The Relationship Between Visual Perspective Taking and Imitation Impairments in Children with Autism

Yue Yu¹, Yanjie Su¹ and Raymond Chan²

¹Peking University, Department of Psychology
²Chinese Academy of Sciences, Neuropsychology and Applied Cognitive Neuroscience Laboratory, Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, China

1. Introduction

Perspective taking (or role taking) refers to the ability of individuals to distinguish their own perspectives from those of others, and to make correct judgments about the latter. It comprises both the visual (perceptual) level at which the individual imagines what another person can see from a contrasting vantage point (Kurdek & Rodgon, 1975) and the social-cognitive (conceptual) level at which he or she assesses another person’s mental state (also referred to as the “Theory of Mind” or ToM; Baron-Cohen, Leslie, & Frith, 1985). Visual perspective taking (VPT) is further differentiated into two levels, namely, the ability to understand that people have selective visual fields (Level-1) and the ability to understand that they have different views of the same object (Level-2) (Flavell, Everett, Croft, & Flavell, 1981).

According to the hypothesis put forward by Baron-Cohen and colleagues (Baron-Cohen, 1988; Frith & Happé, 1999; Leslie & Frith, 1987), individuals with autism may suffer selective deficit to infer mental representations (referred to as meta-representation, Leslie, 1987), which may affect their ability to infer other’s mental state (ToM) and their own mental state (self-consciousness). However, they can still indicate other’s view in VPT tasks. Empirical evidence generally supports the hypothesis of a theory of mind deficit in autism (e.g., Baron-Cohen et al., 1985; Leslie & Frith, 1988; Perner, Frith, Leslie, & Leekam, 1989; Reed, 1994), but the results on VPT tests have been inconsistent. Some studies have demonstrated VPT deficit in autism (Hamilton, Brindley, & Frith, 2009; Reed, 2002; Warreyn, Roeyers, Oelbrant, & de Groote, 2005; Yirmiya, Sigman, & Zacks, 1994), but others have found no such deficit (Baron-Cohen, 1989; Dawson & Fernald, 1987; Hobson, 1984; Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997; Leslie & Frith, 1988; Reed & Peterson, 1990; Tan & Harris, 1991). These contradictory findings may be due to differences in the task demands (Reed, 2002) and age of participants (Warreyn et al., 2005) among the studies, or to such methodological issues as the ceiling effect (Yirmiya et al., 1994). The first aim of the current study is thus to explore the possible reasons for these inconsistent results.
Research has also reported the imitative impairment of individuals with autism, especially with regard to the imitation of non-meaningful gestures (see Rogers & Williams, 2006; Williams, Whiten, & Singh, 2004). One specific feature of these individuals’ gesture imitation is the “difficulty in correctly imitating the orientation of an action in relation to the model’s body” (Rogers & Williams, 2006, pp. 281). For example, individuals with autism tend to reproduce an inward palm when the tester demonstrates an outward palm (Ohta, 1987). This “reverse” pattern has been found to prevail in both high- and low-functioning children with autism/Asperger syndrome (Vanvuchelen, Roeyers, & de Weerdt, 2007), and it is more common in children with autism than in children with other kinds of developmental disorders (Dewey, Cantell, & Crawford, 2007; Smith & Bryson, 1998; Whiten & Brown, 1998).

In imitation tasks involving self-oriented and other-oriented movements, children with autism showed a similar orientation problem. For example, in Meyer and Hobson’s (2004) study, the tester moved objects either close to herself or close to the child, and the child copied the movement. Children with autism tended to copy the geometric orientations of the objects (close-to-tester move after the tester’s close-to-self move), compared to non-autistic delayed children who usually mirrored the tester’s action. In Carpenter, Tomasello, and Striano’s (2005) study the tester demonstrated a movement either towards herself (E1 condition), towards the child (child condition), or towards another tester (E2 condition). Typically developing infants and children with developmental delay showed both self-self role reversal (self-directed movements in E1 condition) and other-other role reversal (E1-directed movements in child condition), but children with autism demonstrated neither of these role reversals.

Ohta (1987) originally interpreted this type of orientation problem as a difficulty in relating different body parts of the model as a whole in mental images. However, later studies indicated that it was more likely to be attributed to difficulty in encoding actions with reference to one’s role or perspective (Barresi & Moore, 1996; Carpenter, et al., 2005; Smith & Bryson, 1998). Meyer and Hobson used the term “geometric repetition” to depict the featured response pattern in the imitation behavior observed in children with autism. Specifically, it refers to the “responding so that the physical movements and locations of the objects acted-upon were replicated” (p.237). In other words, individuals with autism seem not to view actions with reference of the demonstrator’s personal perspective, instead they refer to the frame of non-personal aspects of the context (Meyer & Hobson, 2004). It is thus suggested that the orientation problem seen in imitation constitutes a perspective-taking problem.

