Facilitating joint attention with salient pointing in interactions involving children with autism spectrum disorder

Katja Dindar, University of Eastern Finland, katja.dindar@uef.fi, University of Eastern Finland, Joensuu Campus, School of Educational Sciences and Psychology, P.O. BOX 111, 80101 Joensuu, Finland

Terhi Korkiakangas, University College London, t.korkiakangas@ucl.ac.uk, UCL Institute of Education, University College London, UCL Knowledge Lab, 23-29 Emerald Street, London WC1N 3QS, UK

Aarno Laitila, University of Eastern Finland, Joensuu Campus, School of Educational Sciences and Psychology. Current contact information: aarno.a.laitila@jyu.fi, University of Jyväskylä, Department of Psychology, P.O. Box 35, FIN-40014 University of Jyväskylä,

Eija Kärnä, University of Eastern Finland, eija.karna@uef.fi, University of Eastern Finland, Joensuu Campus, School of Educational Sciences and Psychology, P.O. BOX 111, 80101 Joensuu, Finland

Abstract

Children with autism spectrum disorder (ASD) reportedly have difficulties in responding to bids for joint attention, notably in following pointing gestures. Previous studies have predominantly built on structured observation measures and predefined coding categories to measure children's responsiveness to gestures. However, how these gestures are designed and what detailed interactional work they can accomplish have received less attention. In this paper, we use a multimodal approach to conversation analysis (CA) to investigate how educators design their use of pointing in interactions involving school-aged children with ASD or autistic features. The analysis shows that pointing had specific sequential implications for the children beyond mere attention sharing. Occasionally, the co-occurring talk and pointing led to ambiguities when a child was interpreting their interactional connotations, specifically when the pointing gesture lacked salience. The study demonstrates that the CA approach can increase understanding of how to facilitate the establishment of joint attention.

Keywords: pointing gestures, joint attention, autism spectrum disorder, conversation analysis

Biographical notes

Katja Dindar is currently working toward her PhD in psychology at the University of Eastern Finland. Her thesis focuses on interactional and attentional skills of children with autism spectrum disorder (ASD) during educational interactions with adult co-participants. In her work, she uses multimodally informed CA to examine video recordings and live eye-tracking data.

Terhi Korkiakangas is a social interaction researcher. Her research interests include the use of talk, gaze and body movement in clinical and professional interactions and in interactions involving children with ASD. She played a key role in an ESRC-funded project on teamwork in operating theatres (2012–2013) and was a research fellow in MODE, a node of the ESRC's National Centre for Research Methods (2013–2104). She is currently a British Academy Postdoctoral Fellow (2014–2017), conducting video-based research on communication in operating theatres. Her PhD research examined interactions involving children with ASD using a multimodal approach to CA.

Aarno Laitila currently works as a senior lecturer in psychology at the University of Jyväskylä. His main area of research is the field of psychotherapy, especially in the context of family therapy, and developing therapeutic expertise.

Eija Kärnä works as a professor in special education at the University of Eastern Finland. She has worked in several international research and development projects and

conducted multidisciplinary research with researchers from several fields of science, such as linguistics, psychology, nursing science and computer science. Her research interests are inclusive learning environments, technology for individuals with special needs and communication and interaction of individuals with severe developmental disabilities and ASD.

Facilitating joint attention with salient pointing in interactions involving children with autism spectrum disorder

Abstract

Children with autism spectrum disorder (ASD) reportedly have difficulties in responding to bids for joint attention, notably in following pointing gestures. Previous studies have predominantly built on structured observation measures and predefined coding categories to measure children's responsiveness to gestures. However, how these gestures are designed and what detailed interactional work they can accomplish have received less attention. In this paper, we use a multimodal approach to conversation analysis (CA) to investigate how educators design their use of pointing in interactions involving school-aged children with ASD or autistic features. The analysis shows that pointing had specific sequential implications for the children beyond mere attention sharing. Occasionally, the co-occurring talk and pointing led to ambiguities when a child was interpreting their interactional connotations, specifically when the pointing gesture lacked salience. The study demonstrates that the CA approach can increase understanding of how to facilitate the establishment of joint attention.

Keywords: pointing gestures, joint attention, autism spectrum disorder, conversation analysis

Introduction

Children's ability to follow pointing gestures has been taken as an indicator of their capacity for joint attention. This refers to the ability to share attention and experiences with other people, which constitutes a cornerstone of human development and learning. Joint attention involves the coordination of attention to an object of mutual interest (e.g., Bakeman & Adamson, 1984) or, specifically, the shared 'knowing' between persons that they are attending to the same concern (e.g., Hobson, 2005; Tomasello, 1995). Children with typical development usually learn to initiate or respond to bids for joint attention between the ages of 8 and 15 months (Bakeman & Adamson, 1984; Jones, Carr, & Feeley, 2006). However, children with autism spectrum disorder (ASD), a condition characterised by impairments in communication and social interaction, have apparent difficulties in joint attention (e.g., American Psychiatric Association, 2013; Meindl & Cannella-Malone, 2011).

Joint attention skills have been studied experimentally by measuring children's head turns or gaze shifts towards the location pointed to (see e.g., Clifford & Dissanayake, 2009; Falck-Ytter, Fernell, Lundholm Hedvall, von Hofsten, & Gillberg, 2012; Mundy et al., 2003; Presmanes, Walden, Stone, & Yoder, 2007). In an early study by Baron-Cohen (1989), declarative and imperative pointing were distinguished based on the location pointed to: near the child but far from the tester (imperative function: requesting an object) or out of view of the child (declarative function: sharing a view/an object). Children with ASD tend not to show major difficulties with imperative gestures that serve instrumental aims of obtaining something, e.g., a toy or sweets, with the help of another person (cf. Broekhof et al., 2015). Instead, research suggests that these children have specific difficulties with declarative gestures used for purely social

purposes (e.g., Baron-Cohen, 1989; Bruinsma, Koegel, & Koegel, 2004; Maljaars, Noens, Jansen, Scholte, & Berckelaer-Onnes, 2011; Wetherby, Watt, Morgan, & Shumway, 2007), which might underpin the broader impairments in social interaction.

Ability to follow pointing gestures constitutes a key component of many assessment protocols for ASD, such as the widely used Early Social Communication Scales (ESCS; Mundy et al., 2003) and Autism Diagnostic Observation Schedule (Lord, Rutter, DiLavore, & Risi, 2002). Children's responsiveness to (semi) structured events is usually codified using predefined categories outlined in the observation schedules. For instance, the ESCS pointing trial focuses on whether a child responds to the pointing gesture and name calling by turning his or her head to the location pointed to by the tester. However, the implementation of a coding scheme necessitates the abstraction of these responses from other contextual details, notably the interaction at hand or the child's simultaneous activities. The objective of the experimental approaches to sustain the stability of trials across children builds on an assumption that the assessment interactions can be reproduced in a constant manner over time. This begs the question of how well such structured interactions resemble more spontaneous real-life interactions where talk and gestures are used as part of mundane activities. ASD studies are increasingly examining naturally unfolding interactions such as those in home video recordings (e.g., Clifford & Dissanayake, 2009; Maestro et al., 2005). Yet the codifying of naturalistic data rarely captures the activities of the people interacting with these children. More dynamic approaches could explore gesturing in the streams of naturally occurring interaction and how gestures are designed and used.

Psychological and interactional approaches to gestures

A considerable literature on the psychology of language suggests that gesture and speech together provide an index of mental processing. One substantial argument proposes that gestures and speech are underpinned by a shared psychological structure (e.g., Ekman & Friesen, 1969; Graham & Argyle, 1975; Kendon, 1972, 1980, 2004; McNeill, 1985, 1992). The work of McNeill has been at the core of research on linking gestures and speech as an inseparable unit, suggesting that gestures constitute an integral part of speech rather than an accompaniment to it. For instance, as gestures can depict what is said in order to complement or clarify speech, they can indicate the structure of human cognition as essentially *verbal*: 'sentences and gestures develop internally together as psychological performances' (McNeill, 1985, p. 350). A comprehensive review is beyond the scope of this study, but some evidence that has been used to suggest this unity includes the mutual disruption of gesture and speech in aphasia, a condition that affects the production or understanding of speech following damage to the brain (McNeill, 1985). On the other hand, Goodwin (2004, 2014) has demonstrated how a man who was left with an extremely limited three-word vocabulary because of a severe stroke was nonetheless able to use others' talk and his own gesturing to position himself as a competent interactant. Thus, approaches that locate their analysis in the psychology of an individual seem limited in their accounts for gesturing.

