

**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

Author(s): Dindar, Katja; Korhonen, Terhi; Laitila, Aarno; Kärrä, Eija

Title: An interactional 'live eye tracking' study in autism spectrum disorder : combining qualitative and quantitative approaches in the study of gaze

Year: 2017

Version:

Please cite the original version:

Dindar, K., Korhonen, T., Laitila, A., & Kärrä, E. (2017). An interactional 'live eye tracking' study in autism spectrum disorder : combining qualitative and quantitative approaches in the study of gaze. *Qualitative Research in Psychology*, 14(3), 239-265. <https://doi.org/10.1080/14780887.2017.1290174>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

An interactional ‘live eye tracking’ study in autism spectrum disorder: combining qualitative and quantitative approaches in the study of gaze

Katja Dindar^a, Terhi Korhonen^b, Aarno Laitila^{a,c} and Eija Kärnä^a

^aUniversity of Eastern Finland, Joensuu, Finland; ^bUniversity College London, London, UK;

^cUniversity of Jyväskylä, Jyväskylä, Finland

CONTACT Katja Dindar, katja.dindar@uef.fi, School of Educational Sciences and Psychology, University of Eastern Finland, Yliopistokatu 2, P.O. Box 111, FI-80101 Joensuu, Finland.

Color versions of one or more figures in the article can be found online at www.tandfonline.com/uqrp.

Abstract

Recent studies on gaze behaviours in individuals with autism spectrum disorder (ASD) have utilised ‘live eye tracking’. Such research has focused on generating quantitative eye tracking measurements, which provide limited (if any) qualitative contextual details of the actual interactions in which gaze occurs. This article presents a novel methodological approach that combines live eye tracking with qualitative interaction analysis, multimodally informed *conversation analysis*. Drawing on eye tracking *and* wide-angle video recordings, this combination renders visible some of the functions, or what gaze ‘does’, in interactional situations. The participants include three children with ASD and their adult co-participants during body-movement gaming sessions. The article demonstrates how quantitative eye tracking study can be extended with qualitative microanalytic interaction analysis to recontextualise the gaze shifts identified. The findings in this article show that the co-participants treat a child’s gaze

shifts differently depending on *when* these occur in a stream of other action. The study suggests that introducing this qualitative dimension to eye tracking research could increase its ecological validity and offer new insight into gaze behaviours in ASD.

Keywords

autism spectrum disorder, gaze shifts, functions of gaze, live eye tracking, conversation analysis

1 Introduction

Recent years have seen an increase in the use of eye tracking technologies to examine gaze behaviours in autism spectrum disorder (ASD). Gaze behaviours in ASD have been often described as ‘atypical’, marked by reduced attention to socially salient information (e.g. von Hofsten et al. 2009). Gaze behaviours have been also proposed to underpin difficulties in communication and social interaction associated with ASD. The present study sets out to respond to some of the challenges that have emerged in this primarily quantitative research on gaze which tends to focus on looking time measurements to predefined areas of interest even in naturalistic communicative situations (e.g. Falck-Ytter, 2015). As will be discussed below, gaze-related research can benefit from combining eye tracking with qualitative video-based approach to social interaction to better capture what gaze is used for in the situations in which it occurs.

Eye tracking has been predominantly used to study whether individuals with ASD gaze at socially salient information, including other people, faces and the eye region. While ample research has been generated, the overall evidence is inconclusive. Some studies have shown that

compared to typically developing individuals, individuals with ASD look at non-social objects more than socially salient stimuli (Falck-Ytter & von Hofsten, 2011; Guillon et al. 2014; von Hofsten et al. 2009), have impairments in using gaze to orient to others' communicative cues (Chawarska et al. 2012; Falck-Ytter et al. 2012; Fletcher-Watson et al. 2009; Gillespie-Lynch et al. 2013), fail to predict social events (Ruffman et al. 2001; Senju & Johnson, 2009; von Hofsten et al. 2009), and gaze more frequently at the mouth rather than the eye region of faces (e.g. Hanley et al. 2015; Klin et al. 2002). However, contradictory findings draw a complex picture. For instance, Falck-Ytter (2010) found similar goal-directed eye movements in children with ASD and in typically developing individuals when observing actions performed by other people. Other studies have failed to find increased looking durations at the mouth region or decreased looking durations at the eye area in individuals with ASD (e.g. Dapretto et al. 2006; see Falck-Ytter & von Hofsten, 2011 for a review).

One explanation for such variation in research findings could be the type of stimuli used in eye tracking (Chevallier et al. 2015; Falck-Ytter & von Hofsten, 2011; Freeth et al. 2013; Risko et al. 2012). Videos can provide more dynamic and richer information on social events than static images (Falck-Ytter & von Hofsten, 2011; cf. Falkmer et al. 2011). However, videos can be also problematic insofar as the scenes may not realistically represent everyday situations, and thus the stimuli might not be ecologically valid (Norbury et al. 2009). Video stimuli can also raise concerns beyond the type of scenes viewed. Falck-Ytter and von Hofsten (2011) have stressed that the viewing of videos places participants in the role of a passive receiver of social information and lacks *interactional* opportunities for participation with other people (Falck-Ytter & von Hofsten, 2011; Gobel et al. 2015; Guillon et al. 2014). Recent developments have

attempted to overcome some of these limitations using face-to-face situations in which eye movements are measured using mobile equipment; to date, only a handful of such live eye tracking studies exist (Falck-Ytter, 2015; Falck-Ytter et al. 2015; Freeth et al. 2013; Hanley et al. 2014; Hanley et al. 2015; Magrelli et al. 2013; Nadig et al. 2010; Noris et al. 2012; Thorup et al. 2016; Vabalas & Freeth, 2016).

Live eye tracking can better capture how gaze is used in realistic situations. For instance, some studies of children with ASD have shown a reduced tendency to look at an adult's face during a storytelling situation but not in a cognitive testing situation with an experimenter (Falck-Ytter, 2015; Falck-Ytter et al. 2015). Such evidence suggests that children with ASD might not have a fundamental deficit in their social use of gaze, but that the degree to which gaze is used varies in different contexts. This also links with the notion of ASD related impairments in *declarative* interactions (i.e. the goal is in social sharing) rather than in *imperative* interactions (i.e. attention is shared for instrumental purposes, such as to receive assistance) (e.g. Camaioni et al. 2003; Maljaars et al. 2011; cf. Broekhof et al. 2015).

That gaze behaviours seem to vary in different contexts however calls into question what gaze is used for in the first place: what an instance of gaze 'does' when the eyes are directed to another person in a social situation. Eye tracking research has been mostly concerned with studying 'where' individuals look (the areas of interest [AOIs]) and 'when' individuals look at the AOI, that is, focusing on the timing of gaze. Less attention has been paid to such timing of gaze in naturalistic interactions, which might provide important insight into the 'social' connotations of different gaze behaviours.

1.1 'Where' and 'when' gaze occurs

Quantitatively oriented eye tracking researchers, including Falck-Ytter et al. (2013), have warned that recording overall 'looking times' at a target might be insufficient when measuring competency in the social use of gaze. The argument stands that the attentional differences between children with ASD and those with typical development might be 'related to the exact timing of their eye movements, rather than the distribution of gaze over long periods of time' (Falck-Ytter et al. 2013, pp. 2250). For instance, Hanley et al. (2014) found that children with ASD were slower to gaze at an adult after an unexpected event to check their awareness of the situation. Thus, such *event-related* gaze measurements might be more informative about what makes some instances of gazing 'social'.

However, even in the most novel live eye tracking studies, the actual interactions have not received detailed attention. As Nadig et al. (2010, 12) have put it, 'the time spent looking at a region does not necessarily capture the function of visual attention' – that is, what kinds of functions gaze serves cannot be reduced merely to their quantity. To date, perhaps closest to a naturalistic examination about timing of gaze has been the study by Falck-Ytter (2015), which measured the target and the timing of gaze in a face-to-face situation: as an adult (experimenter) was reading a story to a child. While the findings indicated that the gaze of children with ASD was less frequently aligned with the experimenter's gaze (unlike in interactions with typically developing children), only limited information was available about the stream of action as the alignment happened or was expected to happen. Such details can be crucial in identifying the functions of gaze at a particular moment.

