UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOLOGIA

Jeniffer Evaristo de Souza

Development and Content Validation of a Computer-Based Tasks for Assessing Theory of Mind in Children (ToMT)

Florianópolis, SC
2024

Jeniffer Evaristo de Souza

Development and Content Validation of a Computer-Based Tasks for Assessing Theory of Mind in Children (ToMT)

Dissertação submetida ao Programa de Pós-Graduação em Psicologia da Universidade Federal de Santa Catarina como requisito final para a obtenção do título de Mestre em Psicologia, na área de concentração "Saúde e Desenvolvimento Psicológico", na linha de pesquisa "Avaliação em saúde, desenvolvimento e processos psicológicos básicos".

Orientadora: Profª Drª Chrissie Ferreira de Carvalho

Florianópolis, SC

2024

Souza, Jeniffer Evaristo de
Development and Content Validation of a Computer-Based Tasks for Assessing Theory of Mind in Children (ToMT) / Jeniffer Evaristo de Souza ; orientadora, Chrissie Ferreira de Carvalho, 2024.
81 p.

Dissertação (mestrado profissional) - Universidade Federal de Santa Catarina, Centro de Filosofia e Ciências Humanas, Programa de Pós-Graduação em Psicologia, Florianópolis, 2024.

Inclui referências.

1. Psicologia. 2. Theory of Mind. 3. Naïve Utility Calculus. 4. Early Childhood Assessment. 5. Psychometry. I. Carvalho, Chrissie Ferreira de . II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Psicologia. III. Título.

Jeniffer Evaristo de Souza

Development and Content Validation of a Computer-Based Tasks for Assessing Theory of Mind in Children (ToMT)

O presente trabalho em nível de mestrado foi avaliado e aprovado, em 10/04/2024 membros:

Prof.(a), Dr.(a) Chrissie Ferreira de Carvalho

Universidade Federal de Santa Catarina

Prof.(a), Dr.(a) Natália Martins Dias

Universidade Federal de Santa Catarina

Prof.(a), Dr.(a) Thatiana Lima

Universidade Federal da Bahia

Certificamos que esta é a versão original e final do trabalho

ICP-Edu

Documento assinado digitalmente

Natalia Martins Dias

Data: 24/07/2024 15:45:56-0300

CPF: ***.752.228-**

Verifique as assinaturas em https://v.ufsc.br

Coordenação do Programa de Pós-Graduação

ICP-Edu

Documento assinado digitalmente

Chrissie Ferreira de Carvalho

Data: 24/07/2024 15:16:42-0300

CPF: ***.221.445-**

Verifique as assinaturas em https://v.ufsc.br

Profa. Chrissie Ferreira de Carvalho, Dr.(a) Orientador(a)

Florianópolis, 2024

A meu pai, que sempre me encorajou a seguir meus sonhos, não importa quão grandes ou inusitados fossem.

AGRADECIMENTOS

Primeiramente, gostaria de expressar minha profunda gratidão à minha família: minha mãe, meu pai, meus avós, que, mesmo diante de recursos limitados, acreditaram no poder transformador da educação e me proporcionaram o apoio necessário para trilhar este caminho. Sem vocês, eu não teria chegado até aqui. Agradeço também à querida gata Brisa, cujos ronronares trouxeram leveza aos meus dias. Sou imensamente grata à extensão da minha família: meus estimados amigos Danilo, John John, França, Owen, Elisa, Murilo, Raiza e Marcio. Mesmo distantes ou próximos, permanecemos conectados nos apoiando. Um agradecimento especial à minha amiga e orientadora, Chrissie, pela parceria, ensinamentos e experiências compartilhadas. Agradeço às amigas de mestrado Luiza, Raissa e Amanda, por compartilharem e apoiarem minha jornada neste mestrado. Agradeço também à assistente de pesquisa, Bruna, e às estagiárias, pelo auxílio, comprometimento e responsabilidade demonstrados na coleta de dados desta pesquisa. Sou muito grata a Julian Jara-Ettinger e Liz Spelke pela disponibilidade, conselhos e conversas que me auxiliaram no desenvolvimento e refinamento dos itens. Ao Professor Emílio Takase, que me apresentou ao mundo da pesquisa e, com sua sabedoria, mostrou-me que o ambiente acadêmico pode ser visto fora da caixa. Agradeço a todos os professores e servidores que cruzaram meu caminho durante este mestrado, contribuindo para a minha jornada acadêmica. Finalmente, agradeço à CAPES pelo financiamento da bolsa de mestrado, que possibilitou a realização deste trabalho.

A Smooth Sea Never Made a Skilled Sailor!

ABSTRACT

Theory of Mind (ToM) encompasses various components of social cognition (SC), including the Naïve Utility Calculus (NUC), which emerge during early childhood, enhancing children's abilities to navigate the social world and comprehend and predict others' intentions. This study focuses on the development and validation of computer-based tasks, the Theory of Mind Task (ToMT), to assess ToM in children aged 3 to 8, offering evidence of content validity. The study was conducted in three phases: (1) task development and refinement through an extensive literature review and feedback from social cognition experts; (2) evaluation of content validity, including a qualitative analysis and quantitative review by an expert panel to determine the Content Validity Coefficient (CVC); and (3) a pilot study providing preliminary exploratory analysis, comprehension, and acceptance of the ToMT. The findings suggest that the ToMT has content validity and represents a promising tool for assessing ToM. The ToMT is an innovative computer-based tool for evaluating these constructs in children, with the potential to advance the understanding of the development of these important skills in childhood.

