2022

The Theory of Mind Hypothesis of Autism

Kendall Smith
ksmith64@gardner-webb.edu

The Theory of Mind Hypothesis of Autism

An Honors Thesis
Presented to
The University Honors Program
Gardner-Webb University
18 November 2022

by

Kendall C. Smith

Accepted by the Honors Faculty

Dr. David Carscaddon, Thesis Advisor	Dr. Wilson Hawkins, Director of Univ. Honors
Dr. Meredith Rowe, Honors Committee	Dr. Elizabeth Amato, Honors Committee
Dr. Angelina Smith, Honors Committee	Dr. Abby Garlock, Honors Committee

Abstract

Present theory of mind research suggests that autistic individuals demonstrate deficits in theory of mind capabilities. A literature review was conducted to investigate the claim made by the theory of mind hypothesis of autism that theory of mind deficits are responsible for the social deficits present in autistic individuals. It was concluded that this hypothesis was prematurely accepted as an explanatory model for autism when it is better described as a symptom of autism. Alternative explanations for autistic social deficits such as executive function deficits and the theory of weak central coherence were analyzed. Implications and suggestions for future research are briefly discussed.

Keywords: Autism Spectrum Disorder, autism, theory of mind, cognitive development, social deficits

Table of Contents

The Theory of Mind Hypothesis of Autism

"In saying that an individual has a theory of mind, we mean that the individual imputes mental states to himself and to others (either to conspecifics or to other species as well)" (Premack & Woodruff, 1978, p. 515). In essence, the theory of mind is a cognitive ability that allows people to recognize the mental states of others as well as understand how they affect behavior. Mental states, such as *want, think,* and *believe,* are fundamental motives for many behaviors, both social and private, therefore, recognizing these states in others is a key component in maintaining interpersonal relationships as well as understanding social conduct. Given its significance for social interactions, research on the theory of mind rapidly gained popularity among psychologists who wanted to investigate the extent of its impact. As a result of this research, it was discovered that the theory of mind is not acquired equally among all people. Specifically, deficits in theory of mind capabilities were found among individuals with autism, resulting in the development of the theory of mind hypothesis of autism. This hypothesis stands as an explanatory model for the social deficits present in individuals with autism and posits that these deficits are a result of an underlying inability to cognitively represent the mental states of others. This hypothesis, however, has declined in popularity since its first introduction due to the serious criticism it received for its inability to account for the complex nature of autism. The autistic experience not only varies greatly within the autistic community, but it is also defined by more than an assortment of deficits. Due to this, alternative explanations that incorporate a more comprehensive account of autism have been proposed as replacements to the theory of mind hypothesis. With these alternative explanations, psychologists began to realize that the theory of mind hypothesis had been accepted as an explanatory model for autism before it had been properly researched. As it stands now, the theory of mind continues to be regarded as a universal

cognitive ability that has practical implications for individuals who demonstrate a deficit in this ability; however, it is argued that the theory of mind hypothesis of autism should be redefined as a symptom of autism rather than an explanatory model.

Theory of Mind

History

The concept of theory of mind was initially introduced by an American psychologist by the name of David Premack. In 1978, alongside a graduate student named Guy Woodruff, Premack conducted a research study which aimed to discover whether a fourteen-year-old, African-born chimpanzee named Sarah had the cognitive capacity to infer mental states. To investigate this, Premack and Woodruff created a series of comparative studies whose research designs were specifically developed to assess whether Sarah could demonstrate behaviors indicative of a theory of mind. To maintain uniformity, Premack and Woodruff kept the same research design for each experiment in the series, allowing for only slight variation between the independent variables. In each experiment, Sarah was presented with a series of videos in which a human actor portrayed a variety of problem situations. In the initial experiment, the "problem" in each situation was that the human actor was unable to reach an object of their desire. This inaccessible object, a bunch of bananas, remained constant, but the circumstance preventing the human actor from obtaining it varied in one of four ways: (1) it was out of reach vertically after being attached to the ceiling, (2) it was out of reach horizontally after being placed outside a cage wall, (3) it was within reach outside the cage wall but a box inside the cage obstructed the actor's reach, or (4) the box from problem situation (3) was not only obstructing the actor's reach, but was also weighed down by cement blocks (Premack & Woodruff, 1978).

After being shown each video in turn, Sarah was presented with two photos: one which depicted the actor engaging in a behavior that constituted a solution to the problem, and one which depicted the actor engaging in a behavior unrelated to the problem (Premack & Woodruff, 1978). Using her understanding of behavior and intention, Sarah's task was to identify the correct solution that correlated to each problem situation. Out of twenty-four trials, Sarah successfully chose the correct answer in twenty-one trials ($p < .001$), and all her errors were limited to the condition involving the cement blocks (Premack & Woodruff, 1978).

The subsequent experiments in Premack and Woodruff's series implemented the same research design; however, the content of the videotapes and solution photographs was modified to assess the specificity of Sarah's problem-solving abilities. In their assessment of this specificity, Premack and Woodruff attempted to "escape the usual definition of animal problem" by presenting Sarah with a variety of human-specific problem situations such as dealing with a malfunctioning heater, trying to play an unplugged phonograph, and trying to use a hose that is not properly attached to a faucet (1978, p. 520). These situations give the term "problem" a richer meaning, which provides a more comprehensive assessment of Sarah's theory of mind capabilities (Premack & Woodruff, 1978). The variety of these situations accounted for a range of conditions such as problems not restricted to physical inaccessibility, the modalities "would," "should," and "would like," agent-specific knowledge, observational learning, the "pretend" versus "real" distinction, etc. (Premack & Woodruff, 1978). Taking into consideration the slight variation from study to study, the combined results of these experiments closely reflected the outcome of the initial study.

In response to these outcomes, Premack and Woodruff provided four possible interpretations: identification of matching physical elements, classical associationism, empathy,

and theory of mind (1978). According to the first interpretation, Sarah's success is explained by matching the physical elements shown in the video, such as the props and the human actor, to the corresponding physical elements in the solution photograph (1978). This interpretation, however, is promptly rejected by Premack and Woodruff because their research design was intentionally developed to control for this effect. By ensuring that each physical element was present in every video and solution photograph, the explanation of physical matching is no longer feasible.

The second interpretation, classical associationism, is an explanation of behavior rooted in the idea of sequence familiarity. The concept of sequence familiarity states that if a person, or animal, is presented with an incomplete sequence they recognize, they will attempt to complete it by choosing an element that symbolizes a proper end to the sequence (Premack & Woodruff, 1978). Another interpretation of this concept is the theory of interrupted action, which Premack and Woodruff describe by stating, "If an animal is interrupted when carrying out an act, he will recommence and complete the act as soon as possible. An animal *shown* an interrupted action will do the same, as long as he is intelligent enough to recognize representations of actions" (1978, p. 516). The weakness of this interpretation, however, is its inability to account for an unpredictable future. How can a human, or in this case an animal, anticipate the next event in a sequence they are unfamiliar with? To answer this question, the principle of generalization provides a probable solution. This principle allows a person, or animal, to assume the correct solution to a problem by drawing from a previous experience that is similar in nature to the one presently presented to them. Therefore, this principle increases the probability of accuracy in unfamiliar circumstances by allowing general knowledge to be applied to a broad range of situations. However, despite its suitability it most cases, the principle of generalization is acceptable at best if the data is vague and the sequence in question lacks depth in its description

(Premack & Woodruff, 1978). Thus, by presenting the problems and their solutions explicitly and in acute detail, Premack and Woodruff discredit the principle of generalization as the factor influencing Sarah's behavior, and therefore reject the probability of classical associationism as an explanation.

The third alternative interpretation offered by Premack and Woodruff is the concept of empathy. In this interpretation, the animal "sees the human actor struggling to reach the bananas, 'puts himself in the place of the actor,' and chooses an alternative in keeping with what he would do were he in the actor's predicament" (Premack & Woodruff, 1978, p. 518). In this scenario, the animal's choice is no longer based on inferences made about the actor's knowledge of problem resolution, but rather a prediction of his own behavior if he were in the actor's situation (Premack & Woodruff, 1978). It is important to note that the empathy interpretation can only be supported if the animal's choices are unaffected by the identity of the actor. Once the actor's identity has an influence over the animal's behavior, the empathy view must be dismissed (Premack & Woodruff, 1978). Taking this into consideration, Premack and Woodruff test this influence by replacing the human actors with two trainers Sarah is familiar with: Keith, her favorite trainer, and Bill, a fictitious name for an acquaintance of Sarah's whom she has never shown affection. Throughout the experiment, Sarah was consistently more accurate when the portrayed actor was Keith rather than Bill. Furthermore, when given the choice between "good" photographs, which depicted the correct solutions, and "bad" photographs, which depicted unfavorable mishaps that could happen to the actor in the course of them attempting to solve the problem, she favored Keith by attributing more "good" photographs to him than Bill (Premack & Woodruff, 1978). The results of this experiment suggest that Sarah's behavior is significantly

influenced by the identity of the actor, thereby supporting the rejection of the empathy view as an alternative interpretation (Premack & Woodruff, 1978).