Psychological research on gestures has rarely delved into the dynamic aspects of gesture use. Berger and Rae (2012) have noted that much of McNeill's work draws on narrative elicitation tasks that are conducted in laboratory settings and are thus 'highly constrained by the situation of their production' (p. 1822). Meanwhile, in research

examining naturally occurring interactions, the interest has often been on the physical form of the gestures and their functions in relation to speaker's spoken discourse. For instance, Kendon's work on pointing has focused on the link between the different hand shapes used in pointing, e.g., index finger, thumb and open hand, and the use speakers make in their discourse of the objects referred to when pointing at them. For example, when an object is pointed at to identify it, an extended index finger may be used. On the other hand, if it is pointed at as an exemplar of a category, an open hand may be used (Kendon, 2004: Chapter 11). However, research has also increasingly considered the interactional context of the gesture use and 'the roles they play *for the participants*' (Streeck, 1993, p. 276, emphasis in original), regardless of their physical form (see e.g., Dickerson, Stribling, & Rae, 2007). Likewise, in matters of joint attention, rather than asking whether a child turns to follow a pointing gesture in experimental tasks, we might consider how the person gesturing designs the gesture, how the person assesses the child's response to it and what occurs in the interaction when the gesture is produced.

Murphy (2003) has argued that focusing on gestures outside of interactions can gravely limit our understanding of gesturing as a socially situated activity. As he puts it,

> Examining gestures as communicative tools that people use in interaction, rather than as overly simplified "windows" affording a glimpse inside the brain, we can see that "what matters" is always dependent on the context of a gesture's occurrence, that is to say, "what matters" depends not only on a gesture's placement in the stream of talk, but also on who performs it, who is around to see it, where they are, what types of activities they are involved in, what sorts of

> things are around to help shape the gesture, and an infinite number of other possible contingencies. (Murphy, 2003, p. 30)

The details of naturally occurring interactions are thus important in understanding even such a seemingly simple activity as pointing. Kidwell and Zimmerman (2007) have proposed that research on joint attention has rarely problematised the ways in which gestures are used. Such oversight might simplify the semiotic complexity, structural diversity and socially situated nature of gesturing (e.g., Goodwin, 1986, 2003, 2014; Haviland, 2004; Hindmarsh & Heath, 2000; Mondada, 2014; Streeck, 2009). For instance, gestures that might 'look alike' can undertake different kinds of interactional work depending on how they map onto the unfolding stream of action. Kidwell (2005) has demonstrated this issue well with her study on gaze. She showed how, in a nursery, very young children who are engaged in misbehaviour such as pushing or hitting another child can differentiate the consequences of a caregiver's gazing actions. A 'mere look', i.e., a glance darted in the direction of the child while the adult continues other activities, is not treated as implicating an intervention, and the child usually continues what he or she is doing. However, 'the look', a sustained gaze directed at the child as the caregiver stops other activities in progress, implicates a possible sanction, and the child usually stops the misbehaviour in response to such a gaze. A child can read the consequentiality of 'the look' through *how* the adult looks at the child, with the salience of the gaze emerging as significant.

Kidwell's analysis, like a large strand of social interaction research drawing on the framework of conversation analysis (CA), uses detailed transcriptions of the video-

recorded activities of all parties to demonstrate the interactional work that bodily action and talk can accomplish. This approach has the potential to broaden the research on joint attention, including gesture following, in children with ASD. For instance, psychological research suggests that children with ASD fail to monitor adults' communication channel (Adamson, McArthur, Markov, Dunbar, & Bakeman, 2001; McArthur & Adamson, 1996); thus, they might find gestures presented in silence more difficult to interpret and respond to than gestures accompanied with attention-directing or eliciting verbalisations, e.g., prompting to 'look' or calling the child's name (Leekam, Hunnisett, & Moore, 1998; Presmanes et al., 2007). Hence, co-occurring talk might have some significance for how 'readable' gestures are and what the expected response from the child might be. However, that something has been said or pointed at does not necessarily give away how one is expected to respond. Seemingly simple pointing gestures do not always have a transparent meaning, and one particular gesture can have a range of different potential referents (Goodwin, 2014). Furthermore, the fact that recipients rarely follow a pointing gesture when the referent is not visually accessible or when the gesturer does not point in the actual direction of the referent suggests that pointing gestures are not always designed to be followed (Schegloff, 1984). Previous psychological research has rarely tapped on the issue of how responding to pointing or joint attention bids more broadly gets built and negotiated between the adults and children (cf. Adamson et al., 2001; McArthur & Adamson, 1996).

Other CA-oriented research on gestures in various institutional settings (e.g., Goodwin, 1986, 2003; Hindmarsh & Heath, 2000; Mondada, 2007; Streeck, 2009) and everyday settings across cultures (e.g., Enfield, Kita, & de Ruiter, 2007; Goodwin,

2014; Haviland, 2004; Hayashi, 2005; Mondada, 2014) have further problematised the use of gestures for straightforward 'referential work'. Drawing on workplace interactions, Hindmarsh and Heath (2000) have shown that participants draw upon the activities in progress, not just verbal deixis or gestures, to understand what is being referred to. Speakers use entire body movements—gaze direction, leaning in—in addition to their pointing hand to accomplish referential work. Hindmarsh and Heath have noted that while gestures are often taken to support or clarify deictic references, they can be produced *after* the body movements that project that there is something to be attended to: 'talk reflexively works on behalf of the gesture' (p. 1864).

CA offers a useful frame for considering the organisation of social interactions: how participants order their interactions through turn-taking and the production of sequences of action. One basic sequential structure ordering interactions is the *adjacency pair*, such as a question and an answer. The sequence-initiating action (a question) makes the responsive action (an answer) relevant in the next turn (Schegloff, 2007); should the responsive action be delayed or absent, participants usually orient to some form of interactional trouble in the production of a response. The next turn in a sequence thus provides an important analysis of participants' orientations and understandings of the matters at hand (see *next-turn proof procedure* in Sacks, Schegloff, & Jefferson, 1974). A multimodal approach to CA enables us to consider how bodily action, including the use of gestures, might relate to such interactional concerns. In relation to responsiveness in ASD, one concern might be how co-participants' gestures that accompany questions or instructions might prompt the children to produce a responsive action: how a gesture maps onto on-going action or deictic references and how salient the body movements are. Such features can indicate

why securing joint attention can sometimes prove challenging and other times succeed.

A growing literature has utilised the CA approach to examine interactions involving individuals with ASD in educational and clinical contexts (e.g., Barrow & Tarplee, 1999; Dickerson et al., 2007; Dindar, Korkiakangas, Laitila, & Kärnä, 2016; Dobbinson, Perkins, & Boucher, 1998; Korkiakangas, Dindar, Laitila, & Kärnä, 2016; Korkiakangas & Rae, 2013, 2014; Maynard, 2005; Muskett, Perkins, Clegg, & Body, 2010; Sterponi & Fasulo, 2010). However, to our knowledge, there are no CA studies that examine specifically how pointing gestures are used in interactions with children with ASD, although responsiveness to pointing gestures is extensively described in psychological literature in relation to joint attention (albeit in experimental settings). Thus, the present paper focuses on video-recorded educational interactions between educators, i.e., school teachers and special needs assistants, and children with ASD or autistic features. In these interactions, the children use a computer application, and the educators sit next to them, providing instructions when needed. One practical challenge for the educators seems to be the transitions between different tasks and options, particularly how to best facilitate joint attention and direct a child's attention to the task-relevant objects. The educators routinely direct the children in and through talk and gesture. We examine the design of these gestures and how the children respond to them.

