Spring 2013

Observing Joint Attention in Children with Autism: The Development of a Video Coding Protocol

Jessica Hooker
University of Colorado Boulder

Follow this and additional works at: https://scholar.colorado.edu/honr_theses

Recommended Citation
Undergraduate Honors Theses, 383.
https://scholar.colorado.edu/honr_theses/383

This Thesis is brought to you for free and open access by Honors Program at CU Scholar. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Observing Joint Attention in Children with Autism: The Development of a Video Coding Protocol

Jessica Hooker
Dept. of Speech, Language, and Hearing Sciences
University of Colorado-Boulder

April, 2013

Thesis Advisor:
Amy Thrasher, Dept. of Speech, Language, and Hearing Sciences

Defense Committee:
Amy Thrasher, Dept. of Speech, Language, and Hearing Sciences
Dr. Kathryn Arehart, Dept. of Speech, Language, and Hearing Sciences
Dr. Janette Klingner, Dept. of Education
Dr. Christine Yoshinaga-Itano, Dept. of Speech, Language, and Hearing Sciences
Joint Attention Video Coding Protocol

<table>
<thead>
<tr>
<th>Table of Contents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract ..................................................................</td>
<td>3</td>
</tr>
<tr>
<td>Introduction ................................................................</td>
<td>4</td>
</tr>
<tr>
<td>Background ..................................................................</td>
<td>4</td>
</tr>
<tr>
<td>Methods ....................................................................</td>
<td>15</td>
</tr>
<tr>
<td>Reliability ..................................................................</td>
<td>20</td>
</tr>
<tr>
<td>Discussion ..................................................................</td>
<td>21</td>
</tr>
<tr>
<td>Conclusion ..................................................................</td>
<td>30</td>
</tr>
<tr>
<td>References ..................................................................</td>
<td>31</td>
</tr>
<tr>
<td>Appendix A ..................................................................</td>
<td>36</td>
</tr>
<tr>
<td>Appendix B ..................................................................</td>
<td>45</td>
</tr>
</tbody>
</table>
Joint Attention Video Coding Protocol

ABSTRACT

Children with Autism Spectrum Disorder (ASD) exhibit deficits in joint attention behaviors. As joint attention is considered a pivotal skill, it has been recommended that intervention programs target the development of joint attention, but there remains a lack of comparative research on the efficacy and effectiveness these intervention programs. Informal observational measures have been acknowledged as an ecological valid and necessary measure for assessing intervention outcomes in individual children with ASD and in intervention research, but a standardized methodology has yet to be developed. This paper describes the development and preliminary reliability assessment of the Joint Attention Video Coding Protocol, a systematic behavioral coding protocol designed for observing joint attention behaviors of young children with ASD in natural social contexts.
INTRODUCTION

Children with autism spectrum disorder (ASD) demonstrate wide variety in the presentation of learning profiles. Core areas of challenge include qualitative impairments in social interaction and communication in addition to showing restricted and repetitive patterns of behaviors and interests (American Psychiatric Association [APA], 2000). A unique struggle with social communication, however, is the most distinctive feature of ASD. Joint attention, a foundational component of social communication, has been recognized as a pivotal skill to be targeted in interventions for children with ASD. Due to a lack of appropriate comparative assessment tools, research on the effectiveness of the various joint attention intervention programs available is still lacking. The present paper examines the issues in the development of a protocol for systematic observation of joint attention in natural social environments, designed to serve as a comparative assessment tool for researching intervention programs.

BACKGROUND

Dating back to Kanner’s (1943) seminal case studies, a significant amount of research has been established on the social development and social communication abilities of children with ASD. While there is not any one social communication behavior that is necessary for diagnosis, children with ASD consistently show strong commonalities in the impoverishment of joint attention behaviors (National Research Council [NRC], 2001; American Speech-Language-Hearing Association [ASHA], 2006). Joint attention refers to the coordinating and sharing of attention between an object or event and another communicative partner or partners (Whalen & Schreibman, 2003; Prizant, Wetherby, Rubin, & Laurent, 2003; Mundy, Sigman, Ungerer, & Sherman, 1986; Loveland & Landry, 1986).
Development of Joint Attention

Fundamental behaviors of joint attention include eye contact, gaze shifting between a social partner and an object, and pointing gestures. As a child continues to develop, these behaviors are mastered and advanced through the combination of these behaviors with language. Eye contact, the earliest and most prominent form of joint attention, makes an appearance in the very beginning of an infant’s life. Within 5 days after birth, infants show a significant preference for images of faces with direct eye gaze over faces with an averted gaze (Farroni, Csibra, Simion, & Johnson, 2002). Around 6 months of age, infants begin to reliably follow another’s gaze to the correct side of the room and by 18 months of age, infants are able to consistently follow the direction of another’s gaze and correctly pinpoint the intended target, regardless of its location along the visual path (Corkum & Moore, 1995). Within the first 18 months, infants also develop the ability to follow the points of others, following points to nearby objects first, before being able to follow points to more distant objects (Corkum & Moore, 1995). It is not until about 12 months, though, that infants begin producing their own primitive points, and shortly after begin coordinating eye contact with their points at 15 months (Desrochers, Morissette, & Ricard, 1995).

Within the research literature, joint attention has been segmented into various categories to more specifically describe its different dimensions (Mundy et al., 2007). The ability to follow the direction of a communication partner’s gaze and physical gestures is referred to as responsive (or responding to) joint attention. Initiated (or initiating) joint attention refers to the ability to use gaze shifting and physical gestures to direct the attention of other people as a means of spontaneously sharing experiences. Joint attention has also been categorized as imperative...
Joint Attention Video Coding Protocol

joint attention, that is joint attention used to serve an instrumental purpose, and declarative joint attention, that is joint attention used to serve a social purpose.