Even where the gaze shifts have been examined temporally - that is, showing when (relative to some aspect of interaction or stimuli) children look at an experimenter's facial area - the exact contextual details, including the experimenter's actions, have not been fully explored. Such examination would require at least two methodological considerations to complement the quantified eye tracking measurements: the use of wide-angle video recordings showing the contributions of all participants in an interaction, and a qualitative interactional framework for a structural analysis of these contributions.

1.2 Interactional approaches to gaze

To date, the social scientific approaches to social interaction have been underutilised in eye tracking research. In experimental social psychological research, gaze measurements have been traditionally taken to indicate emotional states and interpersonal attitudes. For instance, mutual gaze has been associated with intimacy between interactants (e.g. Argyle & Dean, 1965), with increased eye contact correlating with interpersonal attraction (Exline & Winters, 1965). However, more naturalistic examinations have considered what gaze accomplishes as part of social interaction, rather than how it reflects on interpersonal psychology. One such approach, *conversation analysis* (CA), is gaining popularity also in the field of ASD research (e.g. O'Reilly et al. 2016). This framework draws on audio-visual recordings of social interactions in naturalistic everyday and institutional settings, which are carefully transcribed for qualitative structural analysis of the organisation of interactions. Such microanalytic research elaborates on sequences of actions through which participants organise their conduct structurally and collaboratively. In short, this means that an action by one party can be responsive to another

party's action; thus, an action can place specific constraints on others to produce a response at a particular moment (Schegloff, 2007). For example, in spoken interactions, a question (i.e. *an initiating action*) projects an answer (i.e. *a responsive action*) to be a relevant contribution from the recipient in the next turn. The responsive actions then provide evidence of the participants' understanding of what (if anything) is expected from them in the following turn (Sacks et al. 1974).

The extension of CA work that examines *multimodal* interaction (e.g. Korhonen & Rae, 2014; Mondada, 2014; Solomon et al. 2016; Stivers & Sidnell, 2005) focuses on the structural organisation of actions beyond speech, including gaze and gesture. Multimodally informed CA has provided evidence for how 'atypical' means of interacting can constitute significant resources for individuals with ASD to engage in social situations. For example, repetitive tapping movements (Dickerson et al. 2007), 'problematic' behaviours (Damico & Nelson, 2005), and 'inflexibility' (Muskett et al., 2010), which could all be glossed as symptomatic behaviours inherent to ASD, can be shown to have interactional relevance when investigated in the context of their occurrence. In order to render visible such relevancies, a multimodal approach to CA does not rely on coding of behaviours *a priori*, but rather examines how individuals *use* resources, including gaze, through video recordings and detailed transcriptions of these recordings. The prior multimodal work around gaze suggests that interactants do not engage in continuous mutual gaze with one another. Rather, gaze towards another person can become relevant at particular moments (Goodwin, 1981; Stivers & Rossano, 2010) and accomplish specific actions: thus the 'function' of gaze varies. Table 1 summarises some of the key functions as generated in social scientific research of interpersonal situations.

[TABLE 1 NEAR HERE]

In the multimodal work of social interaction, gaze has been analysed as part of interactional face-to-face encounters. Stivers and Rossano (2010) have shown that speakers can mobilise a response from others by shifting their gaze to them, and that gaze alone can be sufficient to elicit the response; speech is not always required. Gaze can therefore constitute a social action in its own right and progress the organisation of an interactional exchange. Also young children have been shown to use gaze for interactional purposes, for instance, to appeal for assistance by shifting their gaze to their caregivers (Kidwell, 2009). This suggests that gaze can have multiple functions in interaction, which have to be identified not as decontextualized categories, but as part of other streams of activity.

The existing corpus of multimodally informed CA studies on gaze in ASD is still limited (Dickerson et al. 2005; Dickerson & Robins, 2015; Korkiakangas, 2011; Korkiakangas & Rae, 2014; Tuononen et al., 2016; Wiklund, 2012). However, the key message in these studies has been consistent: a decontextualised examination of gaze (whereby the occurrence of gaze is codified and quantified) cannot fully explain what gaze is used for, or what its significance is for the participants themselves (see Dickerson & Robins, 2015; Korkiakangas & Rae, 2014; Tuononen et al. 2016). For instance, that children with ASD have been shown to be capable for mobilising a response from their co-participants by turning to look at them after a turn-at-talk (Dickerson et al., 2005; Korkiakangas & Rae, 2014) could not be captured should gaze shifts alone be identified for an aggregated distributional analysis.

The placement of gaze in relation to other activities in progress, such as speech, is thus an important measure of the interactional work implicated by an instance of gaze. This requires careful identification of what other parties are doing in these interactions – the actions of the person who is shifting their gaze are to be examined *in relation to*, rather than in isolation from such contributions. However, that the prior interactional studies have based their judgements of eye movements solely on video recordings can pose limitations to the judgements made about those gaze behaviours. An eye tracking study has a greater potential to provide precise measurements of where and when gaze shifts occur. While eye tracking has been traditionally used for quantitative research on gaze, it has the potential to work in harmony with the qualitative multimodally informed CA. This combination of approaches is yet to be utilised in the field of ASD studies (we are only aware of a few studies on typically developing adults' gaze behaviours during conversations in laboratory setups; Auer, in press; Hirvenkari et al. 2013; Holler & Kendrick, 2015), and could provide a novel contribution to the 'event-related' measures of eye movements (Falck-Ytter et al. 2013).

In this article, we demonstrate how this methodological combination works in practice. We will focus on gaze shifts identified in the eye tracking data recorded with three children with ASD during educational gaming sessions. We will limit our focus on such gaze shifts that were directed from the game (a screen) towards other people in the room. First, we code the location and duration of the gaze shifts through eye tracking. We then draw on a multimodal approach to CA to examine these instances in the context of their occurrence, which involves the co-participants' responses to these gaze shifts. This qualitative microanalytic examination of the

previously quantified gaze measurements can render visible some of the different functions that looking at another person has in real-life situations.

2 Methods

2.1 Data

The data for this study includes wide-angle video recordings and live eye tracking data recorded at schools for children with special needs. The data collection took place during weekly ‘activity group sessions’. During the sessions, children with ASD used various technology applications with their familiar teachers and special needs assistants (hereafter referred to as educators) in the presence of researchers. This study focuses on the children’s gazing practices during a Kinect® body-movement game. In this game, the children were required to control an avatar on the screen using their body movements (i.e. hands and feet) to catch moving virtual objects. The children were able to choose the virtual objects they wanted to catch among options presented on the game screen (see Figure 1). The space also included other game-playing areas on the other side of the Kinect area.

[FIGURE 1 NEAR HERE]

The data were collected using two tripod-mounted digital cameras and SensoMotoric Instruments (SMI) mobile eye tracking glasses (sampling rate 30 Hz). The cameras were used to capture a wide-angle view of the game playing setting (including the Kinect screen) from the front and back, thus enabling us to situate the gazing practices within broader social situations.

The eye tracking data were exported using the SMI software as scan path raw video (i.e. no fixation filter was used) to allow the inclusion of quick eye movements, saccades, in the analysis. Each video presented the gaze cursor (seen in orange in Figure 1) on the scene captured by the eye-tracker's video cameras. The data was initially calibrated prior to the playing session using 1-point calibration.¹ To ensure continuing tracking quality, we manually recalibrated the data using the SMI software; for instance, if the children touched the glasses or made rapid head movements that caused the glasses to move.

Using the video annotation tool ELAN (developed by Max Planck Institute for Psycholinguistics), the scan path videos for each participant were synchronised with the video recordings of the participants' whole body movements. The data pool included approximately 232 minutes of scan path video. To quantify the data, we chose a sample of 9 minutes and 45 seconds of video for each child (uninclusive of non-calibrated data), which approximately covered their first time playing the game using the glasses.

2.2 Participants

This article involves data from three children, 11-year-old Matti, 8-year-old Veli and 6-year-old Roope, who represent children of different ages and with different types of ASD. For Matti and Veli, the data sample covers their first session with Kinect, whereas data from Roope's second session is included, as longer periods of non-calibrated data were removed from his data.