Method

Data

The data used in this study consist of approximately 168 minutes of video taken from a larger corpus of interactions between school-aged children and their educators. The children have a diagnosis of ASD or a combination of autistic features and a comorbid

diagnosis, e.g., intellectual disability. The recordings took place during technology-enhanced activity sessions organised by the Children with Autism Spectrum disorders as Creative Actors in a strength-based Technology-enhanced learning Environment (CASCATE) project at the children's school between 2011 and 2014. During these weekly sessions, the children worked together with familiar educators on various educational tasks. The video material for this study involved two boys (whom we call 'Antti' and 'Otto') and was recorded with two tripod-mounted digital cameras. The educators were asked to act as they would in any everyday school setting and to instruct or help the children whenever they wanted or felt the need to do so.

The data for this study were recorded during a specific LEGO® constructing activity. In this activity, the children use plastic LEGO bricks to assemble LEGO models presented on the touch screen. This requires the children to shift attention between various locations and objects, such as between virtual objects on the computer screen and plastic bricks on the table. The activity lasts approximately 10 minutes, depending on the willingness of the children to work. During such an activity, pointing is a potential resource to manage children's participation, to prompt them to focus on certain aspects of the task and to assist them with task transitions. This particular activity was chosen because the preliminary observations indicated that it provided natural opportunities for the educators to perform pointing gestures. As one of the children, Otto, has motor restlessness and was rarely able to sustain his attention for longer periods during the constructing activity, his video recordings are limited to 50 minutes; for Antti, the video material is 118 minutes long. The pictures illustrating the LEGO building activity are presented in Figure 1. The computer application used in the activity was developed in the CASCATE project.

Figure 1. The LEGO® constructing activity

Participants

In the fragments presented in this paper, Antti interacts with his teacher, ‘Kirsi’, and with special needs assistant, ‘Niina’, whereas Otto interacts with his special needs assistant, ‘Helena’. Antti was 13 years and two months and Otto was eight when the data collection for this study began. Antti and Otto were chosen due to their reportedly severe interactional or attentional difficulties. Their clinical reports were available to the researchers. Otto has a main diagnosis of ASD (based on ICD-10 criteria), whereas Antti’s latest diagnosis states the presence of intellectual disability with autistic features. The children also have a variety of other health-related diagnoses that are not specified here for the sake of the children’s anonymity. Both Otto and Antti can be described as ‘autism plus’, i.e., autism with comorbid diagnoses (Gillberg & Fernell, 2014). These comorbidities were often given diagnostic priority in the past; thus, autistic features may have been diagnosed or indicated as ‘additional information’ (see Gillberg & Fernell, 2014), which seems to be the case with Antti. The children’s

diagnostic and other clinical information are presented in Table 1.

Table 1. Antti's and Otto's diagnostic and other clinical information based on their clinical reports

Antti	
Main diagnosis	Severe intellectual disability (F79.1; ICD-10 criteria) with autistic features
Rehabilitation history	Physiotherapy, music therapy, occupational therapy and pivotal response training
Characteristics	Antti's report describes the presence of autistic features. He has motor difficulties, which make his movements appear stiff, including his ability to handle objects. Since childhood, Antti has had occasional grabbing seizures that have made him hurt himself and other people. His ability to maintain social contact varies, and he is easily distracted. He often 'falls into' his thoughts and produces repetitive utterances from a children's cartoon show. Antti has to be verbally instructed to reorient his attention, but he occasionally manages to do so without external instruction.
Otto	
Main diagnosis	Childhood autism (F84.0; ICD-10 criteria)
Rehabilitation history	Music therapy, pivotal response training, occupational therapy, speech therapy
Characteristics	Otto's report states that he enjoys interacting with others, particularly adults. He has difficulties with conceptual language but is able to understand simple and clear instructions. Otto communicates with speech, gesturing, pointing and visual picture cards. Otto also has sensory sensitivities and motor restlessness but is occasionally able to concentrate on tasks when motivated. However, he needs adults' help and instructions in daily activities.

Finnish versions of the Autism Spectrum Screening Questionnaire (ASSQ), the Social Communication Questionnaire (SCQ) and the Strengths and Difficulties Questionnaire (SDQ) were used to gather more information on the children's characteristics (see Table 2). They reached the cut-off scores in all the measurements.

Table 2. Antti's and Otto's scores in ASSQ, SCQ and SDQ

Measurement	Antti's scores	Otto's scores
ASSQ (teacher ratings used)[a]	41	23
SCQ (lifetime version, parent ratings used)[b]	24	21
Subcategories		
Social interaction	6	7[c]
Communication	10	9
Stereotypical behaviour	7	4
Other	1	1
SDQ (parent ratings used)[d]	20	24

Notes: [a] Cut-off score is 22
[b] Cut-off score is 15
[c] One question left unanswered
[d] Cut-off score for 'abnormal' is 16

Antti, Otto, Kirsi, Niina and Helena are all native Finnish speakers. Written consent to participate was obtained from the children's guardians and the educators. Children's willingness to participate was monitored throughout the sessions. All the names of people and places have been changed to prevent the participants from being recognised. The study has been assessed and approved by the ethics committee of the researchers' home institution.

Analytic approach

The study draws on a multimodal approach to CA to analyse the organisation of social interactions. This approach involves a detailed analysis of video-recorded data that are carefully transcribed. Instead of starting off with context-independent, predefined categories for talk, gestures, and other bodily actions, CA seeks to examine actions

structurally as they occur in interactions. While pointing gestures have been commonly investigated in relation to co-occurring talk (e.g., McNeill, 1992), we employ a broader, multimodal perspective (see Stivers & Sidnell, 2005) to consider how participants may draw on different interactional resources, e.g., the use of gaze and body movements, to accomplish social actions.

The analysis began with identifying the pointing gestures performed by the educators. Reliability for this initial procedure was determined by having two researchers working through data from Antti independently. Reliability was considered good based on the intra-class correlation (ICC) of .934, supporting the move for a more detailed analysis of the gestures. The gesture phases were annotated by paying attention to the stroke, hold and withdrawal phases (see Kendon, 2004, pp. 111-113). The movement towards the peak of the gesture is referred to as the stroke. This is the phase where the expressive work of the gesture is usually accomplished (Kendon, 2004, pp. 111-113; refer to the appendix for the illustration of the phases used in the transcripts). Annotations were made frame by frame to determine the timings accurately. We used the ELAN multimedia annotator, developed by the Max Planck Institute for Psycholinguistics, to annotate the data. The initial annotations were reconstructed following the Jeffersonian transcription notations commonly used in CA (see Atkinson & Heritage, 1984). The talk in Finnish was idiomatically translated into English using bold typeface (the children's talk was not always grammatically or lexically correct, resulting in 'rough' translations). The transcription of eye gaze followed the notations developed by Goodwin (1981; see the appendix). Throughout the fragments presented in this paper, the children and the educators sit next to each other, with the educators on the children's left side, both facing the computer screen that provides visual instructions

for the LEGO construction.

Analysis

The analysis considers interactions in which Antti and Otto are using the LEGO application. One concern for the educators facilitating these situations is to direct the children to proceed to a new phase in the LEGO construction or to a new task altogether. Occasionally, when one task has ended, the children might remain engaged with the LEGO bricks rather than return to the screen to select a new model or task. The educators prompt the children to proceed in different ways: through talk and gesture or through gesture alone. Occasionally, the educators succeed in directing a child to progress in the task; sometimes, further work is required. We will consider how gestures occasionally become an issue for both the educators and the children. The analysis particularly focuses on two kinds of instances where 1) the interpretation of co-occurring talk and gesture involves some ambiguity and 2) the pointing gesture is not saliently produced for the child.

Ambiguity in the interpretation of co-occurring talk and gesture

We will begin by considering the implications of the pointing gestures used by an educator, Niina, in the stream of interaction with Antti. It might not always be clear from a gesture alone what interactional work it has been designed to accomplish. Further, talk and pointing might not always occur ‘congruently’ as a gesture might or might not refer directly to the talk produced. In such instances, some ambiguity might develop in interpreting the interactional work underway. We begin to examine this issue with Fragment 1, in which Antti has to move on to the next phase in assembling a

construction. He is holding a LEGO construction in his left hand, which is lifted off the table and close to Niina's face. Antti has brought out an image of a LEGO model on the screen, and we join in as Niina directs Antti to proceed to the next model to continue the construction.

