MOTIONESE: SUBJECT TO PREFERENCE?

By
Natalie Giles Brezack

A THESIS

Presented to the Department of Psychology
And the Honors College at the University of Oregon
In partial fulfillment of the requirements
For the degree of
Bachelor of Science

June 2013

Primary Advisor: Dare Baldwin, Ph.D.
Secondary Advisor: Jennifer Mendoza, M.S.
Clark Honors College Advisor: Louise Bishop, Ph.D.

An Abstract for the Thesis of

Natalie Giles Brezack for the degree of Bachelor of Science

In the Department of Psychology to be taken June 7, 2013

Title: MOTIONESE: SUBJECT TO PREFERENCE?

Approved:_________________________________

Professor Dare Baldwin, Ph.D.

Research by Kuhl, Coffey-Corina, Padden, and Dawson, 2005, demonstrated that typically developing infants prefer to listen to "motherese" speech over a non-speech analog. In contrast, children with autism spectrum disorder show the reverse preference, and the degree to which this is true predicts their progress in phonological development. The current research investigates possible parallels to these findings in children's processing of human action; specifically, whether developmental skills relevant to autism symptomatology (e.g., executive function (EF) and theory of mind (ToM)) predict the degree to which children a) prefer "motionese" versus a non-action analog (or the reverse), and b) their sophistication in extracting structure within intentional action. Preliminary regression results based on participation from 46 preschoolers revealed both EF and ToM skills independently predicted degree of preference for motionese versus a non-action analog. Motionese preference was also a significant predictor of action segmentation skills. Should these findings be borne out in the full sample, they would point to important links between the development of language and intentional action processing, and they may have implications for designing interventions for children developing atypically.

ACKNOWLEDGEMENTS

The author expresses great appreciation to Professor Dare Baldwin for her guidance, support, expertise, and advice through the entire thesis process. In addition, I also thank Jennifer Mendoza for extensive editing, computer programming, and moral support, and Professor Louise Bishop for her continued encouragement and enthusiasm. I also express special thanks to my supportive family, patient boyfriend Bradley Frizzell, and friends Ellen Ingamells, Julia Stafford, and Elise Downing for listening to/putting up with discussions of this project. Particular appreciation is also due to members of the Baldwin lab, without whom this project would not have been possible: graduate students Rose Maier, Robbie Ross, and Jessica Kosie for their expertise and advice; undergraduate research assistants Erica Chesnutt, Alia Mowery, Kelsey Brewer, Camille Estabillo, and Audrey Cloe, who contributed innumerable hours to helping run participants, code behavior, and enter data; and former graduate student Kara Sage, from whom, through observation, I learned how to conduct research studies with children.

TABLE OF CONTENTS

LIST OF FIGURES

Figure *Description* *Page*

GLOSSARY

Action: A gesture, may include an object. Usually has a social component. Generally described as interacting with the world in some way. In this context, all action described is performed by a person and is dynamic (i.e. unfolding in time).

Action Processing: The ability to take in the actions seen in the world, encode them into the brain, and understand their meaning on a social level. Allows interactions between individuals via interpretation of goals and desires through watching humans perform actions. Related to language development.

Autism Spectrum Disorder (ASD): A spectrum of autism symptoms on a continuum of severity. Symptoms include communication/language delays, restricted interests, and unusual behavior patterns. Can be thought of as a social dysfunctional disorder. ASD has heritable genetic components although its cause is unknown.

Baldwin Lab: Developmental psychology lab on the University of Oregon run by Professor Dare Baldwin, Ph.D. Includes graduate students, master's students, undergraduate honors students, and undergraduate research assistants. Work done on action processing, language development, music processing. Studies with infants, children, and adults (college students).

Breakpoint: An important point in an action stream. The start or end of an action.

Coding: Refers to the act of monitoring and recording behavior, often done by RAs (coding looking time, coding accuracy on memory questions, etc.). Also refers to the way in which variables may be defined (gender was coded as male = 0, and female = 1 during data analysis).

Correlation: A statistical relationship between two variables. Can be positive (an increase in one variable predicts an increase in the other) or negative (an increase in one predicts a decrease in the other).

Dependent Variable: Depends on an independent variable. For example, ToM score may be significantly correlated with looking time; looking time is the dependent variable because the looking time score depends on the ToM score. Dependent and independent variables are defined by the study design.

Dwell Time: A paradigm developed by the Baldwin Lab and associates to assess segmentation ability. Involves paging through still frames extracted from a dynamic action video. The computer records looking time to each slide. Overall, adults, children, and infants tend to look longest at breakpoints in action and shortest at within-units in action.

Elasty: Program used to create the non-action analogs. Flipped videos upside down, pixilated videos, and intensified colors.

Event Segmentation Theory (EST): Proposes that individuals look longer to breakpoints than within units because breakpoints represent moments when the action becomes unpredictable. Individuals must attend more to those moments in the action stream because they do not know what the actor performing the action will do next. Within units are more predictable, therefore individuals do not look as long or attend as much to those action units.

Executive Functioning (EF): Skills needed for planning. Include working memory, task switching (the ability to readily apply new instructions to a familiar task), inhibition (preventing self from engaging in a salient response).

Forced-Choice Task: A task with only two options. The participant must choose one option or the other.

Hierarchical Segmentation: Beyond segmenting action into breakpoints versus within units. Segmenting based on different types of breakpoints: coarse, intermediate, and fine. Coarse breakpoints are the most important breakpoints in terms of the overall action stream. Intermediate breakpoints are slightly less important. Fine breakpoints are even less important units. A person segmenting hierarchically using the Dwell Time paradigm will look longest to coarse, less long to intermediate, less long to fine, and shortest to within units. Breakpoints are nested in a hierarchical fashion.

iCoder: Program that RAs use to frame-code looking time to the video preference task. Allows RAs to scroll through the recording of the child's looking behavior to the doubled video frame by frame. RAs observe the direction of the child's gaze and input the direction (Right, Left, or away from the target screen) into iCoder. iCoder output is transformed to generate looking time to the motionese video versus the non-action analog, which is used to calculate preference for motionese.

Independent Variable: A variable that is allowed to vary. It changes the value of a dependent variable.

Log Transformation: Normalizes a skewed distribution. All data analysis must be done on a normal distribution, or one that features the majority of the scores in the central range and fewer scores above and below that central range. Skewed scores may have many values at the positive or negative end of the score distribution, and a log transformation brings the distribution closer to normal.

Looking time: The amount of time that the child is looking at the screen. Also used to describe the percentage of time the child is looking to a particular part of the screen as compared to the entire time spent looking at the screen (percent of target looking time divided by total looking time).

MatLab: The computer program on which the stimuli were played. The program was a modified version of a program used previously in the Baldwin Lab to play stimuli.

Information could be entered on the computer to display stimuli to the participant on a different monitor. Also queued music to play during the stimulus presentation. Collected data input from the participant via an RA. Provided data output in a text file.

Motherese: Infant-directed speech. The natural way in which adults (mothers, fathers, caretakers, etc.) modulate their speech when talking to an infant or child. Features heightened pitch contours, repetition, slowed speed, and often gestures (motionese), heightened positive affect. Aids young children in learning language because the speech modulation functions to accent particular linguistic units (words, phrases, etc.) of interest.

Motionese: Infant-directed action. The natural way in which adults (mothers, fathers, caretakers, etc.) modulate their actions when showing a young child a stream of action (gesturing, displaying how to use a toy, etc.). Features large gestures, slow and repetitive motions, heightened positive affect, increased eye contact. Highlights important points in the action stream and functions as a teaching mechanism; the child is able to use the information gained from motionese learn about the world as well as to infer social information (goals, intentions, internal states, etc.) of the actor.

Neurotypical: A child not diagnosed with any developmental disorder, not delayed in developing language, communication skills, etc.

Non-Action Analog: Derived from the non-action analog in an analogous manner to Kuhl and colleagues' non-speech analog (2005). A motionese video was pixilated and flipped upside down. It matched the motion parameters of the motionese video but disguised the social aspects (human and action) associated with the motionese video. Created on program called Elasty.

Non-Speech Analog: Used in study by Kuhl and colleagues (2005). Derived from motherese speech. Computerized, matched motherese speech for frequency and amplitude. Retained qualities of the original speech sample while disguising the social aspects (human and language).

Oddball Paradigm: Kuhl and colleagues (2005) used this design to measure processing ability of speech stimuli. They measured ERPs (electrical activity in the brain) while participants listened to sequences of syllables, comparing ERPs to standard versus deviant syllables. Typically developing children showed differential ERP responses, while children with ASD did not.

Off-Line Action Processing: This is the stimulus processing that occurs after watching an event unfold. Preference and on-line action processing (the video preference task and the dwell time task) are on-line action processing tasks because we measure preference/processing while the stimulus is presented. Off-line processing refers to the memory questions and imitation done after a stimulus is shown, necessitating participants to recall information gathered earlier and apply that knowledge.

Regression: Statistical method to examine relationships between variables. A regression model can be built with one variable set as the dependent variable and independent variables can be entered in order to see how well the independent variables predict the variability in the dependent variable. A good regression model will use independent variables to predict a significant amount of the variance in the dependent variable.

Research Assistant (RA) or Observer: An undergraduate working in the Baldwin Lab who assisted with the study. There are numerous research assistants associated with the lab and this study. He or she helps to run each session, codes behavior and looking time, and enters data into the computer. Research assistants also watch videos of the sessions and record pertinent data about the child's behavior and responses to questions.

Residuals: Calculated from log transformed dwell time data to account for the fact that participants tend to speed up their clicking rate as a slideshow progresses.

Segmentation: The breakdown of a stream of action into smaller events, paying more attention to more important parts in the action stream and less attention to the less important parts.

Stimuli: What each child sees and/or interacts with during the study. This includes videos played on a computer screen, cards that the participant sorts, pictures on posters and the computer screen, music played during the videos, etc. "Stimuli" is a broad term for anything the child views during the course of the session in the lab.

Subject/Participant: A single child involved in one session of the study.

SPSS: Program used to run all statistical analyses.

Theory of Mind (ToM): Thinking and reasoning about other people's internal states and minds. Awareness that others have a mind that is separate from yours with separate desires, beliefs, and knowledge. Ability to use information to understand others' behaviors or actions.

Variable: A measure of something that changes from participant to participant. The data collected in this study were grouped into variables. For example, looking time (measured in seconds) is a variable because it is different for each participant. Score on a memory question is also a variable, but many participants have the same score on a particular question because a question can either be correct or incorrect.

Video Preference: The task that measures preference for motionese versus a non-action analog using looking time.

Within-Unit: The moments in an action stream between breakpoints. Anything in the stream other than beginnings and endings of actions.

I. INTRODUCTION

Language and Action Processing

Communication through language and action are fundamental to development; together, speech and action processing are necessary to our understanding of the world and people around us. For example, when learning new words, actions such as pointing, gesturing, and eye contact facilitate mapping words to concrete objects in the environment (Baron-Cohen, Baldwin, and Crowson, 1997). Research by Kuhl, Coffey-Corina, Padden, and Dawson (2005) suggests that there are individual differences in young children's speech processing, particularly when comparing neurotypical children and children on the autism spectrum. The current study explores individual and developmental differences that may exist in children's processing of dynamic action, making comparisons to research previously conducted on children's speech processing (Kuhl et al., 2005).

Children's Processing: Typical Development

Speech Processing

Understanding language is crucial to development for young children, and adults fuel communicative progress. Motherese, or infant-directed speech, refers to the innate, natural way in which adults tend to modify their speech when talking to infants and young children (Kemler Nelson, Hirsch-Pasek, Jusczyk, Cassidy, 1989). When using motherese, adults tend to speak more slowly and at a higher pitch, their utterances are shorter and more repetitive, they speak more slowly, and they exaggerate their pitch contour (Fernald and Kuhl, 1987). These modifications function to draw the attention of the child to important parts of the speech, helping children to detect the meaningful words and sentences that are part of the utterance. Research has shown that neurotypical infants and young children prefer to listen to motherese speech compared to adult-directed speech (Cooper & Aslin, 1990; Fernald 1985; Fernald & Kuhl, 1987). Adults naturally modulate their speech towards young children, facilitating communicative development; motherse is an important resource of which neurotypical children take advantage (Kemler Nelson et al., 1989).

Action Processing

In addition to speech processing, infants and young children develop the ability to understand intentional human action within the first months and years of life. This is advantageous because it allows interpretation of goals, intentions, and internal states of the people in a child's life (Shipley & Zacks, 2008; Baldwin, 2012). Understanding why another person performs the actions he or she does, and being able to explain goal-directed motions, is integral to developing social skills. Moreover, understanding action is integral in learning language (Bruner, 1981; Grimminger, Rohlfing, & Stenneken, 2010). For example, non-verbal social cues like eye-gaze and pointing gestures appear to facilitate language learning by indicating objects in the environment to which one refers (Baldwin & Moses, 2001; Baron-Cohen, Baldwin, & Crowson, 1997). A child can pick up on these nonverbal indications and associate them with language spoken simultaneously to begin to map nouns to objects. It would be much more difficult for children to associate language with the world without gestures or actions directed towards objects in space.

In the same way that adults naturally modify their speech to facilitate children's speech processing, they also alter their gestures and motions when performing a sequence of actions for children, a phenomenon known as “motionese” (Brand, Baldwin, & Ashburn, 2002). Adults naturally engage in motionese by exaggerating infant-directed gestures, using simple, repetitive motions, and increasing their level of enthusiasm (Myhr, Baldwin, & Brand, 2004). These features of motionese draw children’s attention to the salient parts of the action stream in the same way that motherese draws children’s attention to important aspects of a linguistic utterance (Baldwin, 2012). Motherese speech and motionese action are often used in conjunction to fuel development; as adults display actions to children using motionese, motherese is often also included. However, motionese can also be used to convey non-linguistic information about human action in general (Brand, Baldwin, & Ashburn, 2002). In another parallel to motherese, typically developing children also prefer to watch action sequences in which the actor uses motionese compared to sequences of adult-directed action (Baldwin, 2012).

Children’s Processing: Atypical Development

While most children benefit from the natural modifications made as part of motherese and motionese, research suggests that there are individual differences in this, especially for atypically developing children. In particular, examining the processing of children with autism spectrum disorder (ASD) can provide useful insight into the question of individual differences. Children with ASD typically have social dysfunctions, and as such, miss out on much of the information conveyed by other people during development (Klin, Jones, Schultz, Volkmar, & Cohen, 2002). Indeed, recent research by Kuhl and colleagues (2005) has demonstrated that children with ASD do not display the typical preference for motherese speech as seen in neurotypical controls. Perhaps a similar aversion to motionese stimuli may be seen in children with ASD.

Autism spectrum disorders are heritable developmental delays characterized by communication and language impairments, atypical motor and sensory behaviors, and restricted interests (Macari, Campbell, Gengoux, Saulnier, Klin, & Chawarska, 2012). The crux of ASD lies in social functioning deficits, which in turn cause many of the other associated symptoms due to the fact that children with ASD neglect crucial social information during development (Klin et al., 2002). Children with ASD are diagnosed on a spectrum of functionality (see *Figure 1*).