Some empirical evidence has supported a relationship between imitation and social-cognitive abilities in individuals with autism. Imitation impairments have been reported to occur in concert with problems in joint attention (Hobson & Hobson, 2007), symbolic play (Brown & Whiten, 2000), or theory of mind (Meltzoff & Gopnik, 1993). However, we have found no study that directly compares imitation and VPT abilities in autism. We have two reasons for investigating such a connection. First, imitation and VPT share the essential element of being able to understand the correspondence between the representation of the self and that of the other. This element is supposed to be impaired in individuals with autism (Hobson & Hobson, 2007; Rogers & Williams, 2006). If individuals with autism are unable to understand the perspective of the demonstrator, they probably will not incorporate the demonstrator’s stance during imitation, thus resulting in reversal error.
Second, imitation tasks capture perspective-taking skills at the perceptual level—to imitate correctly, children need to make an assumption about what the demonstrator sees when making a gesture or movement. More specifically, to map the demonstrator’s action onto one’s own action requires the ability of Level-2 VPT.

Therefore, this study further explores the controversy about VPT among children with autism and discusses the relationships between children’s VPT and imitation skills. We seek to: a) explore whether children with autism show impairments in their Level-2 VPT skills; b) examine the occurrence of reversal error in these children’s imitation performance; and c) investigate whether their VPT skills are related to their imitative skills. We hypothesize that both VPT and imitation skills are impaired in children with autism, and the occurrence of these two impairments are correlated.

2. Method
2.1 Participants
Fifteen children with autism and 15 typically developing children participated in the experiment. Three additional children with autism were excluded because they were not cooperative \((n = 2)\) or because of equipment failure \((n = 1)\). The autism group (12 boys and 3 girls) were diagnosed with either an autistic disorder \((n = 13)\) or an atypical autism \((n = 2)\) by experienced pediatric clinicians from the Institute of Mental Health of Peking University. These diagnoses were made according to the DSM-IV diagnostic criteria of autism (American Psychiatric Association, 1994), and were based on interview with parents, behavioral observation with children and clinical records consultation. School records showed all of the children to have an IQ below 70 at the time they entered school, as measured by the Wechsler Intelligence Scale for Children Revised (WISC-R, Chinese edition; Lin & Zhang, 1986). All showed a preference for the right hand according to the visual-motor test.

Fifteen typically developing children were recruited as controls from a local preschool. None of their parents or teachers had reported any of them to have diagnoses of medical disease or mental illness. These children were matched on a one-to-one basis with the autism group in terms of sex, verbal mental age (VMA), and handedness. The basic characteristics of the group are shown in Table 1. All of the autistic and control children received small toys or stickers after the test. Written informed consent was obtained from the children’s parents or guardians prior to their participation.

2.2 General procedure
The children were tested individually in a quiet room in their own school. To achieve better engagement, the tasks were divided into two sections. The first section lasted for about 30 minutes and included the administration of the Peabody Picture Vocabulary Test Revised (PPVT-R), the visual-motor test, and the gesture imitation task. The second section lasted for about 10 minutes and included the upside-down picture task.

2.3 Measures
2.3.1 General cognitive tests
VMA was estimated using the PPVT-R (Chinese edition; Sang & Miao, 1990). We defined a VMA match between the autistic and control children as a similar PPVT-R original score (a
difference of less than 3). Visual-motor integration ability was assessed by part of the geometric design task, which is a subtest of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI, Chinese edition; Applied Psychology Center of Beijing Normal University, 1986). The tester asked the child to sequentially copy down four figures (a circle, a cross, a hexagon, and a diamond in a square) on a piece of paper. Their performance was scored according to the instruction booklet, out of a total of 12 points.

2.3.2 Upside-down picture task
The upside-down picture task (Guo, Wang, & Wang, 2004) was used to assess children’s VPT ability. Two upside-down pictures were used in this task: a princess-granny picture (see Fig. 1A and 1B) and an elephant-swan picture (reveals an elephant from one side, and reveals a swan after rotating 180°). Both materials were printed on A4 (21 cm × 29 cm) paper. As some of the children demonstrated deficits in naming objects, they were allowed to respond by pointing out the image that resembled what they had seen from several choices (Fig. 1C). The choice images were selected from abstract cartoon sketches so that the child could not get the correct answer merely by perceptually comparing the choice images and the target. For each picture, four choice figures (two representing objects on the upside-down picture and two representing irrelevant objects) were provided and placed on a piece of A5 (15 cm × 21 cm) paper.