Fragment 1

N = Niina; A = Antti

```
                  ((points towards the model on the screen
                  then withdraws hand))
                       .-----.....................
                  screen        Antti              screen
                  ------------..X________________,,-------
1   N:            nyt sulla on jo (.) molemmat siinä
                  now you already have both there

    A:            screen (bottom part)   model pointed at
                  -----------------...------,,,,,,,,,
```

```
                   ((puts hand
                   into the            ((strokes hand
                   brick               to point towards
                   basket))            the screen))
                   .--------.           .-
                   basket              screen
2   A:             ----------...-----
                  (- - - - - - - - - 1)

    N:             .-----.............
                  ((points towards the bottom part
                  of the screen then withdraws hand))
```

```
                  ((points at the bricks that
                  Antti is holding))

                        -----.............

                  bricks in Antti's
                  hand                screen
                  ------------...---------
3   N:            sulla on jo kato molemmat
                  look you already have both
    A:            /.....-------------------
                  ((strokes to touch the model on the
                  screen, withdraws, strokes to touch
                  the navigation button, holds point))

                  screen
                  --------------------------

4   N:            joo
                  yes
```

Antti uses a navigation arrow to bring out an image on the screen. In line 1, Niina says *nyt sulla on jo (.) molemmat siinä* (‘now you already have both there’). Rather than merely acknowledging that Antti already has both bricks in place, as shown in the image, Niina elicits a move forward so that Antti can bring up another image on the screen. A new image would show what kind of additional piece could be added to the construction Antti is holding. This would require Antti to navigate to the next image, yet Niina does not explicitly instruct him to do so. As she speaks, Niina moves her hand towards the model on the screen, with the peak of this move appearing on the lexical item *jo* (‘already’). This gesture occurs in the visual field of Antti, as he is gazing at the image pointed to. Niina leans in conspicuously to monitor Antti’s response, suggesting that he should take some action and move the task forward.

Niina's prompt to Antti implies that he should touch the screen, yet it is a 'declaratively' formatted comment about the construction. The action-implicative nature of her talk is evident from her gesture and posture shift to monitor how Antti responds. However, there is some ambiguity in the prompt in terms of how Antti is to interpret what to do next: To navigate to the screen, Antti should touch the arrow in the bottom-right corner (over which his right hand is hovering), not the image that Niina has pointed to. Yet, her gesture points at the image (the direct 'referent' of her talk), which does not offer an option to navigate (the action pursued). There is thus ambiguity in how to interpret the interactional work of Niina's co-occurring talk and pointing.

Antti responds by dropping his gaze and withdrawing his hand from the navigation arrow, engaging with the basket in front of him (lines 1-2), presumably to proceed by selecting a new LEGO piece. As soon as Antti's hand goes in the basket, Niina points again. She orients to Antti's engagement with the basket as not being pertinent to her prompt and now points towards the navigation arrow in the bottom-right corner of the screen (Figure 2).

Figure 2. Niina points at the navigation arrow and Antti touches the screen

Here, Niina gestures to prompt Antti to proceed, having sequential implications on what Antti should do next. Antti reorients to the screen, and as Niina withdraws her hand, Antti lifts his hand to touch the image on the screen. Antti's conduct is responsive to the pointing gesture, yet he touches the wrong part of the screen in response. Nonetheless, in doing so, Antti responds to both instances of gesturing as a prompt for action—either by engaging with the LEGOs or touching the image on the screen. The timing of his responses right after the pointing indicates that Antti has produced an analysis of the prompt indicating when to proceed.

Pointing here projects what is being made sequentially relevant to the child. In the context of the task transition, then, a gesture can serve as a prompt to go ahead without necessarily providing a clear referential index of exactly 'where' action should be taken. In response to Niina's second gesture, Antti orients beyond the matters of reference and attention sharing by touching the screen.

Fragment 2 illustrates the parallel case of Otto's orientation to the sequential implication of Helena's pointing gesture. We join in as Otto and Helena are about to start a new construction. A virtual 3D construction model on the screen is represented from an angle that does not allow easy visual access to the placement of the bricks in it. Thus, Helena begins to turn the model around using her index finger.

Fragment 2

H = Helena; O = Otto

		((turns the virtual 3D LEGO model using index finger))
1	H:	katotaas (.) mistä alkaa (.) mikä väri
		let's see (.) where it begins (.) which colour

((yawns))

2 O: musta
black

3 H: nii onko iso vai pieni
yes is it big or small

((points at a virtual brick touched by Helena earlier))

4 O: iso
big

5 H: *((grabs Otto's pointing hand and moves it away))*

((points at the lowermost brick in the virtual construction))

6 H: pieni
small

O: *brick basket*

((puts hand in the basket))

screen

7 O: (---------1)

screen

8 O: iso
big

((places a small black brick in front of Otto))

9 H: tässä
here

O: *brick*

((points at the second lowermost brick in the virtual construction))

10 H: sitten mikä on seuraa[vaksi?

then which is next

bricks *screen*
----------..--------------

O: [iso
big
((points again at the previous virtual brick in the model))

((leans towards Otto))
11 H: *((holds the pointing gesture))*
(---------1)

12 O: telt

((withdraws the gesture)) *((places a yellow brick in front of Otto))*
13 H: keltainen (.) siihen päälle
yellow (.) there on top

As Helena is turning the model, she prompts Otto, *katotaas (.) mistä alkaa (.) mikä väri* (‘let’s see (.) where it begins (.) which colour’; line 1). While Helena is turning the model around, her finger is repeatedly touching one particular brick (see Figure 3). What is Helena referring to when she says *mistä alkaa* (‘where it begins’): the brick she is touching (yet apparently not pointing at) with her finger or the lowermost brick, i.e. the first brick of the virtual construction?

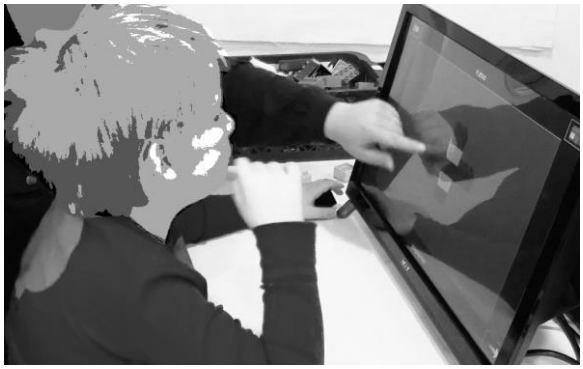

Figure 3. Helena turns the virtual model around

Otto correctly names the colour of the brick being touched by Helena (line 2), which is indeed black in both possible cases. Helena then goes on to ask about the size of the brick (line 3), which gives away that Helena and Otto do not have mutual understanding of which brick is in question: Otto responds by pointing at the brick that Helena touched when she was turning the model around (in Figure 3) and correctly names the size of that brick, *iso* (‘big’; line 4). However, Helena was evidently referring to the lowermost brick of the construction. She attempts to correct the misunderstanding by gently pushing Otto’s hand away (line 5) and pointing at the lowermost brick while giving the correct answer, *pieni* (‘small’; line 6; see Figure 4).

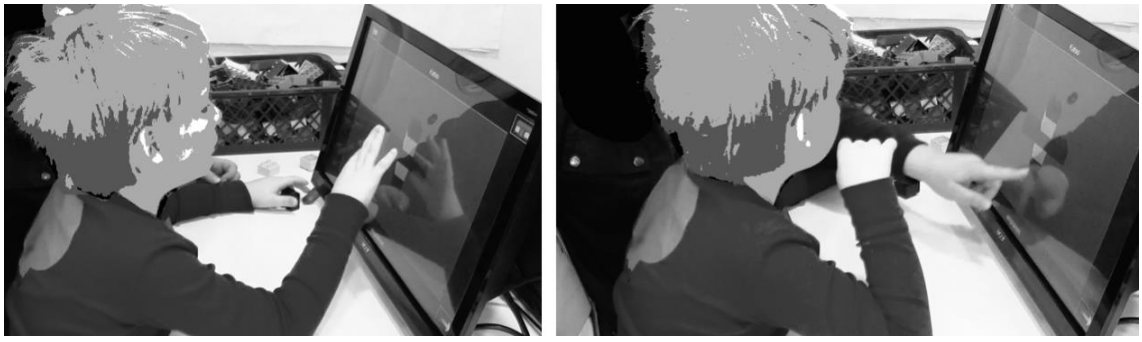

Figure 4. Otto points at a brick and states its size, then Helena manages the misunderstanding and indicates her focus of attention by pointing

However, Otto simultaneously shifts his attention to the brick basket (line 6); thus, he probably does not notice Helena's pointing gesture. While this shift in attention is risking the establishment of joint attention, it also demonstrates that Otto understands the sequential implications of Helena's prompts: she is not merely prompting him to name the colour or the size of the virtual bricks; rather, these prompts inexplicitly refer to the task of finding matching bricks among the plastic ones—exactly to what Otto shows orientation by shifting his attention to the brick basket. Helena's verbal correction, however, redirects Otto's attention to the screen. He looks at the screen for a moment (line 7) and then repeats his answer, *iso* ('big'), suggesting that he is rejecting Helena's correction (line 8), perhaps because he did not observe Helena's pointing gesture that could have solved the misunderstanding.