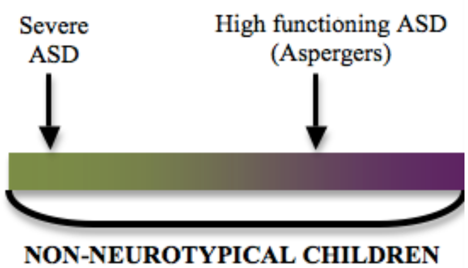


Figure 1. ASD functionality schematic. This figure depicts the scale of functionality for individuals with autism spectrum disorders (ASD), ranging from most severe impairment (left) to least impairment or high functioning (right).

Research suggests that children with ASD show differential processing in comparison to their typically developing peers, especially for processing motherese (Kuhl et al., 2005). In their research, Kuhl and collaborators examined children's preference for motherese and ability to process, or neurologically make sense of, the stimuli. Children listened to speech sequences in motherese as well as sequences of a "non-speech analog," which was generated by "computer warbling" the motherese sequence so it retained the same frequency and amplitude information, but lacked all linguistic elements. These stimuli were played for the participants in a head-turn preference task in which the children learned that turning their head in a certain direction would elicit either the motherese or the non-speech analog to play. This allowed a direct measure of preference for the stimuli; children essentially chose which stimuli they liked more and consequently turned their head in the direction of the preferred stimuli more often. This study also involved a motherese processing task in which the electrical activity in the brain was measured during presentation of motherese speech. The researchers measured brain activity using an “oddball” paradigm, playing standard versus deviant syllabi. Neurotypical infants showed differential brain responses to these two stimulus types, while ASD children did not. This indicates that neurotypical infants successfully differentiated between speech stimuli, showing enhanced processing of motherese, while ASD children did not, displaying reduced processing of motherese.

After collecting data on preference and processing of motherese from neurotypical participants and children with ASD, differential preference and processing patterns were found in each group. In particular, Kuhl and colleagues found that neurotypical children displayed high motherese processing abilities (high syllable differentiation) and that they preferred to listen to sequences of motherese than to sequences of the non-speech analog, replicating previous research (Cooper & Aslin, 1990). Importantly, they also discovered that children with ASD were poorer processers of motherese (failing to differentiate between syllables) compared to neurotypical children. Children with ASD also showed a preference for the non-speech analog compared to motherese, the reverse of the neurotypical sample. Moreover, for children with ASD, their level of ASD symptomatology predicted the degree to which they demonstrated a preference to listen to the non-speech analog in comparison to motherese. In other words, children with more severe ASD symptoms showed increased preference for non-speech relative to those with less severe symptoms.

Through this and other research (Fernald, 1985; Kuhl, 2004), we have gained considerable knowledge about the phenomenon of motherese, although the work by Kuhl and colleagues (2005) was the first of its kind to explore individual and developmental differences in how children process infant-directed speech. However, with respect to motionese, little comparable work has been done. Of particular note is a lack of research investigating individual and developmental differences in children's processing of action. Given the parallels between speech and action, it seems plausible that individual and developmental differences might also exist in the domain of action processing.

Action Processing by Children with ASD

Some literature exists regarding how individuals with ASD process action. For example, children with ASD tend to orient towards non-social action events rather than social events (Klin, Jones, Schultz, Volkmar, & Cohen, 2002). One explanation for this finding is that some action streams contain social information that may be averse to individuals with ASD. The fact that individuals with ASD exhibit different visual scanning patterns than neurotypical controls (Klin, 2002) supports this explanation, suggesting that action processing may indeed be different in ASD individuals, especially when social factors like faces are involved in the stimuli (Chawarska, Volkmar, and Klin, 2010). However, the existing evidence is inconsistent, particularly when considering processing of biological motion. Biological motion, or "point-light displays," is motion stimuli derived from videos of humans performing an action. The video is modified such that the person is no longer shown; instead, the joints on the actor's body are represented by moving lights that match the actor's motion parameters (Blake, 2003). Some studies have suggested that individuals with ASD take longer to recognize biological motion than neurotypical controls (Hubert, Wicker, Moore, Monfardini, Duverger, Fonseca, & Deruelle, 2007; Klin & Jones, 2009). However, other studies have found no difference in biological motion processing skills between ASD and neurotypical participants (Freitag, Konrad, Haberlen, Kleser, von Gontard, Reith, Troje, & Krick, 2008). Further research comparing action processing by typically developing children to that of children with ASD might help clarify some of the inconsistencies in the existing literature.

Dwell Time Methodology

The "dwell time paradigm" may provide further insight into action processing abilities. Recent research employing a novel methodology – dwell time – demonstrates that infants, children, and adults reliably segment dynamic action streams into meaningful units (Hard, Recchia, & Tversky, 2011; Baldwin, Baird, Saylor, & Clark, 2001). During dwell time tasks, participants look at still frame images extracted from a video of dynamic action. Participants advance through the slides at their own pace and the amount of time they spend looking at each slide is recorded. The slides are categorized based on their type: "breakpoint slide" or "within-unit slide." A breakpoint slide depicts the moment when one action ends and another begins, while a within-unit slide shows a moment that occurs in the middle of an action unit. Recent research has revealed that infants, children and adults all tend to "dwell" longer at breakpoint slides compared to within-unit slides, providing evidence of segmentation of dynamic action (Hard et al., 2011; Baldwin et al., 2001; Baldwin, 2012).

The dwell time methodology is relatively new, and provides important insight into how participants understand dynamic action sequences unfolding over time. Participants who look longer to breakpoints than within units are essentially focusing in on the most salient parts of the action stream, while attending less to the units of action less important to the overall goal of the action. Breakpoints have also been described as the moments when predictability declines (Kurby and Zacks, 2007). When watching another person perform an action, we anticipate what motion the actor will make next. Breakpoints in action represent the moments we are unable to anticipate; we focus in on these moments because we are uncertain what the next action will be. This idea underlies Event Segmentation Theory (EST) (Kurby and Zacks, 2007; Hard et al., 2011). In this way, action segmentation is crucial for learning about others, comprehending action sequences, and memory for motions. Dwell time gives us valuable insight into segmentation ability. This methodology may also be useful for examining action processing skills in individuals with ASD; it is possible that they fail to understand the contingency or predictability of ongoing human action, leading these individuals to fail to show typical segmentation patterns. It also may be the case that children with ASD fail to take in information presented socially due to the social aversion inherent to their ASD symptomatology.

Successful segmentation is key to rapid processing of action, an essential skill for developing cognitive functioning (Baldwin, Andersson, Saffran, Meyer 2008). It is possible that young children just developing their action processing skills may be more or less successful at segmentation of action. The current study employs the dwell time paradigm to explore potential individual differences in young children's action segmentation. We hypothesize that some children will be more successful at segmentation, showing longest looking times to breakpoints and shortest times to within-units, while other children may be less successful at segmentation, looking for less time at the breakpoints than do children more successful at segmentation. In other words, we predict that some children will show greater differences in looking times to breakpoints versus within units than will other children. We further predict that segmentation success will be predicted by EF and ToM skills.

The Current Study

The current study explores possible individual and developmental differences that may exist in young children's action processing and preferences for motionese stimuli, and whether those differences are analogous to those discovered for children's speech processing (Kuhl et al., 2005). In other words, would typically developing children and children with ASD respond differently to displays of motionese and a "non-action analog" stimuli? Would they show the same overall pattern of results found by Kuhl and collaborators (2005) in their research on motherese and a non-speech analog?

In a preliminary attempt to investigate these questions, the current study recruited typically developing children, examining individual differences for preference and processing of motionese. Although no children with ASD were included in this phase of the study, our research may still enable us to make inferences about how children with ASD might process action. Recall that children with ASD fall on a spectrum based on their level of functioning (see *Figure 1*). Such a developmental spectrum could be construed as continuous with a spectrum depicting the range of functionality that is considered "typical" (see *Figure 2*), which would be consistent with the current conceptualization of ASD as a continuum of affectedness, rather than a dichotomy between affected and non-affected individuals (Klin et al., 2002).

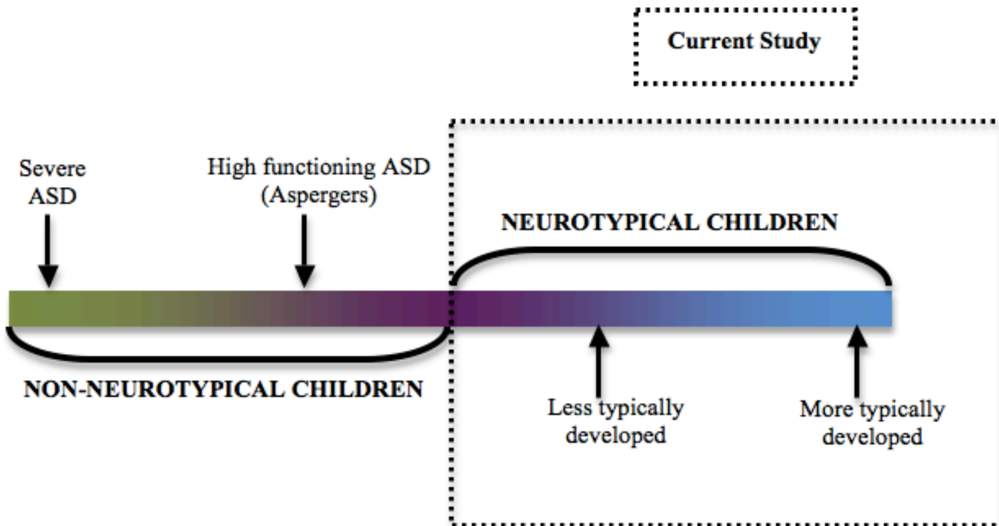


Figure 2. The Autism Spectrum continued through to neurotypical children. This figure shows neurotypical children on the right as continuous with ASD children on the left; shows how less typically developed children could perhaps be used to infer about high functioning ASD children.

In order to determine whether children in the current study displayed individual differences relevant to the spectrum of autism symptomatology, we measured children's executive functioning (EF) and theory of mind (ToM). Executive functioning refers to the skills needed to voluntarily deploy attention and cognitive resources to work through a problem and achieve a goal, and include such elements as working memory, planning, and flexibility (Carlson & Moses, 2001). Theory of mind refers to reasoning about others' mental states (Wellman & Liu, 2004). These two constructs are highly related to each other for both neurotypical and atypically developing populations.

Importantly, children with ASD are particularly impaired in both EF and ToM (Zelazo, Jacques, Burack, Frye, 2002). Given this, the current study uses children's EF and ToM scores to determine level of typicality, which will later enable us to make inferences about how children with ASD might process action.

The current study uses two main tasks to examine children's processing of dynamic action: a video preference task (to measure motionese preference) and the dwell-time paradigm (to measure motionese processing via segmentation ability). In the video preference task, children watch video pairs comprised of videos depicting action sequences with an actor using motionese played alongside videos depicting "non-action analogs," akin to the speech and non-speech analog stimuli used by Kuhl and colleagues (2005). In the dwell-time task, children advance through slides extracted from different videos depicting an actor carrying out sequences of action using motionese. Both tasks measure children's on-line processing of dynamic action processing tasks, meaning that they capture children's processing as it occurs in real time.

We predict that as a group children who are more typically developing (as measured by EF and ToM) will show greater preference for the motionese videos compared to the non-action analog videos. On the other hand, we predict that children who are less typically developing will show a decreased preference for motionese, with a corresponding increased preference for the non-action analog. Similarly, we predict that the more developmentally typical children will show skilled processing of motionese, displaying strong segmentation abilities in the dwell time task, while the less developmentally typical children will show lower processing levels of the stimuli, displaying poorer segmentation abilities in the dwell time task.

In addition to the on-line preference and processing tasks, we also conducted a number of tasks to assess off-line action processing skills. We showed the children live-action demonstrations of toys and gave the children a chance to imitate each action sequence. Children were also asked memory questions about the dwell time and live action displays. Both the imitation task and memory questions provided insight into how well children could remember motionese sequences. We predict that increased ToM, EF, motionese preference, and segmentation ability will predict higher imitation and memory scores. Parents also filled out a number of forms about their child during the study. These measures were parent-reported levels of EF, ToM, language ability, and autism symptomatology (geared towards ASD-like symptoms in the general population). We predicted that these parent report measures would correspond with the data we gathered from the children during the session (for example, high parent-reported EF should correspond with high EF performance during the behavioral tasks).

II. METHODS

1. Participants

Forty-six neurotypically developing children ages 2.5 to 3.5 years old (mean = 34.91 months, SD = 3.35 months; 23 males, 23 females) were included in this study. Two female participants were excluded from the study (the first subject run in the study was excluded because the task order changed after she participated, and the other child was removed because she became ill during the study) for a total of 23 males and 21 females. All children were from the Eugene, Oregon area and were recruited via phone calls. The study was approved through the Institutional Review Board of the University of Oregon. A parent gave informed consent for each child to participate.

2. Equipment Set-Up

Computerized Tasks

Participants sat directly in front of a computer screen where images and videos (the stimuli) were presented. A camera located above the computer screen recorded the participants' face while a ceiling mounted camera simultaneously recorded the computer screen. The camera above the computer screen was connected to a television monitor behind a curtain; a research assistant watched the television monitor to code children's on-line looking behavior. (See Appendix 1, figure 10.) Speakers were located behind the curtain and play classical music during each computerized task.

Non-computerized Tasks

For non-computerized tasks, children sat with the researcher at a small table in the middle of the room. The same cameras recorded children's behavior, although from different angles than during the computerized tasks. Research assistants again watched the television monitor to code participants' performance on these tasks. (See Appendix 1, figure 11.) The set-up for the imitation task differed slightly from the others: the child watches an action sequence at the table by the monitor and imitates the action at the small table (see Appendix 1, figure 12).

3. Procedure

The study took between 40 and 60 minutes per child. The tasks, described below, were done in the same order for each subject. The order of the tasks was as follows:

A. Video Preference
B. Theory of Mind
C. Dwell Time + Memory Questions
D. Executive Functioning
E. Imitation + Memory Questions

A. Video Preference Task (Motionese Preference)

The preference task tests whether children prefer the motionese or the non-action analog when the two videos are played simultaneously side-by-side on the computer screen. One side of the doubled video displays a motionese video and the other side of the screen shows the non-action analog corresponding to a different motionese video (see Appendix 3, figure 21). The video preference test in the current study is modeled on research conducted by Kuhl and colleagues (2005), changing from the language domain to the action domain.

Stimuli

In an attempt to match the design of the speech stimuli used by Kuhl and colleagues (2005), we created two types of videos for this task: one depicting motionese action and the other showing a non-action analog (see figure 3 below). We first filmed the motionese videos; we then generated the non-action analog videos by editing the motionese versions, using a computer program called Elasty to pixilate the motionese stimuli and rotate them 180 degrees.