**Joint Attention and Language in Typical Development**

An important aspect of joint attention is its relationship to language development. First, language is primarily learned indirectly through social exposure, but children must have certain skills in order to adequately learn language. Joint attention provides infants and children with the means to attend to the social stimuli in their environment and the ability to relate and engage with other people. One of the earliest functions of joint attention in language development is “object highlighting,” which Bruner describes as the process in which a caregiver brings objects into a infant’s line of sight and labels them once the infant begin to show consistent eye contact with a caregiver (Bruner, 1983). Another form of object highlighting involves bringing an object the child is already focusing on into the space between the child and caregiver and then labeling it. As the infant develops and the ability to follow another’s line of vision or regard begins to appear, the opportunities for the caregiver to label objects expands greatly (Baldwin, 1991; Bruner, 1983). In both of these processes, the infant’s ability to maintain eye contact and the ability to follow another’s line of regard allows the infant to link the object that is the focus of attention to the word that is used to refer to the object (Kasari, Paparella, Freeman, & Jahromi, 2008). Language continues to develop through joint attention in a bidirectional manner between the child and caregiver. As infants learn what to expect when they follow another person’s joint attention cues, they are able to utilize this ability to learn more about the environment around them. When they point to an unknown object, they know that their communication partner will label the object for them (Baldwin, 1991).
Joint Attention and Social Interaction in Typical Development

In the context of social development, joint attention plays a critical role in a child’s ability to understand and relate to others. Joint attention skills in the first year of life allow an infant to begin understanding other people as intentional agents of communication (Tomasello, Kruger, & Ratner, 1993; Tomasello, Carpenter, Call, Behne, & Moll, 2005). This understanding enables the infant to realize three primary things: (1) others can choose to attend to specific things in their environment while ignoring others, (2) others may want the infant to attend to specific things in the environment, and (3) the infant is able to direct the attention of others to specific things in the environment with certain behaviors (Tomasello, 1995). This, in turn, facilitates more joint attention with others and serves as the foundational precursor to developing theory of mind. More definitely established by the age of four, theory of mind is the understanding of other people as mental agents with their own thoughts, beliefs, and perspectives about the world and their environment and that these are separate from an individual’s own thoughts and beliefs (Tomasello, 1995). Theory of mind is a critical skill necessary for establishing and maintaining relationships with others and for functioning socially within one’s own culture and environment.

Joint Attention in Children with ASD

Numerous studies have found that children with ASD have significant difficulties with many of the fundamental joint attention behaviors mentioned above (McArthur & Adamson, 1996; Mundy, Sigman, & Kasari, 1994; Charman et al., 1997; Hobson & Hobson, 2007; Wetherby et al., 2004). Leekam, Lopez, and Moore (2000) found that preschool children with ASD demonstrated significant difficulty following another’s gaze when compared to developmentally delayed children of similar nonverbal ability. The researchers also reported that
not one of the child participants with ASD gave spontaneous eye contact during the baseline session. Consistent with these results, Swettenham and colleagues (1998) have also found that in semi-naturalistic situations, children with ASD displayed not only fewer gaze shifts between both objects and people and between multiple people, but they spent significantly less time looking at others’ faces in general than both typically developing children and children who were developmentally delayed. In a larger and more recent study, Wetherby and colleagues (2007) used the Behavior Sample measures from *Communication and Symbolic Behavior Scales Developmental Profile* (Wetherby & Prizant, 2002) to describe several joint attention behaviors in children around the age of two. After comparing three cohorts of participants (children with ASD, children who are developmentally delayed, and typically developing children), the researchers observed that children with ASD produced significantly fewer spontaneous 3-point gaze shifts and gaze/point following behaviors than the typically developing children and the developmentally delayed children.

While much of the research has found eye gaze to be a core component of the joint attention deficit in young children with ASD, impairments in other important joint attention behaviors have been acknowledged (Dawson et al., 2004; Sullivan et al., 2007; Werner & Dawson, 2005; Wetherby et al., 2004). Chiang and colleagues (2008) found that young children with ASD displayed significantly fewer pointing behaviors than typically developing children half their age during social communication assessment activities. The children with ASD in this study also displayed significant impairments in the ability to follow points and to initiate joint attention when compared to typically developing children and developmentally delayed children. An earlier study by Stone and colleagues (1997) observed that children with ASD under the age
of four also displayed fewer instances of showing gestures, pointing, and were more likely to manipulate an adult’s hand in expressing requests.

**Impact of Joint Attention on Language in Children with ASD**

The considerable impact of impaired joint attention behaviors, though, is the true hallmark of autism spectrum disorder. As children with ASD demonstrate significant difficulties with fundamental joint attention behaviors, especially eye gaze, it is not surprising that these children demonstrate further difficulties in the other developmental milestones that rely on joint attention skills. Although children with ASD tend to show a more severe impairment in their declarative joint attention abilities than their imperative joint attention abilities, research indicates that the core deficit is likely their general lack of or difficulty in monitoring the attention of other people in relation to objects and events (Charman, 2003; Swettenham et al., 1998; Landa, Holman, & Garret-Mayer, 2007; Kasari, Sigman, Mundy, & Yirmiya, 1990).

The most robust impact of joint attention deficits identified by research has been the impact on language development (Toth, Munson, Meltzoff, & Dawson, 2006; Kasari et al., 2008; Leekam & Ramsden, 2006). In a longitudinal study of 15 children with ASD, Mundy, Sigman, and Kasari (1990) found gestural joint attention skills to be both significantly correlated to language scores during initial assessment and a significant predictor of language ability at the follow-up assessment 13 months later. The authors also reported that neither nonverbal communication scores nor IQ were significant predictors of the language development of the children in the study. Similarly, in the study mentioned earlier, Wetherby and colleagues (2007) found gaze and point following abilities, as well as initiated joint attention scores in children with ASD at about 24-months of age, to be significantly correlated with both nonverbal and
verbal developmental quotients determined by the *Mullen Scales of Early Learning* (Mullen, 1995).

A pivotal longitudinal study by Sigman and Ruskin (1999) also found several significant relationships between joint attention and language in children with ASD. In addition to reporting highly significant concurrent correlations between language skills and both initiative and responsive joint attention during the initial assessment around the age of 3, Sigman and Ruskin also found that responsive joint attention skills at intake were predictive of expressive language gains at follow-up 8 years later.