Next, Helena attempts to move the task forward by giving Otto the small plastic brick and prompting him to focus on the next brick, this time by pointing at the target referred to in her talk (the second lowermost brick in the construction) to reduce the likelihood of yet another misunderstanding. Nevertheless, Otto treats it problematic that Helena has not accepted his previous answer. He points again towards the brick that she had initially touched when turning the construction around and repeatedly states its size, *iso* ('big'; line 10). This time Helena does not say anything but instead orients to the salience of her gesture as an apparent trouble. She continues pointing towards the other brick and slightly leans in towards Otto so that her arm touches his, giving him a tactile sensation that draws further attention to her gesture. Her adjusted body position makes the gesture increasingly visible to Otto (see Figure 5).

Figure 5. Helena and Otto display their conflicting understanding of the referred brick, which Helena deals with by designing her gesture more visible to Otto

Helena does not release her gesture until Otto begins to produce an answer, *telt* (line 12), which Helena takes as an attempt to produce the correct answer, *keltainen* (‘yellow’), allowing them to then move on to placing the correct brick into the construction (line 13). In this interaction, Helena’s talk and the manner in which she touched a brick (gesture which resembled pointing) resulted in a negotiation over which brick should be in their shared focus of attention. Thus, the co-occurring talk and gesture can result in ambiguity concerning their interpretation. Throughout the interaction, rather than treating Otto’s incorrect answers as his inability to understand the question, Helena oriented to the salience of her gestures as the main concern that could be remedied by designing her gesture more accessible to Otto. We now consider the problem of salience as gestures are produced, in more detail.

Pointing and the problem of salience

While pointing can lead to children’s action on the screen or on the bricks in the basket at sequentially relevant junctures, the prerequisite for such accomplishment is for the gesture to be fully visible. We begin by considering an interaction between Otto and Helena that takes place later during the same session as in Fragment 2. In Fragment 3,

Otto is handling a brick and gazing away from the screen, seemingly disengaged from the task.

Fragment 3

H = Helena; O = Otto

1	H:	*((takes the brick from Otto's hand))*
2	O:	*((inaudible talk; changes his body position so as to leave the chair))*
3	H:	*((gently grabs Otto's hand and pulls him back))* minkäslainen pala siel on **what kind of a brick there is**
	O:	*shifts gaze from the right to the left of Helena* . . . . .
4	H:	*((points at the lowermost brick of the construction on the screen))* ensimmäisenä **the first one**
	O:	*shifts gaze from the right to the left of Helena* . . . . .
5	H:	tuolla **there**
	O:	screen .-----

6 (.9)

7 O: *((inaudible talk; nods))*

8 H: *((points at the lowermost brick of the construction))*
mikä väri.
which colour

O: *screen*
---

9 O: *screen bricks*
----...
sininen
blue

H: ———
((holds the pointing gesture))

10 H: *((holds the pointing gesture))*
nii: onko iso vai pieni
yes is it big or small

O: *screen*
---

11 O: *screen bricks*
------..
i- pieni
b- small

H: ———
((holds the pointing gesture))

((withdraws the pointing gesture))
. . .
12 H: mmh:

13 H: *((reaches for the brick basket))*
pi[e-
sma-

O: **[**↑iso

big

14 H: *((pushes the brick basket close to Otto))*
pieni ja si[ninen ensin
small and blue first
O: [iso
big

15 O: *((grabs a big blue brick))*

16 H: nii se on iso
yes that is big

17 H: *((moves the basket slightly))*
mut kat[soppa
but look
O: [((inaudible talk))

18 O: *((reaches for the basket))*
bricks
...---------------,,
(---------1---------2)
H: *((points at the small blue brick))*

19 O: *((reaches and grasps the the small blue brick))*
(---------1---------2)
H: *((attempts to grasp the brick))*

20 H: hyvä
good

In line 1, Helena attempts to re-engage Otto with the task by taking the brick (that is the wrong colour for the present construction) from his hand. Otto says something (inaudible from the video recording) and attempts to leave the chair and end

the task. Helena then gently takes his arm and pulls him back, thus redirecting his attention and creating a 'window of opportunity' to prompt him to focus on the screen (line 3). She points at the lowermost brick of the construction while asking Otto to name its colour (line 4). Otto orients to her prompt by shifting his gaze, but instead of looking at the location pointed to, he shifts his gaze to the left of Helena, which likely hinders him from seeing Helena's gesture. Helena then withdraws her gesture as 'redundant' since Otto may have not seen it. Indeed, Otto does not respond until the withdrawal of Helena's gesture apparently catches his attention, and he shifts his gaze to the screen. Helena uses this opportunity to launch a deictic item, *tuolla* ('there'), to direct Otto's attention. Otto looks at the screen silently for a while until, in line 7, he nods and says something (inaudible from the video recording), which shows Helena his trouble in understanding the referent of her deictic item. The visibility of the gesture has become an issue (Figure 6).

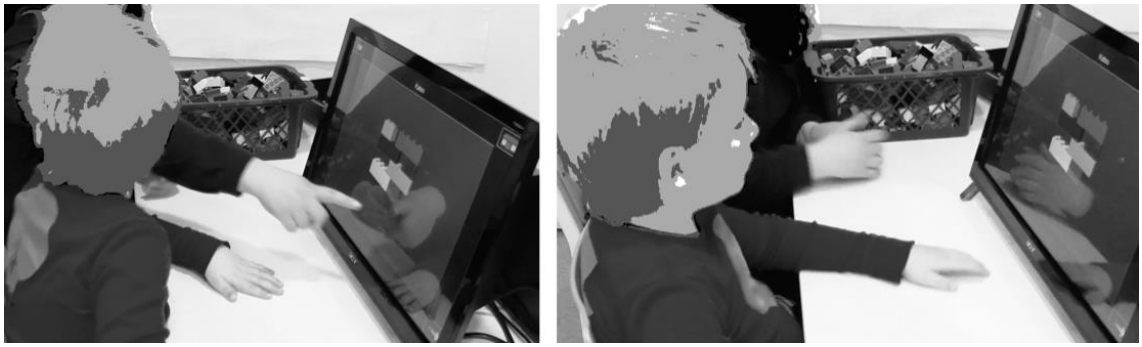

Figure 6. Otto does not orient to Helena's pointing gesture, suggesting the referent of Helena's deictic item, *tuolla* ('there'), is ambiguous

Helena treats Otto's difficulty partially as a referential but also as a sequential matter: without seeing the gesture, he might not be able to disambiguate where to orient

in order to respond. Helena repeats her pointing gesture towards the lowermost brick while reissuing her prompt, now with Otto's full attention (line 8). She continues to hold her gesture, ensuring that they share a mutual focus of attention, while asking Otto about the colour and size of the brick (Figure 7). Here, Otto again shows his understanding of what is implied, i.e., choosing the correct bricks from the basket, and while answering Helena's questions, he shifts his gaze towards the brick basket to orient to locating the correct brick (lines 9 and 11).