Keywords: Theory of Mind. Naïve Utility Calculus. Social Cognition. Early childhood Assessment. Psychometry.

RESUMO

A Teoria da Mente (ToM) abrange diversos componentes da cognição social (CS), incluindo o *Naive Utility Calculus* (NUC), que surgem na primeira infância e aprimoram as habilidades das crianças para navegar no mundo social, compreendendo e prevendo as intenções dos outros. Este estudo tem como objetivo o desenvolvimento e validação de tarefas computadorizadas de Teoria da Mente (ToMT) para avaliar a ToM em crianças de 3 a 8 anos, fornecendo evidências de validade de conteúdo. A pesquisa foi conduzida em três fases: (1) desenvolvimento e refinamento das tarefas por meio de uma revisão da literatura e feedback de especialistas em cognição social; (2) avaliação da validade de conteúdo, incluindo uma análise qualitativa e uma revisão quantitativa por um painel de especialistas para determinar o Coeficiente de Validade de Conteúdo (CVC); e (3) um estudo piloto, que forneceu uma análise exploratória preliminar, verificando a aceitação e usabilidade do ToMT pela população-alvo. Os resultados sugerem que o ToMT possui validade de conteúdo e representa uma ferramenta promissora para avaliar a ToM. O ToMT é uma ferramenta computadorizada inovadora para a avaliação desses construtos em crianças, com o potencial de avançar na compreensão do desenvolvimento dessas habilidades sociais importantes na infância.

Palavras-chave: Teoria da Mente. Naïve Utility Calculus. Cognição Social. Avaliação na Infância. Psicometria.

RESUMO EXPANDIDO

A Teoria da Mente (ToM) é um componente essencial da cognição social que permite às crianças entender que outras pessoas possuem pensamentos, crenças, desejos e intenções diferentes das suas, e utilizar essa compreensão para explicar e prever comportamentos. Um aspecto da ToM é o *Naive Utility Calculus (NUC)*, que sugere que as crianças, desde cedo, aplicam um raciocínio baseado em custos e recompensas para inferir as intenções e preferências dos outros. A ToM se desenvolve significativamente durante a infância, permitindo que as crianças naveguem no mundo social. **Objetivo:** Desenvolvimento e validação de tarefas computadorizadas de ToM, para avaliar a ToM em crianças de 3 a 8 anos, com foco na validação de conteúdo. **Metodologia:** Fase 1: As tarefas foram criadas a partir de princípios da psicologia do desenvolvimento e da ciência cognitiva, baseadas em tarefas existentes como a Escala de Teoria da Mente de Wellman e Liu (2004) e o framework de Naive Utility Calculus de Jara-Ettinger (2016). Fase 2: Validade de Conteúdo. Para a análise qualitativa, um painel de dois acadêmicos forneceu feedback sobre a clareza, pertinência e relevância das tarefas. A análise quantitativa envolveu três acadêmicos brasileiros, que avaliaram as tarefas utilizando uma planilha de validação de conteúdo para calcular o Coeficiente de Validade de Conteúdo (CVC). Fase 3: O estudo piloto incluiu 10 crianças com desenvolvimento típico, entre 3 e 8 anos. O ToMT foi aplicado à população-alvo e as crianças completaram um formulário de feedback para avaliar a aceitação e compreensão das tarefas. Testes de Percepção Social e Reconhecimento de Emoções foram aplicados para análises exploratórias. **Resultados:** O ToMT inclui 12 tarefas animadas em vídeo que abrangem seis componentes principais da ToM: Percepção, Desejo, Conhecimento, Crenças, Emoções e Intenções baseadas no Naive Utility Calculus. A análise qualitativa mostrou que todos os itens foram reconhecidos como relevantes, foi recomendada a inclusão de novos itens e refinamentos no design das animações para melhorar a clareza dos itens. A análise quantitativa indicou CVC totais de 0.84 para clareza, 0.96 para pertinência e 0.95 para relevância, refletindo a validade geral do conteúdo das tarefas. **Conclusão:** O ToMT representa um potencial para avançar a compreensão do desenvolvimento dessas habilidades. A validação de conteúdo e o feedback positivo do estudo piloto indicam que o ToMT possui validade de conteúdo.

Palavras-chave: Teoria da Mente, Naive Utility Calculus, Cognição Social, Avaliação na Infância, Psicometria.

LIST OF FIGURES

LIST OF TABLES

SUMMARY

1. INTRODUCTION

Navigating the social world involves understanding social cues, predicting behavior, recognizing intentions, beliefs, and emotions, building relationships, and adapting to social norms (Hamlin & Bloom, 2008; Spelke et al., 2013). From the moment they are born, children's social cognition begins to develop, and as infants, they recognize goal-directed actions and understand desires (Aschersleben & Jovanovic, 2008; Gergely & Csibra, 2003; Repacholi & Gopnik, 1997; Woodward, 1998). Gradually through childhood, children develop theory of mind (ToM), enabling them to understand that others have distinct mental states and beliefs (Gopnik & Wellman, 1992; Wellman, 2014). A crucial milestone in this development occurs around ages 4-5, when children begin to recognize that others can hold false beliefs (Rackzoky, 2022; Wellman and Liu, 2004; Wellman et al., 2001; Wellman, 2018).

At this stage, children not only understand false beliefs but also develop the ability to reason about others' intentions, actions, and goals, using principles of utility maximization—a component of ToM (Jara-Ettinger et al., 2016b). In this cognitive framework, known as naive utility calculus (NUC), agents are conceptualized as generative models that estimate actions' utilities by balancing costs and rewards. This framework leverages Bayesian inference to reverse-engineer these assessments, allowing agents to predict the behaviors of others based on inferred utilities (Jara-Ettinger et al., 2020; Jara-Ettinger et al., 2019).