The children had been previously assessed for diagnostic and educational purposes, and these clinical documents were available for the researchers. An extensive assessment procedure

was considered unnecessarily distressing for the children; hence, the selected tests were kept to a minimum. These included the ‘Comprehension of Instructions’ test to evaluate comprehension of oral instructions (NEPSY-II) and the ‘Spatial Span’ test to evaluate visual working memory (Wechsler Memory Scale, WMS-III). The Sally-Anne false-belief task was also administered to evaluate children’s theory-of-mind abilities. The teachers completed the Finnish version of the Autism Spectrum Screening Questionnaire (ASSQ), and the children’s parents completed the Social Communication Questionnaire (SCQ) and the Strengths and Difficulties Questionnaire (SDQ).

One of the children, Veli, has a ‘pure’ ASD diagnosis, whereas Matti and Roope have other comorbid diagnoses, representing ‘autism plus’ (see Gillberg & Fernell, 2014). The interactions examined in this article also include the children’s adult co-participants: two special needs assistants (Kaisa and Mirja) and researcher Tommi. All the participants are native Finnish speakers. Table 2 provides information on Matti, Veli and Roope based on their clinical documents.

[TABLE 2 NEAR HERE]

Table 3 presents the children’s scores on NEPSY-II, WMS-III, the Sally-Anne task, ASSQ, SCQ and SDQ.

[TABLE 3 NEAR HERE]

Written consent for children to participate in the study was obtained from the children’s guardians and the educators. The 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 photos captured from the videos have been used with permission, with the facial features of the participants anonymised. The work was fully assessed by the ethics committee of the researchers' home institution and carried out in accordance with the Declaration of Helsinki as revised in 2000.

2.3 Coding of eye tracking data

The videos were coded according to the location of the gaze cursor in relation to specific AOI's using ELAN (see e.g. Holler & Kendrick, 2015). Coding was made on a frame-by-frame basis to determine the exact onsets and durations of the gaze shifts towards people. The AOI's included *people* (when the child gazed at any body part of the people in the environment), which was further specified as *educators* (i.e. familiar teachers and special needs assistants), *researchers* (i.e. researchers who organised the activity group sessions weekly), *children* (i.e. other familiar children who participated the activity group sessions) and *non-present people* (i.e. people who were not present at the Kinect playing station, e.g. another child working at another activity station). We also further differentiated whether the child gazed at others' *head area* or other *body parts*. However, we did not attempt to differentiate between looking at others' eye or mouth region, which would likely yield inaccurate results using live eye tracking (Falck-Ytter et al. 2015). All the non-calibrated or 'off-screen' (looking at an area that fell beyond the bounds of the scene camera) data were identified and removed from the analysis.

2.4 A multimodal approach to CA

We extended the coding of eye tracking data with qualitative multimodal approach to CA. This enabled a structural analysis of how the co-participants treated a child's gaze shifts towards people in the room, which in turn demonstrated what kind of functions they assigned to the gaze shifts. Rather than decontextualising gaze shifts categorically *a priori*, multimodally informed CA is discovery-oriented and proceeds inductively: examining the timing of the gaze shifts in relation to other activities in progress and how the gaze shifts are treated by the gaze-recipients. The interactional relevancy of an instance of gaze shift can be determined by a co-participant's response to it – whether this is by the party who has been gazed at or another party who sees the gaze shift. This interactional framework thus grounds observations empirically to participants' own orientation and responsiveness to each other's conduct in a stretch of interaction (see Schegloff, 2007). The so-called *next-turn proof procedure* (Sacks et al. 1974), wherein a participant's understanding of the preceding action (e.g., talk, gaze, or gesture) is displayed in the next turn-at-talk or other action, is a crucial requirement for the structural analysis of an interactional episode to be accepted as valid (Peräkylä, 2004). Thus, a function of gaze is here measured in terms of a co-participant's treatment of the gaze and what was accomplished with a gaze shift in interaction, rather than through a rendition of what an individual seemed to 'aim for'. A co-participant's response demonstrates how the gaze shift becomes treated functionally: undertaking a specific work (if any) in interaction.

The approach draws on detailed transcriptions of the gaze shifts identified on the recorded data. The transcripts show how gaze moves in relation to speech or during silent

intervals, and in relation to a child's own or a co-participant's actions in the environment. Such information helps to recontextualise the gaze shifts for the analysis of their functions. The rigour of multimodal interaction analysis is produced by co-examining the primary data and transcripts with fellow analysts and following transcription conventions to ensure proper documentation of the primary data (Peräkylä, 2004). The transcriptions also enable third parties to access the observations made from the data. This further contributes to the methodological rigour wherein participants' own orientations to one another are at issue, rather than the orientations imposed by an analyst.

The present study is limited to examining the gaze shifts directed towards other people in the room in a specific game-playing setting (Kinect body-movement game). Thus, from the entire data pool (i.e. 232 minutes of video material), the cases in which a child's gaze moved to another person were transcribed. Talk was transcribed according to the conventions described by Jefferson (in Atkinson & Heritage, 1984), and gaze using the notation developed by Goodwin (1981) (see Figure 11 in Appendix). The latter enables us to represent how gaze moves in time and co-occurs with speech, silence, and other actions.

3 Results

3.1 A distributional examination of eye tracking data

To examine where the children gazed at (i.e. the gaze target), the frequencies, total durations and percentages of their gaze shifts were counted based on the AOI coding. Owing to the small sample size and the illustrative nature of the distributional examination, no statistical tests were employed. Table 4 presents the distributional examination of the location of children's gaze shifts (AOI's).

[TABLE 4 NEAR HERE]

Table 4 shows that during the game playing session, Veli spent more time looking at people (9.15%), compared to Matti (2.68%) and Roope (1.48%). The children also differed in terms of the persons they looked at the most: Matti looked at researchers (58.24%), Veli gazed at other children (37.60%) and Roope gazed at educators (76.94%) the longest. Veli also spent more time looking at people who were not present at the Kinect station (13.53 s) than Matti (1.86 s) and Roope (0.24 s). Of the people that the children gazed at, Matti looked more at their body areas (84.32%), whereas Veli (52.18%) and Roope (62.60%) looked more at their head areas (although for Veli the difference is small).

The coding suggests that the children differed in the distributional pattern of their attention to different AOIs. However, this coding is decontextualized and provides little information about the contingencies under which these gaze shifts occurred (such as what was

said or done just before a child turned to look at an AOI). Thus, the focus on the ‘overall looking times’ on the AOIs does not delineate the social functions (if any) of these gaze shifts. To examine this in detail, we demonstrate how to recontextualise the gaze shifts using a multimodal approach to CA.

In the following section, we draw on the wide-angle video recordings of the gaming sessions to situate the identified gaze shifts within the broader streams of activity. We analyse a selection of gaze shifts to exemplify how the different functions of gaze shifts can be rendered visible through a qualitative structural analysis of the interactional situations: what elicits a child’s gaze to another person, and how a co-participant responds to a child’s gaze shift.

3.2 Analysing gaze shifts through a multimodal approach to CA

The multimodally informed CA approach focuses on the functions of the gaze shifts identified using eye tracking. As the children turned to look at other people, we show how to examine what elicited their gaze, and how the co-participants responded to the gaze, and thus what the gaze accomplished. In line with the principle of sequential organisation of social interactions (Schegloff, 2007), we will focus on examining gaze in both *responsive* and *initiating* environments. We begin with examples in which gaze emerges in a responsive environment within a stream of action (i.e. gaze shift occurred in response to something). We then move to examples of gaze in an initiating environment (i.e. gaze shift initiates an interactional exchange) and elaborate on the different functions within these environments.

Figure 2 outlines the four types of functions found in the analysis. The interactional implications (if any) have been developed from the analysis of how the gaze was elicited and how the co-participants' responded to these gaze shifts. While each example involves a gaze shift directed to a person in the room (indicated by the gaze cursor coloured orange), each instance differs in the degree to which this gaze can be characterised as an *interactional* contribution. In each case, we dissect these functions using detailed transcriptions of what was said and done when the gaze shift occurred, with the analysis unfolding moment-by-moment at the level of tenths of seconds. Drawing on the principles of CA, each function is empirically grounded in the analysis of what the gaze shift accomplishes at the moment it occurs, thus withholding psychologised accounts of participants' internal motivations.