Figure 7. Helena secures a mutual focus of attention by holding her pointing gesture on the screen

Helena holds her gesture until she accepts Otto's answers with *mmh:* (line 12). She then reaches for the brick basket while repeating Otto's answer, when he suddenly overlaps her talk by saying *iso* ('big'), which is an incorrect answer (line 13). Helena goes on to push the brick basket closer to Otto and attempts to correct him by saying *pieni ja sininen ensin* ('small and blue first'). Otto, however, repeats *iso* ('big'; line 14) and grasps a big blue brick from the basket (line 15), suggesting that he is aware of what is requested from him but wishes to pick another kind of brick. By first acknowledging his choice (line 16), Helena then attempts to bring Otto's focus back to

the task at hand by slightly moving the brick basket and saying *mut katsoppa* (‘but look’; line 17). This both demonstrates dispreference for Otto’s selection and makes the brick basket again relevant for the next action. Otto responds by shifting his attention to the basket and reaching for the bricks, which is when Helena uses her index finger to perform a subtle pointing gesture and taps the target brick, i.e., a small blue brick, to help Otto locate it (see Figure 8). Otto ‘reads’ Helena’s gesture within the unfolding course of activity and in light of the prior actions (see Berger & Rae, 2012). Here, the visually salient pointing gesture was produced without talk and was sufficient in securing Otto’s response to the brick. In response, Otto grasps the brick (line 19), demonstrating his understanding of the changing sequential implications of Helena’s pointing: instead of asking him to name the colour or size of the brick, Helena’s pointing requests him to physically grasp it. Helena then accepts his action with praise, *hyvä* (‘good’) (line 20).

Figure 8. Helena points at the correct brick to help Otto locate it

Thus, visual salience is a significant design feature to which both the educator and the child orient. In a parallel example, the pointing gesture initially fails to be seen, and the educator engages in specific work to highlight its salience while sustaining the

gesture. In Fragment 4, just before Kirsi prompts Antti to make a selection on the screen, Antti has dropped his gaze to the basket and picked up some bricks. That Antti's orientation is already directed downwards and away from the screen poses initial challenges for any gesturing to be seen.

Fragment 4

K = Kirsi; A = Antti

```
              [    ((strokes hand towards the screen))
                        ----
1    K:       [ ja sitte
                **and then**
                screen   drops to basket
     A:       [ -------,,

               ((holds the pointing gesture))
                _____
2    K:        (- - - - - - - - - 1)

              [((holds the pointing gesture))
                ______
                 Antti
                .x_____
3    K:       [ valinta
                selection
                basket
     A:       [ -------

              [    ((bends upper body closer to the screen))
                    ______________________________
```

((points towards the screen then withdraws))

4 K: haluatko (.) rakentaa (.) tommosen
do you want to construct one like that

A: *basket screen*

((strokes hand towards the screen then withdraws))

5 K: joo:?
yes

As Antti is gazing down at the bricks, Kirsi produces a continuer, *ja sitte* (‘and then’; line 1), accompanied by a stroke towards the screen. Two options are available on the screen, and Kirsi’s index finger points at ‘figure models’. Kirsi attempts to re-engage Antti in the task by gesturing. She keeps her arm extended near Antti’s line of gaze for maximum visibility (lines 2-3) while uttering an explicit prompt, *valinta* (‘selection’; line 3). The sequential implications of her verbal prompt are clear—Antti should select an option next—yet her gesture is not visible to Antti. This could result in ambiguity in terms of where Antti ought to make a selection: in the basket or on the screen. In response to Kirsi’s prompt, Antti subtly leans more into the basket, presumably to focus on selecting some of the bricks. Kirsi glances at Antti to monitor his response; on seeing that Antti has moved closer to the basket, Kirsi swiftly continues, *haluatko (.) rakentaa (.) tommosen* (‘do you want to construct one like that’; line 4). Rather than producing a question to be answered, Kirsi leads Antti to proceed with a selection on the screen just at the moment when he might start picking up bricks from the basket. Simultaneously with her talk, Kirsi leans her body closer and brings

her arm fully in front of the screen, with her index finger producing a 'beat' in front of the 'figure models' (line 4). Here, Kirsi has changed her strategy to engage Antti.

Her actions escalate in directing Antti's attention to a relevant concern when her initial prompt 'fails' to gain the sought response (see Korkiakangas & Rae, 2013). Kirsi's pointing, which accompanies her speech, has developed into a visually salient gesture designed to interrupt Antti's engagement with the bricks – an activity that Kirsi does not consider pertinent to moving the task forward (see Figure 9). The combination of her leaning in, her suggestion, *haluatko (.) rakentaa (.) tommosen* ('do you want to construct one like that'; line 4) and her salient point at the *screen* facilitate Antti shifting his attention to the 'figure models' and producing a selection on the screen.

Figure 9. The gesture develops into a salient point at 'figure models'

Fragment 4 shows how Kirsi orients to the visibility of her gesture as hindering the establishment of joint attention on the screen as Antti's gaze was directed at the bricks in the basket. As Antti moved his hands to the bricks to select from the basket, Kirsi launched a new verbal prompt and redesigned her gesture by leaning in closer and sustaining her finger pointing at the screen to mark that the issue has not yet been

resolved (Sikveland & Ogden, 2012). Only after Antti reoriented and touched the screen did Kirsi withdraw her gesture. In the current fragment, Kirsi undertakes additional work in redirecting Antti away from the bricks in front of him and to the screen. The trajectory of Kirsi's gesture increases in visual salience through her postural alignment and her concurrent speech, which secure Antti's attention as a prerequisite for moving the task forward.

In the context of these task-related interactions, the observed pointing gesture has been taken as a prompt for action, marking the juncture for a responsive action. The issue of establishing a shared focus of attention that might occur from not seeing the gesture has not been treated only as a problem for joint attention *per se* but as a problem that delays the progression of the task at hand. The strategy to alleviate such trouble is to redo the gesture in full view of the child.

Discussion

This study examined how educators use pointing gestures in interactions with children who have either an ASD diagnosis or autistic features. Much of the previous research on joint attention has been criticised for overlooking how gestures are used as part of interactional situations (see Kidwell & Zimmerman, 2007), raising concerns over their complex nature (e.g., Goodwin, 2003; Streeck, 2009). The analysis showed that the educators used pointing to move the tasks forward, but this work was not always accomplished without issues, including how the gestures were produced. Occasionally joint attention was facilitated using pointing concurrently with speech and postural alignment, which also enabled a pointing gesture to become more salient for the child. The pointing gestures had specific sequential implications beyond mere attention

sharing and made different responses relevant to the children: responding through the touchscreen, naming the characteristics of the virtual bricks on the screen, or picking up actual plastic LEGO bricks. Moreover, the gestures could sometimes prompt the children's response without any accompanying talk yet become understandable in reference to prior actions (see also Berger & Rae, 2012).

As pointing gestures can have multiple candidate referents (see Goodwin, 2014), their interpretation can become complex. For instance, Goodwin (2014) demonstrated how a man who suffered a stroke and was left with extremely limited speech could use gesturing (including pointing) to communicate complex matters. In such interactions, the co-participants often produced candidate understandings of the possible meanings of his gesturing. In the current study, the educators were fully capable of communicating verbally, yet their gesturing posed occasional challenges for the children due to difficulties in establishing mutual focus of attention, which was commonly caused by the children attending to task-irrelevant concerns and thus missing out on the gestural prompts. This finding is in line with the observations of Adamson et al. (2001) and McArthur and Adamson (1996) of children with ASD not attending to the communication channel of their co-participants. In their studies, the adults attempted to manage their 'challenging partners' using more literal acts, e.g., banging objects, to attract their attention. However, in our study, rather than treating children's conduct as problematic, the educators oriented to the salience of their gestures as troublesome and revised the delivery of their gestures. Goodwin (2014) has suggested that such cooperative practices through which the sequential implications of gesturing become understandable should be paid more analytical attention than the underlying psychological processes. The current study suggests that the educators' careful gestural

redesigns play an important role in successfully managing the children's attention and moving the tasks forward, thus securing the establishment of joint attention.

The study also raises a concern that any approach that views gestures as giving access to an individual's intrapsychological processes (e.g., McNeill, 1985) may increase the tendency to consider a lack of responsiveness as an intrapsychological incompetence related to ASD, rather than as a production of the specific interactions taking place. Our analysis demonstrates that it would be troublesome to locate children's responding difficulties as a mere ASD-related pathology when the ways in which the gestures were produced play such a crucial role for both the children and the educators. In a similar vein, Dickerson et al. (2007) have noted that to understand interactional competence in ASD, interactions themselves should become the focus of study rather than individuals with ASD and their actions in isolation from their interactional contexts. This also means giving space to consider more 'atypical' gestures or actions. In their study on mundane tapping actions, i.e., tapping on a board or flashcard, Dickerson et al. (2007) showed that the co-participants treated children's tapping as an expressively communicative gesture rather than a stereotypical, repetitive action symptomatic of ASD. They argued that instead of paying too much attention to the physical form of such actions, the sequential placement of these actions should be considered the key for understanding their interactional relevance. This resonates with the present study in that the pointing gestures 'looked alike' in terms of their form, yet their placement with talk varied, occasionally creating complex situations when talk and gesture co-occurred.