Various factors influence the emergence of ToM, including language development (Astington, 2006; Astington & Jenkins; 1999; Ebert, 2020, Sarmento-Henrique et al., 2019), brain development (Frith & Frith, 2003; Gallagher & Frith, 2003; Schurz et al., 2021), executive function (Aboulafia-Brakha et al., 2011; Fujita et al., 2022), cultural differences (Shahaeian et al., 2011), and socioeconomic factors (Cavadel & Frye, 2017; Kara & Selcuk, 2021; Ruffman et al., 2015). The trajectory of ToM development can vary significantly across different contexts. Studies indicate that children in non-WEIRD (Western, educated, industrialized, rich, democratic) populations often develop ToM through culturally specific social interactions and practices, which can differ significantly from those in WEIRD populations (Dixson et al., 2017; Mayer & Träuble, 2015; Menezes et al., 2014).

Additionally, children with neurodevelopmental conditions such as autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD), and other related conditions often exhibit variations in the developmental trajectories of ToM (Andreou & Skrimpa, 2020;

Ahmed & Miller, 2011; Baron-Cohen, 2000; Baron-Cohen, Leslie, & Frith, 1985; Bora & Pantelis, 2016; Happé & Frith, 2014; Perner et al., 1989; Peterson et al., 2007; Schuwerk et al., 2014; Uekermann et al., 2010; Yu et al., 2020)

During early schooling, ToM significantly enhances social relationships by fostering prosocial behaviors, thereby increasing peer acceptance and reducing rejection (Banerjee et al., 2011; Caputi et al., 2012; Imuta et al., 2016; Slaughter et al., 2015). Academically, ToM plays a crucial role in skills such as reading comprehension and scientific reasoning, enabling children to interpret texts and hypotheses from diverse perspectives (Lecce & Devine, 2021). Furthermore, longitudinal studies indicate that ToM development at age 5 predicts later achievements in reading and mathematics by age 10, partly attributed to improved receptivity to feedback, which enhances learning outcomes (Lecce et al., 2011).

Assessing ToM in clinical and educational settings is essential for understanding individual developmental trajectories and providing tailored, evidence-based interventions (Beaudoin et al., 2020; Fu et al., 2023). For children with neurodevelopmental conditions such as ASD and ADHD, targeted ToM training can significantly enhance social cognition (Hoffman et al., 2016). Assessing ToM requires tools that comprehensively evaluate a wide range of mental states, including beliefs, desires, and intentions, to capture its full complexity (Beaudoin et al., 2020; Fu et al., 2023).

Children's ToM assessments utilize a variety of methods such as spoken stories, comic strips, cartoons, dramatizations, and picture sequences, catering to a range of verbal abilities through diverse presentation and response styles, including spoken language, sign language, and multiple-choice formats (Fu et al., 2023). Video-based assessments are increasingly significant, offering a dynamic and engaging format (Fu et al., 2023). Computer-based neuropsychological assessment (CBNA) tools enhance clinical neuropsychology with precise data acquisition, controlled cognitive measurement, and reduced errors (Galindo-Aldana et al., 2018). Compared to traditional paper methods, CBNA offers reliable, resource-efficient, and accurate assessments(Parsey & Schmitter-Edgecombe, 2013).

When developing new measurement tools, assessing content validity is the foundational step. Content validity ensures that the test items comprehensively cover the construct being measured (Alexandre and Coluci; 2011; Coluci et al., 2015; Delgado-Rico et al., 2012). Following this, tool validity encompasses several critical phases (Reppold & Hutz, 2014). Construct validity assesses the tool's accuracy in measuring theoretical constructs, often through correlational studies with established benchmarks. Criterion validity involves comparing the tool's outcomes with external criteria known to indicate the construct.

Reliability testing ensures consistency across repeated measurements, solidifying the tool's utility in capturing complex constructs (Reppold & Hutz, 2014). These phases are essential for developing scientifically robust measurement instruments (Cohen et al., 2014; Pasquali, 2009).

Existing measures in ToM often focus narrowly on specific components like false beliefs, thereby failing to capture the entire developmental trajectory of ToM (Beaudoin et al., 2020; Fu et al., 2023, Rakoczy, 2022). A literature review by Fu et al. (2023) identified only twelve out of 127 measures that encompass the full spectrum of ToM development, with just four addressing all construct dimensions. Therefore, there is a need for more comprehensive and developmentally sensitive ToM assessment tools. Such tools should capture ToM's multidimensional nature, offering deeper insights into children's social cognitive development (Fu et al., 2023). Both systematic reviews on ToM measures (Beaudoin et al., 2020; Fu et al., 2023) highlight significant gaps—particularly the lack of standardization across diverse populations—and few existing tools are validated using robust methods. Such validation is crucial for accurately measuring targeted psychological constructs (Coluci et al., 2015; Pasquali, 2009).

In the Brazilian context, there are few ToM assessments with verified psychometric properties (Oliveira & Mecca, 2016). The TMEC is a Brazilian instrument that has demonstrated content validity and correlations with external variables (Mecca et al., 2018; Dias et al., 2020). The NEPSY-II, for children aged 3-16, includes a subtest for ToM and assesses emotion recognition (Korkman & Kemp, 1998). Argollo et al. (2009) adapted it for Brazil using back-translation and cultural adjustments, revealing age-related performance differences in psychometric analyses.