[FIGURE 2 NEAR HERE]

3.2.1 Responsive gaze shifts

In the environment of *responsive gaze shifts*, a child's gaze was either spontaneously directed to a person or an event (e.g. sound) or elicited from the child using a prompt (e.g. a request to look). We first consider an instance of the former in Figure 3. Veli has been playing the game, but then something captures his attention, which results in a halt in the game. Veli shifts his gaze from the screen to the door. (Prior to the beginning of the extract, another child, Roope, had caused some distractions by making loud noises in the room. Roope's special needs assistant, Mirja, decided to take Roope outside until he had calmed down. The door from which they had exited a moment earlier is next to Veli and the gaming screen.) We join in as the sound of footsteps apparently catches Veli's attention.

[FIGURE 3 NEAR HERE]

In line 1, Veli orients to the sound of footsteps and shifts his gaze towards the door in anticipation of someone entering the room. Mirja steps in (line 3) and says *noni* ('alright'), notifying their return to the room (line 4). Figure 4 zooms in on Veli's gaze movements upon Mirja's entrance during a silent interval of 2.1 seconds. Veli looks at Mirja and then shifts his gaze to the corridor in an apparent expectation of Roope, who enters the room shortly after Mirja. Veli shifts his gaze to Roope (line 6) and monitors him as Mirja instructs Roope to 'go and sit down'. As Mirja continues speaking to Roope, *odotetaan kun meidän vuoro tulee* ('let's wait until our turn comes'), Veli shifts his gaze between Mirja and Roope, and then re-oriens back to the screen. Even though Veli has been looking at Mirja and Roope during their stretch of interaction, neither of them pays any particular attention to Veli looking at them.

[FIGURE 4 NEAR HERE]

Such instances of *monitoring* were relatively common in the eye tracking dataset: a sound or a movement catches a child's attention, and the child responds by shifting his gaze to that location. In such cases, the people who were being monitored would not notice or respond to the gaze directed at them. Such non-responsiveness to the child's gaze indicates that the gaze shift was not treated as a move to enter an interaction with that person. Thus, while Veli's gaze was directed at both Mirja and Roope, the moment when he turned to look at them indicated his responsive role as an observer of a situation, which did *not* implicate him as a 'ratified' co-participant (see Goffman, 1981; Goodwin, 1981). Indeed, by merely monitoring their return to the room, Veli was not addressing Mirja and Roope in any way; consequently, neither Mirja nor

Roope responded to Veli's gaze directed at them. The primary function of the responsive monitoring gaze was thus to check what was happening in the environment without having interactional implications to the parties being looked at.

Occasionally, a child's gaze shift to another person in the room was responsive to a prompt or an explicit call for attention. In Figure 5, we consider an instance from a different gaming session involving Roope. Roope had been struggling to master the interface of the game: even though he had been moving his hands in front of the screen, he had failed to catch any virtual objects. Such difficulties were relatively common during the first play trials as the children were learning to connect their body movements with the avatar on the screen. Roope managed to complete one game level, but in the next level, the instructions had changed and now required him to use both hands to capture the virtual objects. We join in as Roope stands in front of the screen holding his right index finger extended. As this position would not allow him to catch any objects, Mirja intervenes and provides instructions for Roope.

[FIGURE 5 NEAR HERE]

Mirja waves her hands in the air demonstrating how to catch the virtual objects. She also prompts verbally, *molemmilla käsillä Roope* ('with both hands Roope'). While Mirja speaks, Roope sustains his gaze at the screen. On hearing his name mentioned ('Roope', end of line 1), Roope shifts his gaze towards her. Figure 6 zooms in on Roope's gaze movements from the screen to Mirja precisely when he hears his name being called; the gaze fixates on Mirja's facial area at the end of the lexical item 'Roope'. His gaze then moves to the wall, then to Mirja's hand demonstrating the catching gesture, and back to the wall.

[FIGURE 6 NEAR HERE]

Since Roope did not turn to look at Mirja as she first spoke in line 1, Mirja uses a highly specific address term ('Roope') to redirect his attention to the instructions given. In line 2 Mirja further requests Roope's gaze, *kato* ('look'), and then repeats her instruction while using her hands to demonstrate the required hand movements to catch the objects (line 3). Gaze is a key resource for recipients to show whether or not they are 'acting as hearers' (Goodwin, 1981); here the timing of Roope's gaze shift indicates that he responds to Mirja by attending to her demonstration.

In these two examples, a gaze shift has emerged in response to a change in the environment (i.e. an observed interaction between two people entering the room), or a prompt (i.e. a direct pursuit of attention) from a co-participant. Both examples demonstrate the children's ability to attend to socially relevant information (i.e. other people) in a responsive role. While the monitoring gaze did not have interactionally relevant connotations, the gaze shift following a prompt demonstrated a child's reciprocity to an interactional pursuit, which sought the child's realignment of gaze.

3.2.2 Initiating gaze shifts

The children also directed their gaze from the screen to a person close-by in an initiating environment. The functions of such gaze shifts were rendered visible by examining when the gaze shift occurred in relation to the activities in progress and how the people present responded

to such gaze. In Figure 7, Matti shifts his gaze to a co-present adult after several unsuccessful attempts to catch virtual objects in the game.

[FIGURE 7 NEAR HERE]

Matti halts playing and shifts his gaze from the screen towards Tommi, who is also looking at the screen and does not notice Matti's gaze reaching him. Here Matti's gaze *addresses* Tommi as the selected party among all the other people in the room (see also Dickerson et al. 2005). While looking at Tommi, Matti brings his hands down, and then quietly says something (inaudible from audio). Gaze can be used as a resource to mobilise a response from a co-participant in the environments where talk has not so implicated (Stivers & Rossano, 2010): in sustaining his gaze at Tommi, Matti pursues Tommi to respond in some way. Kaisa, who has been observing the situation on the side of the room, sees Matti's gaze shifting to Tommi, and that Tommi seems unavailable to respond (i.e. Tommi is not 'acting as a hearer', see Goodwin, 1981), and steps in instead. She quickly prompts Matti to use both hands *molemmat kädet saat ottaa käyttöön* ('you can use both hands') (line 2), overlapping with Matti's quiet utterance. Figure 8 demonstrates how Kaisa's instruction starts precisely in response to Matti's head turn as he shifts his gaze to Tommi.

[FIGURE 8 NEAR HERE]

The timing of Matti's gaze shift suggests that he is eliciting assistance with the game: Matti turns to Tommi directly after he has halted playing. The direction of Matti's gaze selects Tommi, who is also a researcher controlling the game using a laptop, as the recipient who would be able to

provide him the assistance he needs. Matti's perseverance with looking at Tommi shows thus competence in speaker-selection (see also Korhonen & Rae, 2014): he does not withdraw his gaze until he receives some instructions, albeit it is Kaisa, rather than Tommi, who provides them to him.

That Kaisa responds to Matti's gaze shift with an instruction provides strong evidence that his gaze was designed to elicit assistance: Kaisa's prompt is produced precisely as Matti shifts his gaze to Tommi; *not* a moment before when Matti was still gazing at the screen. On hearing Kaisa's instructions to Matti, Tommi quickly reorients to Matti to monitor the situation. Matti then shifts his gaze back to the screen to follow Kaisa's instructions (lines 2-3). This re-orientation to the screen underscores that the function of Matti's gaze was to elicit assistance or instructions on the game, since when these had been received, Matti was ready to re-orient back to the game.

Here Matti's gaze has emerged in an initiating environment. It was also treated as initiating an interactional exchange for *imperative* purposes: to mobilise other people to help him with the game. However, some of the gaze shifts to the people in the room occurred when no apparent trouble was taking place. These gaze shifts were also treated differently, and some of them included an element of social sharing. We consider this in Figure 9. The example involves Roope, who has been playing the game rather successfully for a good while.

[FIGURE 9 NEAR HERE]

In line 1, Roope follows a moving virtual object on the screen with his gaze and his hand while uttering *pum* ('boom'). Roope apparently imitates the sound of the virtual objects being 'caught'. Having a slight smile on his face, he turns to Tommi. Tommi notices Roope's gaze to him and looks back at him, thus attending to Roope (line 2). Roope sustains his gaze at Tommi, suggesting that he is pursuing Tommi to respond in some way (Stivers & Rossano, 2010). A silent interval occurs during which Tommi does not immediately say anything, resulting in a momentary halt in the progressivity of their interaction. Tommi's lack of responsiveness suggests his apprehension in *what* to respond: whether to Roope's utterance *pum* ('boom') or to his game-related actions on the screen (Figure 10). Since Roope's gaze remains sustained at Tommi, Tommi responds by appraising his playing, *hyvin menee* ('it's going well') (line 3).