Previous studies examining the role of talk during gesturing have mainly considered how its occurrence facilitates the children's responsiveness (e.g., Leekam et

al., 1998; Presmanes et al., 2007) rather than examining whether talk and gesture might undertake differing interactional work. For instance, the educators' gestures did not always directly refer to their talk and thus made it ambiguous for the children *where* their responsive actions should be taken, e.g., with respect to a navigation button on the screen. Nevertheless, the sequential analysis showed that even though the responses were occasionally incorrect per se, they occurred in sequentially relevant junctures. Future studies could also benefit from detailed sequential examinations on the role of talk in potentially facilitating the children's responsiveness to gestures.

While standardised assessment protocols examine children's responsiveness to (semi) structured events using predefined coding categories, new kinds of competencies can be identified in naturalistic interactions when children demonstrate their understanding of the sequentially implicative nature of gestures and other actions. For instance, Korkiakangas and Rae (2013) have shown how following teachers' gaze, body, and object movements can become a natural test of competency for children with ASD – for example, whether children shift their attention to relevant concerns during educational tasks. Although research is increasingly conducted with recordings of naturally occurring interaction (e.g., Clifford & Dissanayake, 2009; Maestro et al., 2005), the analyses often code the gestures as separate analysis units for the purpose of quantification. Whereas coding schemes require the omission of contextual details for coding 'correct' responses, sequential examinations enable the recognition of naturally emerging interactional competence. Rather than enabling a 'context-independent' view on responding to gestures, the omission of contextual aspects can place severe limits on how the interaction can be understood. If the environment or the activity at hand is not

included in the analysis, vital aspects of the interaction become inaccessible (Goodwin, 2014).

Furthermore, sequential examinations provide insights into what could be considered or coded as a 'correct' response by investigating how the participants orient to these concerns. A particular challenge in the use of predefined coding schemes is deciding on such coding definitions a priori. For instance, studies often record whether children shift their gaze to the location pointed to, which is commonly treated as an interactionally relevant response in joint attention paradigms. Yet as the current study has shown, this is not always the case, as pointing is often used to request actions beyond mere gaze shifts.

This study included only educators' pointing gestures performed with an index finger or using an open hand (although no open-hand pointing was observed). However, it should be noted that there are several other ways to point, such as using the head or eyes (e.g., Goodwin, 2003; Kendon, 2004). We also looked for pointing using objects, but there were no such instances in our data set. The different ways of pointing should be considered in future studies, as there is evidence that the different hand configurations in pointing carry out different actions (Kendon, 2004). In addition, the analysis suggests that educators' gesture use includes a continuous monitoring of the success of their delivery, leading to altering and reissuing the 'failed' prompts. This contrasts with many of the more experimental environments or assessment settings, where keeping the prompts stable across children is prioritised. Research could benefit from focusing on how these children interpret the changing sequential implications of a gesture produced. Moreover, studies using natural settings have the potential to inform the use of communicative strategies, such as how the educators could facilitate the

establishment of joint attention with children with ASD by designing their gestures more salient. Since psychological research discusses impairments not only in how these children respond to gestures but also in how they produce them (Maljaars et al., 2011), interactional examinations could provide valuable detail on how and when these (in)competencies emerge.

Acknowledgements

The authors are grateful to the participants for allowing the video recording and analyses of their interactions. This work was supported by the Academy of Finland [140450], Alfred Kordelin Foundation and Olvi Foundation.

References

Adamson, Lauren B., Duncan McArthur, Yana Markov, Barbara Dunbar, & Roger Bakeman (2001). Autism and joint attention: Young children's responses to maternal bids. *Journal of Applied Developmental Psychology, 22* (4), 439-453.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Atkinson, J. Maxwell & John Heritage (1984). Transcript notation. In J. Maxwell Atkinson & John Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. ix-xvi). Cambridge, MA: Cambridge University Press.

Bakeman, Roger & Lauren B. Adamson (1984). Coordinating attention to people and objects in mother–infant and peer–infant interaction. *Child Development, 55*, 1278-1289.

Baron-Cohen, Simon (1989). Perceptual role taking and protodeclarative pointing in autism. *British Journal of Developmental Psychology, 7* (2), 113-127.

Barrow, Emma & Clare Tarplee (1999). Delayed echoing as an interactional resource: A case study of a 3-year-old child on the autistic spectrum. *Clinical Linguistics & Phonetics, 13* (6), 449-482.

Berger, Israel & John Rae (2012). Some uses of gestural responsive actions. *Journal of Pragmatics, 44* (13), 1821-1835.

Broekhof, Evelien, Lizet Ketelaar, Lex Stockmann, Annette van Zijp, Marieke G. N. Bos, & Carolien Rieffe (2015). The understanding of intentions, desires and beliefs in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45* (7), 2035-2045.

Bruinsma, Yvonne, Robert L. Koegel, & Lynn Kern Koegel (2004). Joint attention and children with autism: A review of the literature. *Mental Retardation and Developmental Disabilities Research Reviews, 10* (3), 169-175.

Clifford, Sally & Cheryl Dissanayake (2009). Dyadic and triadic behaviours in infancy as precursors to later social responsiveness in young children with autistic disorder. *Journal of Autism and Developmental Disorders, 39* (10), 1369-1380.

Dickerson, Paul, Penny Stribling, & John Rae (2007). Tapping into interaction: How children with autistic spectrum disorders design and place tapping in relation to activities in progress. *Gesture, 7* (3), 271-303.

Dindar, Katja, Terhi Korkiakangas, Aarno Laitila, & Eija Kärnä (2016). Building mutual understanding: How children with autism spectrum disorder manage interactional trouble. *Journal of Interactional Research in Communication Disorders, 7* (1), 49-77.

Dobbinson, Sushie, Michael R. Perkins, & Jill Boucher (1998). Structural patterns in conversations with a woman who has autism. *Journal of Communication Disorders, 31* (2), 113-134.

Ekman, Paul & Wallace V. Friesen (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1* (1), 49-98.

Enfield, Nick J., Sotaro Kita, & Jan Peter De Ruiter (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics, 39* (10), 1722-1741.

Falck-Ytter, Terje, Elisabeth Fernell, Åsa Lundholm Hedvall, Claes von Hofsten, & Christopher Gillberg (2012). Gaze performance in children with autism spectrum disorder when observing communicative actions. *Journal of Autism and Developmental Disorders, 42* (10), 2236-2245.

Gillberg, Christopher & Elisabeth Fernell (2014). Autism plus versus autism pure. *Journal of Autism and Developmental Disorders, 44* (12), 3274-3276.

Goodwin, Charles (1981). *Conversational organisation: Interaction between speakers and hearers*. New York: Academic Press.

Goodwin, Charles (1986). Gestures as a resource for the organization of mutual orientation. *Semiotica, 62* (1-2), 29-50.

Goodwin, Charles (2003). Pointing as situated practice. In Sotaro Kita (Ed.), *Pointing: Where language, culture and cognition meet* (pp. 217-241). Mahwah, New Jersey: Erlbaum.

Goodwin, Charles (2004). A competent speaker who can't speak: The social life of aphasia. *Journal of Linguistic Anthropology, 14* (2), 151-170.

Goodwin, Charles (2014). The intelligibility of gesture within a framework of co-operative action. In Mandana Seyfeddinipur & Marianne Gullberg (Eds.), *From gesture in conversation to visible action as utterance: Essays in honor of Adam Kendon* (pp. 199-219). Amsterdam: John Benjamins.

Graham, Jean Ann, & Michael Argyle (1975). A cross-cultural study of the communication of extra-verbal meaning by gesture. *International Journal of Psychology, 10* (1), 57-67.