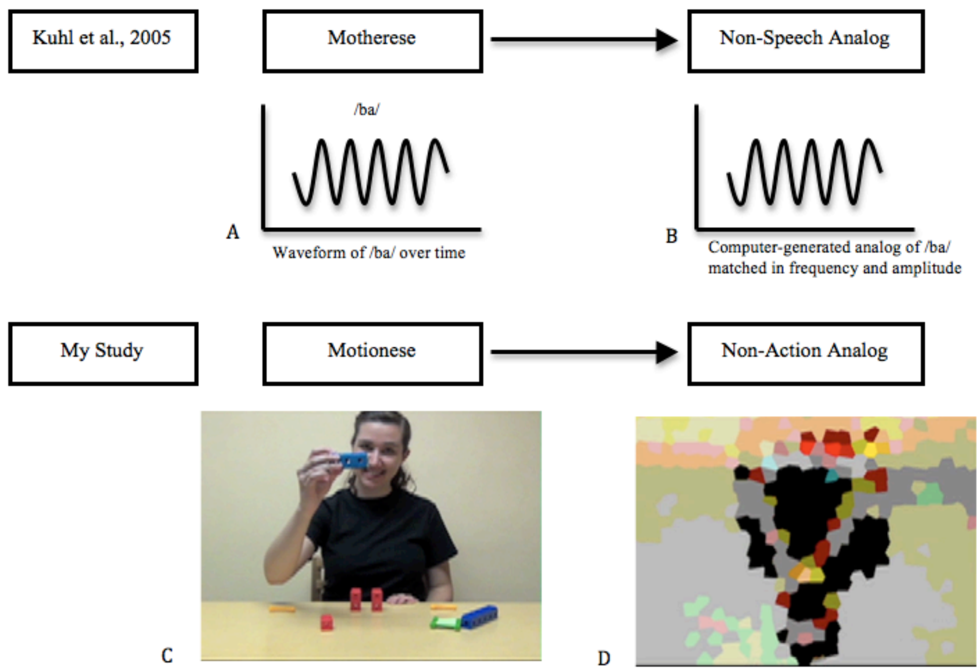


Figure 3. Non-speech and non-action analogs derived from motherese and motionese, respectively. A) Kuhl's motherese stimuli; B) Kuhl's modified non-speech analog; C) "Lego Man" motionese; D) "Sorting" non-action analog.

We created four videos in which the actor performs a series of motionese actions, interacting with a different sets of toys ("Lego Man," "Sorting," "Train," and "Balloon"; see Appendix 2, figures 13-16), modeled on toys used in previous research (Sage and Baldwin, 2012). The four sequences were designed to be visually interesting with multiple parts that could be manipulated. The toys also featured a general goal to be completed by using the parts of the toys in an ordered series of steps.

To depict motionese, the actor displays a positive affect while enhancing facial features; she smiles widely and makes eye contact with the camera; she uses large, exaggerated motions to perform each action in the sequence, and she moves slowly.

Procedure

In this task, children view four doubled videos, each between 40 and 90 seconds. The doubled videos were played in a fixed order but are counterbalanced for motionese vs. non-action analog on each side of the screen (see Appendix 3). During each session, a research assistant uses two different buttons to indicate whether the child is looking to the right or to the left; the computer records this information and calculates the total amount of time the child looked at one side of the screen or the other.

B. Dwell Time (Motionese Processing)

The dwell time task probes how well subjects are able to segment motionese action into breakpoints and within units. Children use a computer mouse to click through still frames of a motionese video played on the monitor and looking times are used to determine segmentation ability. The dwell time paradigm was based on previous research by Hard, Recchia, and Tversky (2011).

Stimuli

The stimuli for this task were created in an identical way to those used for the video preference task. The same actor for the video preference videos filmed two action sequences with the "Puppet" toy and the "Pyramid" toy (see Appendices 2 and 3). The actor used the motionese style while manipulating the toy, as described above. Then, still frames from the two videos were extracted at a rate of two frames per second using iMovie to create a slideshow of images. In total, there were 76 frames per slideshow for each of two slideshows.

Procedure

The still frames are displayed one by one in temporal order on the computer monitor. Participants use a mouse to advance the images; each mouse click advances the slideshow by one slide. Participants first engage in a practice video to learn the contingency of the task. After the practice set, they then advanced through the two slideshows at their own pace. The computer records the amount of time between clicks, or the amount of time the children spend on each slide. A research assistant records looking time to the slideshow as children click through the sequence, which is also recorded on the computer.

C. Imitation Task (Off-Line Motionese Processing)

The imitation task provided a behavioral measure of action processing. For this task, the researcher performed two sets of actions for each participant. The child then had an opportunity to imitate the target actions. The actions that the child performed are scored for accuracy.

Stimuli

The stimuli for this task were two unique toys, “Play-Doh Toy” and “Star Toy” (see Appendix 2, figures 17 and 18). The researcher manipulated the toys in the motionese style in view of the subject following a script of target actions. The researcher also used scripted phrases to draw attention to various aspects of the display.

Procedure

The researcher shows the participant the goal-directed motionese display with each of two toys. After the demonstration of each toy is complete, the child moves to the small table to interact with the toy (see Appendix 1, figure 10). Each participant was allowed 90 seconds in which to imitate the actions he or she observed the researcher perform. A research assistant scores the imitation for accuracy, recording target actions as the child performs them. This methodology is based on work by Sage and Baldwin (2012).

D. Memory Questions (Off-Line Motionese Processing)

Four sets of memory questions were included throughout the study, each set containing four questions about the stimuli (see Appendix 7). One set of questions followed each of the two dwell time task, and one set of questions followed each of the two imitation sessions as part of the imitation task. The memory questions were systematically chosen in order to ensure consistency across each action stream. Each question battery features questions assessing what the toy was, what actions were performed with the toy, and the order in which certain actions were performed. The questions were also formulated in order to probe action segmentation at both coarse and fine grained levels.

Stimuli

The questions were asked verbally and the child was provided with pictures on the monitor corresponding to the possible answers for each question (see Appendix 7, figures 30-33). Each question had two possible answers. Photos were taken of the toys relevant to each question and placed side-by-side on the screen as each question was asked. Answers were counterbalanced for screen side.

Procedure

Questions were presented verbally to the participant and corresponded with the images on the screen. Children were encouraged to select one of the two options in a forced-choice manner. Participants could either answer verbally or indicate the answer with a pointed finger at one of the two images. A research assistant recorded answers.

E. Executive Functioning (EF) Tasks

The executive functioning section of the study included three tasks to probe various parts of executive functioning abilities: Task Switching, Working Memory, and Inhibition. Task Switching is a card sorting task that measures the participant's ability to flexibly modify instructions during a task while inhibiting previous instructions (Diamond, 2005). Working Memory assesses the memory capabilities of each participant. The Inhibition task measures the child's ability to prevent engagement in a salient activity. All three tasks have been found to be impaired in children with ASD (Guerts, 2004).

EF1: Task Switching

The task switching task is a modified version of the classic card sorting task (Diamond, 2005). This task requires that children sort cards by two different dimensions. The classic version of the Dimensional Change Card Sort Task (DCCS) features cards with one of two images on them in one of two colors. Children are instructed to sort each card first by one dimension and then the other (by color and then object, or by object and then color). Participants must be able to understand the changed instructions and implement the change rather than persevere with the original directions. The classic version of the DCCS is difficult for 2.5-3.5-year-olds; thus, a modified version of the task is used in the present study (Diamond, 2005).

Stimuli

For this task, several sets of cards are present for sorting (see Appendix 5, figure 23). The first set is four practice cards: green house, orange bird, orange house, and green bird. For the shape task, a set of 12 cards is used, and a set of identical 12 cards is used for the color task. However, the color and shape sets are ordered differently. The practice, shape, and color cards are ordered the same for every participant. A tray is also provided that has two sides, each side with an example card attached. One side has a orange card with a bird on it, and the other side has an green card with a house on it.

Procedure

Children are first asked to sort four practice cards by shape (bird cards go in one tray, house cards in the other), and corrections are made if the child makes mistakes. If any mistakes are made, the child sorts the four practice cards again. Then the child sorts the set of 12 cards. Following this task, the child is instructed to sort the cards by color instead of shape. The child again practices with the four practice cards, corrections are given, and the child sorts the practice cards again if any mistakes were made during the practice trial. Each participant then sorts the set of 12 cards by color. A research assistant records correctly sorted cards. The researcher does not indicate correct or incorrect trials during the target trials; corrections are only made during the practice trials.

EF2: Working Memory

The working memory task requires that participants remember the faces of different breeds of dogs in order to receive stickers. This task probes memory skills and is based on the trucks task designed by Hughes & Ensor (2005).

Stimuli

Eight white cards are presented for this task (see Appendix 5, figure 24). Each card features two black and white faces of different dog breeds. Cards 1, 3, 5, and 7 feature the same two images; cards 2, 4, 6, and 8 also feature the same images, which differ from the images on the odd numbered cards. Dog faces are counterbalanced for card side. A stack of stickers is also presented, which are awarded as prizes for correct answers. The back of each card is marked with Xs and Os; X's correspond with the same dog on the odd cards, and one of the two dogs on the even cards. O's correspond with the opposite dog on the odd card, and the remaining dog on the even card (see Appendix 5, figure 24).

Procedure

The participant is asked to guess which dog face on each card he or she thinks will be associated with a sticker (see Appendix 5 for full script). The first dog chosen is always correct, and this same dog face must be chosen on all odd cards in order to be correct. A second dog must be chosen on each even card to be correct and receive a sticker. The child receives a sticker for each correct answer, and stickers are placed in an envelope until the end of the task. A research assistant records receipt of a sticker for trials two through eight. Trial one is always correct, and trial two is merely a chance to correctly guess which dog is associated with the sticker. See Appendix 5, figure 24 for more information.

EF3: Inhibition

The inhibition task tests how well the participant can keep from engaging in a salient response. The task is a modified version of a gift delay task in which the child is instructed not to open a gift for a certain amount of time. The child must inhibit his or her initial response to open the present. This task is modeled after the “bow task” from Carlson (2004).

Stimuli

The stimuli is a colorful bag with tissue paper inside. Beneath the tissue paper is a box with a prize (a small slinkey). The bag is festive and inviting (see Appendix 5, figure 25). The bow that is later attached to the bag can be clipped onto the bag with a paper clip.

Procedure

The bag is placed on the table in front of the child and the child is told to remain seated and not to touch the bag or what is inside the bag and not to peek inside until the researcher returns with the bow for the bag (see Appendix 5). The researcher goes behind the curtain for 180 seconds and then returns with the bow. If the child opens the bag early, the researcher returns at that time with the bow. A research assistant records how many times the child touches the bag, peeks into the bag, and opens the present. The RA also records the time at which the child opens the bag and the time at which the child first touches or peeks into the bag (touch/peek latency).

F. Theory of Mind (ToM) Tasks

Four tasks are used to assess each participant’s level of Theory of Mind: Diverse Desires, Diverse Beliefs, False Contents, and Knowledge Access. These tasks probe how well each child understands that other people may have different beliefs, desires, and knowledge than they do. A research assistant records the participant’s responses to each question during each task. All tasks are based on those described by Wellman and Liu (2004). See Appendix 6 for images of the stimuli and scripts used for each task.

ToM1: Diverse Desires

The first task assess whether or not the child understands that other people may have wants and desires that differ from their own.

Stimuli

A poster is provided with images of two snacks (cookies and carrots) and the image of a boy ("Sammy").

Procedure

Each participant is asked which snack they would prefer (cookies or carrots) and are then told that another person (Sammy) would rather have the opposite. They are then asked to choose a snack for Sammy. The child is scored as either responding correctly or incorrectly to the key question: "Which snack will Sammy choose?" If the participant successfully chooses the snack opposite to his or her desired snack for Sammy, the child has gotten the question right. If the child instead chooses the same snack that he or she desires as the snack Sammy would choose, he or she has gotten the question wrong. See Appendix 6 for scripts and figure 26.

ToM2: Diverse Beliefs

The second task probes whether or not the child understands that other individuals may believe different things than what the child believes.

Stimuli

A similar poster is provided for this task. It has images of two locations (a garage and a tree) and an image of a girl ("Linda").

Procedure

This task is very similar to the Diverse Desires task; children are asked to pick one of two locations in which a cat may be hiding (the garage or the tree). They are then told that another person (Linda) believes the cat is hiding in the opposite location. The child is asked where Linda will look for the cat. The score on this task is based on the child's response to the target question, "Where will Linda look for her cat?" A correct score is given if the child answers that Linda will look in the opposite location than the one the child stated previously. See Appendix 6 and figure 27.

ToM3: False Belief

The third task as a classic False Belief task (Wellman & Liu, 2004) in which unexpected objects (ribbons) are hidden inside a Crayon box.

Stimuli

The stimuli are a Crayon box with ribbons inside and the same image of "Sammy" from ToM1.

Procedure

The child is shown the box, and is asked what he or she expects to be inside. The child is asked what another person (Sammy) would think is inside the box if Sammy has not seen inside it yet. The score for this task is based on participant's response to the question, "What does Sammy think is inside the box?" If the child understands that Sammy would think there are Crayons inside the box because he has not yet seen inside, he or she will answer correctly. If the child does not understand that Sammy has less information about the box than the child does, then they will answer incorrectly, stating that Sammy knows there are ribbons inside the box. See appendix 6 and figure 28.

ToM4: Knowledge Access

The final task measures how well children are able to understand that other people may know less than the child because they have not yet had access to the same information that the child has.

Stimuli

In this task, the child is shown a colorful box with unknown contents. Inside is a pink spring. The same image of "Linda" from ToM2 is again used for this task.

Procedure

The participant is shown the box and is asked whether he or she knows what is inside it. The child is then shown the spring inside. The box is closed again, and the child is asked whether another person (Linda) would know what was inside the box if Linda had not yet seen inside the box. Children who answer this question correctly will respond that Linda does not know what is inside the box when asked the target question, "Does Linda know what is inside the box?" Participants who fail this task will respond that Linda does know what is inside, even though she has not seen inside. See Appendix 6 and figure 29.

4. Parent Forms

Parents filled out forms during the study regarding various aspects of their child's development. Forms are listed in the chart below (*Figure 4*).

Form Name	Description	Source
1. CDI-III: MacArthur-Bates Communicative Development Inventory	Word use, sentences and grammar	MacArthur-Bates, 2009; Fenson, 2007
2. Language Form 2	Language understanding and production, specific to this study	Baldwin Lab, 2013
3. EQ-SQ: Empathizing and Systemizing Quotient	Discern autism-like symptoms in the general population	Auyeung, Wheelwright, Allison, Atkinson, Samarawickrema, and Baron-Cohen, 2009
4. Q-CHAT-10: Quantitative checklist for autism	ASD questionnaire	Allison, Auyeung, and Baron-Cohen, 2012
5. CBQ Short Form Version 1: Children's Behavior Questionnaire	Temperament	Rothbart, Ahadi, Hershey, and Fisher, 2001
6. CSUS Long Form: Children's Social Understanding Scale	Social understanding, theory of mind	Putnam and Rothbart, 2006
7. Computer Survey	Computer use, computer mouse use	Baldwin Lab, 2011

Figure 4. Chart of Parent Forms. Forms are listed by title, described, and cited. Parents filled out forms during the study while the child participated in various tasks.