**Impact of Joint Attention on Social Interaction in Children with ASD**

As joint attention fundamentally encourages social engagement, its impact on social interaction is intuitive. In order to develop social relationships with others, a child must be able not only follow and share attention with others, that child must understand what that attention means in relation to the shared engagement. A study of older children with ASD (mean age=12:8) by Travis, Sigman, and Ruskin (2001) reported a significant positive correlation between both prosocial behavior and peer interaction. For this study, prosocial behavior was assessed during a structured laboratory situation and peer interaction was assessed from observations of playground recess time. The previously mentioned study by Sigman and Ruskin (1999) also found initiated joint attention skills of children with ASD at about age 3 to be predictive of the amount of social engagement with peers at follow-up 8 years later. A more recent study by Johnson, Gillis, and Romanczyk (2012) additionally reported significant positive correlations between both initiated and responsive joint attention and social orienting to an adult examiner using eye-tracking technology in children with ASD.
Joint Attention Video Coding Protocol

**Joint Attention in Intervention**

As a result of the research on joint attention and its impact on development, the NRC (2001) and ASHA (2006) have emphasized the importance of targeting joint attention skills in social communication intervention programs designed for children with ASD. More specifically, ASHA (2006) has identified certain types of instruction directly related to joint attention that should take precedence when working with children with ASD, including: (a) increasing spontaneous and functional communication; (b) facilitating play skills for peer interaction; and (c) generalization and maintenance of new skills in natural contexts. Intervention in natural learning environments has been also been recommended as the most effective type of intervention to promote gains in initiations and generalizations of communication (NRC, 2001; ASHA, 2006). These types of environments increase the ecological validity of intervention itself (ASHA, 2006; Merrell, 2001).

**Measuring Joint Attention**

Research supporting efficacy and effectiveness is a critical component of establishing a successful intervention program. Currently, there remains a significant lack of comparative research on intervention programs targeting social communication for children with ASD (Cunningham, 2012). One of the primary reasons for this lack of comparative research is due to a discrepancy between the clinical field and the research field in the types of assessment tools used to measure intervention outcomes (Cunningham, 2012). This has resulted in a large amount of variability in the types the outcome measures used for post-intervention assessment. A primary reason for this variability and inconsistency, Cunningham (2012) reports, is the lack of a “gold standard” for assessing intervention outcomes. In a review of 67 research studies on early interventions for social communication, Cunningham (2012) noted that in single-subject design
Joint Attention Video Coding Protocol

studies, researchers tend to use frequency counts of distinctly designed social behaviors
(behaviors defined specifically for the purposes of an individual study), while in group design
studies researchers have typically used author-developed definitions of behaviors and other
published measures. For example, one study defined an instance of joint attention as a shift in
gaze towards the general direction of the examiner (Macdonald et al., 2006), while another study
defined an instance of joint attention as shift in gaze towards the examiner’s eye region (Dawson
et al., 2004). These inconsistencies in outcome measurement prevent comparisons across
studies, creating a barrier to drawing valid conclusions about optimal intervention programs
(Lord et al., 2005; Cunningham, 2012). ASHA (2006) and the NRC (2001) have recommended
that a multiple-method approach is necessary in order to accurately and validly determine the
impact of intervention on children with ASD. This approach involves both the use of
standardized, norm-referenced tools, gathering information from questionnaires and/or interview,
assessing with formal observation, and observation of natural social environments (informal
observation).

**Formal vs. Informal Observational Measures**

Formal observational measures involve staged communicative contexts designed to elicit
specific behaviors from the child. In these settings an examiner typically follows a strict and pre-
determined routine to encourage both responsive and spontaneous behaviors from a child. One
such measure is the *Early Social Communication Scales* ([ESCS]; Mundy et al., 2003). The
ESCS is a structured observational assessment designed to measure nonverbal communication
skills in young children with ASD or a developmental delay and typically developing children. It
consists of a series of semi-structured tasks; each designed to elicit specific social
communication behaviors from the child, including initiating and responding to joint attention behaviors.

On the other hand, informal observational measures involve the observation of a child’s behaviors within natural contexts. However, Wetherby (2006) has recognized that almost no research is available on measuring the psychometric features of a parent-child interaction in natural contexts, despite its significant ecological validity. Of the few research studies that have begun to look at the discrepancies between staged communicative contexts and more naturalistic contexts, differences in the presentation of behaviors appears to change across contexts. Roos and colleagues (2008) found that young children with ASD exhibited more initiated joint attention behaviors during the administration of the ESCS than in a child-directed play period with an examiner, but exhibited more responsive joint attention behavior during the play period than during the ESCS. In a similar respect, Londono et al. (2010) reported young children with ASD demonstrated more joint attention behaviors overall during an unstructured play period with an adult examiner than during the administration of the ESCS. While this research has reported important findings on the influence of context, further research is needed to explore the interaction.

Advantages of Informal Observational Measures

While much research has been spent on the development and implementation of standardized tools within staged communicative contexts and questionnaire/interviews, little research has been conducted on standardizing methodologies for natural observation, despite the several key advantages it offers in assessing intervention outcomes (Merrell, 2001). One of the most advantageous aspects of behavioral observation is its ability to measure changes in the spontaneity and generalization of the targeted skills in the intervention (Cunningham, 2012;
Yoder & Symons, 2010). Another benefit of this methodology is its applicability in a variety of natural and intervention settings any number of times without influencing the intervention itself or other outcome measures. Cunningham (2012) also notes that behavioral observation is sensitive to subtle changes in behavior during intervention that standardized assessments may miss. Finally, Merrell (2001) additionally comments on the fact that naturalistic observation does not require specific and complex instruments or tests.