Drawing upon existing research and addressing identified gaps in the assessment of ToM, this study introduces the theory of mind task (ToMT) for children aged 3 to 8. This tool, based on Wellman and Liu's Theory of Mind scale (2004) and Jara-Ettinger's naïve utility calculus framework (2016b, 2020), consists of 12 video-animated tasks, with an approximate duration of 15 minutes. Children's comprehension and reasoning of perception, emotions, desires, knowledge, beliefs, and intentions are evaluated with ToMT, incorporating advanced concepts such as preference reasoning, behavior prediction, and moral judgments through the lens of utility calculus.

2. THEORETICAL FRAMEWORK

2.1 Core knowledge and Commonsense psychology based on Naive Utility Calculus

Extensive research in cognitive developmental psychology and cognitive science indicates that infants possess an innate core knowledge system, forming the foundation for early cognitive development (Carey, 2009; Carey & Spelke, 1994; Spelke, 2022; Spelke & Kinzler, 2007; Rule et al., 2020; Ullman & Tenenbaum, 2020). This core system includes distinct cognitive mechanisms for representing objects, actions, number and space, that allow infants to construct mental models of the world (Spelke, 2022; Ullman & Tenenbaum, 2020). Bayesian probabilistic models guide infants' expectations, enabling them to update their mental models with new data, thus refining their understanding and bridging the gap between innate knowledge and learned experiences (Ullman & Tenenbaum, 2020).

Human cognition, based on intuitive theories formed in infancy, helps children understand concepts like gravity, force, and mass, aiding navigation and prediction of the physical world (Baillargeon, 2004; Spelke, 1990). From an early age, infants understand object permanence and expect objects to follow physical principles such as cohesion and continuity (Baillargeon et al., 1985; Baillargeon, 1994; Spelke et al., 1994). Infants initially understand actions by observing their goals and outcomes, a basic form of reasoning (Gergely & Csibra, 2003; Phillips & Wellman, 2005; Woodward, 1998). They also grasp numerical concepts (Feigenson et al., 2004; Pica et al., 2004; Xu & Spelke, 2000) and comprehend geometric principles and spatial relationships (Dehaene et al., 2006; Spelke et al., 2010).

This foundational knowledge serves as a cornerstone for the development of more sophisticated intuitive theories and understanding, enabling children to navigate and make sense of their environment (Aschersleben & Jovanovic, 2008; Carey & Spelke, 1994; Spelke, 2022; Ullman, 2015).Furthermore, studies suggest that infants possess innate cognitive mechanisms for social interaction, including cue recognition and observational learning (Spelke et al., 2013). These mechanisms mature through interaction with the environment, laying the groundwork for advanced intuitive psychology skills essential for social navigation (Spelke et al., 2013; Tenenbaum et al., 2007).

Commonsense psychology, or folk psychology, is the intuitive human ability to attribute mental states—such as intentions, goals, beliefs, desires, and emotions—to oneself and others to understand, predict, and explain behavior (Dennett, 1987; Fletcher, 1984; Heider, 1958; Jara-Ettinger, et al., 2016b; Stojnić et al., 2023; Wellman and Woolley, 1990). Research indicates that infants demonstrate early signs of understanding intentionality and goal-directed behavior in others (Gergely et al., 1995; Liu, 2020; Phillips et al., 2002; Tomasello et al., 2005). By six months, infants can anticipate efficient actions, expecting agents to minimize the cost of their actions (Liu & Spelke, 2017; Scott and Baillargeon, 2013; Skerry et al., 2013). At a young age, children use statistical reasoning to understand others' psychological states, they show surprise by observing unexpected choices, demonstrating their ability to infer psychological states and intentions (Kushnir et al., 2010; Wellman et al., 2016). This early manifestation resembles a rudimentary form of utility calculus (Baker et al., 2009; Jara-Ettinger et al., 2016b). Liu and Spelke (2017) found that infants look longer at inefficient actions, indicating surprise and suggesting an innate expectation of efficiency. By ten months, infants use a naive utility calculus to infer goal value from action costs, expecting agents to prefer more costly goals, reflecting early commonsense psychology (Liu et al., 2017).

Building on this foundation, the NUC refines these attributions by positing that children, from early childhood, interpret actions as efforts to maximize utility by balancing costs and rewards (Jara-Ettinger, 2016b; Jara-Ettinger et al., 2020). They use Bayesian inference to interpret peers' preferences based on statistical patterns and probabilities, similar to how an econometrician analyzes data to make informed predictions (Baker et al., 2017; Lucas et al., 2014). They employ statistical reasoning to infer preferences, distinguish between knowledgeable and ignorant agents, and navigate complex social interactions (Jara-Ettinger et al., 2016b, 2015b). By observing behaviors and outcomes, children build mental models, using these observations as statistical data to infer the mental states and underlying motivations of others. This approach guides them in decoding the reasoning behind decisions, such as the factors influencing a peer’s choice of a toy (Jara-Ettinger et al. 2020).

As infants grow, their basic commonsense psychology evolves into a more sophisticated Theory of Mind (Aschersleben & Jovanovic, 2008; Bartsch, 1996; Bartsch & Wellman, 1989). This involves not only recognizing that others have different desires and goals but also understanding that others can have beliefs that are different from reality and from the child's own beliefs. As they grow and gain more experiences, their understanding of these mental states becomes more sophisticated, allowing them to better navigate complex social situations (Wellman, 2014).