III. RESULTS

A. Data Processing

Video Preference Data

The video preference task was coded using with on-line eye-gaze coding and later using off-line, frame-by-frame coding. During the on-line eye-gaze coding, a trained observer pressed one of two computer keys to indicate when the participant was looking to the right and left sides of the computer screen. The observer was blind to the stimuli being presented and to the hypotheses of the study. During off-line frame-by-frame coding, a different observer used a program called iCoder to determine whether the child was looking to the left side, right side, or away from the screen for each frame of the video. Again, the observer was blind to the stimuli being presented and to the hypotheses of the study. The data derived from off-line coding were utilized in the analyses reported below. Agreement between individual off-line coders was over 95%.

Video preference was measured in terms of the percentage of motionese looking relative to total time looking at either video. Stated another way,

$$\text{Preference} = \frac{look_{motionese}}{look_{motionese} + look_{non-action}} * 100$$

Thus, a higher percentage means a greater the preference for the motionese video (and the lesser the preference for the non-action analog). Across participants, the video preference scores displayed a negatively skewed distribution, so we performed a log transformation to normalized the distribution prior to further analyses with this measure.

Dwell Time Data

Dwell time was coded live by an observer blind to the stimulus. The observer pressed a button when the child looked at the computer screen presenting the stimulus, and released the button when the child looked away from the screen. MatLab played the stimuli on the computer screen and recorded looking time to each slide. Looking time was defined as the amount of time per slide the child looked at the screen, i.e. the amount of time the observer pressed the button during each slide presentation. Looking time to the first slide was excluded for the analysis.

Dwell time scores were then calculated by log transforming the raw data to correct for the negative skew and residualizing the data by fitting a power function to each individual participant's data. Participants were excluded if they looked at less than 50% of slides; 3 participants were excluded in this way. Residualized scores were then Windsorized; outlier scores, defined as scores three standard deviations above the mean, were replaced using with a score three standard deviations above the mean. 41 residualized looking time scores were Windsorized. Residualized scores were sorted by slide type for each participant and average scores were calculated for looking time to each slide type (coarse breakpoint, fine breakpoint, and within). Coarse breakpoints were defined by expert coders to be the start or end of large units of the action stream, while fine breakpoints were defined as the start or end of a smaller action unit.

For this analysis, only the Pyramid dwell time slideshow was used; the Puppet slideshow has not yet been analyzed. The Puppet dwell time slideshow had many slides representing "ambiguous" units of action, appearing neither to be breakpoints or within units, and need to be analyzed further. As such, only the Pyramid slideshow has been included in analysis thus far; Pyramid slides were consistently classified by experts as breakpoints or within units.

Our initial analysis revealed that average looking times were longer to within-unit slides than breakpoint slides. A t-test revealed that averaged across looking times for all participants, looking times were longest to within units compared to breakpoints, $t(39) = -.735$, $p < .01$. This is inconsistent with previous research about how children process dynamic action; prior research has found that mean breakpoint looking time is significantly greater than mean within unit looking time (Hard, Recchia, & Tversky, 2011). As such, we decided to split up the participants into consistent versus reverse segmenters. We created a dummy variable called "Breakpoint Advantage." Breakpoint Advantage was coded as 1 = participants who had longer looking times to breakpoints than within units ("consistent segmentaters," n = 16) or 0 = participants who had longer looking times to within units than breakpoints ("reverse segmenters," n = 23). A dichotomous variable of segmenting/non-segmenting made more sense here than a continuous variable of segmentation as a whole.

EF: Excluded and Composite Scores

We excluded the behavioral working memory task from these analyses because the task did not correlate with other expected variables. It did not correlate with language (CDI: $p = .968$), age ($p = .139$), or other measures of memory (score on memory questions: $p = .668$). Thus, we did not include working memory in our analyses.

Card sorting and gift delay, particularly latency to opening the present in seconds, were correlated with developmental measures and were thus included. Card sorting was correlated with age ($p = .026$), memory questions ($p = .030$), and language measured by the CDI ($p = .071$). Gift delay latency, or the time until the first touch or peek, was correlated with memory ($p = .026$) and language ability (CDI: $p = .089$). As such, the EF score used in analyses was a combined score of card sorting and gift delay scores. The gift delay score was a composite score of total touches (score of: 0 = more than 5 touches, 1 = 1-5 touches, 2 = no touches), total peeks (0 = more than 5 peeks, 1 = 1-5 peeks, 2 = no peeks), time to first touch or peek (score of: 0 = 1-59 seconds, 1 = 60-119 seconds, 2 = 120-179 seconds, 3 = 180 seconds), and time the present was opened (score of 0 = 0 seconds, 1 = 1-179 seconds, and 2 = 180 seconds). The highest possible score was 9, and the lowest score was 0. Scores were converted to percentages to create a gift delay composite score. This percentage was averaged with the percentage on the card sorting task to create an overall EF composite score.

ToM: Behavioral vs. CSUS

Theory of Mind was analyzed both by behavioral tasks (Diverse Desires, Diverse Beliefs, False Contents, and Knowledge Access) and a parent report form (Child Social Understanding Scale, Sort Form, or CSUS). The CSUS had the strongest correlation with video preference while the behavioral tasks, particularly the Diverse Beliefs task, correlated strongly with dwell time segmentation (Breakpoint Advantage). In terms of the video preference task, the CSUS compiled score is almost significantly correlated ($n=41$, $p=.105$). Breakpoint Advantage is significantly correlated with the Diverse Desires task ($n=41$, $p=.012$) and is even more strongly correlated with the median split scored ToM behavioral tasks (1=high ToM, 0=low ToM) ($n=35$, $p=.003$). For this reason, CSUS was used as the measure of ToM for analyses involving video preference, while behavioral measures of ToM were used in analyses of dwell time.

B. Individual Difference Variables and Age

We expected that age would correlate with EF and ToM performance. Age was significantly correlated with Diverse Desires ToM task, $r(41) = .502$, $p = .001$, such that older children were more successful at the Diverse Desires task than younger children. Similarly, older children significantly scored higher when all behavioral ToM tasks were combined, $r(36) = .393$, $p = .018$. Age was also significantly correlated with the EF card sorting task such that older children scored higher on card sorting, $r(43) = .339$, $p = .026$. Age was not significantly correlated with the gift delay composite score, $r(42) = -.087$, $p = .582$, but when the gift delay composite score was combined with the card sorting task, there was a significant relationship with age such that older children had higher EF combined scores, $r(42) = .347$, $p = .025$. Overall, the literature suggests that EF, ToM, memory, imitation, and language should all increase with age (Garon, Bryson, and Smith, 2008; Wellman and Liu, 2004). Our results suggest that our EF and ToM paradigms are working in the standard manner suggested by past studies.

Similarly, we predicted that age would be correlated with imitation ability, memory question performance, and language skills. Age was significantly correlated with imitation ability, $r(38) = .439$, $p = .006$, such that older children performed better on the two imitation tasks. Age was not significantly related to performance on the memory questions, $r(36) = .306$, $p = .069$, indicating that older children (nonsignificantly) scored higher on memory questions than younger children. Older children displayed significantly better language skills both on the CDI and the Language Form 2 (CDI composite score: $r(43) = .526$, $p < .001$; Language Form 2: Language Use: $r(43) = .313$, $p = .041$). Past literature has found that these relationships should hold true (e.g. Sage and Baldwin, 2011), and our corroboration of past age-related findings again suggests that our paradigms are functioning, perhaps with the exception of the memory question tasks.

We also expected EF and ToM to be highly correlated. The only significant relationship between EF and ToM was between EF card sorting and a median split score of ToM, $r(36) = .334$, $p = .046$, such that children who performed better on card sorting also performed well on ToM tasks. This shows (somewhat) that our paradigms are functioning properly; EF and ToM have been correlated strongly throughout the literature (Carlson, Mandell, and Williams, 2004), so our measures of EF and ToM should also be correlated.

C. Video Preference Task

Our main question of interest regarding the video preference task was whether EF and ToM would predict score on the video preference task. We predicted that increased scores on EF and ToM would independently predict increased preference for motionese over the non-action analog.

We examined video preference (log transformed) as the dependent variable in a regression analysis to ascertain the extent to which it was predicted by a number of other individual difference variables, including gender, age in months, language ability (CDI score), ToM score (parent reported on the CSUS), EF score (composite score of card sorting and gift delay), and Breakpoint Advantage.

Model	Unstandardized Coefficients B	Std. Error	Standardized Coefficients β	t	Significance
Gender	-5.841	1.779	-.480	-3.284	**.003****
Age	.177	.308	.098	.573	.571
Language (CDI)	-.161	.048	-.572	-3.351	**.002****
EF (card sort & gift delay)	4.563	1.690	.379	2.699	**.011***
ToM (CSUS)	.451	.135	.439	3.341	**.002****
Breakpoint Advantage	3.230	1.672	.265	1.932	.063
Model	F	Significance			
	4.436	**.001****			

*Note: *p < .05, **p < .01*

Figure 5: Regression model with video preference as the dependent variable. Shows that gender, language (CDI), EF (card sorting task), and ToM (CSUS) are all significant predictors.

This model with the dependent variable of log-transformed video preference explains a significant amount of the variability in preference, $F(7,30) = 5.188$, $p = .001$. There was a significant effect of gender, $B = -5.841$, $t(30) = -3.284$, $p = .003$ such that girls had larger preferences for motionese than did boys. The CDI language scale also predicted video preference, $B = -.161$, $t(30) = -3.351$, $p = .002$, such that children with better language skills looked a shorter length of time at the motionese and longer at non-action. ToM (score on the CSUS) was a significant predictor of video preference $B = .451$, $t(30) = 3.341$, $p = .002$ such that participants with higher CSUS scores showed greater preference for motionese. EF (card sort and gift delay composite) was also a significant predictor, $B = 4.563$, $t(30) = 2.699$, $p = .011$, such that those with higher card sorting scores displayed greater preference for motionese. Breakpoint Advantage was not a significant predictor, $B = 3.230$, $t(30) = 1.932$, $p = .063$, such that participants who looked longer to breakpoints than within units, i.e. participants who segmented action successfully, showed no systematic tendency to look longer to motionese than those who did not segment successfully (although this relationship is nearing significance). Age also was not a significant predictor, $B = 0.177$, $t(30) = .573$, $p = .571$. This means that older children did not show higher preference for motionese than did younger children; the variance in motionese preference was unrelated to age of participants.

From this regression analysis, we found that EF (card sorting and gift delay) and ToM (parent report on the CSUS) independently predicted video preference such that increased EF and ToM both predicted increased preference for motionese over the non-action analog. Additionally, these variables predicted motionese preference above and beyond age. Consistent segmentation nearly significantly predicted greater preference for motionese.

D. Dwell Time Task

Our question of interest for the dwell time task was whether EF and ToM scores independently predicted segmentation ability. We predicted that increased EF and ToM scores would predict higher segmentation ability. Specifically, we predicted that increased EF and ToM would correspond with increased looking time to breakpoints over within units, as well as “consistent segmentation,” or Breakpoint Advantage, and decreased EF and ToM would correspond with decreased looking time to breakpoints relative to within units.

A regression model run with Breakpoint Advantage as the dependent variable did not yield any significant results; thus, graphs, t-tests, and ANOVAs were run on EF and ToM with Breakpoint Advantage. Participants were divided into two groups based on a median split of their EF scores. The EF score used in this case was a composite of gift delay latency to opening the present and card sorting. Participants with high EF were compared to those with low EF on their dwell times by slide type. As shown in the figure below, looking times in the high EF group display the pattern found in previous dwell time research: highest average looking time to coarse breakpoints, lower looking time to fine breakpoints, and lowest looking time to within units (Hard, Recchia, & Tversky, 2011). In contrast, the low EF group displays the reverse pattern with longest looking time to within units, shorter looking time to fine breakpoints, and shortest looking time to breakpoints.

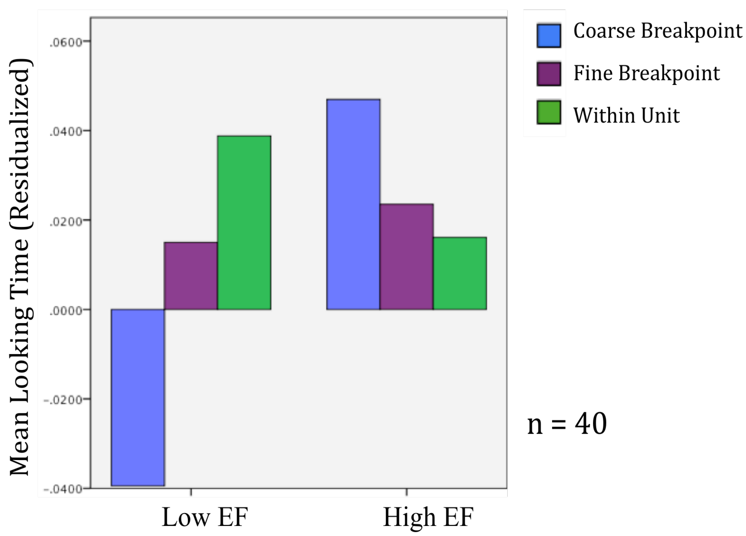


Figure 6: Pyramid dwell time scores by median split EF groups. EF split into high and low scores (based on a composite of card sorting and gift delay tasks) predicts hierarchical segmentation in the high EF group and the reverse segmentation in the low EF group. Looking time to coarse breakpoints are blue bars, time to fine breakpoints are purple bars, and time to within units are green bars.

We ran a one-way ANOVA to test differences in looking times between the high/low EF groups. We found a nearly significant difference in looking times to coarse breakpoints in low versus high EF, $F(1,38) = 2.978$, $p = .093$, such that participants with high EF scores tended to look longer to coarse breakpoints than did those with low EF. There was a nonsignificant difference in looking time to fine breakpoints, $F(1,38) = .171$, $p = .681$. There was a significant difference in looking time to within units, $F(1,38) = 4.351$, $p = .044$, such that those in the high EF group looked shorter to within units than did those in the high EF group.

We then created a graph to compare average looking times to breakpoints (a combination of coarse and fine breakpoints) versus within units. (See figure 7 below). Even when represented with breakpoints combined, an independent samples t test revealed no significant difference between looking times to breakpoint slides, $t(38) = 1.198$, $p = .238$.

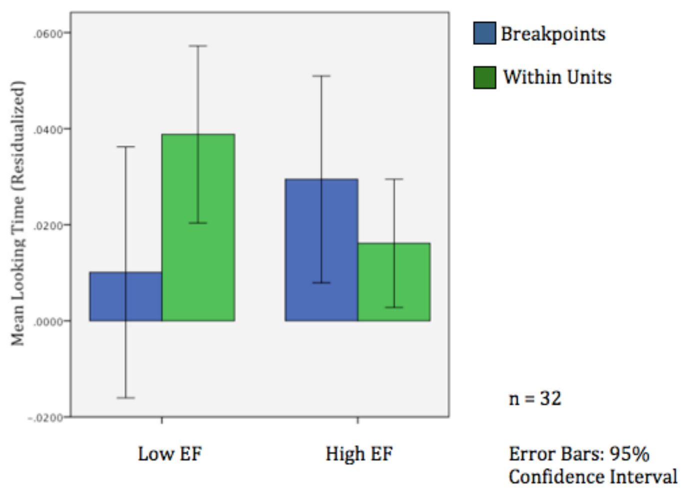


Figure 7. Looking times to pyramid breakpoints versus within units by EF group. The blue bars represent breakpoints and the green bars represent within units. Error bars set to significance of $p = .05$ are included.