[FIGURE 10 NEAR HERE]

While Roope's utterance *pum* ('boom') did not by its design or delivery make a response relevant, Roope's *gaze* was designed to invite Tommi to respond (see also Korhonen & Rae, 2014; Stivers & Rossano, 2010). The timing of Roope's gaze shift right after his utterance suggests that, rather than initiating interaction for any instrumental purpose, Roope seeks to share an observation of successful object catching with Tommi (presumably in response to the sound of virtual objects) and elicits a response to that observation. Thus, Roope's gaze to Tommi suggests a *declarative* purpose – yet Tommi's delayed response indicates that mutual understanding was not immediately reached in terms of *what* Roope wanted to share with Tommi.

3.2.3 Summary

The multimodal approach to CA can complement quantitatively oriented (AOI-based) eye tracking measurements with a qualitative examination, which helps examining the functions of gaze behaviours identified. The present study, which was based on precise measurements of gaze, rendered visible four types of functions within initiating and responsive environments that varied in their interactional implications, i.e. the degrees to which the gaze shifts were taken to contribute to some stretch of social interaction. Monitoring the activities of others without addressing them appeared to be broadly devoid of interactional connotations, while those gaze shifts that were interactionally geared either responded to an interactional bid by another person or initiated an interactional exchange with them. The timing of the gaze shifts with respect to other activities in progress was crucial in delineating the interactional connotations by the co-participants, who either (a) responded by engaging in interaction with the child (i.e. with gaze and some talk), or (b) did not respond to the gaze shifts, indicating that the child's looking was not treated as betokening interactional participation from the child.

4 Discussion

This study set out to demonstrate how to combine live eye tracking and a multimodal approach to CA to study the functions of gaze shifts, i.e. what gaze accomplishes in an interactional context (here limited to a game-playing activity). The co-participants responded to the children's gaze shifts differently depending on when these occurred during a stream of activity; that is, how a gaze shift was timed with the on-going game, talk, and body movements or other people's

activities. In this study, the children's gaze shifts were identified in responsive or initiating environments, and their functions varied in the degrees to which they provided an interactional contribution to a situation at hand.

Previous experimental psychological studies in ASD have rarely considered the use of gaze beyond a child's responsive role, and even less examination has been conducted on the actions of co-interactants, especially on their responses to a child's gaze. However, the responses from others are important in the analysis of what gaze does in real-time exchanges, what makes some instances of gaze socially oriented, and precisely *interactional*. Streeck (1993) talks about 'contractual' nature of mutual gaze, which requires an act of recognition what is 'going on' between participants whose eyes meet. In the present study, the co-participants made judgements on the degree a child's gaze implicated 'ratified' interactional participation, not on the basis of a child's gaze shift alone but in response to its timing with other activities in progress.

The children in the present study shifted their gaze to a co-participant occasionally to initiate interaction. This aligns with previous interactional research on children with ASD that have showed competency in the use gaze in social interactions, for example, to address another person (Dickerson et al. 2005) and to elicit feedback from a co-participant (Korkiakangas & Rae, 2014) in a similar manner as participants with typical development (see e.g. Goodwin, 1981). Such competencies can be rendered visible when behaviours such as gaze are examined in detailed interactional contexts and as tied with the contributions of other people. Dickerson and Robins (2015) have stressed that interactional nuances are likely to be lost when gaze behaviours are coded in a decontextualised manner, rather than transcribed and analysed sequentially. Thus,

while eye tracking research traditionally codifies and quantifies the overall looking times at an AOI, it can remain as a limited approach in rendering the different functions of gaze visible. The current study has demonstrated how the functions of gaze shifts can be identified through an analysis of line-by-line transcription of the events that occurred in wide-angle video recordings as the gaze shifts emerged.

The current study has limitations that should be acknowledged. The study was focused only on three children in a specific educational setting. Thus, the children's production of gaze shifts should be primarily understood in this specific game-playing context. However, as the identified functions of gaze here align with prior multimodally informed CA work of gaze in everyday interactions (e.g., Korhonen & Rae, 2014; Stivers & Rossano, 2010), they seem possible in different contexts (see Peräkylä, 2004 for generalizability in CA), and offer hypotheses for further research with a larger sample of children. Despite the context-specificity of this study, it foregrounds the benefits of qualitative interaction analysis on quantitative eye tracking measurements. Extending this methodological combination to diverse interactional contexts, and comparisons with typically developing children, would be a fruitful line of inquiry.

While the use of mobile eye tracking glasses offers potential for researching interaction in naturalistic settings with multiple participants wearing the eye tracking glasses (as in Holler & Kendrick, 2015), our study was limited to the Kinect gaming session interactions during which only children wore the glasses. Eye tracking equipment warrants some consideration also from a research perspective. While this technology can provide more accurate measurements of gaze behaviours than is possible in merely video-based analysis (Guillon et al. 2014), there has been a

tendency to over-rely on quantified eye tracking measurements in isolation. In this article, we have suggested that a more detailed and ecologically valid approach in the study of gaze would entail eye tracking in combination with the qualitative analysis of wide-angle video recordings in recontextualising the gaze shifts measured.

Mobile eye tracking equipment can also interfere with the interactions taking place. Although the technology allows participants to move freely in space, the participants are usually actively aware of the situations in which they are wearing the glasses. In the present study, the game-playing activity began with calibrating the device and the children were asked to refrain from touching the glasses unless necessary due to its effect on the calibration. Thus, live eye tracking using mobile equipment brings in elements from experimentally oriented research and does not result in 'fully natural' situations, as the gaze-related research is usually apparent to all participants. Despite these limitations, mobile eye tracking enables ecologically valid research compared with computer-presented stimuli (e.g. using table-mounted eye-trackers), since gaze can be analysed in actual interactions between co-participants.

The interactional framework can nevertheless complement live eye tracking research with a 'bottom-up', data-driven approach (as has been called for by Falck-Ytter et al. 2013) in at least three ways. First: the setting. The naturalistic environments where children are allowed to act rather freely (although here the setting was restricted within the parameters of the Kinect game) can render visible the different purposes for which gaze is used in social settings: those designed for social interaction and those devoid of such implications. For instance, the gaming activity meant that children sometimes required help with catching the virtual objects; when faced with

difficulties, children were found to request assistance by turning their gaze to the adults in the room (see also Kidwell, 2009). The gaming took place in a large room with several people present, which naturally provided opportunities to monitor their actions. While experimental rigour requires that procedures are kept 'stable' and 'reproducible', there can be enormous merit in examining more free-flowing naturalistic interactions with other people.

Second, the wide-angle video footage is crucial for locating the gazing practices within the on-going streams of activity. While synchronising eye tracking data with video recordings can be time-consuming and laborious, it provides a valuable resource for the interactional exploration of gaze-in-context. Third, the combination of eye tracking with an interactional framework is paramount for dissecting the social qualities of gaze. These can be examined empirically by analysing the actions of all parties on a moment-by-moment basis. For instance, a co-interactant can respond to an instance of gaze from a child differently in terms of *when* the gaze occurs during a stretch of interaction. This, in turn, can indicate what kind of work, if any, the gaze accomplished in that interaction.

Since the primary aim of live eye tracking is to study gaze behaviours in real social situations, such research could benefit greatly from the framework that situates the gaze observations within the interactions in which they have emerged. This combination of live eye tracking and qualitative microanalytic study of social interaction enables us to study individuals with ASD as active social participants, rather than as passive observers of information, which has been the approach taken in traditional eye tracking studies to date.

Funding

The authors received funding from Alfred Kordelinin Säätiö, the Olvi Foundation, and Suomen Akatemia, grant number 140450.

Notes on contributors

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 ASD during educational interactions with adult co-participants. She applies multimodally informed CA in her research. She has conducted her PhD research in the Children with Autism Spectrum Disorders as Creative Actors in a Strength-Based Technology-Enhanced Learning Environment (CASCATE) project funded by the Academy of Finland (2011-2014).

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 works as a professor of psychology at the University of Eastern Finland. 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.

Notes

¹ 1-point calibration was chosen for quick calibration; initial attempts showed a 3-point calibration procedure was slow and challenging for the participating children.