**Concerns with Informal Observational Measures**

Of the concerns regarding the use of behavioral observation as an outcome measurement identified by Merrell (2001), two appear to be most relevant to this paper: context specificity and the lack of social comparison data. Context specificity, or the impact a specific situation has on behavior, limits the ability to interpret the analyzed behavior in regards to intervention. It is very likely that children with ASD with behave much differently in a situation in which they are forced to interact with strangers than in situation in which they are interacting with familiar people. Along this same line, a child with ASD is likely to behave different in a group setting (*e.g.*, preschool) than at home with a caregiver. If only a single specific context is used for the observation of target behaviors, it is likely that these target behaviors will not be accurately represented in relation to all contexts (Cunningham, 2012). The lack of an appropriate social comparison group prevents the establishment of suitable criteria from which intervention success can be based. While the foremost importance of using behavioral observation as an outcome measures is to assess the changes in the behaviors of the child with ASD being observed, it is also critically important to have a social norm to assess where the child with ASD stands in relation to the behaviors of typical peers.
Joint Attention Video Coding Protocol

The Need for an Informal Observational Measure

Despite its potential limitations, using natural behavioral observation as a tool to assess intervention outcomes for children with ASD is still recognized as a necessary component for ecologically valid research on joint attention interventions and for assessing intervention outcomes in the clinical setting in addition to other standardized and norm-referenced measures (Merrell, 2001; ASHA, 2006; NRC, 2001; Cunningham, 2012). This paper describes the development of a video coding protocol designed specifically to observe and analyze the joint attention behaviors of children with ASD across a variety of natural and intervention contexts. This protocol, the Joint Attention Video Coding Protocol (JACP; Appendix A), intends to serve as a first step in creating an operationally defined behavioral coding system to be used in social communication intervention research. The details of the development of the JACP and the initial steps towards establishing reliability will be described.

METHODS

Review of the Early Social Communication Scales

The JACP was largely modeled after an abridged version of the Early Social Communication Scales (ESCS; Mundy et al., 2003). Through the use of the ESCS, three main dimensions of nonverbal communication, Joint Attention, Behavioral Requests, and Social Interaction, are measured through the observation of behaviors that are typically demonstrated by children between 8 and 30 months of age. Behaviors are segmented into Initiating or Responding behaviors and then further divided into “higher-level” or “lower-level” behaviors. For example, the behavior for alternating gaze between an object and the examiner, Alternate, is categorized as a lower-level Initiating Joint Attention behavior within the ESCS. Behaviors are observed during a series of structured tasks administered in a staged communicative context.
Joint Attention Video Coding Protocol

between the examiner and a young child. Each task involves various toys placed strategically on
a table or posters on the walls next to and behind the child in order to elicit interaction from the
child. The interactions are video recorded and coded later by the examiner. Basic video coding
for the ESCS consists of noting the frequency of each behavior defined within the manual.

This manual was chosen as a foundational model for the JACP for two primary reasons. First, the ESCS is one of the more widely used and supported measures for assessing the joint
attention behaviors of young children with ASD in research (Cunningham, 2012). Its widespread
use has also given it an established track record of reliability (Chiang et al., 2008; Toth et al.,
2006; Kasari, Freeman, & Paparella, 2006; Mundy et al., 2007). Second, this manual includes
core joint attention behaviors that children with ASD tend to have the most difficulty with,
including eye contact and following line of regard. Behaviors selected for the JACP were chosen
among those that were included in the ESCS.

Developing the JACP

The first step in the development of the JACP involved selecting behaviors to be defined
and implemented for the preliminary testing of reliability. A preliminary version of the protocol
included a total of twelve joint attention behaviors in an attempt to fully capture the possible
behaviors that might occur in a natural setting. In order to make initial reliability coding more
feasible, the number of behaviors was significantly reduced to four; three from Initiating Joint
Attention: Eye Contact, Alternate, and Point; and one from Responding to Joint Attention:
Following Line of Regard. For the initial development of the JACP, these behaviors were not
distinguished between “higher-level” and “lower-level” behaviors as they are in the ESCS. The
behaviors selected from the Initiating Joint Attention behaviors in the ESCS were chosen for two
reasons: (a) they are foundational joint attention behaviors that continue to be demonstrated
Joint Attention Video Coding Protocol

throughout the lifespan, and (b) they are the most commonly used behaviors in research assessing joint attention. *Following Line of Regard* was selected from the *Responding to Joint Attention* behaviors in the ESCS for slightly different reasons: (a) of the two possible behaviors, it appeared to more readily transferable to natural behavior observation and (b) the similarity to other responsive joint attention behaviors used in research (*e.g.*, *following eye gaze/head turn*).

Once the initial target behaviors were chosen, the next step was to adjust the original ESCS definitions to be more applicable to natural social settings. A critical step in making natural observation an effective research and assessment tool is the construction of operational definitions to be applied in a systematic and controlled way. As the ESCS uses structured tasks specifically designed to elicit joint attention interaction, the operational definitions for the selected behaviors directly reflect the expected behavioral response of the related task. An example is *Follows Line of Regard*. In the ESCS this behavior is only coded during one of the nine possible tasks in which the tester points to designated posters in the testing room. Basic adjustment of the definitions for the JACP consisted of removal of any task-specific components and broadening the definitions such that they were more representative of natural behavior.

After this preliminary round of definition adjustment, video clips of young children with ASD were used to informally assess the applicability and appropriateness of the new operational definitions. These video clips included children with ASD and typically developing children similar in age participating in different group contexts during the Story of Friendship program. The Story of Friendship program is an eight-session social communication intervention program for children with ASD at the University of Colorado Boulder. This program supports social communication and joint attention skills in young children with ASD by incorporating: (a) typically developing peers, (b) naturalistic play environments, (c) video modeling and (d)
systematically planned opportunities for joint attention in turn taking interactions (Thrasher et al., 2012). Each child with ASD participating in the program receives both individualized support from a graduate clinician within the systematically planned turn taking interactions and opportunities to generalize skills in free play and small group activities.

After several observations of multiple children during various sessions of the Story Friendship program, operational definitions were refined further to more closely describe target behaviors as they occurred in the natural settings. Even after several revisions, many of the definitions in the JACP do not differ significantly from the original definitions from the ESCS, with the exception of some key modifications:

**Eye Contact**

In the ESCS, *Eye Contact* may only be coded when a child is touching or manipulating an inactive object. The JACP broadened this definition to allow *Eye Contact* to be coded for instances in which the child is holding an object in general and during an active event involving the child and at least one other person. Refer to Appendix A for complete definition.