Children sat beside a table and the tester sat next to them. The tester showed them the elephant-swan picture first, followed by the princess-granny picture. Take the princess-granny picture as example. The frequency with which the princess or the granny was shown first was balanced between participants. The tester asked children “Who do you see in the picture?” and repeated the question if necessary. If children remained silent, then he would take out the sheet of paper with choice of images and asked them to point out the one that resembled the person they see from the picture. After children had given their answer, they were asked to point out the mouth of that person to confirm what they had seen. For the elephant-swan picture, children were asked to point out the animal’s head as confirmation. After viewing the two pictures, children moved to the other side of the table and named the person or animal on the opposite side. The same confirmation process followed. The VPT test was administered only after children had identified both of the objects in at least one picture. In the VPT test, the tester asked children to go back to their original seat, and placed a bear puppet opposite to them. After making sure that children understood the puppet was looking at the picture from the opposite side of the table, the tester asked them “What do you see in the picture?” and “What would the puppet see in the picture?” Children were then asked to “point out the head/mouth” of the figure in the picture. Each child completed four such trials (two pictures in two orientations).

Each single answer was scored as correct, opposite, irrelevant, or no response. Correct answers were taken as exact or proximate naming (e.g., naming a princess as a girl or young lady) or the correct selection from the choice images. Opposite answers referred to naming or pointing to the object on the reverse side (e.g., naming the princess as granny). If children gave other answers or if their oral answers conflicted with what they had pointed to, then their response was scored as irrelevant. In the VPT test, two correct responses revealed the ability to infer another’s view, whereas one correct and one opposite response (i.e., answering according to the children’s view in both questions) suggested an egocentric standpoint (Piaget & Inhelder, 1963).
Fig. 1. Princess-granny picture and its choice figures: A. the princess; B. the granny (reverse of A); C. the choice picture, including a princess, a granny, and two irrelevant figures.

2.3.3 Gesture imitation test

The imitation test was adapted from Ohta’s (1987) Gesture Imitation Test. To optimize the test for younger children, we reduced the number of trials and selected gestures that were easy to perform. Only non-meaningful gestures were included, as such gestures best capture the “reproduce the specific action” component of imitation (Williams et al., 2004). Six gestures were chosen after a pretest among typically developing 3 to 4 year olds. The unimanual gestures included a fist, a “V” posture with the index and middle fingers pointing upward, and a “bull’s horn” posture with the thumb and little finger pointing upward. Each gesture could be made with the palm held either outward or inward. Three bimanual gestures were developed from Ohta’s stimuli of T signs and were used in the test (Fig. 2).

The tester and children sat across the table facing each other. For each trial, the tester attracted children’s attention by saying “watch me” before demonstrating the gesture. In the training trial, the tester held up the index finger of his left hand with the palm facing out, and asked children to “copy my action.” Children who held up their right index finger facing out were considered to be correct, and those who did not make this gesture were corrected until they understood the kind of mirror copying that was expected. The main part of the test then began with the imitation of unimanual gestures, followed by bimanual gestures. The procedure was similar to the training trial except for the tester’s feedback. If children did wrong, the tester would give them a verbal hint (“Are you sure?”) and see if
they can correct by themselves. For those who stayed wrong, the tester would demonstrate the gesture a second time, and children were given a second attempt. The order of the gestures and the direction of the palm were balanced across participants.

Fig. 2. Bimanual gestures used in gesture imitation.

Each child’s imitation performance was videotaped and scored for correctness and error type. Correctness for each gesture was scored between 0 and 3 (3: correct in the first attempt; 2: self-corrected after the verbal hint; 1: correct in the second attempt; 0: failed). A simplified version of the six-type error system of Vanvuchelen et al. (2007) was used to determine the error types. Errors were categorized as behavioral, content, spatial, or no error. Behavioral errors referred to no response or unwilling to copy. Content error referred to using the wrong hand or fingers to perform gestures, or positioned the fingers wrongly. If the fingers were correctly positioned, but the spatial direction was wrong (rotated from the correct position), then this was regarded as a spatial error. For each gesture, error types were scored for both the first and second attempts (if children were correct at the first attempt, both attempts would be scored as “no error”). Twenty percent of the videotapes (six participants, three from each group) were randomly selected and scored by an independent observer who was blind to research hypotheses. Interrater agreement was assessed using Cohen’s kappa statistic. For correctness, $\kappa = 0.76$, and for error type, $\kappa = 0.75$, both of which are above the excellent agreement level defined by Fleiss (2003).
3. Results