After identifying and rejecting the three alternative interpretations for Sarah's behavior, Premack and Woodruff offer one final interpretation: the presence of a theory of mind (1978). This interpretation is recommended to the reader as the most viable explanation for Sarah's behavior because her accuracy in choosing the correct answer suggests that she imputes two mental states to the human actor: intention and knowledge (1978). The first mental state, intention, is imputed to the human actor as soon as Sarah recognizes the objective of the actor's actions in each problem scenario. The second mental state, knowledge, is imputed by Sarah when she assumes the human actor knows how to solve the problem that prevents them from accomplishing this objective (Premack & Woodruff, 1978). This interpretation, according to Premack and Woodruff, is the most logical explanation for Sarah's behavior because it remains valid even when the independent variable is manipulated to control for various influences that could affect behavior. Where the other alternative explanations failed to remain valid across different circumstances, the validity of the theory of mind explanation remained constant. Therefore, Premack and Woodruff conclude their study with the claim that chimpanzees, and by extension humans, have the cognitive ability known as the theory of mind, which allows them to impute mental states and use them to understand and predict behavior (1978).

Development as a Cognitive Construct

Following Premack and Woodruff's 1978 publication, the theory of mind became a trending topic within the cognitive sciences. Its rapid increase in popularity prompted scholars to investigate the characteristics and consequences of this theory. Due to this, a new era of cognitive research emerged that focused on how the ability, or inability, to understand other

minds affects social functioning as well as interpersonal relationships. From a modern perspective, the development and subsequent acceptance of the theory of mind as a cognitive construct is oftentimes solely attributed to the research done by Premack and Woodruff; however, in retrospect, current theory of mind research owes much of its success to the extensive groundwork laid by former cognitive psychologists.

As is true with many cognitive theories, the theory of mind traces its origin to Jean Piaget, a cognitive psychologist whose research on childhood development significantly contributed to the field of developmental psychology (Flavell, 2004). A predominant principle endorsed by Piaget was the idea of cognitive egocentrism, a concept simply defined as the "embeddedness of one's own view" that manifests itself during the early stages of a child's cognitive development (Rubin, 1973, p. 102). This principle signifies that a child, at this stage of development, is "unable to shift mental perspective in order to differentiate among several aspects of an event and between his own and others' points of view" (Rubin, 1973, p. 102). This understanding of egocentrism has played a crucial role in the foundation of theory of mind research because, as this definition reveals, the nature of an egocentric mind directly reflects the functional consequences of one which lacks the theory of mind. The direct correlation of these two principles not only provides insight into the cognitive mechanisms behind the theory of mind, but also affirms Piaget's role as a founding figure in the history of the theory of mind.

As an extension of Piaget's work, research regarding metacognitive development came to the forefront of cognitive sciences (Flavell, 2004). The central focus of metacognitive research is the concept of metacognition, which is defined as "any knowledge or cognitive activity that takes as its object, or regulates, any aspect of any cognitive activity" (Flavell, 2004, p. 275). This principle has been applied as the basis for a variety of research regarding a child's cognition,

such as comprehension, problem solving, communication, and perception, which all are skills that are interconnected with the theory of mind (Flavell, 2004).

Prior to the introduction of the theory of mind as an independent concept, Piagetian tradition and metacognitive research were the leading theories that were widely accepted as the basis for understanding a child's cognition. However, following Premack and Woodruff's historic 1978 publication, the opinion of Piaget's work regarding egocentrism and the resulting work on metacognition shifted from them being leading cognitive theories to primitive theory of mind research. This reconceptualization of Piaget's work and the work on metacognition as foundational theory of mind research demonstrates how theory of mind principles have been present in cognitive psychology decades before their identification in Premack and Woodruff's publication. Thus, this research, regardless of how it may develop moving forward, recognizes that its origins are rooted in the work done by cognitive psychologists who preceded it.

Key Features of the Theory of Mind

As the theory of mind began being researched as an independent construct, discoveries were made about its role in social interactions and interpersonal relationships that evolved the understanding of this construct from a theoretical state to one with practical applications. As part of this evolution, an important addition was made to Premack and Woodruff's original definition of the theory of mind. As it was first introduced, the theory of mind encompassed the ability to recognize or infer another's thoughts or beliefs. To this skill, a second criterion was added, which incorporated the ability to understand how these thoughts not only may differ from one's own, but also how they affect behavior (Lewis et al., 2017). It is with this modified understanding of the theory of mind that subsequent research has based its findings.

Embedded within this definition are several underlying features that are essential to the theory of mind as a psychological construct. Of these features, a person's capacity for meta-representation is often identified as one of the most essential concepts to the understanding of the theory of mind. Specifically, the capacity for meta-representation is significant because it is a required skill that must be developed within a child before the theory of mind can be actualized (Lewis et al., 2017). Meta-representation, according to the cognitive sciences, is defined as a person's mental capacity to not only represent a situation, but also concurrently represent their own and another person's relationships to this situation (Astington & Dack, 2020). This capability is particularly present in philosophical discussions regarding the theory of mind because it is widely recognized as a "necessary precondition for moral responsibility, self-consciousness, and social interaction," each of which are diversely affected by a person's capacity, or lack thereof, for the theory of mind (Wimmer & Perner, 1983, p. 104). The representations referred to within meta-representation, also known as "mental states," are the personal beliefs and desires that serve as the internal motivations that drive our external behaviors and thus, "mediate our activity in the world" (Astington & Dack, 2020, p. 366). Throughout the research, these mental states are also referred to as "intentional states," however, in this context, the term "intentional" takes the technical meaning "aboutness" derived from philosophical literature rather than its everyday usage to mean "deliberate" (Astington & Dack, 2020, p. 366). In keeping with this terminology, the expression used to describe the number of individual mental states involved in a single condition is known as the order of intentionality (Lewis et al., 2017). Strictly speaking, the orders of intentionality reflect a degree of awareness. Beginning with zero, this order represents entities, such as computers and some species of insects, which are "not aware of the contents of their 'minds,'" and thus lack a sense of self-

awareness (García Landa, 2013, p. 2). In these instances, there are zero mental states being accounted for, thus the term zero-order intentionality is used to describe this state of awareness. In contrast, the first order of intentionality involves one mental state and indicates the beginning of self-awareness. This order includes "beliefs and desires (etc.) but no beliefs and desires *about* beliefs and desires" (Dennett, 1983, p. 345). This order takes the form "x *believes that* p," as in the sentence, "Amanda believes that today is Wednesday" or "y *wants that* q," as in the sentence, "Jonathan wants that chocolate bar" (Dennett, 1983, p. 345). To build off this, second-order intentionality "has beliefs and desires (and no doubt other intentional states) about beliefs and desires" (Dennett, 1983, p. 345). Meaning, second-order intentionality includes two mental states: a personal mental state and the mental state of another person. This order takes the form "x *wants* y to *believe* that x is …" as in the sentence "Jeremy wants Kristina to believe that Jeremy is sick" or "x *believes* y *expects* x to …" as in the sentence "Margaret believes Christopher expects Margaret to cook dinner" (Dennett, 1983, p. 345). Each subsequent order builds off the previous order by adding an additional mental state to the condition. Theoretically, there are an infinite number of intentional states, however, normal adults typically reach their cognitive limit around the fifth order, with only a small percentage of people able to perform successfully past that point (Lewis et al., 2017). Recent theory of mind research has focused specifically on second-order intentionality because it is the first order in which a child has developed the ability to understand the mental states of others well enough to recognize a false belief (Lewis et al., 2017). It is with this understanding of second-order intentionality that further research into the measurement of theory of mind capabilities has based its findings.

Theory of Mind Measurement

After its debut as a psychological construct in Premack and Woodruff's study, the significance of the theory of mind was challenged through a series of commentaries written by scholars who claimed that the original research had been poorly conducted. The main critique ascribed to Premack and Woodruff's publication was that their conclusions were unfounded because they had been drawn from insufficient findings (Lewis et al., 2017). Commentators and critics alike believed Premack and Woodruff failed to clearly identify how Sarah's ability to choose the correct photograph signified the presence of a "theory of mind," therefore, the critics concluded that Premack and Woodruff's research was well-intentioned, but not well-supported. As support for their claim, commentators identified the distinction between the recognition of intention and the attribution of belief as the main flaw in Premack and Woodruff's research design. The original 1978 study focused mainly on Sarah's ability to recognize the goal, or intention, of the person in the problem scenario. When she was repeatedly able to identify the correct solution, it supported Premack and Woodruff's claim that Sarah was practically applying a theory of mind. However, as the theory of mind began to take shape outside of the original research, it became increasingly clear that the attribution of belief, rather than the recognition of intention, is a superior indicator of the theory of mind. Only in the case of belief attribution can one confidently conclude that Sarah was inferring a mental state when choosing the proper solution rather than simply responding as they would if they were in the situation themselves (Astington & Dack, 2020). Therefore, because the claims made in Premack and Woodruff's original research were founded upon the recognition of intention, it was concluded that in order to adequately represent the theory of mind within the research, a more accurate way to identify it needed to be established.