**Follows Line of Regard**

As mentioned earlier, the ESCS limits the *Following Line of Regard* behavior to only be coded during one specific task. Another feature of this operational definition that narrows its use is the requirement of a type of cue used to elicit this response. According to the ESCS, in order to elicit this behavioral response, the tester must turn her entire torso towards the intended stimulus while simultaneously calling the child’s name three times and using a “short-arm point” in which her elbow is anchored to her side. In constructing the definition of *Follows Line of*
Joint Attention Video Coding Protocol

*Regard* in the JACP, the potential behaviors of the interaction preceding the child’s response were expanded to allow for any clear physical gesture, with or without a vocalization, that are used to direct the child’s attention toward the intended object or event (*e.g.*, *pointing* or *head turn*). Refer to Appendix A for complete definition.

**General Coding Procedure**

As in the ESCS, an event-based coding scheme is used for observation with the JACP. Event-based coding schemes measure the total frequency of occurrence of target behaviors in a specified time period. This type of measurement system is recommended in behavioral observation because it provides more information about a child’s behaviors than rating scales and checklists, it is one of the more commonly used systems in behavioral observation research, and most importantly, it is sensitive enough to detect small changes in behavior between observations (Yoder & Symons, 2010; Merrell, 2001). It has been largely acknowledged and reiterated that meaningful outcome measures, including gains in spontaneous initiations and generalizations across environments, are crucial to the assessment of a child’s response to intervention (NRC, 2001; ASHA, 2006; Cunningham, 2012). The other commonly used measurement system used in behavioral observation research is an interval-based measure, in which the total duration of each target behavior is the focus of observation. This coding method was not chosen for the JACP due to the very brief durations of some of the behaviors (*e.g.*, *eye contact*) included in the JACP.

A major difference between the ESCS and the JACP is the differentiation of interaction partners. Most research on joint attention in children with ASD has typically focused on the behaviors occurring within adult-and-child interactions. A critical aspect of accurately assessing any changes in the initiations and generalizations of joint attention behaviors of children with
Joint Attention Video Coding Protocol

ASD is the consideration of “with whom” the children are expressing their joint attention behaviors in their natural social environments. Because research has been largely focused on measuring the frequency of specific joint attention behaviors, structured situations involving adult examiners and caregivers or observations of child-parent interactions have been studied are most commonly used. While these methodologies provide a significant amount of information about a child’s abilities in elicited situations with one adult and in natural interactions with a parent, they fail to assess two essential components of social communication: peer interaction and interaction within group settings. The JACP seeks to address these aspects of social communication by distinguishing to whom joint attention behaviors are directed: Adults or Peers. For this initial development of the JACP, an Adult was considered as any persons that are significantly older than the child being observed and a Peer was considered as any other child within the observational context.

RELIABILITY

An essential step in establishing the JACP as a valid tool for assessing the joint attention behaviors for children with ASD is obtaining high reliability. As this paper seeks to detail the development of the JACP, a description of a preliminary intraobserver agreement assessment follows.

Test Period of Observation

One period of observation from the Story of Friendship program was chosen for the purpose of determining an initial intraobserver reliability score. This period, Songs and Goodbye, was video-recorded during the sixth session of the program. Songs and Goodbye follows the last period of free play before the children are released to their caregivers at the end of the session. The main component of this group interaction setting is the children’s song, “If You’re Happy
and You Know It.” During this song, each child was given the opportunity to pick from one of six emotion cards for her turn and then chose how she would like the group to express that emotion (e.g. during “happy,” jumping up and down or clapping). This period was chosen for video observation because it very closely resembles a small group setting and activity that is frequently found within preschool or childcare settings. Although this period is part of the intervention program itself, it is not a period of direct, systematically planned opportunities for intervention for the children with ASD. Rather, it primarily functions as an opportunity to generalize within a group setting the skills in joint attention previously systematically addressed in a peer dyad context.

Results

Intraobserver agreement was assessed via video coding of one seven-minute segment of video during the group activity described above. The joint attention behaviors of one child with ASD and one typically developing child were coded from the video by the same video coder on two separate occasions. Agreement scores were calculated for each of the four JACP behaviors individually. A high intraobserver agreement score indicates coder consistency. Agreement scores for the child with ASD were as follows: Eye Contact, .93, Alternating Gaze, .89, Point, .98, Follows Line of Regard, .98. Agreement score for the typically developing child were as follows: Eye Contact, .98, Alternating Gaze, .79, Point, 1.0, Follows Line of Regard, .98.

DISCUSSION

Reliability

Intraobserver agreement scores were quite high for the initial round of reliability coding. While these high reliability scores are very promising, it is possible that the scores are slightly inflated for two possible reasons. First, it is likely that many of the behaviors with a perfect or
near-perfect reliability score obtained such a score due to very low frequency counts during the video. For example, the reliability score for Follows Line of Regard for the child with ASD was a .98 because this behavior was observed only 1-2 times during the entire video clip. This was also observed for the frequency counts of Eye Contact, Point, and Follows Line of Regard during video coding for the typically developing child. See Appendix B for frequency counts and reliability scores of each behavior during both coding sessions. Second, the video coding was done by the author of the JACP, possibly influencing the reliability outcomes. However, with these concerns in mind, the high reliability scores should still be considered a positive outcome for the initial development of the JACP.

Although the purpose of video coding was to determine an initial intraobserver reliability score, differences in the frequencies of certain behaviors between the child with ASD and the typically developing child should be noted. The most obvious difference in behaviors between the two children occurred for Alternating Gaze, with the typically developing child exhibiting roughly 30 more Alternating Gaze behaviors than the child with ASD. This considerable difference in frequency of occurrence may provide important information about the joint attention abilities of the child with ASD within a peer-group context and Alternating Gaze may be an important behavior that needs to be observed and assessed further for this child. However, because the child with ASD and the typically developing child were not matched for chronological age or mental age, any possible inferences about the child with ASD’s joint attention ability have insufficient support.