3.1 Data analyses
The between-group contrasts were analyzed using independent \( t \) tests, the Mann-Whitney \( U \) test, and mixed-design ANOVA in SPSS 13.0, with a threshold of \( \alpha = .05 \) for statistical significance. Fisher’s exact test was used to analyze the frequency data. The effect sizes of planned contrasts were assessed with Cohen’s \( d \) and unbiased \( r \) (Field, 2005). A summary of results is listed in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Autism group ((n = 15))</th>
<th>Control group ((n = 15))</th>
<th>( t )-value</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronological age (y; m)</td>
<td>8; 11 SD 0; 7 Range 8-0-9; 11 SD 3; 3y 9; 0; 8 Range 2; 8-5; 1</td>
<td>21.86</td>
<td>&lt; .001</td>
<td></td>
</tr>
<tr>
<td>Verbal mental age (y; m)</td>
<td>4; 10 SD 0; 8 Range 3; 2-6; 8 SD 4; 9y 0; 11 Range 3; 2-6; 8</td>
<td>0.01</td>
<td>&gt; .01</td>
<td></td>
</tr>
<tr>
<td>Time of training (y; m)(^a)</td>
<td>1; 6 SD 1; 1 Range 0; 1-2; 7</td>
<td>-- -- --</td>
<td>-- -- --</td>
<td>-- -- --</td>
</tr>
<tr>
<td>Visual-motor score (max:12)</td>
<td>6.0 SD 3.2 Range 2-11</td>
<td>2.7 SD 1.6 Range 1-5</td>
<td>3.56</td>
<td>.003</td>
</tr>
<tr>
<td>Imitation score (max: 18)</td>
<td>7.1 SD 3.9 Range 2-15</td>
<td>11.5 SD 2.7 Range 5-15</td>
<td>-3.58</td>
<td>.001</td>
</tr>
<tr>
<td>Imitation error (max: 18)</td>
<td>10.7 SD 3.4 Range 3-15</td>
<td>5.8 SD 2.6 Range 3-13</td>
<td>4.39</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>VPT score (percent correct)(^b)</td>
<td>13.9 SD 28.3 Range 0-75</td>
<td>82.8 SD 33.7 Range 0-100</td>
<td>-8.50 (^c)</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

\(^a\) Time of training is defined as the time period of training in a special education school, and applies only to the autism group.

\(^b\) VPT: visual perspective taking. Seven children in the autism group and nine in the control group participated in this test.

\(^c\) Intergroup difference in the VPT score is measured by the Mann-Whitney \( U \) value.

Table 1. Characteristics and descriptive statistics for the autism group and the control group.

3.2 General cognitive tests
Autism group and control group were not different in the original PPVT-R score (for autism \( M = 34.3 \), for control \( M = 34.1 \), \( t(28) = 0.04 \), \( p = .97 \), \( d = 0.01 \)). According to the Chinese norm (Sang & Miao, 1990), their scores were converted into VMAs of 4 years 10 months and 4 years 9 months, respectively. The children with autism exhibited better performance than did the control children in the visual-motor task, \( t(20.5) = 3.56 \), \( p = .002 \), \( d = 1.30 \).

3.3 Upside-down picture task
The results for the upside-down picture task are presented in Table 2. Six children in the autism group were excluded from this analysis because they failed the initial recognition part. Of these failures, one child failed to recognize all of the objects presented to him, and five children stuck with one object and could not figure out the competing ones (e.g., they could not recognize “granny” after they saw “princess” on the other side). The latter event also occurred to four typically developing children. Overall correctness in figuring out the competing objects did not differ between the groups (\( t(28) = -1.08 \), \( p = 0.29 \), \( d = -0.32 \)). Of the nine children with autism who participated in the VPT test, only two were able to infer the puppet’s view (although their performance was not consistent across the trials), and seven gave egocentric answers in 11 trials. In contrast, 10 of the 11 children in the control group were able to correctly infer the puppet’s perspective (two were inconsistent), and only one gave egocentric answers. The Fisher’s exact test shows an unequal distribution of the
number of children capable of taking another’s perspective versus the number of those taking an egocentric view between the groups ($p = .003$). The average VPT scores of the autism and control groups confirmed this difference ($U = 8.5, p < .001, d = -2.18$).

<table>
<thead>
<tr>
<th>Not included in VPT test</th>
<th>Autism group ($n = 15$)</th>
<th>Control group ($n = 15$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>fail to name either object</td>
<td>6 (40%)</td>
<td>4 (27%)</td>
</tr>
<tr>
<td>fail to recognize the alternative object</td>
<td>1 (7%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Included in VPT test</td>
<td>9 (60%)</td>
<td>11 (73%)</td>
</tr>
<tr>
<td>fail to take the perspective of the doll</td>
<td>7 (47%)</td>
<td>1 (7%)</td>
</tr>
<tr>
<td>inconsistently take the perspective of the puppet</td>
<td>2 (13%)</td>
<td>2 (13%)</td>
</tr>
<tr>
<td>consistently take the perspective of the puppet</td>
<td>0 (0%)</td>
<td>8 (53%)</td>
</tr>
</tbody>
</table>

Table 2. Percentage of each type of performance in the autism ($n = 15$) and control ($n = 15$) groups during the upside-down picture task. VPT: visual perspective taking.