5 References

- Argyle, M and Dean, J 1965, 'Eye-contact, distance and affiliation', *Sociometry*, vol. 28, pp. 289–304.
- Argyle, M, Ingham, R, Alkema, F and McCallin, M 1973, 'The different functions of gaze', *Semiotica*, vol. 7, pp. 19–32.
- Atkinson, JM and Heritage, J 1984, 'Transcript notation', in JM Atkinson & J Heritage (eds.), *Structures of social action: Studies in conversation analysis*, Cambridge University Press, Cambridge, MA, pp. ix–xvi.
- Auer, P in press, 'Gaze, addressee selection and turn-taking in three-party interaction', in Brone (ed.), *Eye-tracking in interaction. Studies on the role of eye gaze in dialogue*, Amsterdam, Benjamins.
- Broekhof, E, Ketelaar, L, Stockmann, L, van Zijp, A, Bos, MG and Rieffe, C 2015, 'The understanding of intentions, desires and beliefs in young children with autism spectrum disorder', *Journal of Autism and Developmental Disorders*, vol. 45, pp. 2035–2045.

- Camaioni, L, Perucchini, P, Muratori, F, Parrini, B and Cesari, A 2003, 'The communicative use of pointing in autism: Developmental profile and factors related to change', *European Psychiatry*, vol. 18, pp. 6–12.
- Chawarska, K, Macari, S and Shic, F 2012, 'Context modulates attention to social scenes in toddlers with autism', *Journal of Child Psychology and Psychiatry*, vol. 53, pp. 903–913.
- Chevallier, C, Parish- Morris, J, McVey, A, Rump, KM, Sasson, NJ, Herrington, JD and Schultz, RT 2015, 'Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters', *Autism Research*, vol. 8, pp. 620–628.
- Damico, J. S. and Nelson, RL 2005, 'Interpreting problematic behavior: Systematic compensatory adaptations as emergent phenomena in autism', *Clinical Linguistics & Phonetics*, vol. 19, pp. 405–417.
- Dapretto, M, Davies, MS, Pfeifer, JH, Scott, AA, Sigman, M, Bookheimer, SY and Iacoboni, M 2006, 'Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders', *Nature Neuroscience*, vol. 9, pp. 28–30.
- Dickerson, P, Rae, J, Stribling, P, Dautenhahn, K and Werry, I 2005, 'Autistic children's coordination of gaze and talk: Re-examining the 'asocial' autist', in K Richards & P Seedhouse (eds.), *Applying conversation analysis*, Palgrave Macmillan, Hampshire, pp. 19–37.
- Dickerson, P. and Robins, B 2015, 'Looking or spotting: A conversation analytic perspective on interaction between a humanoid robot, a co-present adult, and a child with an ASC', in M

O'Reilly & J N Lester (eds.), *The Palgrave handbook of child mental health*, Palgrave Macmillan, London, pp. 59–78.

Dickerson, P, Stribling, P and Rae, J 2007, 'Tapping into interaction. How children with autistic spectrum disorders design and place tapping in relation to activities in progress', *Gesture*, vol. 7, pp. 271–303.

Exline, RV and Winters, LC 1965, 'Affective relations and mutual glances in dyads', In S Tomkins & C Izzard (eds.), *Affect, cognition and personality*, Springer, New York, pp. 319–350.

Falck-Ytter, T 2010, 'Young children with autism spectrum disorder use predictive eye movements in action observation', *Biology Letters*, vol. 6, pp. 375–378.

Falck-Ytter, T 2015, 'Gaze performance during face-to-face communication: A live eye tracking study of typical children and children with autism', *Research in Autism Spectrum Disorders*, vol. 17, pp. 78–85.

Falck-Ytter, T, Carlström, C and Johansson, M 2015, 'Eye contact modulates cognitive processing differently in children with autism', *Child Development*, vol. 86, pp. 37–47.

Falck-Ytter, T, Fernell, E, Hedvall, ÅL, von Hofsten, C and Gillberg, C, 2012, 'Gaze performance in children with autism spectrum disorder when observing communicative actions', *Journal of Autism and Developmental Disorders*, vol. 42, pp. 2236–2245.

Falck-Ytter, T. and von Hofsten, C 2010, 'How special is social looking in ASD: A review', *Progress in Brain Research*, vol. 189, pp. 209–222.

- Falck-Ytter, T, von Hofsten, C, Gillberg, C and Fernell, E 2013, 'Visualization and analysis of eye movement data from children with typical and atypical development', *Journal of Autism and Developmental Disorders*, vol. 43, pp. 2249–2258.
- Falkmer, M, Bjällmark, A, Larsson, M and Falkmer, T 2011, 'The influences of static and interactive dynamic facial stimuli on visual strategies in persons with Asperger syndrome', *Research in Autism Spectrum Disorders*, vol. 5, pp. 935–940.
- Fletcher-Watson, S, Leekam, SR, Benson, V, Frank, MC and Findlay, JM 2009, 'Eye-movements reveal attention to social information in autism spectrum disorder', *Neuropsychologia*, vol. 47, pp. 248–257.
- Freeth, M, Foulsham, T and Kingstone, A 2013, 'What affects social attention? Social presence, eye contact and autistic traits', *PLoS One*, vol. 8, pp. e53286.
- Gillberg, C and Fernell, E 2014, 'Autism plus versus autism pure', *Journal of Autism and Developmental Disorders*, vol. 44, pp. 3274–3276.
- Gillespie-Lynch, K, Elias, R, Escudero, P, Hutman, T and Johnson, SP 2013, 'Atypical gaze following in autism: A comparison of three potential mechanisms', *Journal of Autism and Developmental Disorders*, vol. 43, pp. 2779–2792.
- Gobel, MS, Kim, HS and Richardson, DC 2015, 'The dual function of social gaze', *Cognition*, vol. 136, pp. 359–364.

- Goffman, E 1963, *Behavior in public places. Notes on the social organization of gatherings*, Free Press, New York.
- Goffman, E 1981, 'Footing', in E Goffman (ed.), *Forms of Talk*, University of Pennsylvania Press, Philadelphia, pp. 95-124.
- Goodwin, C 1981, *Conversational organisation: Interaction between speakers and hearers*, Academic Press, New York.
- Goodwin, MH 1980, 'Processes of mutual monitoring implicated in the production of description sequences', *Sociological Inquiry*, vol. 50, pp. 303–317.
- Guillon, Q, Hadjikhani, N, Baduel, S and Rogé, B 2014, 'Visual social attention in autism spectrum disorder: Insights from eye tracking studies', *Neuroscience & Biobehavioral Reviews*, vol. 42, pp. 279–297.
- Haddington, P 2006, 'The organization of gaze and assessments as resources for stance taking', *Text & Talk*, vol. 26, pp. 281–328.
- Hanley, M, Riby, DM, Carty, C, McAteer, AM, Kennedy, A and McPhillips, M 2015, 'The use of eye-tracking to explore social difficulties in cognitively able students with autism spectrum disorder: A pilot investigation', *Autism*, vol. 19, pp. 868–873.
- Hanley, M, Riby, DM, McCormack, T, Carty, C, Coyle, L, Crozier, N, Robinson, J and McPhillips, M 2014, 'Attention during social interaction in children with autism: Comparison to specific

language impairment, typical development, and links to social cognition', *Research in Autism Spectrum Disorders*, vol. 8, pp. 908–924.

Heath, C 1986, *Body movement and speech in medical interaction*, Cambridge University Press, Cambridge.

Hirvenkari, L, Ruusuvoori, J, Saarinen, VM, Kivioja, M, Peräkylä, A and Hari, R 2013, 'Influence of turn-taking in a two-person conversation on the gaze of a viewer', *PloS One*, vol. 8, pp. e71569.

Holler, J and Kendrick, KH 2015, 'Unaddressed participants' gaze in multi-person interaction: Optimizing reciprocity', *Frontiers in Psychology*, vol. 6, <http://dx.doi.org/10.3389/fpsyg.2015.00098>.

Kendon, A 1967, 'Some functions of gaze-direction in social interaction', *Acta Psychologica*, vol. 26, pp. 22–63.

Kidwell, M 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*, vol. 38, pp. 417–449.