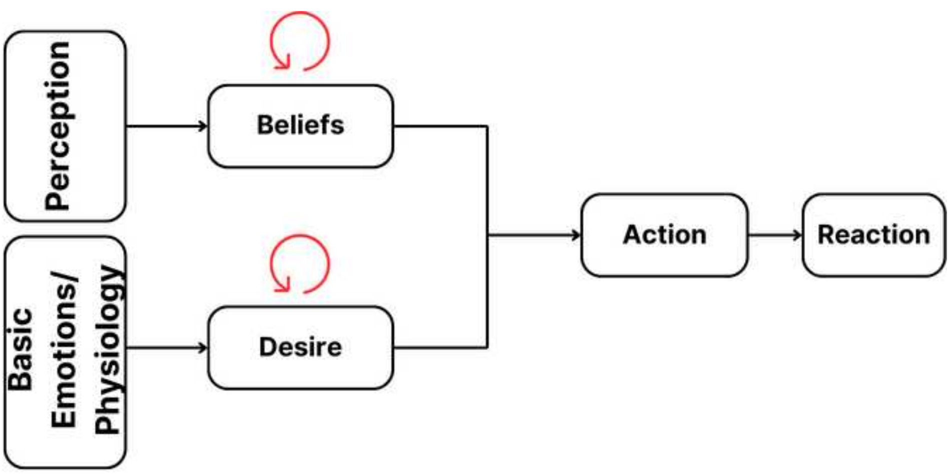
2.2 Theory of Mind Development

The development of ToM occurs gradually during childhood (Wellman, 2018). It begins to emerge towards the end of the first year, when infants develop an understanding of perception-goal psychology (Rakoczy, 2022). Furthermore, infants as young as 14 months can discern and reason about the desires of others, which marks a significant milestone in ToM development (Repacholi and Gopknik, 1997). This stage of acknowledging various desires and the ability to perceive from different visual perspectives, known as Level I Perspective Taking with evidence suggesting that starts from 24 months (Moll & Tomasello, 2006), indicates an initial awareness of the diversity in mental states. Though this stage does not yet demonstrate the child's full capacity for meta-representation, it serves as a precursor to the more advanced belief-desire psychology (Bartsch, 1996; Bartsch & Wellman, 1989; Rakoczy, 2022; Wellman et al. 2001; Wellman & Bartsch, 1988).

Figure 1, from Wellman and Bartsch (1988) presents a scheme for belief-desire reasoning, starting with perception—processing sensory input such as sight and sound. These perceptions inform beliefs, including knowledge and assumptions, while basic emotions and physical states like love or fear parallel this process. Such feelings fuel desires, which, along with beliefs, prompt actions. These actions then elicit emotional reactions. This framework outlines the progression from sensory input to emotional response, guided by the interplay of beliefs, desires, and actions (Wellman, 2014).

Figura 1

Scheme for Depicting Belief-desire Reasoning



Note. This figure illustrates the belief-desire reasoning process, highlighting how beliefs and desires lead to actions and subsequent emotional reactions. Adapted from Wellman and Bartsch (1988).

Level II Perspective Taking evolves from the prior stage, as children begin to understand that an object can appear differently from various viewpoints (Flavel et al. 1981). This understanding is more complex and sophisticated, as it requires the child to appreciate that not only do people see from different positions but that these positions can lead to different interpretations or beliefs about the same object (Flavel et al., 1981). Children start to recognize that others may hold different beliefs based on their unique experiences and information, paving the way for understanding false beliefs (Masangkay et al. 1974).

At four years old, children reach a significant milestone in cognitive development with the maturation of ToM. They gain the ability to recognize that beliefs and desires can differ from their own, a concept solidified by their understanding of false beliefs (Wellman and Liu, 2004, Rakoczy, 2022; Wellman, 2018). This key stage, evidenced by their success in false-belief tasks, represents a major advancement in ToM, highlighting their newfound ability to anticipate that others' actions may be driven by distinct, even erroneous, beliefs (Baron-cohen et al., 1985; Wellman, 2018; Mayes et al. 1996).

Children's understanding of ToM continues to evolve beyond age four, with progressively developing abilities to comprehend complex emotions and higher-order recursive thoughts throughout adolescence. The gradual mastery of advanced ToM skills, such as interpreting irony and indirect speech, unfolds over time. These nuanced differences in ToM proficiency persist into adulthood and are often accompanied by a reciprocal decline observed in older age (Derksen et al., 2018; Rakoczy, 2022; Wellman, 2018).

The development of ToM in non-WEIRD populations shows diverse patterns influenced by cultural and environmental factors. For instance, Samoan children show delayed performance on false-belief tasks compared to their Western counterparts, often failing both false-belief and true-belief tasks (Mayer and Träuble, 2015). Moreover, Dixson et al. (2017) found that rural Melanesian children from Vanuatu exhibited a marked delay in false-belief performance compared to urban counterparts, highlighting the impact of factors like access to formal education and mental state talk. Research on Asurini children in Brazil's Amazon region showed variability in ToM development, with children displaying ToM abilities between ages 3 and 10, influenced by unique cultural practices and environmental interactions (Menezes et al., 2014).

These findings underscore the intricate and multifaceted nature of ToM development, highlighting the significant influence of both neurodevelopmental and sociocultural factors in shaping children's understanding of mental states.

2.2.1 Theory of Mind Assessment

Traditionally, ToM assessments have centered around false belief tasks, which have played a pivotal role in understanding mental state attribution in children (Wellman, 2018). Fu et al. (2023) conducted a systematic review encompassing a broad array of methodologies for assessing ToM in children. They identified various approaches including scenarios, spoken stories, comics, films, cartoons, and more, with traditional tasks like false belief being most prevalent. A growing body of research suggests that ToM is a multifaceted construct, necessitating assessments that capture its cognitive and affective dimensions comprehensively (Beudoin et al., 2020; Fu et al., 2023; Wellman, 2018). Table 1 provides an overview of the developmental sequence in Theory of Mind (ToM) acquisition among preschoolers, highlighting the diverse desires, beliefs, and emotions tasks and their corresponding success rates.