### 3.4 Gesture imitation

The total imitation score of children with autism was significantly lower than those in the control group ($t (28) = -4.40, p = .001, d = -1.31$). This discrepancy was evident in both the unimanual gestures ($t (14) = -2.62, p = .01, d = -0.96$) and bimanual gestures ($t (14) = -2.69, p = .01, d = -0.98$). Fig. 3 showed the percentages of correct imitation for the first and second attempt, as analyzed in a $2 \times 2 \times 2$ repeated measures ANOVA. The results again revealed the main effect of group, $F (1, 14) = 24.33, p < .001, \eta^2 = .64$, the main effect of uni/bimanual, $F (1, 14) = 9.79, p < .01, \eta^2 = .41$, as well as the main effect of number of attempt, $F (1, 14) = 62.42, p < .001, \eta^2 = .82$. Both groups performed better in unimanual gesture and in the second attempt. An interaction between group and number of attempt was also observed, $F (1, 14) = 4.51, p = .05, \eta^2 = .24$. Autism group made less improvement between the first and second attempt, $t (14) = 2.12, p = .05, d = 0.54$.

Fig. 3. Correct percentages of the first and second attempts in gesture imitation
We analyzed the types of errors in a 2 (group) × 2 (uni/bimanual) × 3 (error type) ANOVA. Results (Fig. 4) showed a main effect for all three factors. The number of errors in the autism group was significantly higher than that in the control group ($F(1, 28) = 18.70, p < .001, r = 0.63$), and the error for bimanual gestures was higher than that for unimanual gestures ($F(1, 28) = 4.82, p = .02$). Main effect for error type was also significant ($F(2, 13) = 46.94, p < .001, \eta^2 = 0.63$). A post hoc multiple-comparison Bonferroni test revealed greater spatial error than content error ($MD = .32, p < .001$) and behavioral error ($MD = .30, p < .001$), but no difference between the latter two ($MD = .02, p = .28$). The only interaction observed was group × error type ($F(2, 56) = 5.03, p = .02$). The autism group made significantly more spatial errors than the control group, $t(21.6) = 2.95, p = .008, d = 1.08$. There were no group differences for the behavioral and content errors.

Fig. 4. Mean percentage frequency of different types of errors (+SE of total error) for the autism ($n = 15$) and control ($n = 15$) groups in unimanual and bimanual gesture imitation.

Spatial error accounted for the majority (81.4%) of total errors, so we investigated spatial error in detail. All of the observed spatial errors took the form of “reversal errors” — that is, instead of mirroring the demonstration, children rotated the hand position by 180° to replicate the actual position of the demonstration. The results showed that children with autism were more prone to this kind of error, $t(21.6) = 2.95, p = .008, d = 1.08$, and this was true for both the unimanual ($t(28) = 2.82, p = .009, d = 1.03$) and bimanual gestures ($t(28) = 2.06, p = .049, d = 0.75$).

3.5 Correlation analysis

We used correlation analysis to investigate the potential relationship between imitation and VPT deficiency in autism. The results of this analysis (Table 3) showed that VMA is correlated with imitation score ($r = .38, p = .04$); this correlation was significant for bimanual gestures ($r = .41, p = .03$), but not for unimanual gestures ($r = .19, p = .33$). Visual-motor ability and educational level showed no significant correlation with imitation performance. The VPT score was highly correlated with the imitation score ($r = .74, p < .001$), and this correlation remained significant after controlling for VMA ($r = .67, p = .002$). The fact that children with autism performed poorly in both VPT and imitation tasks may contribute to the correlation, so we analyzed the two groups separately. In the autism group, VMA was
not associated with scores for imitation, but the VPT score showed marginal correlation with imitation score ($r = .60$, $p = .09$). Children who passed the VPT test in at least one trial scored higher in imitation than those who failed the VPT test ($t(7) = 2.66$, $p = .03$, $d = 2.13$). Correlation between VPT and imitation scores was also observed in the control group, but it was not statistically significant ($r = .46$, $p = .15$).