Kidwell, M 2009, 'Gaze shift as an interactional resource for very young children', *Discourse processes*, vol. 46, pp. 145–160.

Klin, A, Jones, W, Schultz, R, Volkmar, F and Cohen, D 2002, 'Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism', *Archives of General Psychiatry*, vol. 59, pp. 809–816.

- Korkiakangas, TK (2011), *Eye-gaze in multimodal interactions involving children with autism spectrum disorders*, Doctoral dissertation, University of Roehampton.
- Korkiakangas, T and Rae, J 2014, 'The interactional use of eye-gaze in children with autism spectrum disorders', *Interaction Studies*, vol. 15, pp. 233–259.
- Lerner, GH 2003, 'Selecting next speaker: the context-sensitive operation of a context-free organization', *Language in Society*, vol. 32, pp. 177–201.
- Magrelli, S, Jermann, P, Basilio, N, Ansermet, F, Hentsch, F, Nadel, J and Billard, A 2013, 'Social orienting of children with autism to facial expressions and speech: a study with a wearable eye-tracker in naturalistic settings', *Frontiers in Psychology*, vol. 4, <http://dx.doi.org/10.3389/fpsyg.2013.00840>.
- Maljaars, J, Noens, I, Jansen, R, Scholte, E and van Berckelaer-Onnes, I 2011, 'Intentional communication in nonverbal and verbal low-functioning children with autism', *Journal of Communication Disorders*, vol. 44, pp. 601–614.
- Mondada, L 2014, 'Pointing, talk, and the bodies: Reference and joint attention as embodied interactional achievements', in M Seyfeddinipur & M Gullberg (eds.), *From gesture in conversation to visible action as utterance: Essays in honor of Adam Kendon*, John Benjamins, Amsterdam, pp. 95–124.
- Muskett, T, Perkins, M, Clegg, J and Body, R 2010, 'Inflexibility as an interactional phenomenon: Using conversation analysis to re-examine a symptom of autism', *Clinical Linguistics & Phonetics*, vol. 24, pp. 1–16.

- Nadig, A, Lee, I, Singh, L, Bosshart, K and Ozonoff, S 2010, 'How does the topic of conversation affect verbal exchange and eye gaze? A comparison between typical development and high-functioning autism', *Neuropsychologia*, vol. 48, pp. 2730–2739.
- Norbury, CF, Brock, J, Cragg, L, Einav, S, Griffiths, H and Nation, K 2009, 'Eye- movement patterns are associated with communicative competence in autistic spectrum disorders', *Journal of Child Psychology and Psychiatry*, vol. 50, pp. 834–842.
- Noris, B, Nadel, J, Barker, M, Hadjikhani, N and Billard, A 2012, 'Investigating gaze of children with ASD in naturalistic settings', *PLoS One*, vol. 7, pp. e44144.
- O'Reilly, M, Lester, JN and Muskett, T 2016, 'Discourse/conversation analysis and autism spectrum disorder', *Journal of Autism and Developmental Disorders*, vol. 46, pp. 355–359.
- Peräkylä, A 2004, 'Reliability and validity in research based on naturally occurring social interaction', in D Silverman (ed.), *Qualitative research: Theory, method and practice* (2nd ed.), Sage, London, pp. 283–304.
- Risko, EF, Laidlaw, KE, Freeth, M, Foulsham, T and Kingstone, A 2012, 'Social attention with real versus reel stimuli: Toward an empirical approach to concerns about ecological validity', *Frontiers in Human Neuroscience*, vol. 6, <http://dx.doi.org/10.3389/fnhum.2012.00143>.
- Ruffman, T, Garnham, W and Rideout, P 2001, 'Social understanding in autism: Eye gaze as a measure of core insights', *Journal of Child Psychology and Psychiatry*, vol. 42, pp. 1083–1094.

- Sacks, H, Schegloff, EA and Jefferson, G 1974, 'A simplest systematics for the organization of turn-taking for conversation', *Language*, vol. 50, pp. 696–735.
- Schegloff, EA 2007, *Sequence organization in interaction: A primer in conversation analysis* (vol. 1), Cambridge University Press, New York.
- Senju, A and Johnson, MH 2009, 'Atypical eye contact in autism: Models, mechanisms and development', *Neuroscience & Biobehavioral Reviews*, vol. 33, pp. 1204–1214.
- Solomon, O, Heritage, J, Yin, L, Maynard, DW and Bauman, ML 2016, "What brings him here today?": Medical problem presentation involving children with autism spectrum disorders and typically developing children', *Journal of Autism and Developmental Disorders*, vol. 46, pp. 378–393.
- Stivers, T and Rossano, F 2010, 'Mobilizing response', *Research on Language and Social Interaction*, vol. 43, pp. 3–31.
- Stivers, T and Sidnell, J 2005, 'Introduction: Multimodal interaction', *Semiotica*, vol. 2005, pp. 1–20.
- Streeck, J 1993, 'Mutual gaze and recognition: Revisiting Kendon's "Gaze direction in two-person conversation"', in M Seyfeddinipur & M Gullberg (eds.), *From Gesture in Conversation to Visible Action as Utterance: Essays in honor of Adam Kendon*, John Benjamins, Amsterdam, pp. 35-58.

- Thorup, E, Nyström, P, Gredebäck, G, Bölte, S and Falck-Ytter, T 2016, 'Altered gaze following during live interaction in infants at risk for autism: An eye tracking study', *Molecular Autism*, vol. 7, <http://dx.doi.org/10.1186/s13229-016-0069-9>.
- Tuononen, K, Korhonen, T, Laitila, A and Kärrä, E 2016, 'Zooming in on interactions: A micro-analytic approach examining triadic interactions between children with autism spectrum disorder and their co-participants', *SAGE Methods Cases*, London, SAGE Publications Ltd, <http://dx.doi.org/10.4135/978144627305015595395>.
- Vabalas, A and Freeth, M 2016, 'Brief report: Patterns of eye movements in face to face conversation are associated with autistic traits: Evidence from a student sample', *Journal of Autism and Developmental Disorders*, vol. 46, pp. 305–314.
- von Hofsten, C, Uhlig, H, Adell, M and Kochukhova, O 2009, 'How children with autism look at events', *Research in Autism Spectrum Disorders*, vol. 3, pp. 556–569.
- Wiklund, M 2012, 'Gaze behavior of pre-adolescent children afflicted with Asperger Syndrome', *Communication and Medicine*, vol. 9, pp. 173–186.
- Zwaigenbaum, L, Thurm, A, Stone, W, Baranek, G, Bryson, S, Iverson, J, Kau, A, Klin, A, Lord, C, Landa, R and Rogers, S 2007, 'Studying the emergence of autism spectrum disorders in high-risk infants: Methodological and practical issues. *Journal of Autism and Developmental Disorders*, vol. 37, pp. 466–480.

Appendix

Figure 11. Transcription notations.

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

Symbol	Meaning
↑ (up arrow)	following talk goes up suddenly
(in parentheses)	uncertain transcription
((word in double parentheses))	transcription comment or non-vocal action
<u>word</u> (underlining)	spoken with emphasis
[text]	adjacent lines overlap
[text]	

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.

Table 1. Some identified functions of gaze shifts to other people.

Source	Gaze function
Goffman (1963), (1981), Heath (1986)	Goodwin Displaying attention and (dis)engagement
Argyle et al. (1973), (1981), Kendon (1967)	Goodwin Displaying participation roles
Argyle and Dean (1965), et al. (1973)	Argyle Displaying intimacy
Exline and Winters (1965)	Displaying interpersonal attitudes
Heath (1986), Lerner (2003)	Kendon (1967), Regulating turn taking or turn allocation

Argyle et al. (1973), Goodwin Monitoring others' expressions
(1980), Kendon (1967)

Stivers and Rossano (2010) Soliciting a response

Haddington (2006), Kidwell Implementing social actions (e.g. appealing
(2005, 2009) for assistance, displaying a stance)

Table 2. Information on Matti, Veli and Roope based on their clinical documents.

Matti	
Age	11 years and 3 months
Diagnoses	Mental retardation (F79.0; ICD-10 criteria) Childhood autism (F84.0; ICD-10 criteria)
Rehabilitation history	Occupational therapy, physiotherapy, music therapy and speech therapy
Characteristics	Matti's documents describe him as being collaborative and happy when interacting with others. He is described as responsive but not active in making initiations. Matti's language development is delayed, and he occasionally uses repetitive phrases. He benefits from clear, context-related instructions and augmentative and alternative communication methods. Matti often needs prompting to act. He has challenges in motor coordination and balance.