Challenges with Developing the JACP

Developing a video coding protocol to be used in a variety of social settings, including intervention settings, presented many challenges. These challenges were realized through the
Joint Attention Video Coding Protocol

complex process of the protocol development and reliability testing phases of the preliminary
and current versions of the JACP. One of the greatest challenges was developing operational
definitions that were both broad enough to be used in a variety of settings and narrow enough to
capture the essential features of the behavior. Many operational definitions used in research to-
date are designed explicitly for situations in which the target behavior is directly elicited through
cuing. However, many of the cues used to elicit joint attention from a child with ASD in these
situations are not necessarily found in a natural social setting, especially if there is a group of
people (e.g., preschool/classroom setting). For example, a commonly measured joint attention
behavior in research is the ability to follow a head turn. In a research setting, the child is usually
placed in a face-to-face environment with an experimenter most commonly referred to as the
“joint attention paradigm” (Corkum & Moore, 1995). In this paradigm the experimenter will
complete a sequence of multiple trials in which she first establishes eye contact with the child
and then turns her head toward the designated stimulus in an attempt to elicit a head turn from
the child. While this is a direct measure of the child’s ability to respond appropriately to a head
turn cue, it is not necessarily an accurate representation of following a head turn in a natural
social setting.

There are several other influential factors within a natural social setting that affect a
child’s ability to follow the head turn of another interaction partner. First, it cannot necessarily
be expected that a communication partner is going to establish eye contact directly prior to
making a head turn. Rather, after observing numerous interactions in a natural setting, it was
observed that it is much more likely that an interaction partner will periodically check in with the
child to confirm joint attention, and assume that the child will follow her head turn if she looks at
a new stimulus out of the child’s line of sight. Also, unlike the typical research setting, an
interaction partner is likely to accompany her head turn with a vocalization, reducing the need to establish eye contact prior to her joint attention cue.

Another major challenge in developing the JACP was the impact of a group setting on joint attention. This brought up many unique challenges in considering how to revise the current operational definitions to be more applicable to a setting with multiple people. After its preliminary development, the current definitions in the JACP still do not accurately capture the joint attention behaviors as they occur in a group context. Instead, some of the joint attention behaviors appear to take on a similar, yet noticeably different form. This was most apparent for two of the behaviors in the JACP: *Alternating Gaze*, and *Eye Contact*.

The preliminary definition of *Alternating Gaze* from the JACP, modeled after the ESCS’s definition, was a gaze shift between an object and another person when the object is currently the focus of attention between the child and that other person. During a group activity though, an object may be active (currently the focus of attention) between two other interaction partners. When the object is active in the hands of another, the target child of observation is placed into a situation of *indirect* joint attention. Indirect joint attention occurs when the child is not currently the center of attention in a group interaction, but must continue to rely on joint attention behaviors to maintain social engagement. An example of this occurred in the *Songs and Goodbye* video used for reliability analysis. As each child is given the opportunity to choose an emotion card, the cards are active in his hands, while the rest of the group waits for him to choose an emotion. During this period of time, the rest of the children are more indirectly involved in the activity. A child’s gaze shift between the emotion cards in the hands of a peer and his face is an important behavior that the current definition in the JACP does not wholly capture. In its current form, the definition in the JACP implies that *Alternating Gaze* may only occur when an object is
active between the child being observed and another person, not necessarily when it is active between two other people or even between another person and the group as a whole.

Another aspect of alternating gaze that the current JACP does not address is the shifting of gaze between the faces of multiple interaction partners. This was also observed during the *Songs and Goodbye* activity, especially in the behavior of the typically developing child. Shifting of gaze between the faces of other group members is a critical aspect of maintaining joint attention and participation in the group activity. In larger group settings, and as children grow older in general, social interactions involving objects are likely to become less frequent than interactions involving an activity or conversation. This results in an increased importance of attending to the faces of all the interaction partners as a means of maintaining joint engagement. Further research is needed to more accurately assess these behaviors and their relative importance in intervention for younger children with ASD.

In relation to the point above, the behavior of *Eye Contact* presents similar challenges to reflect upon. Depending on the type of group activity, the amount of eye contact a child engages in may vary greatly. During a group activity, it is likely that the typical child will show a higher frequency for looking at the faces of the other group members than the child will for making actual eye contact with other people. This was observed in several interactions during informal video observation. According to the JACP definition, this would not be coded as eye contact. However, this behavior, commonly referred to as social referencing, still provides important information about the child’s joint attention during a group activity. In order to remain engaged with the other members during a group activity, the child needs to be aware of the intentions and attitudes of others, and that attention to faces of other group members may provide this information. In a group activity, it is likely that each member with be making only brief eye
contact with the other members to confirm joint attention and much more time will be spent reading the faces of the other members to maintain a joint engagement in the activity.

A possible solution for assessing and incorporating the differences in behaviors between group settings and individual settings is to construct a separate set of operationally defined behaviors to be implemented specifically for group contexts. These definitions would provide much more information about a child’s level of engagement during group activities, such as the frequency of gaze shifts between the faces of group members. However, further research on the dynamics of group interaction is needed in order to pursue this line of development of the JACP.

A third challenge with attempting to systematically observe joint attention in natural social setting is the time intensiveness of video coding. This was demonstrated in the numerous hours spent coding for intraobserver reliability from a single seven-minute video. In a single coding a session, a significant amount of time must be spent on both deciphering the behavior of the target child as well as the behaviors of the other people in the observational setting. If the observation is of a group setting, this markedly increases the amount of time needed to code a single video. The coder must be aware of the behaviors and intent of the interaction partners in the video in order to accurately understand how the child is using joint attention to interact with others in that environment. A possible way to alleviate this issue and increase the practicality of using natural observation would be to choose social settings in which the child is only interacting with one to two other people. However, limiting natural observation to only these settings creates a skewed representation of the child’s social interactions and cannot completely describe the status of the child’s joint attention abilities. An alternate solution is the implementation of multiple coders, each coding a small number of videos individually. This is the current practice
of video observation research, but the drawback is that this procedure will require a significant amount of time to train observers before they can code with acceptable reliability (Merrell, 2001).