Taking the error type into account, we found that spatial error, but neither of the other two types of errors, was inversely correlated with the imitation score ($r = -.64$, $p < .001$). Moreover, while VMA, visual-motor ability, and educational level were not found to be correlated with spatial error, significant correlation was found between spatial error and the VPT score across the two groups ($r = -.59$, $p < .01$). Within the autism group, children who passed the VPT test in at least one trial made fewer spatial errors than did those who failed ($t(7) = 2.82$, $p = .02$, $d = 2.34$). The correlation between VPT and spatial error was not significant in the control group ($r = -.12$, $p = .73$).

<table>
<thead>
<tr>
<th></th>
<th>Verbal mental age</th>
<th>Visual-motor score</th>
<th>Education level</th>
<th>VPT score$^a$</th>
<th>Total spatial error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual-motor score</td>
<td>0.33</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education level</td>
<td>-0.22</td>
<td>-0.03</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VPT$^a$ score</td>
<td>0.42</td>
<td>-0.19</td>
<td>0.09</td>
<td>-0.59***$^†$</td>
<td></td>
</tr>
<tr>
<td>Spatial error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>-0.13</td>
<td>0.32</td>
<td>0.12</td>
<td>-0.59***$^†$</td>
<td></td>
</tr>
<tr>
<td>unimanual</td>
<td>-0.01</td>
<td>0.34</td>
<td>0.05</td>
<td>-0.49*</td>
<td>0.59***$^†$</td>
</tr>
<tr>
<td>bimanual</td>
<td>-0.16</td>
<td>0.23</td>
<td>0.12</td>
<td>-0.52*</td>
<td>0.92***$^†$</td>
</tr>
<tr>
<td>Imitation score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>0.38*</td>
<td>-0.15</td>
<td>-0.3</td>
<td>0.74***$^†$</td>
<td>-0.64***$^†$</td>
</tr>
<tr>
<td>unimanual</td>
<td>0.19</td>
<td>-0.16</td>
<td>-0.16</td>
<td>0.73***$^†$</td>
<td>-0.49**</td>
</tr>
<tr>
<td>bimanual</td>
<td>0.41*</td>
<td>-0.08</td>
<td>-0.34</td>
<td>0.53*</td>
<td>-0.53**</td>
</tr>
</tbody>
</table>

$^a$ VPT: visual perspective taking.

$p < .05$, **$p < .01$, ***$p < .001$.

$^†$ Significant correlations after Bonferroni correction for multiple comparisons (Abdi, 2007), $α' = .05/35 = .0014$.

Table 3. Pearson correlations between individual variables and performance in gesture imitation.

4. Discussion

In this study we compared children with autism and typically developing children for their performance in a VPT task and an imitation task. The results confirmed our three hypotheses: a) children with autism showed impairment in their Level-2 VPT skills, as they were mostly unable to determine what a puppet might see in the picture that they themselves were looking at; b) these children also performed poorly in gesture imitation compared to the controls and committed more spatial errors in the form of reversal error; and c) the VPT and imitation scores and the occurrence of spatial error were correlated with one another. However, we should interpret these findings cautiously. It should be noted
that the comparisons in this study are much younger typically developing children, and we
did not recruit a developmental delay control with both chronological age (CA) and VMA
matched. Therefore, the findings are still preliminary to suggest that the impairments are
specific to children with autism.

The finding of VPT impairment in children with autism is in line with some previous
studies (Hamilton, et al., 2009; Reed, 2002; Warreyn, et al., 2005; Yirmiya et al., 1994), but
contradicts the findings of others (Baron-Cohen, 1989; Hobson, 1984; Leekam et al., 1997;
Leslie & Frith, 1988; Reed & Peterson, 1990; Tan & Harris, 1991). One of the possible reasons
for this inconsistency is the different designs of tasks. Most of the aforementioned studies
had applied Level-1 VPT tasks. These include the “puppet seeing” paradigm, the “hide and
seek” paradigm, and the “cube” paradigm (seminal design see Hobson, 1984). In each of
these paradigms, children are required to predict whether a puppet (or the tester) would see
certain object that can be either in sight/ not shielded or out of sight/shielded.

The accomplishment of such Level-1 VPT tasks requires an understanding of whether there
is an unbroken line-of-sight between the agent’s eyes and the target. This ability is generally
believed to be unaffected by the pathology of autism (Leslie & Frith, 1988). In fact, five out
of the eight studies testing Level-1 VPT (Baron-Cohen, 1989; Hobson, 1984; Leslie & Frith,
1988; Reed & Peterson, 1990; Tan & Harris, 1991) found children with autism performed
more than 90% correct on average. One (Leekam et al., 1997) found 66% children with
autism pass the test, which is comparable to children with Down syndrome and normal
children. One (Reed, 2002) found that children with autism were impaired in VPT tasks with
high working memory load. One (Warreyn, et al., 2005) found Level-1 VPT deficiency in a
much younger sample of autism (average CA 5:6, VMA 4:4).