Veli

Age 8 years and 7 months

Diagnoses Childhood autism (F84.0; ICD-10 criteria)

Rehabilitation Speech therapy and occupational therapy
history

Characteristics Veli's documents state that the development of his spoken language has been delayed since he was 2 years old. His speech is occasionally repetitive, and he has difficulties in multiple language-related skills. New situations and people are also difficult for him to handle, and his use of eye contact is scarce. The documents describe him as easily distracted. He understands clear and short sentences and benefits from adults being supportive and providing instructions.

Roope

Age	6 years
Diagnoses	Moderate mental retardation (F71.0; ICD-10 criteria) Childhood autism (F84.0) Other specified behavioral and emotional disorders with onset usually occurring in childhood and adolescence (F98.8) Specific developmental disorder of motor function (F82)
Rehabilitation history	Physiotherapy, music therapy and speech therapy
Characteristics	Roope's speech is described as markedly delayed. He can however use sentence-length speech to communicate. He has difficulties understanding instructions but benefits from clear guidance. Roope has perceptual and attentional difficulties, and transitions are described as difficult for him. His fine motor skills are delayed. Roope easily engages in eye contact and interactions with others.

Table 3. Assessment scores on NEPSY-II, WMS-III, Sally-Anne Task, ASSQ, SCQ and SDQ for Matti, Veli and Roope.

Measurement	Matti's scores	Veli's scores	Roope's scores
NEPSY-II (‘Comprehension of Instructions’) ^a	0	1	1
WMS-III (‘Spatial Span’) ^b	4	10	0
Sally-Anne task	Passed	Passed	Did not pass
ASSQ (teacher ratings) ^c	24	38	21

SCQ (lifetime 21
 version, parent
 ratings)^d *No information available* 8

SCQ subcategories

Social interaction 7 1

Communication 5 3

Stereotypical behaviour 8 4

Other	1	0	
SDQ (parent ratings) ^e	27	20	12

^a Raw points out of 20

^b Raw points out of 32

^c Cut-off score: 22

^d Cut-off score: 15

^e Cut-off score for 'abnormal': 16

Accepted Manuscript

Table 4. The frequencies, total durations and percentages of the children's gaze shifts with respect to the AOI's.

	Matti			Veli			Roop		
	<i>f</i>	Duratio n ^f	% ^g	<i>f</i>	Duratio n	%	<i>f</i>	Duratio n	%
People	5	13.65	2.68	18	51.86	9.15	3	8.37	1.48
(total)	7			6			5		
Educator	2	3.84	28.1	56	12.23	23.5	2	6.44	76.9
	1		3			8	0		4
Researcher	2	7.95	58.2	47	6.60	12.7	1	1.69	20.1
	2		4			3	2		9
Child	0	0	0	50	19.50	37.6	0	0	0
						0			0

Non-	1	1.86	13.6	33	13.53	26.0	3	0.24	2.87
present	4		3			9			

Body part

Head	1	2.14	15.6	11	27.06	52.1	2	5.24	62.6
	4		8	5		8	1		0
Body	4	11.51	84.3	93	24.80	47.8	1	3.13	37.4
	9		2			2	4		0

^f Total durations in seconds

^g Percentages based on total gaze durations

Figure 1. Video camera view of the broad social situation (frame 1), a child playing the Kinect game (frame 2) and an eye-tracker view (frame 3).



Figure 2. Four functions of gaze shifts identified.



Figure 3. Transcribed interaction of Mirja and Roope's entrance.

1 ((sounds of footsteps))

2 V: screen shifts towards the door

3 M: [((steps in))
(-----1-----2)
M (face)
V: [x_____

4 M: [noni
alright
V: [M (body)

5 V: M (body) corridor
(-----1-----2-----3)

6 R: [((R enters the space))
M: meehän istumaan
go and sit down
V: [R (face and body)
x_____

7 M: [odotetaan kun meidän vuoro tulee
let's wait until our turn comes
V: [M (body) R (body) screen
x , , , -----

Figure 4. Veli monitors Mirja and Roope's entrance.

$$M: \left[\begin{array}{l} ((steps\ in)) \\ (-----1-----) \\ V: \begin{array}{l} M\ (face) \\ x \end{array} \end{array} \right.$$

$$V: \begin{array}{l} M\ (body) \quad \text{corridor} \\ (-----1-----2-----) \end{array}$$



Veli's eye-tracker view

Veli's eye-tracker view

Accepted Manuscript

Figure 5. Transcribed interaction between Roope and Mirja.

1 M: *((waves hands in the air))*
 ↑molemmilla käsillä Roope
 with both hands Roope

R: *screen* *M*
 -----x

2 M: *((waves hands))*
 kato
 look

R: *M (face)*

3 M: *((waves hands))*
 ota nappaa molemmilla käsillä
 take catch with both hands

R: *wall M (hand) wall*
 ----x-----

Figure 6. Roope responds to Mirja's prompt by shifting his gaze to her.

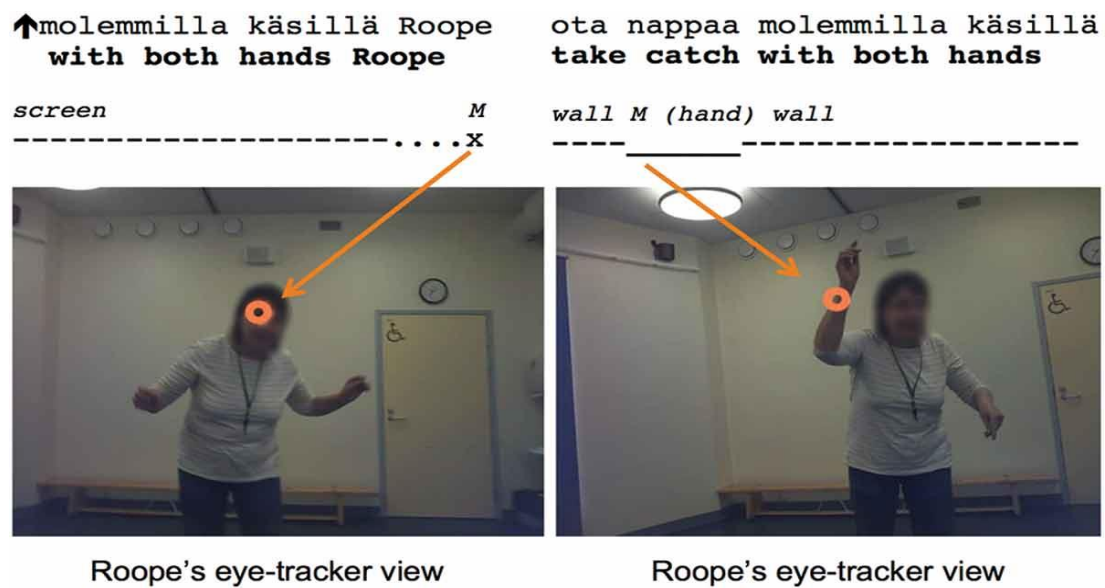


Figure 7. Transcribed interaction between Matti, Kaisa and Tommi.

1 M: *screen*
 ---.....
 (-----1)

2 M: *T (body)* *screen*
 X-----
 (**unclear talk**)

T: *screen M* *screen*
 ---...X-----

K: molemmat kädet saat ottaa käyttöön
you can use both hands

3 M: *((puts hands together to catch objects))*

Accepted Manuscript

Figure 8. Matti shifts his gaze to Tommi.

$\left[\begin{array}{l} T \text{ (body)} \\ x \\ M: \text{ (unclear talk)} \\ \text{screen} \\ T: \text{ --...} \\ K: \text{ molemmat} \\ \text{both} \end{array} \right.$



Camera view of Matti



Matti's eye-tracker view

Figure 9. Transcribed interaction between Roope and Tommi.

((smiles))

 screen
 --.

1 R: pum
boom

2 R: [*T* (*body*)
 x_____

(-----1)

T: [R_____

3 T: hyvin menee
 (it's going) well

Figure 10. Roope utters pum ('boom') and shifts his gaze to Tommi.