A fourth major potential challenge of implementing the JACP is context control. While the JACP is designed as a tool for assessing joint attention behaviors in a variety of natural social contexts in order to measure intervention outcomes for children with ASD, observational context is a variable that must remain relatively controlled. The observational context must remain consistent in order to maintain the validity of the JACP in not only measuring the changes in the behaviors of a single child across time, but as well as in comparing changes in the behavior of children across studies (Merrell, 2001). For the assessment intervention outcomes for an individual child with ASD, at least two observational contexts are recommended to be selected for coding in order to more accurately capture the child’s behavior. These contexts should be representative of the child’s typical social routine (e.g., preschool) and remain as consistent as possible across time. When considering the use of the JACP in comparative research, observational contexts that are likely to be similar across most children’s experiences should be selected. Preschool or classroom settings are great examples of natural social settings that many children typically participate in. Within those settings, a classification system of group or small group contexts may be necessary in order to be able to compare across sufficiently similar contexts.

Consideration of video camera positioning and angles for recording is another logistical challenge needing to be further addressed within the JACP. As previously mentioned, within group settings all of the people within the observational context are likely to convey important information about the joint attention behaviors of the child being observed. Therefore it is necessary to maintain clear visibility for all participants in the group. This challenge was
highlighted during reliability coding. Because of the spacing of the people within the video, the location of the camera, and the capabilities of the camera, not all of the group members were clearly visible within the viewing frame. Some of the group members were completely outside of the viewing frame or positioned in such a way that their faces were not clearly visible. This resulted in a loss of information and certain points within the video, negatively influencing the ability to accurately code the child’s behaviors. In thinking about observing a group setting, it is important to consider how the visibility of other people impacts the coding a child’s behavior. Situations involving a particularly large group of people (e.g., classroom) or those involving a large amount of movement by all people (e.g., playground) are going to be more difficult to film in regards to camera placement and viewing frame, as opposed to dyadic settings involving only two people (e.g., caregiver and child).

**Next Steps**

The first major next step towards improving the ecological validity of the JACP will be editing and refining the definitions of its current behaviors. As previously described, there are several influential variables that need to be addressed. One influential factor the JACP does not appropriately account for is the impact of another person’s physical movement on the child’s behavior. In order to obtain an accurate depiction of a child’s joint attention behavior, a coder must be certain that the child is actually engaging in joint attention and not simply responding to random visual stimuli. For example, a child is likely to glance at the face of another person, simply because that person has started scratching her nose, not because she is engaging in some type of joint attention behavior. This especially affects coding of *Eye Contact* and *Alternating Gaze*. The JACP procedures need be revised to explicitly inform coders how to distinguish between true joint attention behaviors and behaviors that may have been elicited by movement.
Another possible alteration needed specifically for definition of *Point*, is accounting for ambiguous points. It was observed during the preliminary video coding that during some of the points exhibited by the child with ASD the intended interaction partner was unclear. These points were considered ambiguous because the child did not share eye contact with another person while using the point gesture. Possible adjustments to this definition include incorporating a symbol that demarcates ambiguous points from points with a clear intended interaction partner. Ambiguous points should still be included because they provide important information about the quality of a child’s joint attention. In observing a child’s ambiguous points, one is able to infer that it is likely the child needs support in making eye contact with another person during points, but that child does not necessarily need support with the actual pointing gesture.

An important second step in the ongoing development of the JACP is the inclusion of more joint attention behaviors, both initiated and responsive. This is necessary in order to more completely capture a child’s joint attention behaviors in natural social settings. Unlike most staged communicative contexts, not every natural social setting necessarily elicits all types of joint attention behavior. Certain joint attention behaviors aren’t necessary or appropriate in specific situations. For example, the typically developing child observed for assessing reliability did not express any points during the video for coding. This was likely because she was not in a situation in which she needed to direct another’s attention through pointing. In considering this, it is important to include a variety of possible joint attention behaviors in the JACP.

Additional potential behaviors from the ESCS to be redefined and incorporated into the JACP include *Shows* or *Bid to Caregiver*. These behaviors involve the child presenting an object to another person or her caregiver. Other possible behaviors not included in the ESCS, but
Joint Attention Video Coding Protocol

previously included in the preliminary version of the JACP and observed during informal observations include spontaneous gives, gives in response to requests, and shows in response to requests. The definition for Follows Line of Regard in the current JACP may also be broken down into multiple responsive joint attention behaviors with gesture-specific definitions, such as Follows Point and Follows Head Turn. A fourth potential adjustment to the JACP would be to distinguish between behaviors demonstrated with coordinated eye contact and behaviors expressed without eye contact (e.g., Point or Point + Eye Contact).

Lastly, one of the most important steps for establishing the JACP as an ecologically valid tool for assessing the joint attention behaviors of children with ASD in regards to intervention outcome is establishing high inter-rater reliability in a variety of social contexts with several other children. Classroom and playground settings would be preferred settings to further develop the JACP and high inter-rater reliability in regards to both peer interaction and group settings; however, these settings also create difficulties in consideration of camera positioning and visibility. Observing intervention settings and dyadic contexts are also important for assessing validity and reliability.

CONCLUSION

Despite the previously mentioned challenges with developing the JACP and the suggested next steps, further research and development of the JACP will likely prove worthwhile in future intervention research for children with ASD. Addressing these challenges and completing these steps will occur in an ongoing process of revisions to the JACP as more information is obtained from the informal observations of children with ASD and typically developing children in diverse social interactions in various natural environments.
REFERENCES


Joint Attention Video Coding Protocol


Joint Attention Video Coding Protocol


Joint Attention Video Coding Protocol


Appendix A:

Joint Attention Video Coding Protocol
Joint Attention Video Coding Protocol

Jessica Hooker
University of Colorado-Boulder
2013
BRIEF BACKGROUND

This protocol has been developed for the purpose of systematically observing the initiated and responsive joint attention behaviors of young children with Autism Spectrum Disorder (ASD) using video analysis. It has been designed for use in both natural social settings and intervention settings in order to serve as a supplemental tool to be employed when assessing social communication intervention outcomes for young children with ASD. In order to gain the most information from natural social settings, this protocol includes measures for identifying two different types of communication partners: adults and peers. Behaviors of interest included in this protocol are nonverbal communication behaviors that typically appear in children between 8 and 30 months of age.* The operational definitions in this protocol have been modeled after those found in the Early Social Communication Scales* and similar operational definitions found in previous research.