The Level-2 VPT, as tested by our upside-down picture task, could not be simply solved by
the line-of-sight rule. Children needed to further infer what the puppet would conclude
from the properties it observed (e.g., whether the puppet would perceive a granny face with
a mouth at the top or a princess face with the mouth at the bottom), which involved a
mindreading component. Five studies (Dawson & Fernald, 1987; Hamilton, et al., 2009; Reed
& Peterson, 1990; Tan & Harris, 1991; Yirmiya et al., 1994) had administrated Level-2 VPT
tasks in children with autism, of which two (Reed & Peterson, 1990; Tan & Harris, 1991)
found no impairment.

In Reed and Peterson’s (1990) task, children were instructed to "turn (an object) round so I
(the tester) can see the nose/tail/front/back". However, this task could be accomplished
with the line-of-sight rule, if children focus on one part of the object (e.g., the nose of a teddy
bear). Tan and Harris’s (1991) Level-2 VPT task is susceptible to a similar problem because it
required children to answer what the dolls would say that was “in front” of them. In
addition, the autism sample in their study was much older than our sample in both CA (12:8
versus 8:11) and VMA (7:7 versus 4:10), and those children performed as well as typically
developing children in all of the tasks administrated.

Dawson and Fernald’s (1987) study employed five perceptual perspective taking tasks, of
which three (Upside down, Face and Spontaneous) resembled the tests used in our study.
Sixteen children with autism (6-14 years old) scored an average of 7.8 out of 18 combining
these five tasks. In comparison their average score on conceptual perspective taking was 7.4
out of 11. Unfortunately this study did not employ a control group so we do not know
whether the impairment was due to delayed general intelligence or it was specific to autism.

Yirmiya and colleague’s (1994) employed two turntables that have 3 or 10 items on them.
Children were requested to rotate one turntable to replicate the tester’s view on the second. Children with autism (age 9;3-16;10) performed worse than the CA and IQ matched normal controls. Finally, Hamilton and colleagues (2009) included a large and well-controlled sample (23 children with autism, 60 typical children in three age groups) and tested them on a turntable-type Level 2 VPT task and a closely matched mental rotation task. The result showed that children with autism have difficulty in VPT task compared to the mental rotation task. Taken into the background of these studies, our results indicate that although children with autism can infer another’s line of sight, they may suffer problem in Level-2 VPT tasks which involve a mindreading component.

In the current study, we also confirmed that children with autism display reversal error during imitation, which is consistent with previous discoveries (Ohta, 1987; Meyer and Hobson, 2004; Smith & Bryson, 1998; Vanvuchelen, et al., 2007). However, the current test differs from the previous studies in that we used a training trail where children were trained to perform mirror-image imitation. Instead of measuring instinctive response (as emphasized in Meyer & Hobson, 2004), we expected to elicit a predominant response before the main test. The results show that both groups made considerable amount of spatial error in their initial response, but autistic children were less willing to change their response style in their second attempt. Considering other studies which revealed that geometric repetition also occurred in children with learning difficulties (Meyer & Hobson, 2004) and its occurrence declined with age in typically developing children of 3-6 years old (Ohta et al., 1987), it might indicate that learning disability play a role in the imitation style of autistic children.

The current study has found the score of VPT and imitation to be correlated. Both of them also inversely correlated with the occurrence of reversal error. These findings may have implications for the mechanisms of imitative impairment in autism. According to intersubjectivity theories of autism, such as the self-other mapping theory (Rogers & Pennington, 1991) and identification theory (Hobson & Meyer, 2006), the imitative impairment in autism is rooted in difficulties in coordinating the representations of self and other, i.e., the ability to “see the other as a template of the self” or to “identify with the other.” These theories explain reversal error nicely because this type of error appears to be a specific problem related to an understanding of the self-other relationship (Meyer & Hobson, 2004; Smith & Bryson, 1998), as opposed to general motivational, perceptual, or executive problems during imitation.

The current finding of a correlation between reversal error and VPT scores further supports intersubjectivity theories. The VPT and gesture imitation tasks differ in perceptual and motor requirements, but they share a common requirement to coordinate perspectives. In both tasks, children sat face to face with the tester/puppet and focused on an object (gesture/picture) placed between them. To succeed, the children needed to assume that the tester/puppet seated across from them was an agent just like themselves with a distinct perspective. In other words, children need to attribute perspectives to both the puppet in the VPT task (“he would see an upright face facing towards him”) and the tester in the imitation task (“he held his palm facing inwards towards himself”). If the children failed to perceive the other as having a distinct perspective, they would probably use their own viewpoint to assume what the other would see, and at the same time perceive the demonstrator’s gestures as having no personal context. The majority of our autism sample displayed such failure to appreciate the other’s perspective, which fits in with intersubjectivity theories in explaining imitative and social-cognitive impairment in autism.
This study has a number of methodological limitations. Due to limited access to standard assessment tools in mainland China, the diagnoses of autism in this study were based on the checklist of symptoms from the DSM-IV, therefore lean heavily on the clinical experiences of paediatricians in daily clinical routine. We were not able to recruit a more stringent criterion for making such a diagnosis. Further study adopting more stringent diagnostic criteria according to the international standard should be conducted in the near future to cross validate the current findings.