Table 1

Developmental Sequence in ToM Acquisition Among Preschoolers

ToM Components	Description of the Tasks	% of Correct Responses in Preschoolers (3 to 6 years)
Diverse Desires	Children are presented with a toy figure (Mr. Jones) and two snacks (a carrot and a cookie) on a sheet of paper. After choosing their preferred snack, they learn Mr. Jones prefers the opposite snack. To pass, they must predict Mr. Jones's choice, recognizing his desire differs from their own. This task stems from studies by Wellman and Woolley (1990) and Repacholi and Gopnik (1997).	95%
Diverse Beliefs	Using a toy figure (Linda) and drawings of bushes and a garage, after the child guesses where Linda's cat is hiding, they're informed Linda believes the opposite. They must then predict where Linda will look for her cat. The task is adapted from Wellman and Bartsch	84%

	(1989) and Wellman et al. (1996).	
Knowledge Access	A nondescript box with a hidden toy dog inside is shown to the child, who is then asked about another figure's (Polly) knowledge of the contents without Polly having seen inside. Modified from Pratt and Bryant (1990) and Pillow (1989) to align with the format of the Contents False-Belief task.	73%
Contents False Belief	A Band-Aid box, supposedly containing Band-Aids but actually containing a toy pig, is used to assess whether the child understands that another person (Peter) can have a false belief about its contents. Adapted from Perner, Leekam, and Wimmer (1987), this task has been widely modified and used (Wellman et al., 2001).	59%
Explicit False Belief	This scenario involves a toy figure (Scott), who mistakenly believes his mittens are in a different location than where they actually are. Children are asked where Scott will look for his mittens, assessing understanding of false beliefs. Inspired by Wellman and Bartsch (1989) and Siegal and Beattie (1991).	57%
Belief – Emotion	Involving a Cheerios box that actually contains rocks, children predict Teddy's emotional reaction upon receiving the box, then after discovering its true contents. This task is derived from Harris, Johnson, Hutton, Andrews, and Cooke (1989).	52%
Real – Apparent Emotion	Through a story about Matt, who hides his true feelings about a mean joke, children distinguish between Matt's real feelings and his facial expression. The task originates from Harris, Donnelly, Guz, and Pitt-Watson (1986).	32%

Note. Data adapted from Wellman and Liu (2004).

The seminal study by Henry M. Wellman and David Liu in 2004, titled Scaling of ToM Tasks, aimed to empirically establish a developmental progression in preschoolers' understanding of ToM. The study involved an assessment of 75 children aged between 2 years 11 months to 6 years 6 months across seven tasks designed to tap into different aspects of ToM understanding. These tasks ranged from understanding diverse desires to discerning real from apparent emotions, encapsulating a developmental trajectory from simpler to more complex mental state understandings, as we can see at Table 4.

The findings confirmed a consistent developmental progression where children systematically acquired the ability to pass more complex tasks based on their age and cognitive development (Wellman & Liu, 2004). The Wellman and Liu (2004) study provides a structured framework for understanding the progression of ToM in early childhood, highlighting the gradual development from basic to more complex understandings of mental states (Beaudoin et al., 2020).

The Diverse Desire experiment by Repacholi and Gopnik (1997) elucidates toddlers' recognition that others can possess desires differing from their own. In this study, participants, toddlers aged 14-18 months, were randomly assigned to conditions following a baseline assessment of their food preferences. The methodology involved a food-request procedure where toddlers observed the experimenter's reactions - expressions of pleasure or disgust - toward crackers and broccoli, respectively, before deciding which food to offer her. This approach tested the children's capacity to infer the experimenter's desires based on emotional cues.

Furthermore, within the ToM tasks, Visual perspective-taking tasks evaluate children's understanding that different people may perceive the same object differently due to their physical positions, as discussed by Masangkay et al. (1974) and Flavell et al. (1981). Their experiments were designed to assess children's understanding of direct visibility of objects to others (Level 1), where participants discerned which animal was visible to the experimenter on a dual-sided card, probing their grasp of straightforward visibility without the need for perspective-taking. Additionally, they assess children's insight into perspective differences in perception (Level 2) by having them determine the orientation (upside down or right side up) of a turtle picture from the experimenter's viewpoint across the table. These studies provide evidence regarding the developmental shift from recognizing direct visibility to a nuanced appreciation of perspective differences (Rakoczy, 2022).

False belief tasks serve as critical assessments for ToM (Wellman, 2018). Established studies, such as those by Wimmer and Perner (1983) and Baron-Cohen et al. (1999), have demonstrated that by the age of four, children typically acquire the ability to discern these divergent beliefs. For instance, in the "Sally-Anne" task, a child's accurate prediction that Sally will look for her marble in the original location, despite its relocation, signifies an important milestone in the development of understanding false beliefs. Furthermore, a significant component of false belief assessments involves evaluating second-order false beliefs (Astington et al., 2002). These tasks evaluate an individual's ability to comprehend that one person can hold incorrect beliefs about a third party's perspective or belief, offering insight into the sophisticated layers of ToM (Perner & Wimmer, 1985; Astington et al., 2002).