GENERAL CODING GUIDELINES & PROCEDURES

Coding will be conducted via video recorded observations. Basic coding consists of noting the frequency of occurrence of each Initiated Joint Attention (IJA) and Responsive Joint Attention (RJA) behavior that occurs. These behaviors are defined in the next section and examples of each behavior have been provided.

Code behavior using the coding sheet provided below. For each behavior, mark each corresponding box with one of two letters to establish the communication partner: A for adult or P for peer. The coding sheet has been provided below.

An adult should be considered as any person who is significantly older than the child being observed and any other child within the observational context should be considered a peer.

Do not code any behaviors in which the other person’s line of sight is ambiguous. It is better to not rate a behavior than to code it with insufficient information.

If the child leaves the viewing frame or is obscured, make a note on the coding sheet, as well as the timecode of occurrence. If the child re-enters the viewing frame, record the timecode of occurrence.

During video coding, only one behavior should be coded at a time in order to allow for the highest sensitivity to the child’s behavior.

A description of the context in which the video observation is taking place should be recorded in the Notes section of the Coding Sheet (e.g., Number of people present).

BEHAVIORS OF INTEREST

Initiated Joint Attention

1) **Eye Contact:** The child makes eye contact with another person (adult or peer) during an *active event,* in response to another person’s vocalization directed at the child, or while touching an object. The person the child is looking at must also be looking at the child in order for the child to receive this code.

An *active event* will be defined as any activity involving the child and at least one other person that is the current focus of attention (e.g., *playing a game*).

Code *Eye Contact* if the child makes eye contact while touching another person. Do not code *Eye Contact* if the child’s touch is in protest or aggressive.

Eye contact may occur quickly and briefly. Therefore, vigilance and an alert state are required for reliable coding. Do not code *Eye Contact* if the child’s behavior was elicited by movement or by talking that does is not directed towards the child.

Code *Eye Contact* even if it occurs in combination with another joint attention behavior (e.g., *Point*), except for when it occurs during a gaze shifting sequence involving an object. This should be coded as *Alternating Gaze*.

Note: The video recording should enable coders to reference the general position of the eyes of the other people and reliably determine when the child is looking at the upper region of another person’s face (definition of eye contact) as opposed to looking at the lower portion of the person’s body. If the child’s behavior is ambiguous, be conservative and do not count it as eye contact.

Eye Contact Example 1:
Eye Contact Example 2:

2) *Alternating Gaze*: The child alternates looking between an *active object* and another person’s face. To receive this code, the child must shift his/her gaze from the object to the other person’s face or vice versa. The child does not need to then gaze back at the object (or person) to receive this code.

An *active object* will be defined as any toy or object that is currently the focus of attention between the child and at least one other person.

When the child repeatedly shifts his/her gaze between an object and other person’s face, count each gaze back to the other person’s face as an *Alternating Gaze*.

*Alternating Gaze* will most typically occur when an object is active on the floor or in the hands of another person, but can also be coded if the child looks up to another person’s face after an object becomes active in his/her own hands.
3) **Point**: With a clear articulation of the index finger, the child points to an active object or to any other unobtainable object or event.

*Point* should only be coded when the index finger is extended and adjacent fingers are noticeably inclined downward, or away from the index finger toward the palm. Reaching with fingers extended and an upward inclination of the index finger should not be coded as a Point. Pushing or scratching an object with one finger should not be considered a Point. However, touching an object with an index finger with the hand in a pointing configuration should be considered a Point.

Points to objects beyond the frame of view of the camera should be coded as a **Point**.

Points occurring as an imitation of another person’s point to an object or event should not be coded as a **Point**.

A **Point** may occur with or without simultaneous eye contact.

---

*Mundy et al., 2003*
Point Example 1:

Point Example 2:
Responsive Joint Attention

1) *Follows Line of Regard:* The child clearly follows another person’s line of regard by immediately (within several seconds) turning his/her head to the appropriate area.

*Line of Regard* will be defined as any instance in which another person makes a physical gesture (e.g., *pointing* or *head turn*) to direct the child’s attention toward an active object or event. This physical gesture may or may not be accompanied by a vocalization (e.g., “*look*”).

*Follows Line of Regard* Example 1:
# CODING SHEET

**Subject:**

**Video:**

**Observation Setting:**

**Coder:**

**Date Scored:**

A = Adult  
P = Peer

<table>
<thead>
<tr>
<th>Eye Contact (IJA)</th>
<th>Alternate (IJA)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Eye Contact Icons" /></td>
<td><img src="image2.png" alt="Alternate Icons" /></td>
</tr>
</tbody>
</table>

Total A: ____  
Total P: ____  
Total: ____

<table>
<thead>
<tr>
<th>Point (IJA)</th>
<th>Follows Line of Regard (RJA)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image3.png" alt="Point Icons" /></td>
<td><img src="image4.png" alt="Follows Line of Regard Icons" /></td>
</tr>
</tbody>
</table>

Total A: ____  
Total P: ____  
Total: ____

Notes:
Appendix B:

Intraobserver Reliability
## INTRAOBSERVER RELIABILITY SCORES

<table>
<thead>
<tr>
<th></th>
<th>Eye contact</th>
<th>Alternating Gaze</th>
<th>Point</th>
<th>Follows Line of Regard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child with ASD</td>
<td>.93</td>
<td>.89</td>
<td>.98</td>
<td>.98</td>
</tr>
<tr>
<td>(Frequency Code 1)</td>
<td>(7)</td>
<td>(8)</td>
<td>(7)</td>
<td>(2)</td>
</tr>
<tr>
<td>(Frequency Code 2)</td>
<td>(8)</td>
<td>(9)</td>
<td>(8)</td>
<td>(2)</td>
</tr>
<tr>
<td>Typically Developing Child</td>
<td>.98</td>
<td>.79</td>
<td>1.0</td>
<td>.98</td>
</tr>
<tr>
<td>(Frequency Code 1)</td>
<td>(1)</td>
<td>(38)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>(Frequency Code 2)</td>
<td>(2)</td>
<td>(39)</td>
<td>(0)</td>
<td>(1)</td>
</tr>
</tbody>
</table>