From this line of reasoning, an entirely new line of research that focused on the measurement of the theory of mind began to develop. This facet of theory of mind research sought to answer questions regarding the extent to which a person had developed the theory of mind, and how it could be qualitatively measured as an independent cognitive skill. Two influential figures whose research significantly contributed to the answering of these questions were Heinz Wimmer and Josef Perner. In 1983, Wimmer and Perner conducted a series of experiments whose purpose was to investigate children's cognitive capacity to represent another person's definite belief, even if the child recognizes it to be inconsistent with the truth. Simply put, they wanted to scientifically investigate at what age a child could represent a false belief in another person. Their justification for such an endeavor lies within the practical importance of representing a false belief, which Wimmer and Perner believed to consist in the "use of this representation as a frame of reference for interpreting or anticipating the other person's actions" (1983, p. 106). By accurately representing a false belief, a child has the ability to not only predict the behavior of another person, but also to manipulate the situation through deception to accomplish a personal agenda. It is for this reason that the capability to deceive is often regarded as reasonable evidence for the presence of a theory of mind. Deceptive action, according to Wimmer and Perner, indicates the theory of mind for two reasons: (1) deceptive strategies require a high level of adaptability for flexible application, and (2) they require the conceptualization of a person's wrong belief as an implicit objective in their planning strategy (1983). Despite its significance as a constituent of the theory of mind, Wimmer and Perner wanted to conduct research that could provide insight into how the theory of mind could be identified aside from behaviors such as deception.

To compose their research questions and create their research design, Wimmer and Perner took inspiration from studies that investigated developmental characteristics that are retrospectively identified as precursors to the theory of mind. These studies, such as those conducted by Hood and Bloom (1979) and Shultz et al. (1980), demonstrated that children have a basic, innate capability to represent a mental state and express the relation in which they, and another person, stand to this representation (Wimmer & Perner, 1983). The findings of these studies, despite their significance in identifying innate instances of theory of mind, are insufficient in representing a fully actualized theory of mind because they exclusively investigate how children represent mental states that mirror their own, without requiring them to represent states that are distinctly different. In stating that a "more complicated meta-representational problem arises when one has to explicitly represent the *difference* between one's own and somebody else's relation to the *same* propositional content," Wimmer and Perner identified the critical component that previous theory of mind research lacked: the ability to "account for the lack of knowledge in another person" (1983, p. 105). With this additional component to consider, Wimmer and Perner implemented a false belief paradigm they believed could accurately test a subject's comprehension of a false belief held by another person. The specifics of this paradigm are described as follows: "The subject is aware he/she and another person observe a certain state of affairs x. Then, in the absence of the other person the subject witnesses an unexpected change in the state of affairs from x to y. The subject now knows that y is the case and also knows that the other person still believes that x is the case" (Wimmer & Perner, 1983, p. 106). In order to make this paradigm realistically functional as a comprehension test for a false belief, it needed to be adapted into story format. From this, various stories, such as the story of Maxi, have been constructed. Maxi, the lead protagonist, places a chocolate bar into cupboard x. Following this,

Maxi leaves the scene. Then, in his absence, his mother removes the chocolate bar from cupboard x and relocates it to cupboard y. At this point, test subjects are asked to indicate via verbalization or pointing gestures in which cupboard Maxi will look for his chocolate when he returns (Wimmer & Perner, 1983). If the test subject is capable of representing Maxi's false belief (the chocolate is in cupboard x) separate from what they understand to be true (the chocolate is in cupboard y), they will give the accurate response by choosing cupboard x. Several variations of this false belief task have been adapted from this original paradigm to investigate how changing the conditions in the scene affects the accuracy with which the test subjects respond, however, in any form, this paradigm is still the most widely accepted assessment to measure the theory of mind.

Despite its popularity as a theory of mind measure, the reliability of false belief tasks has not been extensively researched. Furthermore, the few studies that have been conducted have produced mixed results (Hughes et al., 2000). One of the first studies that examined the test-retest reliability of false belief tasks, conducted by Mayes et al. (1996), suggests "unacceptably poor results for first-order false-belief tasks" (Hughes et al., 2000, p. 483). However, upon further investigation, it was discovered that the word choice, question sequence, and targeted skill of the question affected the accuracy with which the child answered (Hughes et al., 2000). In addition to these factors, the child's cognitive and verbal processing abilities were also found to be influences affecting the child's success (Hughes et al., 2000). Taking this into consideration, researchers Hughes et al. (2000) conducted a second study of test-retest reliability that included modified questions that controlled for the issues faced in the first study. In contrast to the study conducted by Mayes et al. (1996), the results of this study indicated good reliability for tasks that assess theory of mind capabilities. The mean kappa score of .52 indicated fair to

moderate reliability for any single false belief task, and a high Cronbach's alpha score indicated "significant internal consistency for aggregate scores" (Hughes et al., 2000, p. 487). From their findings, Hughes et al. also concluded that intellectual ability does not affect the reliability of task performance.

This conclusion is of particular significance because theory of mind tasks are administered to a wide variety of children, many of whom have developmental disorders such as Down syndrome and autism spectrum disorder (Hughes et al., 2000). Thus, strong within-child correlations allow theory of mind tasks to be used to cross assess normally developing children with developmentally delayed children without compromising their findings due to flawed reliability. This conclusion is specifically pertinent to autism research because historically, researchers have questioned the validity of false belief tasks for "individuals with both ASD [autism spectrum disorder] and limited verbal skills because 'they lack the cognitive and verbal skills necessary to answer the control questions, success on which is usually an inclusion criterion'" (Hutchins et al., 2008, p. 204). However, with the improved reliability of false belief tasks due to the inclusion of factors that control for intellectual differences, false belief tasks are now believed to be sufficient for both clinically normal and developmentally delayed children.

Alternative assessments for measuring theory of mind abilities, such as deceptive container tasks and appearance-reality tasks, may not be as widely practiced as the original false belief paradigm, but they are essential in providing a new perspective on how theory of mind abilities manifest in various scenarios (Perner & Lang, 1999). As an extension of the false belief paradigm, the deceptive container task assesses the extent to which a child has an awareness of others' false beliefs as well as their own, however, its use of a different assessment method provides a new perspective on theory of mind capabilities. To execute the deceptive container

task, a child is shown an ordinary container, such as a Smarties box, and is asked what its contents might be. The use of context clues allows most children to respond correctly by identifying the item on the container, in this example, by saying "Smarties." However, following their response, the child is shown the real contents of the box, which are typically items, such as pencils, which have no logical connection to the anticipated contents. Once the child is made aware of their own false belief, they are asked two questions: (1) what would a person who has not seen the real contents believe is in the box? and (2) what did you believe was in the box before you were shown the real contents (Perner & Lang, 2001)? The way the child answers these questions is indicative of whether they have an advanced theory of mind, or if it is still in the early stages of development.

In addition to the deceptive container task, another beneficial theory of mind assessment is the appearance-reality task (Perner & Lang, 1999). Success on the appearance-reality task is significant not only because this distinction is apparent in everyday emotional, perceptual, conceptual, and social activity, but it also seems to "presuppose the explicit knowledge that human beings are sentient, cognizing *subjects* whose mental representations of objects and events can differ, both within the same person and between persons" (Flavell, 1986, p. 419). Therefore, because of its direct correlation to theory of mind abilities, this task is highly valued among scholars and commonly administered as a theory of mind assessment. The model for the appearance-reality assessment involves the presentation of an illusory stimulus, such as a sponge that resembles a rock, to a child followed by two questions: one about appearance, and one about reality (Flavell, 1986). Although the nature of the stimulus is meant to be challenging, it is presented in a nondeceptive manner. For example, before being asked either of the questions, the child is encouraged to investigate the stimulus to discover its characteristics. This allows for a

more accurate assessment of the child's theory of mind because it eliminates the possibility of inaccuracy based on deception. The manner in which a child responds to these theory of mind measurements, in addition to the original false belief paradigm, provides a comprehensive, but not exhaustive, perspective on how theory of mind capabilities manifest themselves in children's cognition.

Course of Development in a Child

Following its introduction in Premack and Woodruff's 1978 publication, the theory of mind rapidly gained popularity among scholars. These scholars sought to expand the theory of mind from its rudimentary form to a fully developed psychological construct. To achieve this, a range of empirical questions regarding its typical course of development needed to be answered. The first logical step toward achieving this was to administer theory of mind tests in experimental settings to pinpoint the factors that affect theory of mind development in a child.

As is typical of other cognitive functions, scholars agree that there is not a singular moment of acquisition when the theory of mind is instantaneously developed in a child. Rather, its development is a process of evolving social understanding that occurs during a child's first formative years, where the theory of mind results as a culmination of this process (Astington & Dack, 2020). To further understand this process, Henry Wellman and David Liu (2004) created a five-stage scale that ranks theory of mind tasks in ascending order based on level of difficulty and age of acquisition. This scale follows the structure and scoring requirements of a Guttman scale, which means the items on the scale are "ranked in difficulty such that if a person responds positively to a given item, that person must respond positively to all easier items. Thus, theoretically, a given score on a Guttman scale can only be reached with one pattern of response, and if we know a person's score, we know how that person responded to all items in the scale"

(Wellman & Liu, 2004, p. 532). Furthermore, a given score on this type of scale not only indicates the cumulation of skills required to get that score, but it also allows for an accurate assumption of which skills the person lacked that would cause them not to progress past that point. Beginning with the lowest level of difficulty, the five stages on Wellman and Liu's scale are (1) diverse desire, (2) diverse belief, (3) knowledge access, (4) contents false belief, and (5) real-apparent emotion (Wellman & Liu, 2004). The first stage, diverse desire, requires a child to recognize that another person may have differing desires in response to the same stimuli. For example, the assessment model for this stage typically involves presenting a child with a toy figure and two separate images depicting food items, such as cookies and carrots. The child is asked to choose which item they would prefer for snack time. Following their choice, the child is informed that the toy figure likes the opposite choice. The test in this scenario is to ask the child which food the toy figure would choose at snack time. To pass the test, they would need to respond with the food item opposite to their preference (Wellman & Liu, 2004). The second stage, diverse belief, requires a child to not only recognize that another person's beliefs may differ from their own, but also requires them to understand how these different beliefs affect their behavior. The assessment model for this stage follows the classic false belief paradigm proposed by Wimmer and Perner (1983), in which the child observes a story character place an item in a specific location, then, in the character's absence, the item is relocated to a different location. (Wellman & Liu, 2004). At this point, the child is tested on their ability to recognize how their knowledge differs from the character's belief, by identifying where they think the character will look for their item when they return. To pass this test, the child must choose the first location, where the character remembers putting it, rather than the second location, where the child knows it to be.

The third stage, knowledge access, focuses on the idea that seeing leads to knowing. During this stage, a child must realize that the opposite is also true, that not seeing an event leads to a lack of knowledge about said event. Therefore, the child must understand that knowledge is subject to experience, thus, their knowledge may differ from other people who have different experiences (Wellman & Liu, 2004). The assessment model for this stage involves presenting the child with an object, such as a plastic box, with an unidentified object inside. Following this, the child is asked what they think could be inside the box. In this scenario, the child's answer to this question is of no particular importance because immediately following their response, they are presented with the actual contents of the box. Once this knowledge has been established, a character, who was not present when the contents of the box were revealed, is introduced to the scene. At this point, the child is tested with two questions: (1) did this character see inside the box, and (2) does this character know the contents of the box (Wellman & Liu, 2004)? To successfully complete this task, the child must understand that the character's knowledge differs from their own because they did not experience the moment when the box was opened, and thus answer "no" to both questions (Wellman & Liu, 2004).

The fourth stage, contents false belief, is similar to the third stage in the sense that it requires the child to be aware of how experience affects knowledge. However, in this stage, the child is first made aware of their own false belief. The assessment model for this stage parallels the deceptive container task identified in Perner and Lang's 1999 study mentioned above. In this assessment model, a child is presented a container with a distinct label on it, and subsequently asked what they believe is in the box. Following their response, the child is shown the true contents of the box, which are different from what the label indicates. To be successful in this task, a child must recognize that another person will experience the same false belief as they did

about the contents of the container before being made aware of the truth. Therefore, when asked what someone else will think is in the box, the child must respond with the item indicated on the label rather than the item they know is in the box.

The fifth, and final stage of the theory of mind scale, is the real-apparent emotion distinction. In this stage, a child must be able to understand that a person can feel one emotion, their real emotion, while simultaneously displaying a different emotion, a façade. The assessment model for this stage begins with emotion identification questions to assess whether the child has the base ability to label emotions. Following this, the test administrator tells the child a story about a fictional character. In the story, the character's friends are mean to him, but he chooses to hide his feelings about it. At this point, the child is shown a picture of the character expressing an emotion that is unrelated to his true emotion. After this, they must answer two questions: (1) how does the character actually feel in this situation, and (2) how did they try to make it look like they felt in the situation (Wellman & Liu, 2004)? To be successful in this task, the child's response to the first question must be distinctly more negative than their response to the second question. If a child successfully passes this final stage, due to the nature of a Guttman scale, it reveals their cumulative knowledge of mental representations, and thus indicates the presence of an advanced theory of mind. The identification of these five stages significantly contributed to theory of mind research because it drastically changed how this skill was recognized in a child. The preexisting belief about the theory of mind was that it was an all-or-nothing skill, where the child either had it or they lacked it; however, the introduction of this scale changed that belief and replaced it with the idea that the theory of mind is a gradation of ability. This redefinition allows researchers to more accurately identify a child's level of cognitive skill as well as pinpoint the degree to which they possess the theory of mind.

Research suggests that normally developing children typically progress through these stages and reach its culmination by the end of the preschool years, when the theory of mind, unless hindered by certain circumstances, becomes fully realized (Astington & Dack, 2020). This claim is supported by the findings of many false belief studies, most notably, those conducted by Wimmer and Perner (1983). In their initial studies implementing a myriad of false belief tasks on children between the ages of three and nine, Wimmer and Perner conclude by stating, "The convergence of findings from such a variety of different tasks suggests that around the ages of 4 to 6 years the ability to represent the relationship between two or more persons' epistemic states emerges and becomes firmly established" (1983, p. 126). These findings are historically important in the development of the theory of mind as a construct because it pinpoints the age range in which the theory of mind typically manifests itself in a child. From this, an even more significant conclusion can be drawn: failure to develop a fully realized theory of mind within said age range could be indicative of a developmental delay or cognitive impairment.

The potential implications for children who lack a theory of mind introduced questions about the factors that affect the development of the theory of mind within a child. One of these factors that has proven to be particularly influential is the role of culture. In their original 1978 publication, Premack and Woodruff claim that the theory of mind is an innate, universal cognitive ability in human adults, and as previously mentioned, the typical age of acquisition is between four and six. However, many researchers question whether these claims remain true cross-culturally. To investigate this question, a psychologist by the name of Penelope Vinden conducted several research studies that explored how theory of mind abilities develop in children from non-Western cultures (Garfield et al., 2001). Based on the results of these studies, Vinden concluded that the developmental pattern of the theory of mind is not culturally universal

(Garfield et al., 2001). The support for this claim stems from her study conducted on a group of children, aged four to eight years old, who were from the remote Junín Quechua dialect group of Peru. In this study, Vinden administered standard theory of mind tests and found that a majority of the oldest children repeatedly answered incorrectly to questions assessing their understanding of their own and others' false beliefs (Garfield et al., 2001). Within the same study, Vinden also found that increasing age had no statistical significance on the improvement of the children's performance (Garfield et al., 2001). Following this, Vinden conducted a similar study in which she concludes that children from Western cultures (Europe, North America, and Australia) significantly outperform children from non-Western cultures (Tolai and Tainae cultures of Papua New Guinea, and Mofu culture from Northern Cameroon) in their level of success on theory of mind assessments (Garfield et al., 2001). Potential explanations for the variation in the development of the theory of mind across cultures pertains to linguistic differences, cultural discrepancies regarding behavior and the understanding of its causes, or a combination of both (Garfield et al., 2001). Regardless of its explanation, the importance of this distinction lies within its interpretation and how it affects the collective understanding of the theory of mind. Although the findings of Vinden's studies support the notion that the theory of mind is innate in all human beings, they also suggest that it is more complex than originally believed. Furthermore, her studies also suggest that the theory of mind is a cognitive process that operates according to its own autonomous dynamics, rather than being driven solely by biology (Garfield et al., 2001). In other words, the theory of mind is a complex process that is susceptible to the influence of factors such as culture and environment. This autonomy is particularly evident in the comparison of its development between normally developing children and children with intellectual, social, or neurodevelopmental disorders, such as autism spectrum disorder.

Autism Spectrum Disorder

Definition and History

Autism spectrum disorder (ASD) is a "category of neurodevelopmental disorders characterized by challenges concerning social skills, speech development and behavior" that affect a person's aptitude for communication and social interaction (Sobieski et al., 2022, p. 1). Before it was identified as a set of neurodevelopmental disorders and given the name autism spectrum disorder, descriptions of autistic behaviors have been unknowingly reported throughout history. One of the earliest recorded cases of what is now believed to be autism was the case of Hugh Blair (Wolff, 2004). In 1747, Blair was brought before an Edinburgh court to determine the state of his mental capacity before being granted a marriage certificate. As part of his case, behaviors such as tactlessness, abnormal gaze, odd motor mannerisms, echolalia, and insistence on sameness were brought to the court's attention, where he was subsequently described as having a "silent madness" (Wolff, 2004, p. 202). Half a century after Blair's case, a man by the name of John Haslam published a report in which he unknowingly contributed to the development of autism research by describing a case of autism in a seven-year-old boy. In his 1809 publication entitled "Observations on Madness and Melancholy," Haslam describes this boy's behavior in the chapter called "Cases of insane children" (Wolff, 2004, p. 202). According to this report, the child was slow to walk and late to talk, had a poor grasp of distance, had several obsessive preoccupations, and preferred solitude over social interactions (Wolff, 2004). The preference for solitude described in this case spurred curiosity about how extreme isolation during early infancy affects behavior. From this, extensive research regarding the speculated disabilities of "wolf children," that is children who were "discovered in the wild, who had supposedly been reared by wolves or other wild animals," came into the forefront of preliminary

autism research (Wolff, 2004, p. 202). The most notable wolf child case was a boy named Victor who was found in the woods in 1798 (Wolff, 2004). According to the reports of Jean Itard, the physician who attempted to educate and humanize him, Victor's behavior included expressionless gazing, no behavioral imitation of adults, difficulties distinguishing emotions, intense moods that changed quickly, and the inability to communicate using language other than guttural vocalizations (Wolff, 2004). At present, cases of wolf children are no longer believed to be instances of autism because they stem from childhood trauma. Cases such as Victor's are obscured by the effect of extreme isolation in childhood, a factor which is now known to cause "quasi-autistic patterns" (Wolff, 2004, p. 203). Regardless, cases of wolf children were historically associated with autism, and are considered important in the development of autism as a disorder because of the similarities in maladaptive social behaviors witnessed in wolf children and autistic children. Furthermore, Itard's work with Victor is historically significant because it inspired further investigation into mentally handicapped children and the efficacy of behavioral modification techniques.

Inspired by Itard's work with Victor, Leo Kanner published a historic article in 1943 in which he identifies and describes autism as a distinct disorder in a way that had never been done before him (Wolff, 2004). In his publication, "Autistic Disturbances of Affective Contact," Kanner describes autism by stating, "The outstanding 'pathognomonic,' fundamental disorder is the children's *inability to relate themselves* in the ordinary way to people and situations from the beginning of life" (Kanner, 1943, p. 242). For those who fit this description of behavior, Kanner labels their diagnosis as "early infantile autism." He goes on to describe this disorder as "an *extreme autistic aloneness* that, whenever possible, disregards, ignores, shuts out anything that comes to the child from the outside" (Kanner, 1943, p. 242). Interestingly, in his description of

autism, Kanner paid significant attention to the role of the parents. Later in life he commented on this role by stating, “There is a resemblance between their make-up and that of their children, except that their aloofness has not reached the gross proportions of a psychotic illness” (Wolff, 2004, p. 203). The connection Kanner made between the parents and their children is historically significant because it alludes to the genetic heritability of autism, a factor that would be researched more extensively decades later. Based on Kanner’s work in identifying the behavioral characteristics and sequelae of autism, future researchers were able to make significant discoveries about the nature of autism, thus the term “Kanner syndrome” is oftentimes retrospectively used to reference his account of autism (National Autistic Society, n.d.).

Alongside Kanner, another important figure in the history of autism is Hans Asperger. In the year following Kanner’s publication, Asperger published an account of children with similar behavioral disabilities as those described by Kanner as having early infantile autism (National Autistic Society, n.d.). Asperger’s work in identifying trends in disordered behaviors brought about an entirely new diagnosis. Asperger’s account of the disordered children differed from Kanner’s in that his cases demonstrated various abilities, such as superior grammatical language, which were missing in Kanner’s cases. The combination of behavioral disabilities and abilities described by Asperger came to be known as “Asperger’s syndrome” (National Autistic Society, n.d.). In retrospect, the exact distinction between behavior characteristics and cognitive impairments between Kanner’s syndrome and Asperger’s syndrome is disputed. However, regardless of this dispute, their combined efforts toward autism research remain significant because their accounts of children with similar behavioral disabilities contributed to the establishment of autism as a disorder with specific characteristics and diagnostic criteria.

Modern perspectives of autism incorporate the contributions made by individual scholars, such as Kanner and Asperger, and of various genres of research, such as the research on wolf children, by presenting autism on a spectrum. In doing so, the autistic spectrum has become an umbrella term to encompass several disorders that used to merit a separate diagnosis such as "early infantile autism, childhood autism, Kanner's autism, high-functioning autism, atypical autism, pervasive developmental disorder not otherwise specified, childhood disintegrative disorder, and Asperger's disorder" (American Psychological Association [APA], 2022, p. 60). By combining each of these disorders under one diagnosis, it demonstrates the shift in how neurodevelopmental disorders are recognized, from individual type to degree of severity. Furthermore, the creation of the autistic spectrum subsequently establishes ASD as a heterogeneous disorder with a dimensional nature, however, it also presents a multitude of problems that manifest themselves in the diagnosis of ASD.

Diagnostic Procedures and Criteria

The diagnostic procedure for ASD, as evidenced by the vast variation in prevalence estimates, is a flawed process that would benefit from improved reliability as well as more diligence in accurately diagnosing children who historically have been underdiagnosed and misdiagnosed. Regardless, by following the current diagnostic procedure for ASD and by considering the factors known to correlate with misdiagnosis, clinicians are attempting to improve accurate diagnosis of ASD. The current diagnostic procedure is a "long-term and multi-stage process aimed at recognizing existing disorders and assessing a child's functioning on many levels" (Sobieski et al., 2022, p. 2). Seeing as the symptoms of ASD typically become evident in early childhood, the critical first step in identifying ASD begins with parents, guardians, or other individuals who have frequent contact with the child (Sobieski et al., 2022).

Specifically, general practitioners play a key role in the early identification of ASD because they not only recognize the behavioral signs of ASD, but they also assess the need for further consultations with specialists. Unfortunately, several problems arise when the responsibility of recognizing ASD falls to general practitioners. Problems such as the insufficient knowledge of doctors about ASD, along with the failure to devote enough time to allow the symptoms of ASD to present themselves within an appointment, as well as instances when general practitioners marginalize, minimalize, or ignore parents' concerns all contribute to the oversight of ASD among medical professionals (Sobieski et al., 2022). Once the identification of ASD symptoms has been made by a person who is familiar with the child, the second step towards diagnosis is to rule out other medical conditions that may be causing the symptoms. For this reason, consultations with specialists such as audiologists or geneticists are essential (Sobieski et al., 2022). The final, and most definitive, step is to receive an official diagnosis from a team of specialists such as psychiatrists, psychologists, special educators, or speech therapists (Sobieski et al., 2022).

As a guideline for diagnosis, the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision* (DSM-5-TR) provides a set of five diagnostic criteria that when exist concurrently, indicate the probable presence of ASD (APA, 2022). The first, and most ubiquitous, diagnostic criterion is the persistent impairment of "social communication and social interaction across multiple contexts" (APA, 2022, p. 56). The presence of this criterion is illustrated by deficits in social-emotional reciprocity, nonverbal communicative behaviors, and/or in developing, maintaining, or understanding relationships (APA, 2022). The second criterion, defined as "restricted, repetitive patterns of behavior, interests, or activities," is manifested by a combination of stereotyped behaviors (APA, 2022, p. 56). These behaviors can

include, but are not limited to, repetitive motor movements, the insistence on sameness, inflexible attachment to routines, highly fixated interests that exceed normality in intensity or focus, hyperactivity or hypoactivity to sensory input, or unusual interest in sensory features in the environment (APA, 2022). The potential behavioral manifestations of the first two criteria, are so broad in scope that the DSM-5-TR only provides an illustrative, not exhaustive, list of examples. Given this scope, the third and fourth criteria provide parameters for these manifestations. The third criterion limits the potential for diagnosis by stating that the symptoms of ASD must, to some extent, be present in the early developmental period, though they may manifest later in life when social demands exceed social capacity (APA, 2022). The fourth criterion continues these parameters by stating that the symptoms must cause "clinically significant impairment in social, occupational, or other important areas of current functioning" (APA, 2022, p. 57). The fifth, and final, criterion states that the occurrence of these disturbances cannot be better explained by a differential diagnosis of another disorder such as intellectual developmental disorder, global developmental delay, or social (pragmatic) communication disorder (APA, 2022).

Due to the lack of clear borders between the diagnostic criteria of each disorder under the umbrella of ASD, and the resulting difficulties in diagnostically distinguishing between them, the DSM-5-TR, provides specifiers and modifiers to differentiate individuals (Lord et al., 2020). The severity specifiers range from level one severity, "requiring support," to level three severity, "requiring very substantial support," and distinguish between intellectual disabilities, social communication difficulties, language impairment, and comorbid genetic or medical conditions (APA, 2022, p. 58). It is important to note that while these specifiers describe an individual's

current symptomology, they also recognize that symptom severity may vary by context or fluctuate over time (APA, 2022).

Prevalence, Etiology, and Comorbidity

According to the World Health Organization, the estimated worldwide prevalence of ASD is 1 in 160 children; though, this estimate varies greatly depending on the research method used and the country of interest (Sobieski et al., 2022). Within the United States specifically, the prevalence is reported to be between one and two percent of the population (APA, 2022). Regarding ethnicity, there seems to be a consistently lower prevalence among African American and Latinx children with rates of 1.1% and 0.8%, respectively, than among White children with a rate of 1.3%. Furthermore, the male to female ratio across several epistemological samples appears to be 3:1 (APA, 2022). These estimates, though gathered through empirically valid research methods, may be skewed due to the misdiagnosis, delayed diagnosis, and underdiagnosis of women and individuals from some ethnoracial backgrounds (APA, 2022).

Given its prevalence worldwide, the etiology of ASD has become a topic of extensive research within the past few decades. From this research, several environmental and genetic risk factors that may precede the presence of ASD have been identified. Environmental factors such as advanced parental age, extreme prematurity, or in utero exposure to certain drugs or teratogens have each been identified as risk factors for ASD and other neurodevelopmental disorders (APA, 2022). Further environmental factors include short intervals between pregnancies, gestational diabetes mellitus, and valproate use during pregnancy (Lord et al., 2020). Despite the evidence correlating these factors to the presence of ASD, it is important to recognize that this relationship cannot be considered causal, but rather the factors are reactive or contributory to autism (Lord et al., 2020). Furthermore, it is also important to recognize studies

that have found significant zero associations between certain environmental factors and ASD. These zero associations include vaccinations, delivery by cesarean section, delivery by assisted vaginal birth, prolonged labor, and the use of assisted reproductive technologies (Lord et al., 2020). This suggests that these environmental factors, contrary to the information provided in non-scholarly media, are not empirically supported as risk factors for ASD. Genetically speaking, one of the most significant factors regarding the presence of ASD is heredity. ASD is arguably one of the most heritable common medical conditions, with heredity estimates ranging from 37% to over 90% (APA, 2022). A second consideration regarding genetic risk factors is genetic makeup, and whether there are genetic mutations within that makeup. Currently, an estimated 15% of all cases of ASD seem to be associated with a genetic mutation (APA, 2022), and within those mutations, over 100 genes and genomic regions have been confidently associated with ASD (Lord et al., 2020). It is important to note, however, that the same causal restriction applied to the environmental risk factors also applies to genetic mutations. Meaning, not all individuals with the same genetic mutation will develop ASD (APA, 2022).

As with many neurodevelopmental disorders, ASD has a high rate of comorbidity with other disorders of similar nature. Specifically, intellectual developmental disorder and language disorder are the most frequently associated with ASD, while specific learning difficulties and developmental coordination disorder are also common (APA, 2022). Furthermore, ASD also has a high rate of comorbidity with psychiatric disorders. It is estimated that around 70% of individuals with ASD have one comorbid mental disorder, and around 40% have two or more comorbid mental disorders (APA, 2022). Of these, the most common are anxiety disorders, depression, and attention-deficit hyperactivity disorder (APA, 2022). Further medical conditions such as epilepsy also frequently co-occur in individuals with ASD (APA, 2022). The frequent

comorbidity of mental disorders in individuals with ASD, in combination with the similarities in symptoms among neurodevelopmental disorders and the general underdiagnosis of ASD, results in extensive difficulties in accurately diagnosing ASD. Therefore, a more accurate diagnostic procedure for ASD as well as clearer distinctions between disorders need to be established in order to properly represent the autistic community.

The Connection Between the Theory of Mind and Autism

Following the publication of Premack and Woodruff's original study on Sarah the chimpanzee, many scholars wanted to expand on their research to investigate how their findings, which were based upon an animal subject, could be applicable to the human experience. Premack and Woodruff touched on this concern briefly in their conclusion by stating, "In assuming that other individuals *want, think, believe,* and the like, one infers states that are not directly observable and one uses these states anticipatorily, to predict the behavior of others as well as one's own. These inferences, which amount to a theory of mind, are, to our knowledge, universal in human adults" (1978, p. 525). This conclusion prompted extensive research into the true universality of the theory of mind. The findings of this research suggest that the theory of mind, by nature, is applicable to all people; however, its acquisition seems to be inherently more difficult for people with preexisting neurodevelopmental disabilities. Most notably, research suggests that people with autism lack an advanced theory of mind (Baron-Cohen et al., 1985). Furthermore, it is argued that this specific cognitive deficit may explain the significant social impairment people with autism face, as well as other functional consequences of autism such as difficulties developing and maintaining social relationships, and general poverty in pretend play (Baron-Cohen et al., 1985). The consequential relationship between the theory of mind and autism was first presented by a psychologist by the name of Simon Baron-Cohen and his

colleagues Alan Leslie and Uta Frith. In their publication, “Does the autistic child have a ‘theory of mind’?,” Baron-Cohen et al. hypothesize that autistic children not only lack a theory of mind, but also that this deficit is independent of general intelligence or mental retardation and specific to autism (1985). To test and support their hypothesis, Baron-Cohen et al. designed a research study that incorporated the false belief paradigm developed by Wimmer and Perner (1983). The test subjects were children with autism (n = 20), children with Down syndrome (n = 14), and clinically normal children of preschool age (n = 27) (Baron-Cohen et al., 1985). To control for intelligence as a potential confounding variable, the three diagnostic groups were chosen from subgroups with varying mental age and general intelligence (Baron-Cohen et al., 1985). The autistic test subjects were chosen from a high-functioning subgroup with a relatively high mean IQ of 82 and a mean mental age that was higher than that of the Down syndrome group. In comparison, the Down syndrome group had a significantly lower IQ, with a range of 42 to 89 and a mean of 64. As the control group, the mental ages of the clinically normal children were assumed to roughly correlate with their chronological ages, which were lower than both the autistic and Down syndrome groups (Baron-Cohen et al., 1985).

Following the false belief task paradigm, the children were presented with a scenario in which two doll protagonists, Sally and Anne, place a marble in a basket, then, while Sally is off-stage, Anne relocates the marble to a box (Baron-Cohen et al., 1985). Following this, the same scenario was presented to the children for a second time, however, in this scenario, a third possible location for the marble, the experimenter’s pocket, was introduced. Within this research procedure, the test subjects were asked four critical questions: (1) the naming question, (2) the belief question, (3) the reality question, and (4) the memory question. The naming question, which assessed whether the children could identify, by name, which doll was which, was

answered correctly by each test subject (Baron-Cohen et al., 1985). Furthermore, the reality question, which asked the children to identify where the marble was after it had been relocated, and memory question, which asked the children where the marble was before it was relocated, were also answered correctly for both trials without a single exception. These results allowed the researchers to confidently conclude that the answers provided by the children were a product of their knowledge and implicit beliefs, and not from a misunderstanding or confusion about the scenario. The difference in responses appeared only with the belief question: "Where will Sally look for her marble" (Baron-Cohen et al., 1985, p. 41)? It is significant to note that the way in which each child answered this question, apart from one Down syndrome child, was consistent on both trials, regardless of accuracy. The results for the Down syndrome group and the clinically normal group were notably similar, with 12 out of 14 (86%) Down syndrome children, and 23 out of 27 (85%) clinically normal children, answering correctly on both trials by indicating the false belief location of the marble (Baron-Cohen, 1985). In comparison, 16 out of 20 (80%) autistic children answered incorrectly on both trials by indicating the actual location of the marble, a difference that proved to have strong statistical significance ($p < 0.001$) (Baron-Cohen et al., 1985).

Due to the significant group difference in how the children responded to the belief question, the findings of this study supported the original hypothesis that autistic children lack an advanced theory of mind, and that this specific deficit is independent of intelligence. However, since there were four autistic children who answered the belief question correctly on both trials, the study concludes by hypothesizing that these children were in fact capable of forming second order representations and predict that their social impairments "would show a rather different pattern from those autistic children who fail to use a theory of mind" (Baron-Cohen et al., 1985,

p. 43). These autistic children who seem to have a capacity for the theory of mind represent exceptions to the findings addressed in Baron-Cohen et al.'s study. However, regardless of this subset of autistic children, in subsequent studies that replicated Baron-Cohen et al.'s 1985 study, similar results were consistently found, which appear to be "highly robust, in that despite wide variations in techniques used to assess it, comparable results are still obtained" (Baron-Cohen, 1991, p. 301). Regardless of the chosen stimuli to represent story characters or the person administering the test, similar results have been consistently recorded across a multitude of studies. This finding prompted further investigation into how specific this deficit is, not only to the autistic community, but also within the symptomology of the disorder itself (Baron-Cohen, 1991). To further investigate the specificity of the theory of mind deficit within the symptomology of autism, Simon Baron-Cohen created a research study that tested three skill areas: (1) relationship recognition, (2) interpersonal reciprocity, and (3) understanding the animate-inanimate distinction (1991). Before conducting this research, the baseline understanding of the specificity of the impairment regarding mental states was that "understanding the propositional mental states believe, know, think and dream are all impaired in people with autism, whilst their understanding of non-propositional mental states such as 'simple' emotions (e.g. happiness and sadness as outcomes of situations) is unimpaired" (Baron-Cohen, 1991, p. 302). This gives reason to believe that the theory of mind deficit may be "confined to (or at least most severe in) their understanding of propositional mental states" (Baron-Cohen, 1991, p. 302). The results of this study found that the test subjects performed exceptionally well in all three of the targeted cognitive abilities, which indicates that "subjects with autism are neither impaired in their ability to recognize simple relationships, nor in their ability to show simple reciprocity, nor in their understanding of the animate-inanimate

distinction, relative to normal subjects and subjects with mental handicap" (Baron-Cohen, 1991, p. 311). Therefore, the results of this study support the original hypothesis that the theory of mind deficit in autism is highly specific to second order mental representations. The significance of the specificity of the theory of mind deficit, according to Baron-Cohen and his supporters, is that it provides an explanation for the impairments in social communication that are characteristic within autistic children. This significance, however, would later be challenged by psychologists who believed it was logically flawed and was too incomprehensive to be accepted as an explanation for the social behavior of autistic children.

Counterargument Against the Theory of Mind Hypothesis of Autism

In response to the consequential claims made in Baron-Cohen et al.'s landmark 1985 publication, a new era of autism research was initiated (Tager-Flusberg, 2001). For nearly a decade following this publication, there was an avid enthusiasm to investigate the theory of mind hypothesis of autism, as it came to be known in the literature, because it "not only explained the failure of children with autism on tasks tapping theory-of-mind abilities, but also provided a unified explanation for the primary diagnostic impairments in pretend play, social functioning, and communication" (Tager-Flusberg, 2001, p. 173). However, this honeymoon period only lasted for a short while before skepticism began to arise and its significance as an explanatory model for autism began to be challenged. Criticisms of this explanatory model addressed several factors such as comprehensiveness, applicability, universality, and specificity. In an examination of the historical development of the theory of mind hypothesis, Tager-Flusberg conceptualized these concerns into a variety of theoretical questions (2001, pp. 174-175):

- Are deficits on theory-of-mind tasks universal among individuals with autism?
- Are deficits on theory-of-mind tasks unique to individuals with autism?

- How can the theory of mind hypothesis explain the impairments that are evident in infants with autism, long before the emergence of a representational theory of mind?
- How can the theory of mind hypothesis explain some of the other features of autism, such as repetitive behaviors and interests or savant abilities?
- Can failure on theory-of-mind tasks be interpreted in terms of other constructs, such as executive functions or language?

From these questions, it is evident that the theory of mind hypothesis, and those who supported it as an explanatory model for autism, faced opposition that became increasingly difficult to dismiss. This opposition, however, was not unfounded. From the first study investigating the relationship between autism and the theory of mind, it was evident that a minority of the research participants with autism (4 of 20, or 20%) were able to successfully complete the theory of mind task presented to them (Baron-Cohen et al., 1985). Though this percentage varied from study to study, its presence alone is significant because theories, to be considered comprehensive, should account for the population that exists as an exception to itself, a feature which the theory of mind hypothesis of autism lacks. Furthermore, the specificity of this hypothesis to autism is challenged because research studies have found evidence of similar theory of mind impairments in nonautistic children and adolescents with mental retardation at a "higher rate than would be expected given their age and developmental level" (Tager-Flusberg, 2001, p. 175). This same finding has also been identified in oral deaf children, blind children, children with specific language impairment, and people with schizophrenia (Tager-Flusberg, 2001). Thus, if this impairment extends to numerous populations, it calls into question why this hypothesis is exclusively interpreted as the distinctive deficit in autism.

The third criticism mentioned in the bulleted list above is directly concerned with the temporal precedence of the theory of mind deficit. The central claim of the theory of mind hypothesis of autism is that the social and communicative impairments that are characteristic of autism can be explained by the absence of a theory of mind. However, this idea fails to consider a key characteristic of autism: symptoms are typically recognized within the second year of life, if not earlier (APA, 2022). Empirical studies and anecdotal evidence suggest that infants with autism exhibit a combination of deficits in social behaviors such as "social responsiveness, empathy, play, joint attention, and imitation" (Tager-Flusberg, 2001, p. 175). Of these deficits, not all are dependent on the cognitive ability to understand representational minds; therefore, using the theory of mind deficit to explain these behaviors is an inaccurate interpretation of the mechanisms that underlie autistic social difficulties.

Social difficulties, although often emphasized as the central defining feature of autism, do not wholly describe the characteristics of this disorder. There are many other characteristics of autism that, when recognized in combination with social difficulties, more accurately represent the autistic experience. Characteristics such as repetitive behaviors and fixated interests are not only required by the DSM-5-TR for diagnosis, but also play a significant role in the daily functioning of a person with autism. Outside of the diagnostic criteria, there are several other features that are commonly associated with autism, such as savant abilities and remarkable visual perceptual skills, which are not accounted for by deficits in representing mental states. Therefore, the fact that these characteristics of autism cannot be interpreted within the theory of mind framework suggests that the theory of mind hypothesis of autism is an incomplete explanatory model that provides a limited perspective of autism (Tager-Flusberg, 2001). Finally, the fifth criticism from the bulleted list above begs the question of whether the theory of mind hypothesis

of autism is the best explanation for autistic social deficits, or if there are other possible explanations that are more fitting. In response to this, two alternative explanations have been provided through research: fundamental deficits in executive function, and deficits in the capacity for language (Tager-Flusberg, 2001). Regarding language, it has been argued that theory of mind performance "results from the nature of language needed to understand the tasks and test questions," thus, failure on theory of mind tasks may be a direct result of "limitations and impairments in the linguistic knowledge of children with autism" (Tager-Flusberg, 2001, pp. 176-177). Regarding executive function, it has been argued that classic theory of mind tasks, such as the false belief task, require a child to rely on executive functions such as action monitoring or self-regulation. Therefore, if a child's executive function capabilities are impaired, it may explain their failure on these tasks. With this understanding, an alternative explanation for autism arises that provides an arguably more thorough account of a "range of both social and nonsocial problems, including the repetitive behaviors and interests as well as play deficits that define autism" (Tager-Flusberg, 2001, p. 166).

Rival Cognitive Theories for Social Deficits in Autism

As a result of the criticisms raised against the theory of mind hypothesis of autism, rival cognitive theories have been proposed as alternative explanations for the social deficits in autism. Of these alternative explanations, the theory of impaired executive function has proven to be a prominent hypothesis that stands in opposition to the theory of mind hypothesis. Simply stated, executive function is an umbrella term for a variety of functions that are relevant to metacognitive processes. It is "the ability to maintain an appropriate problem-solving set for attainment of a future goal; it includes behaviors such as planning, impulse control, inhibition of prepotent but irrelevant responses, set maintenance, organized search, and flexibility of thought

and action" (Rajendran & Mitchell, 2007, p. 232). In studies that cross examine executive function in autistic children, clinically normal children, and children with other disorders such as attention-deficit hyperactivity disorder, schizophrenia, obsessive compulsive disorder, and Tourette syndrome, findings indicate that impairments in executive function manifest in varying degrees of severity according to the type of disorder a child suffers from (Rajendran & Mitchell, 2007). These results indicate that the theory of impaired executive function, rather than being a unique deficit specific to autism, manifests itself as a distinct profile within autism that distinguishes it from other neurodevelopmental disorders (Rajendran & Mitchell, 2007). The autistic profile of impaired executive function includes difficulties in mental flexibility, inhibition of prepotent responses, and challenges in self-monitoring. These functions are key components in a variety of everyday behaviors; thus, their impairment may begin to explain the social dysfunction demonstrated in autistic children. The strength of this theory in comparison to the theory of mind hypothesis is that it accounts for a wider range of autistic characteristics such as the non-social aspects of autism, and it is the "only theory that acknowledges both the cognitive and motor (repetitive hand flapping, rocking) characteristics of autism" (Rajendran & Mitchell, 2007, p. 237). Therefore, as an alternative explanation for autism, the theory of impaired executive function provides a more comprehensive perspective of autistic social dysfunction than the theory of mind hypothesis.

A second cognitive theory that challenges the theory of mind hypothesis of autism is the theory of weak central coherence (WCC). The principal idea of this theory is that "typically developing individuals process information by extracting overall meaning or gist," however, it is postulated that this process is weak or absent in autistic individuals (Rajendran & Mitchell, 2007, p. 237). This suggests that individuals with autism "process things in a detail-focused or

piecemeal way – processing the constituent parts, rather than the global whole" (Rajendran & Mitchell, 2007, p. 237). When this theory is applied to autistic individuals, it begins to explain both social and non-social behaviors that are characteristic to this disorder. The extreme attention to detail and a general disregard of the overall gist of a social interaction, according to WCC research, explains why autistic individuals struggle with identifying and utilizing deception, identifying social faux pas, generalizing across different social settings, and recognizing disguised emotions (Rajendran & Mitchell, 2007). Furthermore, this theory is also reflected in nonsocial characteristics such as the presence of highly fixated interests that exceed normal intensity. According to the WCC theory, highly fixated interests allow autistic individuals to focus, or expend their energy, on one stimulus without having to divide their attention, a skill that is oftentimes impaired in autistic individuals. The significance of this theory parallels the significance of the theory of impaired executive function in the fact that it accounts for a wider variety of autistic characteristics, both social and nonsocial, than the theory of mind hypothesis. Therefore, given its ability to provide a more extensive perspective of autistic characteristics, it is oftentimes regarded as a superior explanation for the social deficits in autism than the theory of mind hypothesis.

Limitations of Current Research

In addition to the criticisms the theory of mind hypothesis of autism has received regarding its application as an explanatory model, this hypothesis has also faced criticism regarding its practical limitations. In their review of the theory of mind hypothesis of autism, Klin et al. claim that the significance of this approach "lies in its increased attention to social development in autism;" however, its limitation lies in its usefulness (1992, p. 862). They claim, "its usefulness is still somewhat limited at present due to the fact that various predictions

regarding the nature of autistic social dysfunction which are apparently implied by this hypothesis have not, as yet, been fully examined empirically" (Klin et al., 1992, p. 862). By this, Klin et al. insinuate that the theory of mind hypothesis, because of the unexplored predictions that naturally result from this hypothesis, may have been accepted before the research was able to support it. Due to this lack of empirical support, two predictions derived from this hypothesis have garnered substantial attention regarding their questionable validity. The first of which is the "*primary*" versus "*secondary*" issue concerning the nature of the social deficits that are characteristic in autism (Klin et al., 1992, p. 863). In this context, the terms "*primary*" and "*secondary*" indicate symptoms which occur as a direct result of an established etiology or stem from other "symptomatic features of the syndrome," respectively (Klin et al., 1992, p. 863). From their first identification by Kanner (1943), the social deficits associated with autism were regarded as primary and presumably innate (Klin et al., 1992). However, the theory of mind hypothesis of autism redefines the nature of the social deficits from primary to secondary. According to the theory of mind hypothesis, social deficits "result from a failure of maturation of the metarepresentational skills presupposed by normal social functioning" (Klin et al., 1992, p. 863). For this to be true, however, it suggests that social functioning in autistic children should develop normally until the point in development when theory of mind abilities become present in children's cognition. This raises questions about how theory of mind capabilities are operationalized and when they are expected to develop.

The postulated age at which a child begins demonstrating theory of mind capabilities varies greatly throughout the research. This is, in part, due to the controversy about which infant behaviors constitute theory of mind abilities. It is known that babies, as early as birth, "prefer social over nonsocial stimuli," and exhibit a multitude of behaviors that expose their gradual

development of social awareness (Astington & Dack, 2020, p. 368). Early social interactions include reactive and interactive smiling and vocalizations, which begin around two months of age. These vocalizations first appear in dyadic interactions, meaning the infant and one adult are the sole participants involved, and gradually develop into triadic interactions, meaning both individuals are focused on a third-party stimulus (Astington & Dack, 2020). One of the skills that emerges in triadic interactions is the ability to "coordinate attention with others by following gaze or pointing gestures" (Astington & Dack, 2020, p. 368). This skill is also known as "joint attention," and is a "major feature of infant social behavior" (Astington & Dack, 2020, p. 368). As infants become increasingly adept at joint attention, they begin engaging in what is known as social referencing. Social referencing is the behavior in which infants will look to an adult, usually their mother, when they are in an ambiguous situation and are unsure of how to react. They will then base their reactions on the positive or negative expression of the adult (Astington & Dack, 2020). The research regarding these behaviors provided cognitive science with a more comprehensive understanding of how the theory of mind develops because it identified innate social behaviors that act as precursors to the actualization of the theory of mind (Astington & Dack, 2020).

However, this research has limitations of its own. Within this field of study, there has been much debate regarding the methods, data, and interpretations of the research. Outside of observational research, most theory of mind testing relies on false belief tasks as the basis for assessing theory of mind capabilities. Due to this, the validity of theory of mind research has been frequently challenged by scholars who believe there is not sufficient evidence to support their claims when most of the preliminary research was solely based on a child's performance in a singular skill set. Admittedly, false belief tasks are generally reliable, and it is widely accepted

that “successful performance on the false-belief task demonstrates an understanding of the idea that people’s relationship to the world is mediated by their mental representation of it;” however, the focus of this research seemed too narrow to convince many researchers of its comprehensiveness (Astington & Dack, 2020, p. 366). Wellman and Liu specifically call attention to this shortfall by stating that theory of mind research “faces measurement limitations by typically using single tasks, essentially false-belief tasks to assess children’s understanding” (2004, p. 537). The exclusive reliance on false belief tasks as the determiner of whether a child possesses theory of mind capabilities undermines the applicability of the results because they are based on research that did not exhaust multiple avenues of investigation. Due to this, the interpretation of the data’s significance, even if it has been replicated across several studies, is debatable because the methods with which it was collected allowed for gaps in the research (Astington & Dack, 2020). Consequently, further research investigating developments in infants that may allude to the theory of mind must, for more applicable results, be investigated without solely depending on the use of false belief tasks (Astington & Dack, 2020).

The second prediction derived from the theory of mind hypothesis of autism is the “*generalized*” versus “*discrete*” distinction of the nature of autistic social dysfunction (Klin et al., 1992, p. 863). In this context, the term “*generalized*” implies that autistic social dysfunction “disables all skills necessary for interacting with others,” and the term “*discrete*” implies that autistic social dysfunction “disables only certain social skills as a result of a rather specific, for example cognitive, impairment” (Klin et al., 1992, p. 863). The determination of which side of this distinction is accurate is a complex issue for several reasons: (1) social disabilities found in autism are highly unique to each autistic individual, and vary greatly in severity and range, (2) some social skills, as a result of social skills training received in school, or diligent parents, will

emerge later during the course of development, and (3) there is a sense of interrelatedness between social, cognitive, and communicative skills (Klin et al., 1992). According to Kanner (1943), the social dysfunction experienced among autistic people is a generalized phenomenon that not only affects all social skills, but is also present from birth (Klin et el., 1992). In comparison, the theory of mind hypothesis postulates that autistic social dysfunction is more discrete because it only affects behaviors that "presuppose the need to attribute beliefs, intentions or desires to others, i.e. using a 'theory of mind"' (Klin et al., 1992, p. 864). Accordingly, one should expect that "autistic children should fail to exhibit social behaviors mediated by the metarepresentational capacity, whereas social behaviors with no such demands should be observed in their social functioning" (Klin et al., 1992, p. 864). This expectation of autistic social functioning proves to be inaccurate because there is evidence of social impairments in behaviors that do not require the use of a representational mind. An example of this is the impaired social response of autistic babies to social contacts such as their atypical "anticipatory posture upon being picked up and body moulding to caregivers" (Klin et al., 1992, p. 864). Therefore, the *"discrete"* prediction of autistic social behavior, when compared to research investigating impaired social behaviors that do not require a representational mind, loses its plausibility, therefore partially discrediting the theory of mind hypothesis of autism as an explanatory model.

Directions for Future Research

The two predictions Klin et al. derived from the theory of mind hypothesis of autism, at present, seem to be neglected among experimental research. Thus, to resolve this issue, it is recommended that future research consider reconfiguring the focus of the research by "examining normative rather than deviant social behaviors, and real-life rather than experimentally-created social situations" (Klin et al., 1992, p. 864). In doing so, it may be

possible to gain a clearer understanding of the ecological validity of the theory of mind hypothesis of autism. Furthermore, it is also recommended that future research provides more explicit operational definitions for what behaviors constitute theory of mind precursors. With a more established and agreed upon description of these behaviors, it is possible to more accurately "evaluate the developmental account of autistic social deficits put forward by the hypothesis" (Klin et al., 1992, p. 864).

Conclusion

As a psychological construct, the theory of mind describes a person's cognitive ability to represent the mental states of others as well as understand how these mental states affect their behavior. As a continuation of this idea, the theory of mind hypothesis of autism posits that individuals with autism lack the ability to represent the mental states of others. This hypothesis, since its introduction in 1985, has been widely regarded as an explanatory model for the social deficits that are characteristic in autism. However, upon further research, this hypothesis proved to be insufficient to explain autistic social dysfunction because it lacked comprehensiveness, applicability, universality, and specificity. Due to this, alternative cognitive theories have been introduced as replacements for the theory of mind hypothesis. The fatal flaw identified regarding the theory of mind hypothesis lies in the fact that it was confused as an explanatory model, when at best, it is a symptom of autism. Regardless of its validity, the theory of mind hypothesis of autism introduces the idea that an unspecified cognitive deficit may underlie the social dysfunction in autism, an idea that has significant potential consequences for the future of autism research.

References

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

Astington, J. W., & Dack, L. A. (2020). Theory of mind. *Encyclopedia of Infant and Early Childhood Development*, 365-379. https://doi.org/10.1016/B978-0-12-809324-5.21284-4

Baron-Cohen, S. (1991). The theory of mind deficit in autism: How specific is it? *British Journal of Developmental Psychology 9*(1), 301-314. https://doi.org/10.1111/j.2044-835X.1991.tb00879.x

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37-46. https://doi.org/10.1016/0010-0277(85)90022-8

Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *The Behavioral and Brain Sciences 6*(1), 343-390. http://doi.org/10.1017/S0140525X00016393

Flavell, J. H. (1986). The development of children's knowledge about the appearance-reality distinction. *American Psychologist, 41*(4), 418-425. https://doi.org/10.1037/0003-066X.41.4.418

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly, 50*(3), 274–290. http://www.jstor.org/stable/23096166

García Landa, J. A. (2013). *Hierarchically minded: Levels of intentionality and mind reading.* SSRN. https://ssrn.com/abstract=2250379

Garfield, J. L., Peterson, C. C., & Perry, T. (2001). Social cognition, language acquisition and the development of the theory of mind. *Mind & Language, 16*(5), 494-541. https://doi.org/10.1111/1468-0017.00180

Hughes, C., Adlam, A., Happé, F., Jackson, J., Taylor, A., & Caspi, A. (2000). Good test-retest reliability for standard and advanced false-belief tasks across a wide range of abilities. *Journal of Child Psychology and Psychiatry, 41*(4), 483-490. https://doi.org/10.1111/1469-7610.00633

Hutchins, T. L., Prelock, P. A., & Chace, W. (2008). Test-retest reliability of a theory of mind task battery for children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *23*(4), 195–206. https://doi.org/10.1177/1088357608322998

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*(3), 217-250.

Klin, A., Volkmar, F. R., & Sparrow, S. S. (1992). Autistic social dysfunction: Some limitations of the theory of mind hypothesis. *Journal of Child Psychology and Psychiatry, 33*(5), 861-876. 10.1111/j.1469-7610.1992.tb01961.x

Lewis, P. A., Birch, A., Hall, A., & Dunbar, R. I. M. (2017). Higher order intentionality tasks are cognitively more demanding. *Social Cognitive Affect Neuroscience, 12*(7), 1063-1071. 10.1093/scan/nsx034

Lord, C., Brugha T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones E. J. H., Jones, R. M., Pickles, A., State, M. W., Taylor, J. L., & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature Reviews: Disease Primers, 6*(1), 1-23. https://doi.org/10.1038/s41572-019-0138-4

National Autistic Society. (n.d.). *The history of autism.* https://www.autism.org.uk/advice-and-guidance/what-is-autism/the-history-of-autism

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3*(9), 337-344. https://doi.org/10.1016/S1364-6613(99)01362-5

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515-526. https://doi.org/10.1017/S0140525X00076512

Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review, 27*(2), 224-260. https://doi.org/10.1016/j.dr.2007.02.001

Rubin, K. H. (1973). Egocentrism in childhood: A unitary construct? *Child Development, 44*(1), 102-110. https://doi.org/10.2307/1127685

Sobieski, M., Sobieska, A., Sekułowicz, M., & Bujnowska-Fedak, M. M. (2022). Tools for early screening of autism spectrum disorders in primary health care – a scoping review. *BMC Primary Care, 23*(46), 1-26. https://doi.org/10.1186/s12875-022-01645-7

Tager-Flusberg, H. (2001). A reexamination of the theory of mind hypothesis of autism. In J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 173-193). Lawrence Erlbaum Associates Publishers.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*(2), 523-541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Wimmer, H., Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103-128. https://doi.org/10.1016/0010-0277(83)90004-5

Wolff, S. (2004). The history of autism. *European Child & Adolescent Psychiatry, 13*(1), 201-208). https://doi.org/10.1007/s00787-004-0363-5