In the study by Arslan et al. (2017), the Chocolate Bar Task narrative is utilized to assess children's understanding of second-order false beliefs. In this narrative, Ayla and

Murat, siblings engaged in play, encounter a situation where their mother gives chocolate to Murat while excluding Ayla due to her misbehavior. Subsequently, Murat hides the chocolate in a drawer, unbeknownst to Ayla. Following Murat's departure, Ayla moves the chocolate to a toy box without Murat's knowledge. The control questions proposed by Arslan et al. (2017) for the "Chocolate Bar" story ensure children comprehend the storyline and characters' beliefs. These questions evaluate whether children recognize Murat's and Ayla's awareness regarding the chocolate's location, testing memory and perspective-taking abilities. This foundational understanding sets the stage for more intricate inquiries into beliefs, thereby preparing children for deeper exploration of ToM concepts.

During later childhood, typically around 8-10 years old, children demonstrate significant advancements in ToM abilities through various tasks that reflect their developing social cognitive skills (Fu et al., 2023). In the systematic review by Fu et al. (2023), studies suggested that at this stage, children exhibit proficiency in tasks such as recognizing lies, enabling them to differentiate between truthful statements and deceptive ones. Furthermore, they acquire an understanding of sarcasm and irony, discerning the disparity between literal meanings and intended messages, thereby enhancing their interpretation of subtle social cues. ToM development unfolds throughout childhood, as evidenced by diverse tasks measuring various components of this construct (Arslan et al., 2017; Beaudoin et al., 2020; Wellman & Liu, 2004). Children progressively comprehend mental states, enhancing social cognitive skills like interpreting sarcasm and irony by late childhood.

2.2.2 Naive Utility Calculus Experiments

The concept of the naive utility calculus posits that individuals act to maximize their utilities by balancing rewards and costs (Jara-Ettinger et al., 2020). This intuitive framework helps predict and explain the behavior of others, forming a core component of commonsense psychology (Jara-Ettinger et al., 2016b). By integrating this model with a Bayesian framework, researchers aim to understand how even young children infer others' beliefs, desires, and motivations from their actions, providing a computational basis for social reasoning (Jara-Ettinger et al., 2020).

Tables 1, 2, and 3 below present a series of experiments from Jara-Ettinger et al. (2016a, 2016b, 2015a, 2015b) that explore children's ability to infer others' actions through the lens of naive utility calculus. These experiments examine children's reasoning about costs and rewards, differentiate between knowledgeable and ignorant agents, and assess their judgments on the moral status of agents.

Understanding Costs and Rewards Underlying Rational Actions

Research by Jara-Ettinger et al. (2015a) investigates children's understanding of rational action using the NUC, focusing on whether children can infer individual differences in behavior based on unobservable costs and rewards. They explore children's advanced reasoning beyond basic Theory of Mind through experiments with five and six-year-olds. Experiments on Table 2 the authors test three key concepts: whether children understand that agents do not always pursue the highest rewards due to high costs, recognize that costs and rewards vary between agents and are not directly observable, and predict how changes in costs and rewards affect an agent's actions.

Table 2

Experiments on Children's Understanding of Costs and Rewards Underlying Rational Action

Experiment	Ability Tested	Hypothesis	Description	Key Findings
Experiment 1	Understanding cost-reward trade-offs	Children understand that agents maximize utility, not just rewards.	32 children (mean age: 5.85 years, range 5.0 - 6.9 years) observed a puppet choosing between two treats with different costs in separate trials.	Children inferred the puppet's preferences based on the cost-reward trade-offs, indicating they consider both costs and rewards.
Experiment 2	Inferring agent-specific competencies	Children can use differences in agents' preferences to infer their competencies.	32 children (mean age: 5.8 years, range 5.0 - 6.9 years) saw two puppets choosing between treats, with one puppet preferring a more costly treat.	Children identified the puppet unable to perform a high-cost action, showing they infer competencies based on preferences.
Experiment 3	Manipulating costs to infer competencies	Children can manipulate external costs to learn about an agent's subjective costs.	16 children (mean age: 6.0 years, range 5.1 - 6.8 years) placed treats in locations with different costs to see how the puppet	Children placed high-reward treats in high-cost locations to infer the puppet's competence, demonstrating

			would choose.	understanding of cost-reward effects.
Experiment 4	Identifying agents based on subjective rewards	Children can identify agents whose preferences reveal their competencies.	16 children (mean age: 6.0 years, range 5.0 - 6.9 years) chose which puppet to test based on their preferences for treats in high-cost locations.	Children selected the puppet with the higher-cost preference, indicating they can use preferences to infer competencies.

Note. Data adapted from Jara-Ettinger et al. (2016a).

The experiments reveal young children's nuanced grasp of rational action, weighing costs and rewards in decisions. This research illuminates early economic reasoning and cognitive underpinnings of social judgments (Jara-Ettinger et al., 2015a).

Children's Understanding of Action Under Uncertainty

Through a series of experiments, Jara-Ettinger et al. 2016a investigate whether children can differentiate between knowledgeable and naive agents based on the stability and outcomes of their choices, understanding that knowledgeable agents achieve higher rewards and make more stable decisions. As shown in Table 3, these experiments reveal that children as young as 4 years old grasp the impact of uncertainty on decision-making. They predict outcomes based on agents' knowledge, infer prior knowledge from action outcomes and stability, highlighting a sophisticated early understanding of expected utility theory in social cognition.