We then examined ToM. A one-way ANOVA showed no significant differences in looking time to coarse or fine breakpoint slides (coarse: $F(1,32) = .854$, $p = .362$; fine: $F(1,32) = 2.172$, $p = .150$). There were significant differences in looking times to within slides between ToM groups, $F(1,32) = 4.339$, $p = .045$, such that those with high

ToM scores looked shorter to within unit slides than did those with low ToM scores. As such, we graphed looking time to breakpoint slides (coarse and fine combined) versus within slides as a function of ToM group (see Figure 8 below). ToM scores were median split for this analysis.

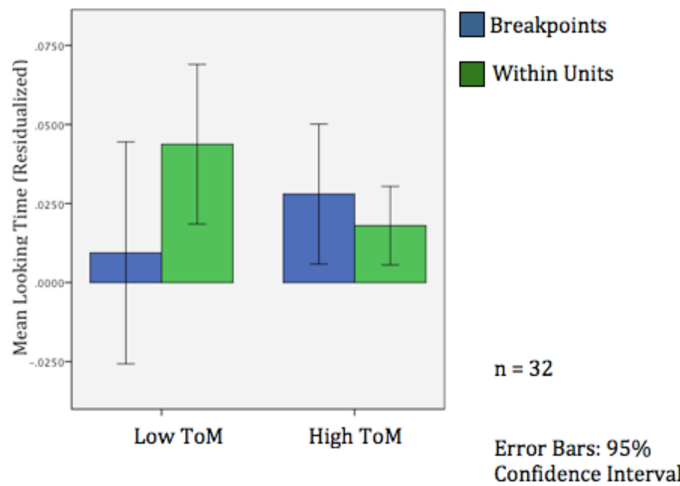


Figure 8. Looking times to pyramid breakpoints versus within units by ToM group. The blue bars represent breakpoints and the green bars represent within units. Error bars set to significance of p = .05 are included.

Overall, we were somewhat able to confirm our hypotheses regarding dwell time. We found that EF and ToM independently predicted looking time to within unit slides such that those higher on EF and ToM measures tended to look less long at within unit slides than those low on EF and ToM, who looked longer at within unit slides. We also found a nonsignificant trend of longer looking to coarse breakpoint slides in those with high EF and ToM than those with lower EF and ToM. This shows a nonsignificant trend towards better segmentation in high EF/ToM children and poorer segmentation in low EF/ToM children.

E. Imitation and Memory Questions

For the off-line action processing measures of imitation ability and memory question performance, we were interested in whether imitation or memory could be predicted by video preference or segmentation ability. We predicted that increased preference for motionese would correspond with increased imitation ability and increased segmentation. Similarly, we predicted that better action segmentation would also predict increased imitation and memory scores.

The two imitation tasks and four sets of memory questions were coded by RAs live and again off-line from the recording of the session. Off-line coding was used for the present analyses and will be later compared to on-line coding for reliability purposes. Imitation tasks were coded for accuracy on a pre-determined set of target actions. The memory questions were forced-choice and were coded for accuracy. Composite imitation scores and composite memory scores were created.

See scatterplots below for correlations of imitation and memory tasks with the dependent variable of video preference. The scatterplots show a positive correlation between imitation and video preference (r = 0.283), and also a positive correlation between memory questions and imitation (r = 0.305). See bar graphs below for trends in imitation and memory ability relative to segmentation. The bar graphs show higher scores on imitation and memory in the consistent segmentation group than in the reverse segmentation group.

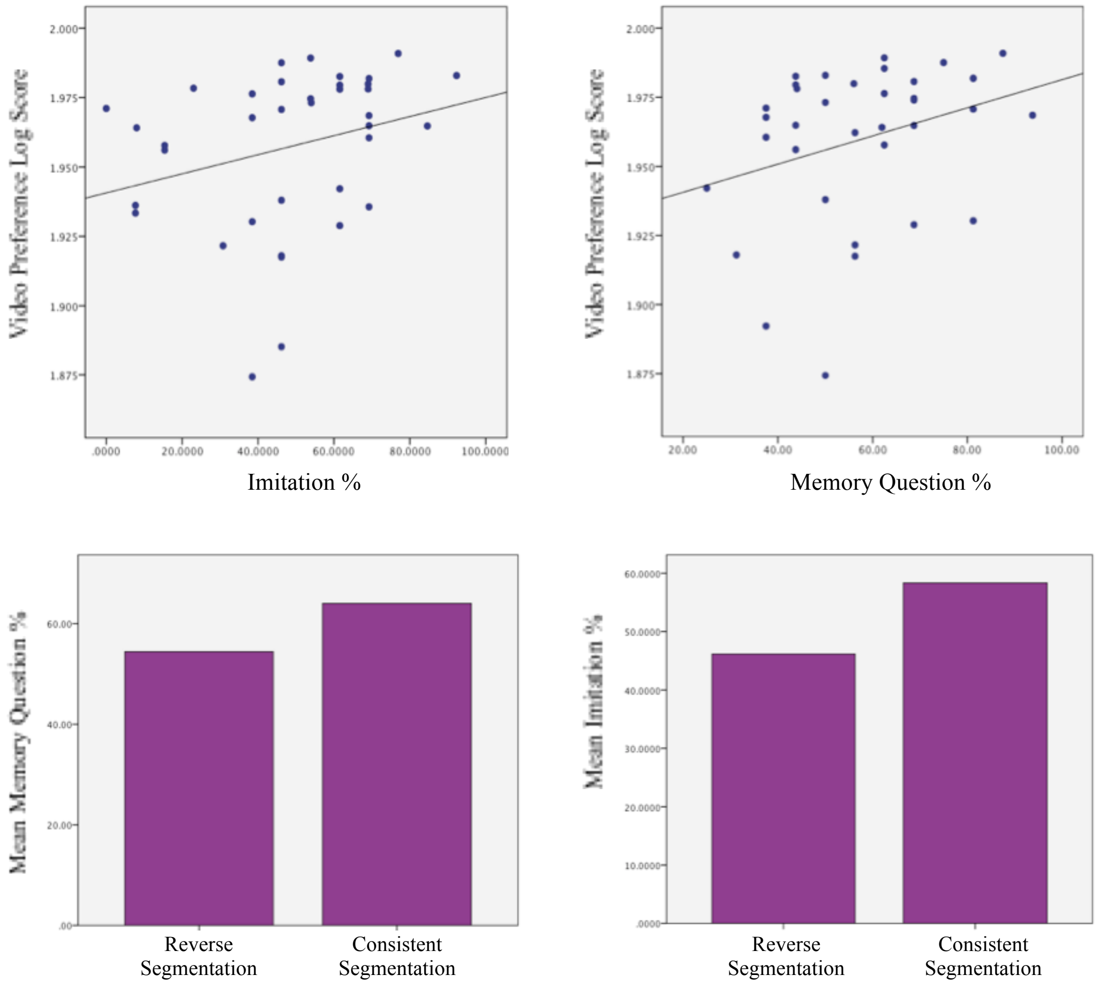


Figure 9: Video Preference and Breakpoint Advantage predict trends in memory question accuracy and imitation ability. Increased imitation and memory scores independently predict increased motionese preference. Consistent segmenters (who looked longer to breakpoints than within units) showed higher average memory and imitation scores than did reverse segmenters.

IV. DISCUSSION

Overview of Results

The current study investigated whether young children show individual and developmental differences in their processing of dynamic, intentional action. Based on preliminary data from 46 neurotypical children, it appears that children's levels of EF and ToM do predict differences in their action processing. Specifically, children who displayed increased performance on EF and ToM tasks showed greater preference for motionese over non-action. Moreover, children with higher EF scores displayed hierarchical segmentation patterns, looking longer to breakpoint slides than to within-unit slides. This pattern of segmentation is consistent with previous research on action processing by adults, children, and infants (Hard, Recchia, & Tversky, 2011; Baldwin, Baird, Saylor, & Clark, 2001), and such increased segmentation typically confers more sophisticated processing of action streams. This suggests that children with high EF and ToM skills are processing action more effectively. These children also performed well on imitation and memory tasks. Children with lower EF and ToM scores, on the other hand, displayed a different pattern of results. These children showed decreased preference for motionese, and a corresponding increased preference for the non-action analog. Additionally, these children showed poorer segmentation skills, decreased performance on imitation tasks, and reduced memory question scores.

Implications

Our results suggest that there may be individual and/or developmental differences in the ways in which children process intentional human action. Children more advanced in the developmental areas of EF and ToM may be processing action in a more adult-like fashion, providing evidence for an evolving mechanism of action processing in which some children are better equipped to understand unfolding action streams than others. This research also has implications for children's social development. Much of the information provided by action displays, particularly motionese, is inherently social, allowing children who are effectively processing action to pick up on important social cues from other people. These social cues are important for children's learning. More efficient action processing earlier in life may confer more social understanding to children, allowing them to interpret the goals and intentions of other people more effectively, in turn allowing them to learn about the world and people around them at an earlier age.

Of our participants, those who were more attracted to exaggerated displays of human action may take in the information, encode it into memory more effectively, and replicate it more accurately than those who are less attracted to motionese. This suggests that motionese preference may grant greater knowledge of action streams. Prior research has shown that action processing abilities are crucial for learning; it could be that children who display low ToM and EF skills are displaying latent developmental skills because they are missing out on important information from the surrounding world. Lesser attention to motionese seems to correspond with reduced processing, imitation, and memory, suggesting that children not attuned to motionese are perhaps less perceptually aware of important events in unfolding action streams.

These results are reminiscent of a hypothesis about ASD by Klin (2002). Klin and colleagues propose that the social deficit in ASD is key to the ASD sympromatology pattern; it may cause these individuals to avoid social sources of information, failing to gather information crucial for learning. It is likely that the natural way in which adults modify their action (motionese) and speech (motherese) have evolved to aid children in learning to communicate with others, inferring intentions through action understanding and speech through language processing. When children lack the ability to focus on social cues of motherese and motionese, they may be missing key learning opportunities. It is possible that Klin's social hypothesis of ASD (2002) extends through the spectrum to typically developing children as well. Any social aversion young children display may lead to later development of important skills; such delayed development may contribute to symptoms of ASD at different levels of functionality, both within the neurotypical population and through the autism spectrum.

These results can be interpreted more broadly; the individual differences seen in this study in processing and preference for motionese may be related to the study by Kuhl and colleagues (2005) and may therefore have implications for children with ASD. Kuhl and colleagues found that a child's aversion for motherese corresponded with ASD symptoms, and through future studies, we may find a similar trend with motionese. As yet, nothing can be said about the ASD population in terms of action processing and preference for motionese, but our results indicate that within the general population, differences in preference and processing exist. Such differences may carry over into the autism spectrum, shedding light on how children with ASD process action, just as Kuhl and colleagues (2005) discovered more about how these children process language.

Limitations

This study also has various limitations that should be addressed through future research. The fact that no participants with ASD were included appears to be a large drawback, but in fact it has enabled the study of individual and developmental differences within the general population. In terms of the study design, we have found several limitations: discrepancies emerged in measuring ToM with parent report versus behavioral tasks, issues arose with the EF working memory task, an unexpected negative correlation with language was found, and we encountered problems with coding one of our dwell time stimuli.

The two measures of ToM that we used, the behavioral tasks performed during the study session and the parent report CSUS, did not account for the same variance in preference or processing of motionese. The preference task was better predicted by the CSUS, while the processing task was better predicted by the behavioral tasks. The behavioral tasks showed the most variability on the easiest tasks (Diverse Desires, Diverse Beliefs, and Knowledge Access), which are the first types of ToM skills that develop in life. Since these lower-level tasks were related to segmentation, it suggests that action-processing ability may be contingent on basic understanding of mentalizing about others. The CSUS taps all types of ToM, suggesting that preference for motionese may be related more to all levels of ToM, perhaps the later developed mentalizing skills in particular. It could also be the case that parents have a mismatched understanding of their child's ToM capacities, leading to reporting bias. However, since the CSUS has been verified in previous research, as have the behavioral ToM tasks, it seems more likely that the two measures are indeed picking up on different nuances of ToM; the behavioral task gleaning information about simple ToM understanding and the CSUS tapping overall ToM understanding.

The other task that did not seem to function as expected during our study was the working memory task, which was a modified version of the “Trucks Task” (Hughes & Ensor, 2005). This task failed to correlate with other developmental measures such as age and ToM. It also did not correlate with the other EF measures, inhibition and task switching. One possibility as to why this task did not correlate with the other measures could be that the modification we made – using pictures of dogs instead of trucks – disrupted the validity of the task. We changed the images from trucks to dogs to have more gender-neutral stimuli; perhaps trucks are more memorable or more distinguishable than dogs are, or maybe children are more familiar with trucks than dogs. Another possibility is that working memory is simply not predictive of ASD symptomatology. This may indeed be the case; Klin and colleagues have found that the skills of planning, flexibility, and set shifting are the areas in which children with ASD are thought to be most impaired, rather than in working memory skills (2002). Still, the fact that working memory did not correlate with developmental measures or other EF measures suggests a flaw in the task design.

The CDI language parent report form also created issues; it was negatively correlated with both preference and segmentation ability, such that children with poorer language skills showed increased preference for motionese and enhanced segmentation skills. This result was unexpected, and should be explored in future studies. We will examine the CDI results and see if they verified after collecting data from a larger sample. If replicated, this result would point to unexpected developmental trends.

One of the dwell time videos used in this study also posed challenges for data analysis; the Puppet video was difficult for expert coders to categorize based on slide type. This slideshow should be further explored because it may yield information about types of motionese less suited for presentation in studies, which may in turn affect types of motionese that are easier or harder for individuals to segment. Many of the slides in the Puppet slideshow were coded as “ambiguous” since coders could not agree on action boundaries. It would be interesting to explore the looking times to the ambiguous slides in the slideshow in neurotypical children of various ToM/EF skills, children with ASD, and adults. Groups with high agreement in looking times to the ambiguous slides may provide important information about how individuals with different developmental levels segment unusual moments in a stream of unfolding action.

Future Directions

After completing data collection for the current study, attaining a larger sample size (60 to 80 participants), we plan to run a sample of children with ASD for comparison. These children will be matched to the mental age of the typically developing 2.5- to 3.5-year-olds in the current study, so the children with ASD will be between 7 and 8 years of age. We predict that children with ASD would show a similar, though more extreme, pattern of results to the lower EF/ToM group in the current study – looking to the non-action analog for a greater percentage of time and not displaying the typical dwell-time pattern indicative of hierarchical processing. This would be analogous to the results found by Kuhl and colleagues (2005) for the speech domain.

Another intriguing modification to this study would be employing a combined stimuli including both motionese and motherese. This way, links between language and action processing could be examined within the same individual. It would be interesting to see whether degree of preference for motionese predicts preference for motherese, and to what extent the speech and action domains are similar or different in their ability to predict ASD symptoms. A combined speech and action study could also show differences in the development of speech and action processing. Since both are crucial for development, observing when one develops different preferences and processing patterns could aid in teasing apart these related areas.

A longitudinal approach would be useful for future studies. It is possible that preference and processing of motionese and motherese stimuli may be predictive of ASD symptomatology later in life. Thus, we plan to investigate samples of adults and infants for comparison. Recruiting a sample of college-aged participants could yield findings about whether preference and processing differences occur later in development. Additionally, investigating an infant sample would yield similar information, and it would be interesting to see if adults, children, and infants all show the same range of preference and processing differences. Samples of children, adults, and infants could act as a starting point for a longitudinal study, tracking infants' preference and processing of motherese and motionese through adulthood, and comparing these variables to measures of ASD. For example, if an infant looks longer to the non-action analog, would that then predict development of ASD symptoms? This leads to broader clinical and diagnostic implications that could from further work on this topic.

Furthermore, if this future research on children with ASD yields the predicted results, then it may be possible to use the video preference and dwell-time tasks as the basis for a diagnostic tool to help identify children with ASD. Since these tasks can be easily adapted to use with infant participants, it maybe even be possible to use them for earlier detection of ASD. Given the connection between speech and action, it might be most effective for such a diagnostic tool to assess infants' and young children's processing of both motionese and motherese.

Conclusions

Taken as a whole, this study has potential to aid our understanding of how ToM and EF contribute to action processing. Decreased ToM and EF abilities have been documented in individuals with ASD, and both contribute to what Klin and colleagues (2002) have proposed as the underlying deficit in ASDs: social dysfunction. Low ToM skills are associated with decreased ability to mentalize about other individuals, leading to difficulties interpreting others' speech and actions, which, accoding to this framework, causes social disengagement. Decreased EF abilities have also been associated with areas of the brain involved in social functioning (Klin et al., 2002). Motionese, like motherese, is a social form of communication; therefore, children with ASD may find it aversive. Reduced attention to motionese would be consistent with evidence about children with ASD displaying low attention to motherese (Kuhl et al., 2005); together, future analogous findings about motionese aversion combined with known motherese aversion could provide further evidence for Klin's model of the social deficit in individuals with ASD.

Further information about how individuals, particularly developmentally delayed children, process action and speech has potential to facilitate the creation of specialized teaching and learning programs. Such programs could be tailored for the individual processing styles of neurotypical and developmentally atypical children with differential processing and preference patterns. In this way, we might be able to aid children's processing of action and their resulting communication skills. For example, if a child has trouble understanding instructions for a task, perhaps a different instruction method could be created to eliminate potential aversive social factors.

The knowledge gleaned from this and similar studies could have impacts both in clinical and educational settings, as well as explaining how humans may differ in understand the world around them. Exploring the ways in which individuals take in, process, store, and interpret information, as well as how these mechanisms evolve over the lifespan, is key to developmental psychology. Understanding the processes that drive cognitive understanding can aid in accommodating those who comprehend information differently.

V. APPENDICES

Appendix 1: Study Set-Up

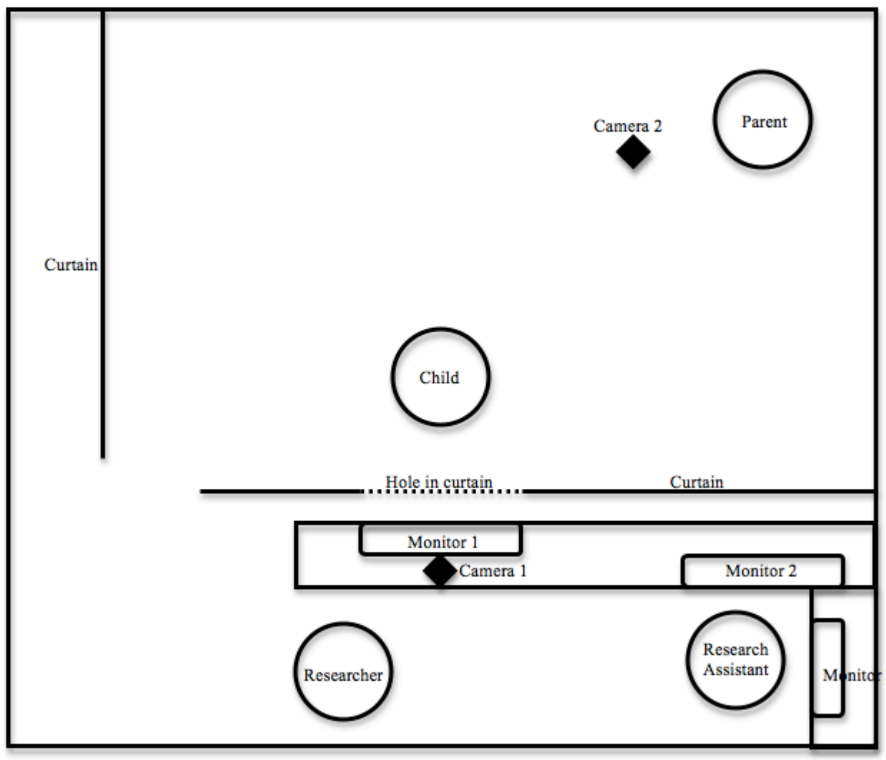


Figure 10. Set-Up for Computerized Tasks

- Child sits in front of hole in curtain and watches stimuli on Monitor 1 through hole in curtain.
- Research assistant controls stimuli presentation with Monitor 2 and codes live feed from Monitor 3.
- Parent fills out forms.
- Camera 1 records child, Camera 2 records Monitor 1.

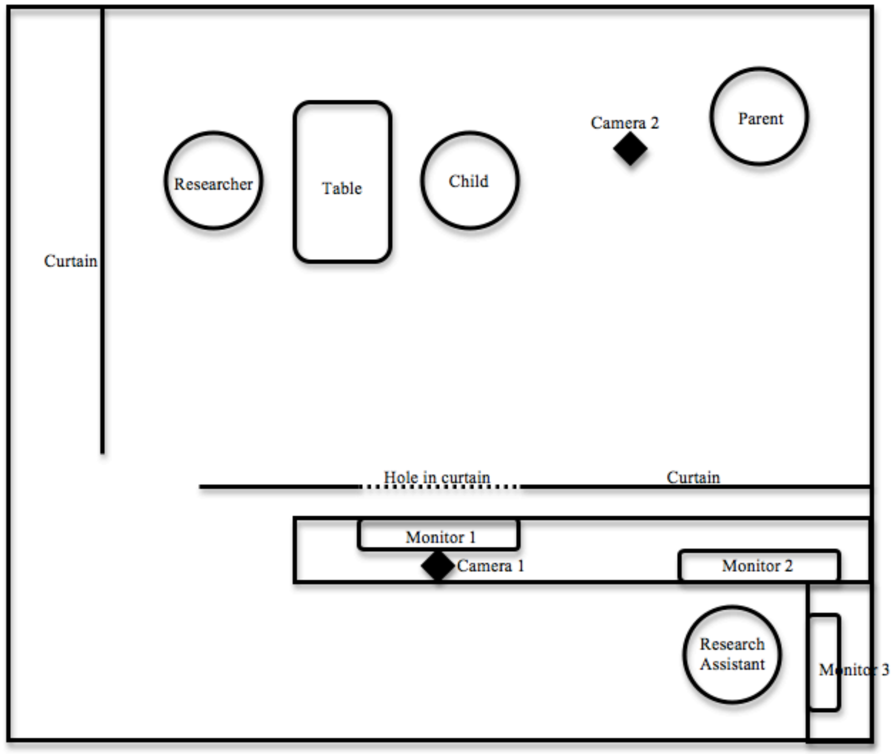


Figure 11. Set-Up for Non-Computerized Tasks (Executive Function and Theory of Mind).

- Child sits with researcher at Table to do tasks.
- Camera 1 records Table, Child, and Researcher.
- Research Assistant codes behavior from Monitor 3.
- Parent fills out forms.

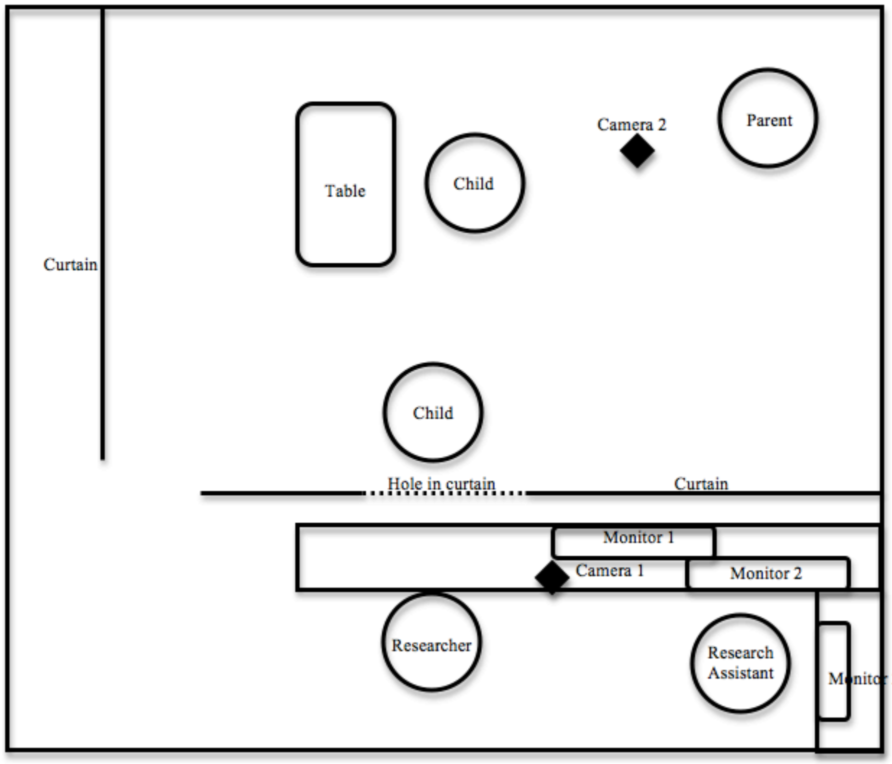


Figure 12. Set-Up for Non-Computerized Task (Imitation).

- Monitor 1 and Camera 1 are moved to the right.
- Researcher sits behind Hole in Curtain and demonstrates toy actions.
- Child sits in front of Hole in Curtain and watches the demonstration.
- Camera 1 records Child, Camera 2 records Researcher.
- After demonstration, Child moves to Table. Researcher brings Child the toy, Child imitates at the table.
- Research Assistant codes behavior from live feed on Monitor 3.
- Parent fills out forms.

Appendix 2: Toy and Motionese Descriptions

<u>Toys for Video Preference Task</u>

Video Toy 1: "Lego Man"

"Lego Man" consisted of the actor putting together large Lego-like pieces to create the figure of a person. The actor holds out each successive piece while smiling before putting each piece in place and emphasizes pressing one piece into the other.

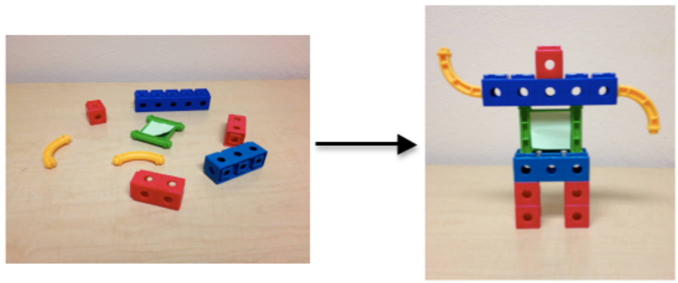


Figure 13. Lego Man toy. Pieces before and after actor puts toy together.

Video Toy 2: "Sorting"

The "Sorting" task involves putting red and yellow balls and vehicles into one side or the other of a plastic container. The objects are sorted by functionality, either a ball or a vehicle, rather than by color, red or yellow. The actor uses exaggerated gestures to indicate the functionality of the toy (rolling the ball between her hands and "driving" the cars across her open palm) before placing it into the appropriate side of the container.

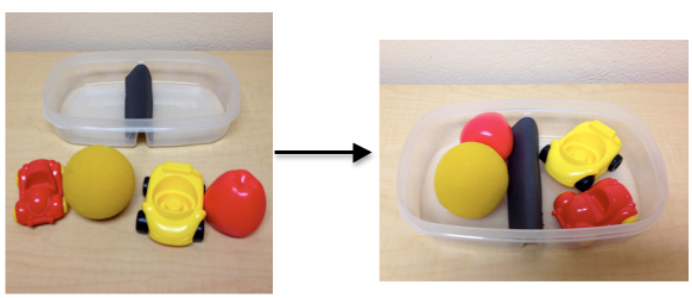


Figure 14. Sorting toy. Objects before and after sorted into bins.

Video Toy 3: "Balloon"

The "Balloon" task uses an air pump to fill a balloon with air. The actor places the balloon on the pump and uses exaggerated pumping motions to demonstrate the action sequence until the balloon is filled. The actor then waves at the balloon as the air is released.

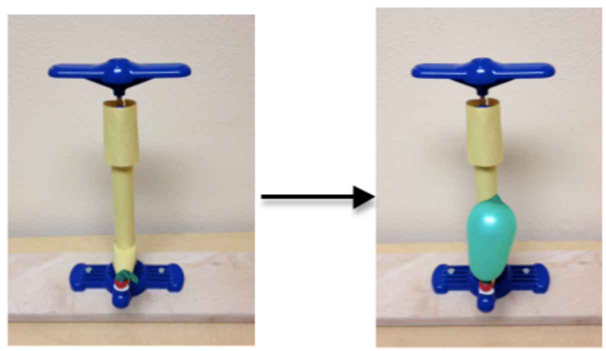


Figure 15. Balloon toy. Balloon before and after inflation.

Video Toy 4: "Train"

"Train" involves the actor demonstrating various actions with a toy train. The actor twists the front of the train down to form the "cannon." She puts a plastic lion into the front of the train and presses a button to release the lion from the cannon. The actor then opens middle of the train to reveal a platform. She places a plastic boy on the platform and presses a button to make the boy spin. All actions (twisting, opening, placing, and pressing buttons) are accompanied by exaggerated motions and facial expressions.

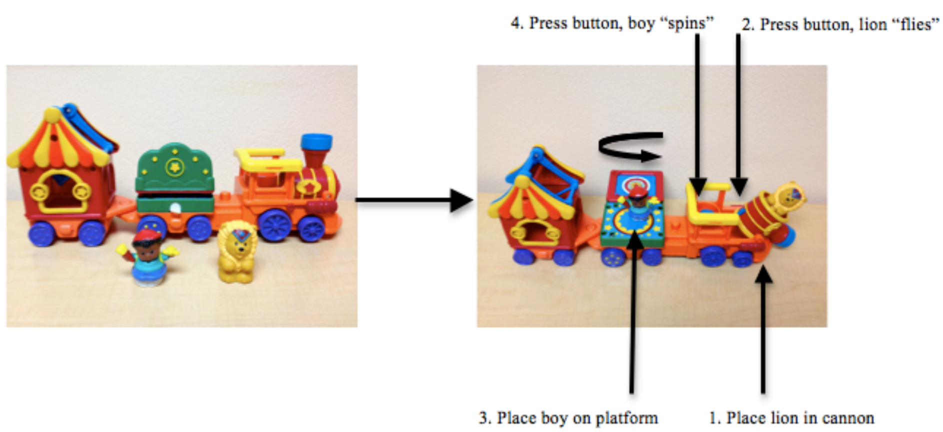


Figure 16. Train toy. Shows characters before and after being placed on train. Arrows and descriptions on right indicate actions done to characters and train.