The autism group and control group were only matched on their VMA, whereas their CA and visual-motor ability differed significantly. These differences may cause group asymmetry in many aspects, including life experience, education, etc. Therefore, the current finding is insufficient to lead to strong conclusion that the VPT and imitation impairment are specific to autism as opposed to other kinds of developmental delay. A better design would include an additional group of children with learning difficulties or a mental handicap, at the same time having similar CA and VMA with the autistic children. However, it should be noted that VMA provides a conservative estimate of mental age in an autistic sample (Reed & Peterson, 1990). Because we matched the two groups on VMA, the autism group was much older than the control group, and possessed better visual-motor skills. Therefore the lower VPT and imitation scores of the autism group is unlikely to be attributable to mental retardation or poor motor skills.

The sample size was relatively small, with only 15 participants in each group, and the sample is rather heterogeneous as the variation in VMA is large. Nevertheless, we observed consistent differences between the two groups, which resulted in acceptable statistical power despite of the adverse effect of small sample size. A replication with large sample size is needed in the future to confirm the current findings.

Notwithstanding these limitations, the current findings suggest the deficiency in Level-2 VPT in children with autism may be related to their imitative impairment, which supports intersubjectivity theories. These results point to the possibility that autistic children’s impairments in social cognition and imitation may be rooted in the lack of awareness to see other people “as a template of the self” (Rogers & Pennington, 1991). If future studies could confirm the current findings, several practical implications may be derived. First, the fact that children with autism were more prone to reversal errors and that the number of reversal errors correlated with VPT and imitation performance may contribute to the identification of autism. Since imitation starts in early infancy (Meltzoff & Gopnik, 1993), reversal error in imitation or other kinds of turn-taking play may serve as an indictor in the screening of infants at risk for autism and in the early diagnose of autism. Second, the current results also suggested that fostering intersubjective engagement in children with autism might enhance both their imitation skills and their ability to appreciate events in the world from multiple points of view. Further studies are needed to explore the underlying mechanisms that cause VPT and imitation impairments in children with autism, and to examine the feasibility of improving these skills through fostering intersubjectivity experience.

5. Conclusion

This study compared the performance of 15 children with autism and 15 matched typically developing children on a Level-2 VPT task and a gesture imitation task. The children with autism performed worse in both tasks compared to control group, and the scores for the two
tests correlate with each other and with the amount of reversal error during imitation. These findings suggest a problem in coordinating the perspective of self and others underlies both the imitation and VPT impairments of children with autism.

6. Acknowledgment

This study was supported by a grant from the Natural Science Foundation of Beijing (7072036) and the NSFC (30770728, 30970907) to Dr. Yanjie Su and grants from the Project-Oriented Hundred Talents Programme (O7CX031003) and the Knowledge Innovation Project of the Chinese Academy of Sciences (KSC X2-YW-R-131 & KSCX2-EW-J-8) to Dr. Raymond Chan. We are also indebted to the children and their parents and the teachers of the Haidian Peizhi Central School for the Mentally Retarded and the Preschool Center affiliated to Peking University.

7. References


Applied Psychology Center of Beijing Normal University. (1986). Wechsler Pre-school and Primary Scale of Intelligence [Chinese ed.] [in Chinese], Beijing Normal University Press, Beijing, China


The Relationship Between Visual Perspective Taking and Imitation Impairments in Children with Autism


The aim of the book is to serve for clinical, practical, basic and scholarly practices. In twenty-five chapters it covers the most important topics related to Autism Spectrum Disorders in the efficient way and aims to be useful for health professionals in training or clinicians seeking an update. Different people with autism can have very different symptoms. Autism is considered to be a spectrum disorder, a group of disorders with similar features. Some people may experience merely mild disturbances, while the others have very serious symptoms. This book is aimed to be used as a textbook for child and adolescent psychiatry fellowship training and will serve as a reference for practicing psychologists, child and adolescent psychiatrists, general psychiatrists, pediatricians, child neurologists, nurses, social workers and family physicians. A free access to the full-text electronic version of the book via Intech reading platform at http://www.intechweb.org is a great bonus.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: