able to integrate speech and gestures than TD children. This result is similar to previous studies, which have also reported integration impairments in language production during early childhood (Sowden et al., 2008, 2013). Our results suggest that the delay in overall speech–gesture integration continues through middle and late childhood.

One suggestion is that children with ASDs simply have delayed, but otherwise typical gesture development (Sowden et al., 2008). Speech and gestural production are two independent systems in early language development. Children with ASDs may not automatically learn to integrate these two systems as they age, but may be taught to do so through intervention (Sowden et al., 2008). Another possibility is that children with ASDs have difficulty with sensorimotor coordination between hands and mouth (Iverson, 2010), which may cause difficulty in the integration of speech and gestures.

As opposed to TD children, children with ASDs did not produce utterances containing co-speech supplementary gestures (e.g. pointing to a block and saying “rotate”). While the absence of supplementary co-speech gestures among 2- to 3-year-old children with ASDs is also reported in Sowden et al.’s (2013) study, it is surprising to find that such absence persists among children with ASDs in their middle to late childhood. Sowden et al. suggested that the absence of supplementary co-speech gestures might be attributed to reliance on deictic gestures among children with ASDs. However, in our study, half or more of gestures produced by both children with ASDs and TD children were deictic gestures, and close to 30% of co-speech gestures produced by TD children were supplementary to the accompanying speech. Therefore, the absence of supplementary co-speech gestures among children with ASDs could not be accounted for by reliance on deictic gestures.

Research findings from neuropsychological studies may help explain the delay in speech–gesture integration in children with ASDs. Hubbard et al. (2012) examined the neural processing of co-speech beat gestures during social communication in children with ASDs and TD children in late childhood. Their results indicated that autistic brains could not integrate cross-modal semantic information effectively. Since production of supplementary co-speech gestures requires integration of cross-modal information, children with ASDs may find difficulty in producing supplementary co-speech gestures because of their inability to coordinate information from different modal systems.

In fact, the absence of supplementary co-speech gesture among children with ASDs is also found in children with DS. Iverson et al. (2003) found that children with DS gestured less often than TD children, but they produced gesture–word combinations at a rate comparable to their TD counterparts. However, most of the co-speech gestures produced by children with DS were reinforcing co-speech gestures.

By contrast, children with specific language impairment (SLI) seem to be able to supplement speech with gestures. Evans et al. (2001) found that children with SLI gestured at a similar rate to their TD peers. Interestingly, children with SLI were more likely to produce supplementary gestures than TD children.

Previous research has shown that supplementary gestures play a prominent role in the transition from one-word speech to multi-word combinations in TD children during infancy (Özçalışkan and Goldin-Meadow, 2005; Rowe and Goldin-Meadow, 2009). Then, as children age, speech becomes the dominant modality for communication (Capirci et al., 1996; Pizzuto and Capobianco, 2005). Similar to TD children, children with ASDs were able to produce multi-word combinations when interacting with their caregivers. In fact, the proportion of speech-only utterances was comparable in both groups. However, children with ASDs had lower scores on the narrative measure of cohesion. It is possible that the absence of supplementary gestures impairs the ability to form complex sentences using connectives, but we could not address this issue due to the limited data on supplementary co-speech gestures. To address the causal relationship between supplementary gestures and the use of connectives in children with ASDs, longitudinal studies will be necessary.

Regarding other types of co-speech gestures, we found that only one TD child produced utterances containing reinforcing co-speech gestures, while most of the TD children produced utterances containing gestures that
disambiguated the accompanying speech. These findings are consistent with previous research which has shown that disambiguating and/or supplementary gestures are more prevalent than reinforcing gestures in TD children as early as 18–20 months (McEachern and Haynes, 2004; Pizzuto and Capobianco, 2005). In fact, TD children in early childhood first develop reinforcing co-speech gestures, followed by disambiguating and supplementary co-speech gestures (Rowe and Goldin-Meadow, 2009). A similar pattern was found in this study among children with ASDs: they produced utterances containing disambiguating co-speech gestures more often than those containing reinforcing gestures. This suggests that there is a transition from reinforcing co-speech gestures to disambiguating co-speech gestures among children with ASDs in their middle to late childhood. These children can use co-speech gestures to specify referents that are ambiguous in speech. However, they are not yet able to produce supplementary gestures, which convey additional information not expressed in the accompanying speech.

Conclusion

To summarize, our findings showed that children with ASDs have delay in gestural production in a naturalistic context through their middle to late childhood. Instead of language and cognitive skills which were found to be related to gestural production in TD children, severity of socio-communication impairments was the variable to be related to gestural production in children with ASDs in our study. Gesture impairment also appeared to include a failure to integrate speech and gestures, with supplementary gestures being absent in children with ASDs. The findings extend our theoretical understanding of gestural production and its relation to speech in school-aged children with ASDs during spontaneous interaction, which may help to refine therapeutic strategies (e.g. producing supplementary co-speech gestures) tailored to the specific needs of children and adolescents with ASDs.

This study was approved by an ethics committee and has therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. All parents gave their informed consent prior to their children’s inclusion in the study.

Acknowledgements

We acknowledge the help of our research assistants Ben Ka-Ho Choi, Wing-Lam Amy Chong, Sheera Chan, and Hiu-Man Lavender Chiu. We thank two anonymous reviewers and Stephen Matthews for their comments. Special thanks to all the children and their parents for their help and dedication to education.

Funding

This research has been fully supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (Project no. 449813) and grants from the Chinese University of Hong Kong (Project nos CUHK4930017; CUHK4058005).

Notes

1. A proposition is the meaning content of units within the clause.
2. A predicate is the portion of a clause, excluding the subject that expresses something about the subject.
3. Cantonese examples are given in the Jyut6Ping3 romanization system, developed by the Linguistic Society of Hong Kong (Tang et al., 2002) to meet both linguistic criteria and the constraints imposed by computer applications. The International Phonetic Alphabet (IPA) and Yale correspondences are given in Matthews and Yip (2011: 461–463). The numbers at the end of each syllable represent the tone, from 1 (high level) to 6 (low level).
4. We defined an utterance based on speech. Therefore, gestures which were produced in the absence of speech were not considered utterances. Yet, this situation did not occur. In other words, both groups of children produced gestures while talking.
5. The purpose of a hold-up gesture is not to manipulate the object, but to draw the interlocutor’s attention to it.
6. All children gestured while talking (see Note 4). However, there were some utterances in which spatial information was expressed by gestures only but not speech (e.g. a child says, “I like it” while pointing to the block).

References