Haviland, John B. (2004). Gesture. In Alessandro Duranti (Ed.), *A companion to linguistic anthropology* (pp. 197-221). Malden, MA: Blackwell.

Hayashi, Makoto (2005). Joint turn construction through language and the body: Notes on embodiment in coordinated participation in situated activities. *Semiotica, 156*, 21-53.

Hindmarsh, Jon & Christian Heath (2000). Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics, 32* (12), 1855-1878.

Hobson, R. Peter (2005). What puts the jointness into joint attention? In Naomi Eilan, Christoph Hoerl, Teresa McCormack, & Johannes Roessler (Eds.), *Joint attention: Communication and other minds. Issues in philosophy and psychology* (pp. 185-204). New York: Clarendon Press/Oxford University Press.

Jones, Emily A., Edward G. Carr, & Kathleen M. Feeley (2006). Multiple effects of joint attention intervention for children with autism. *Behavior Modification, 30* (6), 782-834.

Kendon, Adam (1972). Some relationships between body motion and speech. An analysis of an example. In Aaron Siegman & Benjamin Pope (Eds.), *Studies in dyadic communication* (pp. 177-210). Elmsford, New York: Pergamon Press.

Kendon, Adam (1980). Gesticulation and speech: Two aspects of the process of utterance. In Mary Ritchie Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207-227). The Hague: Mouton and Co.

Kendon, Adam (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kidwell, Mardi (2005). Gaze as social control: How very young children differentiate "the look" from a "mere look" by their adult caregivers. *Research on Language and Social Interaction, 38* (4), 417-449.

Kidwell, Mardi & Don H. Zimmerman (2007). Joint attention as action. *Journal of Pragmatics, 39* (3), 592-611.

Korkiakangas, Terhi & John Rae (2013). Gearing up to a new activity: How teachers use object adjustments to manage the attention of children with autism. *Augmentative and Alternative Communication, 29* (1), 83-103.

Korkiakangas, Terhi & John Rae (2014). The interactional use of eye-gaze in children with autism spectrum disorders. *Interaction Studies, 15* (2), 233-259.

Korkiakangas, Terhi, Katja Dindar, Aarno Laitila, & Eija Kärnä (2016). The Sally–Anne test: an interactional analysis of a dyadic assessment. *International Journal of Language and Communication Disorders*. Advance online publication. doi: 10.1111/1460-6984.12240.

Leekam, Susan R., Emma Hunnisett, & Chris Moore (1998). Targets and cues: Gaze-following in children with autism. *Journal of Child Psychology and Psychiatry, 39* (7), 951-962.

Lord, Catherine, Michael Rutter, Pamela C. DiLavore, & Susan Risi (2002). *Autism diagnostic observation schedule (ADOS).* Los Angeles: Western Psychological Services.

Maestro, Sandra, Filippo Muratori, Maria Cristina Cavallaro, Chiara Pecini, Alessia Cesari, Antonella Paziente, Daniel Stern, Bernard Golse, & Francisco Palacio-Espasa (2005). How young children treat objects and people: An empirical study of the first year of life in autism. *Child Psychiatry and Human Development, 35* (4), 383-396.

Maljaars, Jarymke, Ilse Noens, Rianne Jansen, Evert Scholte, & Ina van Berckelaer-Onnes (2011). Intentional communication in nonverbal and verbal low-functioning children with autism. *Journal of Communication Disorders, 44* (6), 601-614.

Maynard, Douglas W. (2005). Social actions, gestalt coherence, and designations of disability: Lessons from and about autism. *Social Problems, 52* (4), 499-524.

McArthur, Duncan & Lauren B. Adamson (1996). Joint attention in preverbal children: Autism and developmental language disorder. *Journal of Autism and Developmental Disorders, 26* (5), 481-496.

McNeill, David (1985). So you think gestures are nonverbal? *Psychological Review, 92* (3), 350-371.

McNeill, David (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

Meindl, James N. & Helen I. Cannella-Malone (2011). Initiating and responding to joint attention bids in children with autism: A review of the literature. *Research in Developmental Disabilities, 32* (5), 1441-1454.

Mondada, Lorenza (2007). Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies, 9* (2), 194-225.

Mondada, Lorenza (2014). Pointing, talk, and the bodies: Reference and joint attention as embodied interactional achievements. In Mandana Seyfeddinipur & Marianne Gullberg (Eds.), *From gesture in conversation to visible action as utterance: Essays in honor of Adam Kendon* (pp. 95-124). Amsterdam: John Benjamins.

Mundy, Peter, Christine Delgado, Jessica Block, Meg Venezia, Anne Hogan, & Jeffrey Seibert (2003). *Early social communication scales (ESCS).* Retrieved from http://healthsystem.ucdavis.edu/mindinstitute/ourteam/faculty_staff/ESCS.pdf

Murphy, Keith M. (2003). Building meaning in interaction: Rethinking gesture classifications. *Crossroads of Language, Interaction, and Culture, 5*, 29-47.

Muskett, Tom, Mick Perkins, Judy Clegg, & Richard Body (2010). Inflexibility as an interactional phenomenon: Using conversation analysis to re-examine a symptom of autism. *Clinical Linguistics & Phonetics, 24* (1), 1-16.

Presmanes, Alison G., Tedra A. Walden, Wendy L. Stone, & Paul J. Yoder (2007). Effects of different attentional cues on responding to joint attention in younger siblings of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37* (1), 133-144.

Sacks, Harvey, Emanuel A. Schegloff, & Gail Jefferson (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50* (4), 696-735.

Schegloff, Emanuel A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 266-296). Cambridge: Cambridge University Press.

Schegloff, Emanuel A. (2007). *Sequence organization in interaction: Volume 1: A primer in conversation analysis*. Cambridge: Cambridge University Press.

Sidnell, Jack (2005). Gesture in the pursuit and display of recognition: A Caribbean case study. *Semiotica, 156* (1/4), 55-87.

Sikveland, Rein Ove & Richard Ogden (2012). Holding gestures across turns: Moments to generate shared understanding. *Gesture, 12* (2), 166-199.

Sterponi, Laura & Alessandra Fasulo (2010). "How to go on": Intersubjectivity and progressivity in the communication of a child with autism. *Ethos, 38* (1), 116-142.

Stivers, Tanya & Jack Sidnell (2005). Introduction: Multimodal interaction. *Semiotica, 156* (1/4), 1-20.

Streeck, Jürgen (1993). Gesture as communication I: Its coordination with gaze and speech. *Communications Monographs, 60* (4), 275-299.

Streeck, Jürgen (2009). Forward-gesturing. *Discourse Processes, 46* (2-3), 161-179.

Tomasello, Michael (1995). Joint attention as social cognition. In Chris Moore & Philip J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103-130). Hillsdale, NJ: Erlbaum.

Wetherby, Amy M., Nola Watt, Lindee Morgan, & Stacy Shumway (2007). Social communication profiles of children with autism spectrum disorders late in the second year of life. *Journal of Autism and Developmental Disorders, 37* (5), 960-975.

Appendix

Transcription conventions for speech (according to Jefferson in Atkinson & Heritage, 1984)

Symbol	Meaning
. (period)	Preceding talk is falling, stopping
? (question mark)	Preceding talk is rising
: (colon)	Preceding sound is lengthened
- (dash)	Indicates a cut-off
↑ (up arrow)	Following talk goes up suddenly
(in parentheses)	Uncertain transcription
((word in double parentheses))	Transcription comment or non-vocal action
word (underlining)	Spoken with emphasis
(.)	Pause: cannot be timed (less than 0.2 seconds)
[text] [text]	Adjacent lines overlap

Transcription conventions for gaze (based on.Goodwin, 1981)

Symbol	Meaning
———	Continuous line indicates that a party is gazing towards the co-participant.
---	Cut line indicates that a party is gazing at an object or direction described above the cut line.
x	X marks the specific point where gaze reaches the co-participant or another specified target.
,,,	Commas indicate dropping or withdrawing gaze.
...	Dots mark the movement that brings a party's gaze towards the other.

Transcription conventions for pointing gestures

Symbol	Meaning
	Stroke phase = The movement towards the peak of the gesture
	Hold/post-stroke hold phase = The gesture may be sustained following its stroke
	Withdrawal phase = The gesture is withdrawn following its stroke phase or