<u>**Toys for Dwell Time Task**</u>

Dwell Time Toy 1: "Pyramid"

The "Pyramid" video consisted of an actor manipulating toys attached to a cardboard pyramid covered in felt. The attached toys were a plastic duck, three pieces of plastic fruit, and a foam dog puzzle. The toys attached to the pyramid could all be manipulated in some way (the small duck in front of the large duck could be pulled, the fruits could be affixed to the pyramid, and the foam dog figure could be moved to finish the puzzle. Each type of toy was affixed to a different side of the pyramid such that when the pyramid was rotated, a new toy could be displayed and manipulated. This toy was based off of a study done by Kara Sage for her dissertation (Sage, in publication). Again, all aspects of the pyramid toy manipulation were done in the motionese style.

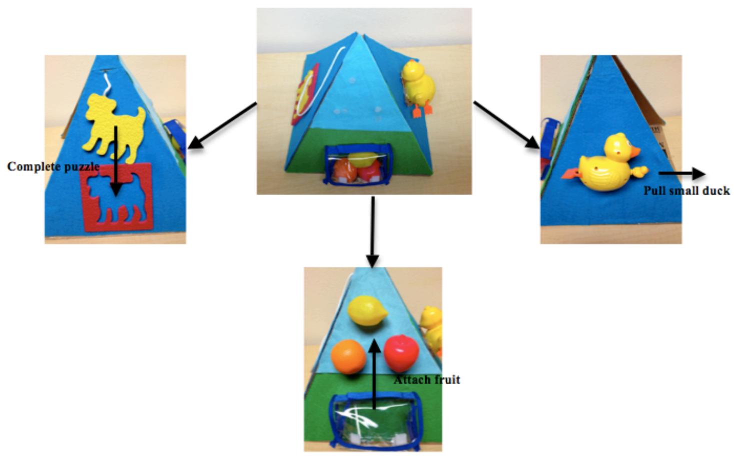


Figure 17. Pyramid toy. Central image depicts entire toy, each of three other images depict each side of the toy. Arrows and descriptions explain actions done on each side of toy.

Dwell Time Toy 2: "Puppet"

The second toy was a panda puppet created out of cardboard with detachable parts (arms, eyes, and mouth). The actor attached each of the parts to the puppet in turn. She then placed the puppet on her hand and made it open and close its mouth. The actor engaged with the toy in a the motionese style, including waving at the completed puppet.

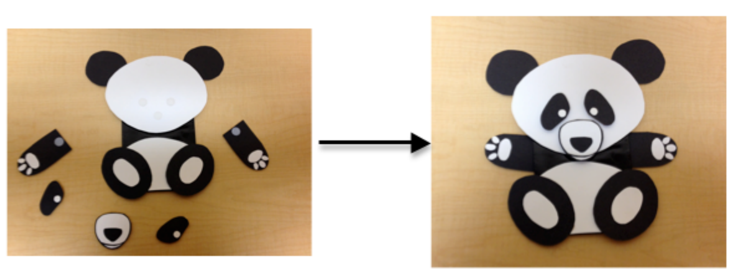


Figure 18. Puppet toy. Shows before and after puppet toy is put together.

Toys for Imitation Task

Imitation Toy 1: "Play-Doh"

The "Play-doh" toy involves a small plastic elephant that comes apart. Then, play-doh is inserted between the pieces of the elephant and the pieces are pushed together in order to extrude play-doh out the nose of the elephant. Then, the top piece is twisted to slice through the play-doh. The end result is small pieces of play-doh shaped like butterflies. The pushing and twisting motions were especially emphasized though motionese.

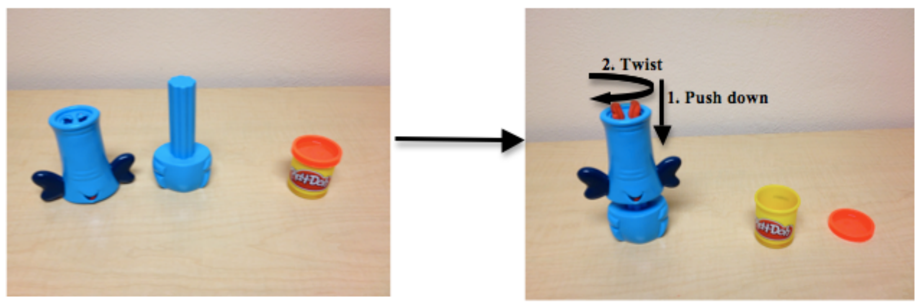


Figure 19. Play-doh toy. Shows before and after play-doh is extruded. Arrows and descriptions on right image explain actions performed on toy.

Imitation Toy 2: "Star"

The "Star" toy is a wooden disc with five pegs sticking upright from the surface. Strings with colored popsicle sticks are attached to each peg. Each stick matches the color of one of the pegs. The actor places each stick on the appropriate peg and the result is a star shape made from the strings. Each color is indicated in turn via motionese.

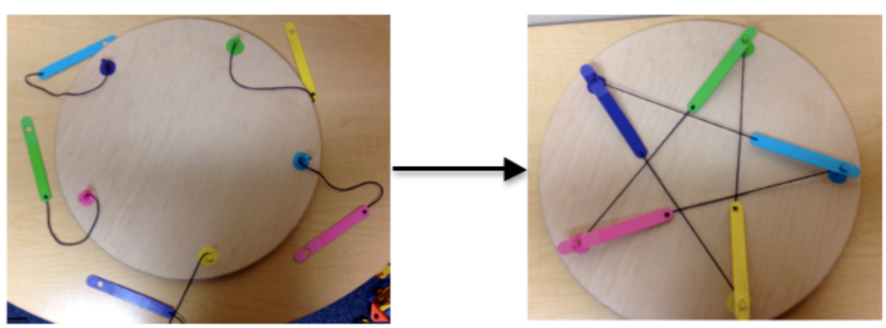
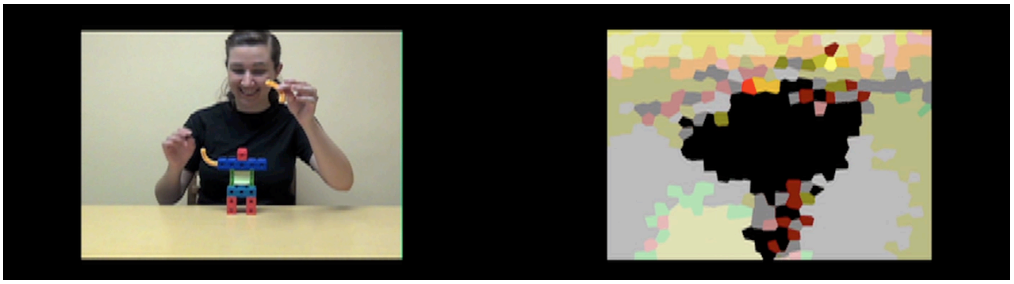


Figure 20. Star toy. Shows before and after sticks are attached to form yarn star.

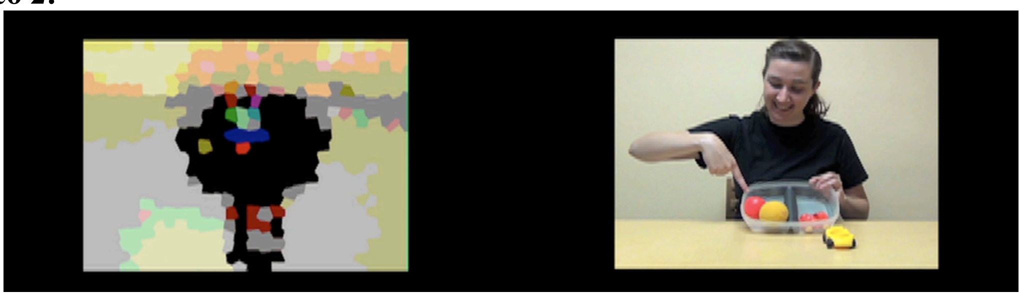
Appendix 3: Video Preference "Doubled Videos"

Video 1:



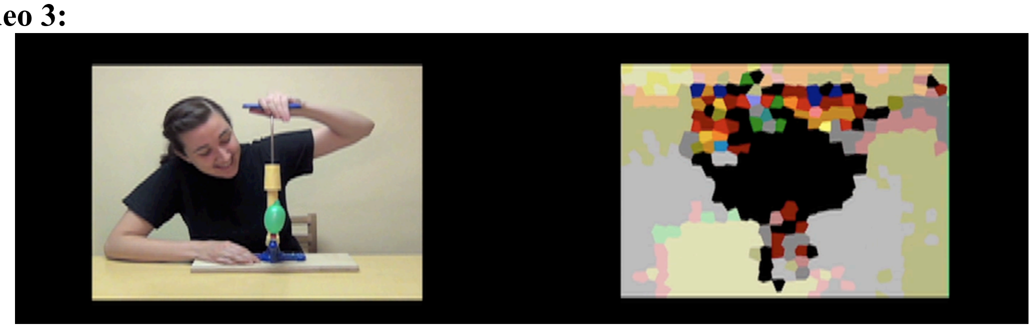
Lego Man Motionese | Sorting Pixelated & Rotated

Video 2:



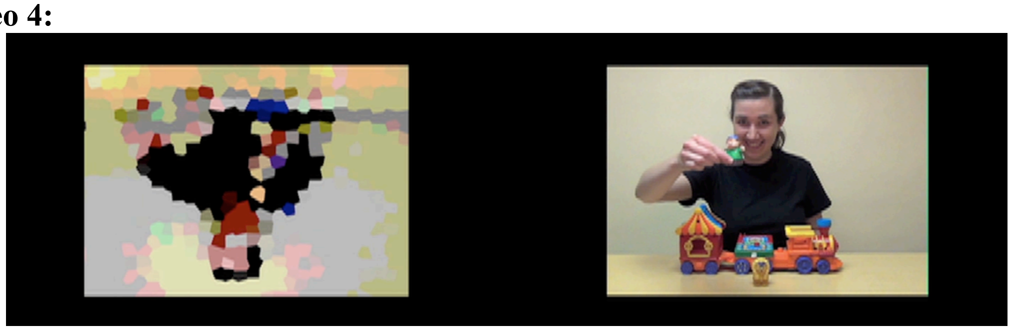
Lego Man Pixelated & Rotated | Sorting Motionese

Video 3:



Balloon Motionese | Train Pixelated & Rotated

Video 4:



Balloon Pixelated & Rotated | Train Motionese

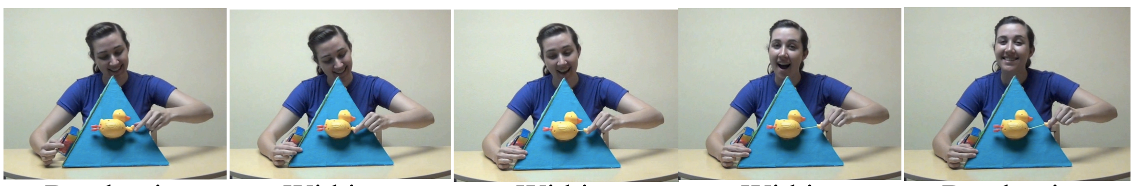
Figure 21. Screen shots of four video preference stimuli.

Appendix 4: Dwell Time Slides

Practice Slides

Practice slides: not goal-directed, merely displays the contingency of picture change with mouse click

Pyramid Toy



Breakpoint Within Within Within Breakpoint

Puppet Toy

Breakpoint Within Within Breakpoint Within

Figure 22. Screen shots of dwell time stimuli. Practice slides, pyramid slides, and puppet slide examples. Pyramid and puppet slides are marked with breakpoint vs. within unit designations.

Appendix 5: Executive Functioning Scripts

EF1: Task Switching—Modified Card Sort (Diamond, 2005)

Introduce shape task:*"Ok, this is a house* [point to display card] *and this is a bird* [point]. *Can you point to the house? And can you point to the bird? Good job! We're going to play a shape game, ok? You're going to put these cards into the trays. You're going to put the houses in this side and the birds in this side. Can you point and show me where the house cards go? Great! And can you point and show me where the bird cards go? Great! Now it's your turn!"*

Sort 4 practice cards with corrections (orange house, green bird, green house, orange bird). Retry all 4 cards again if any mistakes are made. Hold up one by one, say *"Where does this one go?"* If correct, *"Great job! Let's try another one! Where does this one go?"* If incorrect, say *"Oh, this one is a bird (house), so it goes here. Let's try another one! Where does this one go?"* Once all the way through practice one, if mistakes were made, redo the practice. For each practice card, hold it up and ask where it goes, then hand the card to the child so they can place it in the tray.

After practice is done, say *"Great job! Let's play."* Hand to kid one by one. If kid doesn't place card in a tray, say, *"Where does this one go?"* If they ask if they are correct, say, *"You show me this time!"*

Introduce color task: *"Ok, great job with that game! Now we're going to play a different game. This game is a color game! This card is green* [point] *and this card is orange* [point]. *Can you point to the green card? And can you point to the orange card? Good job! For this game, all the green cards go here and all the orange cards go here. Can you point and show me where the green cards go? Great! And can you point and show me where the orange cards go? Great! Now it's your turn!"*

Sort 4 practice cards with corrections (orange house, green bird, green house, orange bird). Retry all 4 cards again if any mistakes are made. Hold up one by one, say *"Where does this one go?"* If correct, *"Great job! Let's try another one! Where does this one go?"* If incorrect, say *"Oh, this one is green (orange), so it goes here. Let's try another one! Where does this one go?"* Once all the way through practice one, if mistakes were made, redo the practice.

After practice is done, say *"Great job! Let's play."* Hand to kid one by one. If kid doesn't place card in a tray, say, *"Where does this one go?"* If they ask if they are correct, say, *"You show me this time!"*

Figure 23. Stimuli for card sorting task. Above: tray with 2 sides and attached example cards. Below: the four practice trial cards.

<u>EF 2: Working Memory</u><u>—Modified Truck Task (Hughes & Ensor, 2005)</u>

Introduce task: *"This game is a guessing game, and you can win lots of stickers! Here are some cards; each has two pictures of dogs. One dog is the right one that will give you a sticker, and one is not. You have to guess which dog will give you a sticker."*

Card 1: This card is always correct. Say, *"Which dog do you think will give you a sticker? Good job! That's the one with the sticker!"*

Cards 2-8: *"Which dog do you think will give you a sticker?"* Follow the marks on the backs of the cards (either X's or O's) for the "correct" pattern of dogs.

If **incorrect**, say *"Oh, that wasn't the dog with the sticker. Try to remember the other dog for next time."*

If **correct**, say *"Good job! That's the one with the sticker!"*

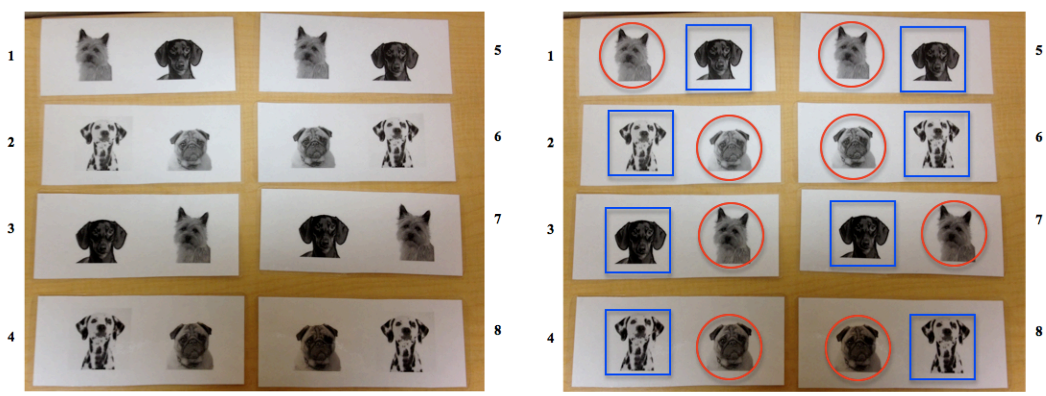


Figure 24: Stimuli and answer key for working memory task. Left image: card stimuli. Right image: Key of correct answers. If child picks dog on left (with red circle) for trial 1, all subsequent correct answers follow the red circle path. If child picks card on right (with blue square) for trial 1, all subsequent answers follow the blue square path. Trial 1 is always correct, and trial 2 is at chance. Blue square answers and red circle answers are counterbalanced for side of card (left vs. right).

EF 3: Gift Delay—Modified Bow Task (Carlson, 2004)

Place bag with box and slinky inside on table. Say, *"I've got a prize here in this bag for you. But before you open it, I need to go get the bow for the bag. Just sit in this chair until I get back. I'm going to go get the bow now. So just stay in that chair; don't touch the bag or what's inside it, and don't peek inside the bag until I come back with the bow."*

Return with bow after 3 minutes or when child opens bag. If they have waited the entire time, say, *"Great job waiting!"* Or if they open the bag early, say, *"That's ok, it's really hard to wait."*

Figure 25: Stimuli for gift delay task. Gift bag with bow attached**.**

Appendix 6: Theory of Mind Tasks Scripts

ToM1: Diverse Desires (Wellman & Liu, 2004)

"This is Sammy. It's snack time, so Sammy wants a snack to eat. Here are two different snacks: cookies and carrots. Which snack would you like best? Would you like cookies or carrots?"

If child selects cookies [*same*], Sammy will want carrots [*opposite*]. If child selects carrots [*same*], Sammy will want cookies [*opposite*].

"Well, that's a good choice, but Sammy really likes [opposite]. He doesn't like [same]. What he likes best are [opposite]. So, now it's time to eat. Sammy can only choose one snack, just one. Which snack will Sammy choose? Cookies or carrots?"

Figure 26. Diverse Desire stimuli. Sammy is pictured in the middle with cookies on one side and carrots on the other.

ToM2: Diverse Beliefs (Wellman & Liu, 2004)

"This is Linda. Linda wants to find her cat. Her cat might be hiding in the garage or it might be hiding in the tree. Where do you think the cat is? In the garage or in the tree?"

If child selects garage [*same*], Linda will think the cat is in the tree [*opposite*]. If child selects tree [*same*], Linda will think her cat is in the garage [*opposite*].

"Well, that's a good idea, but Linda thinks her cat is in the [opposite]. She doesn't think it's in the [same]. Linda thinks her cat is in the [opposite]. So, now it's time for Linda to look for her cat. Where will Linda look for her cat? In the garage or in the tree?"

Figure 27. Diverse Beliefs stimuli. Linda is pictured in the middle with an image of a garage on one side and an image of a tree on the other.

ToM3: False Contents (Wellman & Liu, 2004)

"Here's a crayon box. What do you think is inside the crayon box? [*You don't know? That's ok, take a guess! You don't know? Ok, let's find out!*] *Let's find out!* [open] *Wow, it really has ribbons inside!* [close]*"*

Memory Question 1: "*Ok, do you remember what is in the crayon box?"* If child answers memory question incorrectly, they are excluded from task.

[hold up Sammy] *"Here's Sammy again. Sammy hasn't seen inside this crayon box. Now here comes Sammy."*

Target Question: *"So, what does Sammy think is inside the box? Does Sammy think there are crayons or ribbons inside the box?* [*You don't know? That's ok, take a guess!*]"

Memory Question 2: *"Did Sammy see inside the box?"* If child answers memory question incorrectly, they are excluded from task.

Figure 28. False Contents stimuli. Crayon box with ribbons inside. During task, box is closed and then opened to reveal ribbons.

<u>ToM4: Knowledge Access</u> <u>(Wellman & Liu, 2004)</u>

“*Here’s a box. What do you think is inside the box?* [*You don’t know? That’s ok, take a guess! You don’t know? Ok, let’s find out!*] *Let’s find out!* [open] *Wow, it really has a spring inside!* [close]

Memory Question 1: *“Ok, do you remember what is in the box?* [Hold up Linda] *Here’s Linda again. Linda hasn’t seen inside the box. Now here comes Linda.”* If child answers memory question incorrectly, they are excluded from task.

Target Question: *“So, does Linda know what is inside the box?* [*You don’t know? That’s ok, take a guess!*]*”*

Memory Question 2: *“Did Linda see inside the box?”* If child answers memory question incorrectly, they are excluded from the task.

Figure 29. Knowledge Access stimuli. Box with pink spring inside. During task, box is closed and then opened to reveal spring.

Appendix 7: Memory Questions

Correct answers are *italicized*.
Question types are as follows:

1. Coarse-grain question about overall goal of action

2-4. Fine-grain questions about details of action including one question about HOW I did something, what ORDER I performed the actions, and WHAT part of the toy was.

<u>Dwell Time Question Set 1: Pyramid Toy</u>

1. What did I do with the toy?
   Did I take it apart or *play with it*?

2. What fruit did I put on the toy?
   Did I use a *lemon* or grapes?

3. What did I do with the duck?
   Did I squeeze the large duck's beak or did I *pull on the small duck*?

4. Which part of the toy did I play with last? Was it the *puzzle* or the duck?

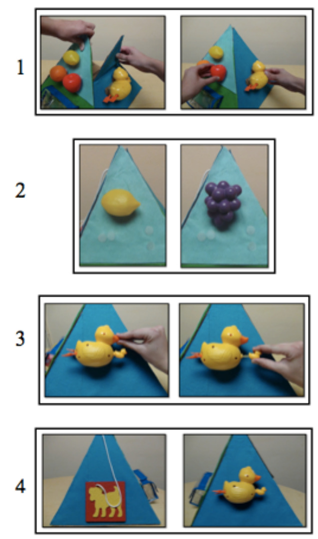


Figure 30. Pyramid toy question pictures. Screenshots of question stimuli.

<u>Dwell Time Question Set 2: Puppet Toy</u>

1. What did I do with the pieces of the puppet?
   Did I take the puppet apart or did I *put the puppet together*?

2. What color were the eyes of the puppet?
   Were they *white* or blue?

3. What did I do with the puppet when I was done?
   Did I *put it on my hand* or did I flip it upside down?

4. Which piece did I put on right before I put the puppet on my hand? Was it the eyes or the *arms*?

Figure 31. Puppet toy question pictures. Screenshots of question stimuli.

<u>Imitation Task Question Set 1: Play-Doh Toy</u>

1. What did I do with the elephant? Did I *use the elephant to make butterflies* or did I use the elephant to roll out the play-doh on the table?

2. What came out of the elephant? Play-doh *butterflies* or play-doh worms?

3. What did I do with the play-doh butterflies that I made? Did I stick them back inside the toy or did I *set them on the table*?

4. What did I do after the butterflies came out of the top of the elephant? Did I twist the toy or did I *take the butterflies out*?

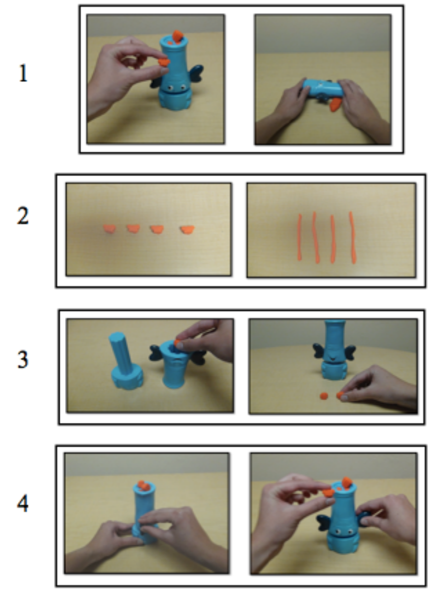


Figure 32. Play-doh toy question pictures. Screenshots of question stimuli.

<u>Imitation Task Question Set 2: Star Toy</u>

1. What did I make with the yarn? Did I make a circle or a *star*?

2. What did I do with the blue tool? Did *I put the blue tool on the blue peg* or did I put the blue tool on the yellow peg?

3. What did I do with the tools and the yarn? Did I *stick the tools on the pegs* or did I wrap the yarn around the pegs?

4. Which tool did I move first? Was it the green tool or the *pink tool*?

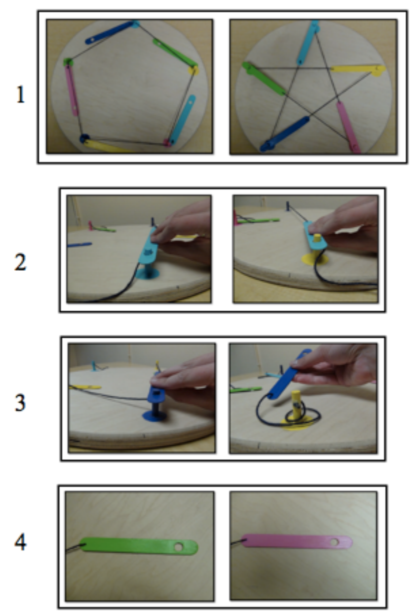


Figure 33. Star toy question pictures. Screenshots of question stimuli.

VI. REFERENCES

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(2), 202-212.

Auyeung, B., Wheelwright, S., Allison, C., Atkinson, M., Samarawickrema, N., & Baron-Cohen, S. (2009). The Children's Empathy Quotient and Systemizing Quotient: Sex Differences in Typical Development and in Autism Spectrum Conditions. *J. Autism Dev Disord, 39*(11), 1509-1521.

Baldwin, D.A. (2012). Redescribing action. M. Banaji & S.A. Gelman (Eds.), *Navigating the Social World: A Developmental Perspective.* Oxford University Press.

Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition, 106*, 1382-1407.

Baldwin, D.A., Baird, J.A., Saylor, M.M., & Clark, M.A. (2001). Infants parse dynamic action. *Child Development, 72*(3), 708-717.

Baldwin, D.A. & Moses, L.J. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development, 10*(3), 309-329.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze (SDG) strategy to crack the code of language? *Child Development, 68,* 48–57.

Brand, R.J., Baldwin, D.A., & Ashburn, L.A. (2002). Evidence for "motionese": Modifications in mothers' infant-directed action. *Developmental Science, 5*(1), 72-83.

Brand, R.J. & Shallcross, W.L. (2008). Infants prefer motionese to adult-directed action. *Developmental Science, 11*(2), 853-861.

Bruner, J. (1981). The social context of language acquisition. *Language & Communication, 1*(2/3), 155-178.

Carlson, S.M. & Moses, L.J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032-1053.

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: stability and prediction from ages 2 to 3. *Developmental psychology*, *40*(6), 1105.

Chawarska, K., Volkmar, F., & Klin, A. (2010). Limited attentional bias for faces in toddlers with autism spectrum disorders. *Arch Gen Psychiatry, 67*(2), 178-185.

Cooper, R.P. & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61,* 1584-1595.

Fenson, L. (2007). *MacArthur-Bates Communicative Development Inventories: user's guide and technical manual*. Paul H. Brookes Publishing Company.

Fernald, A. & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*, 279-293.

Freitag, C.M., Konrad, C., Haberlen, M., Kleser, C., von Gontard, A., Reith, W., Troje, N.F., & Krick, C. (2008). Perception of biological motion in autism spectrum disorders. *Neuropsychologia, 46*, 1480-1494.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological bulletin*, *134*(1), 31.

Grimminger, A., Rohlfing, K.J., & Stenneken, P. (2010). Children's lexical skills and task demands affect gestural behavior in mothers of late-talking children and children with typical language development. *Gesture, 10*(2-3), 251-278.

Hard, B.M., Recchia, G., & Tversky, B. (2011). The shape of action. *Journal of Experimental Psychology: General, 140*(4), 586-604.

Hubert, B. Wicker, B., Moore, D.G., Monfardini, E., Duverger, H., Da Fonseca, D., & Deruelle, C. (2007). Brief report: Recognition of emotional and non-emotional biological motion in individuals with autism spectrum disorders. *J Autism Dev Discord, 37*, 1386-1392.

Jones, W. & Klin, A. (2009). Heterogeneity and homogeneity across the autism spectrum: The role of development. *J Am Acad Child Adolesc Psychiatry, 48*(5), 471-473.

Kemler Nelson, D.G., Hirsh-Pasek, K., Jusczyk, P.W., Wright Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *J. Child Lang, 16*, 55-68.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Defining and quantifying the social phenotype in autism. *Am J Psychiatry, 159*, 895-908.

Kuhl, P.K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews: Neuroscience, 4*, 831-843.

Kuhl, P.K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Developmental Science, 8*(1), F1–F12

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in cognitive sciences*, *12*(2), 72-79.

Macari, S.L., Campbell, D., Gengoux, G.W., Saulnier, C.A., Klin, A.J., & Chawarska, K. (2012). Predicting developmental status from 12 to 24 months in infants at risk for autism spectrum disorder: A preliminary report. *J Autism Dev Discord*, 42, 2636-2647.

Myhr, K.L., Baldwin, D.A., & Brand, R.J. (2004). Probing the benefits of infant-directed action. (*Unpublished Manuscript*).

Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment*, *87*(1), 102-112.

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child development*, *72*(5), 1394-1408.

Sage, K., & Baldwin, D. (2011). Disentangling the Social and the Pedagogical in Infants' Learning about Tool‐Use. *Social Development*, *20*(4), 825-844.

Sage, K., & Baldwin, D. (2012). Exploring Natural Pedagogy in Play Preschoolers: Cues Parents Use and Relations Among Them. *Education Research and Perspectives*, *39*(1), 153-181.

Shipley, T.F. & Zacks, J.M. (2008). Understanding events: From perception to action. *Oxford: Oxford UP.*

Wellman, H.M. & Liu, D. (2004). Scaling theory-of-mind tasks. *Child Development, 75*(2), 523-541.

Zelazo, P.D., Jacques, S., Burack, J.A., & Frye, D. (2002). The relation between theory of mind and rule use: Evidence from persons with autism-spectrum disorders. *Infant and Child Development, 11*, 171-195.