Table 3

Experiments on Children's Understanding of Action Under Uncertainty

Experiment	Ability Tested	Hypothesis	Description	Key Findings
Experiment 1	Understanding effect of knowledge on rewards	Knowledgeable agents are more likely to accrue high actual rewards than naive agents.	16 children (mean age: 5.09 years, range 4.13-5.89 years) observed	Children inferred the knowledgeable puppet was more likely to say "Yum!",

			two puppets choose the same fruit, one said "Yum!" and the other "Yuck!"	showing they understand the impact of knowledge on expected rewards.
Experiment 2	Understanding stability of choices based on knowledge	Knowledgeable agents are more likely to make stable choices over time than naive agents.	16 children (mean age: 5.16 years, range 4.01-5.96 years) observed two puppets choose the same fruit, one changed their mind.	Children inferred the naive puppet was more likely to change their mind, indicating an understanding of choice stability related to knowledge.
Experiment 3	Inferring knowledge based on reward outcomes	Agents who obtain high rewards are more likely to have been knowledgeable prior to making their choice.	32 children (mean age: 5.12 years, range 4.12-5.98 years) saw two puppets choose the same fruit, one said "Yum!" and the other "Yuck!"	Children inferred that the puppet who said "Yuck!" was the naive one, showing they use reward outcomes to infer prior knowledge.
Experiment 4	Inferring knowledge based on stability of choices	Agents who make stable choices are more likely to be knowledgeable.	16 children (mean age: 5.64 years, range 4.04-5.93 years) observed two puppets choose the same fruit, one changed their mind.	Children inferred the puppet who changed their mind was the naive one, indicating they use choice stability to infer knowledge.

Note. Data adapted from Jara-Ettinger et al. (2016a).

These findings significantly enhance our understanding of how children develop cognitive skills related to decision-making and social reasoning, demonstrating their early ability to evaluate and infer the knowledge and stability of others based on observed behaviors and outcomes (Jara-Ettinger et al. 2016b).

Toddlers' Reasoning About Costs, Competence, and Culpability

The experiments proposed by Jara-Ettinger et al. (2015b) investigate how toddlers use the naive utility calculus to make social judgments about agents' actions. Table 4 details these experiments, which aim to test whether toddlers prefer more competent agents who achieve goals with lower costs and whether they judge less competent agents, who incur higher costs, as nicer when both fail to help. These studies seek to determine if toddlers show a preference for agents who demonstrate higher competence and efficiency in achieving goals, and whether they exonerate or judge more sympathetically agents who incur higher costs and fail to help, attributing their failure to high effort rather than lack of motivation.

Table 4

Experiments on Toddlers' Social Reasoning: Costs, Competence, and Culpability

Experiment	Ability Tested	Hypothesis	Description	Key Findings
Experiment 1	Preference for competent agents	Toddlers prefer agents who incur lower costs to achieve goals	24 toddlers (mean age: 21.58 months) saw two puppets activate a toy; one did so easily, the other struggled. Children chose which puppet to play with.	93.75% preferred the more competent agent.
Experiment 2	Social evaluation based on cost of actions	Two-year-olds prefer less competent agents who fail to help	17 two-year-olds (mean age: 2.64 years) saw two puppets activate a toy and then refuse to help a parent. Children chose which puppet to play with.	68.75% preferred the more competent agent despite refusal to help.
Experiment 3	Moral judgment based on agent competence and cost	Toddlers judge less competent agents as nicer	66 two-year-olds (mean age: 2.48 years) saw two puppets activate a toy and refuse to help. Children judged which puppet was nicer or which they preferred to play with.	81.25% preferred the competent agent for play; 68.75% said the less competent agent was nicer.

Note. Data adapted from Jara-Ettinger et al. (2015b)

These experiments demonstrate toddlers' nuanced social reasoning, showing their sensitivity to action costs and ability to distinguish competence from moral goodness. They apply the naive utility calculus in moral judgments, excusing less competent agents from helping duties due to perceived higher costs, and viewing them more favorably (Jara-Ettinger et al., 2015b).

The naive utility calculus offers a unifying framework for understanding social cognition, demonstrating that even young children possess sophisticated inferential abilities regarding others' actions (Jara-Ettinger et al., 2020). Through various experiments, it has been shown that children can infer competencies, preferences, and motivations by observing the costs and rewards associated with actions. This framework not only enhances the understanding of early cognitive development but also bridges intuitive decision-making with formal economic theories, suggesting a deep-rooted connection between commonsense psychology and scientific models of human behavior (Jara-Ettinger et al., 2020).

3. OBJECTIVES

3.1 General Objective

To develop and validate a Theory of Mind computer-based tasks (ToMT) to evaluate ToM abilities in children aged 3 to 8.

3.2 Specific Objectives

- Design and formulate a comprehensive set of tasks for the target age group to capture the constructs of ToM and NUC.
- Verify evidence of content validity of the ToMT.
- Verify target population comprehension and acceptance in the tasks of the ToMT.

4. METHOD

The method of this study is developed in three distinct steps. First, the development of tasks; this is followed by a content validity analysis involving both qualitative and quantitative expert panels; and finally, a pilot study is conducted with the target population.

4.1 Development of a computer-based Theory of Mind Task (ToMT)

4.1.1 Study Procedures

A comprehensive literature review was conducted to define the construct and establish its theoretical foundation. The ToMT was developed based on principles from the development of social cognition theories in Developmental Psychology and Cognitive Science (Baker et al., 2011; Flavell, 1991; Jara-Ettinger et al., 2016a; Jara-Ettinger et al., 2019; Spelke, 2013; Wellman & Liu, 2004). The tasks were chosen and adapted from Wellman and Liu's Theory of Mind Scale (2004) and Jara-Ettinger's Naive Utility Calculus framework (2016), creating a comprehensive measure of ToM.

The development process involved creating tasks to assess children's comprehension and reasoning of perception, emotions, desires, knowledge, beliefs, and intentions. The construction of these tasks was guided by experts in the field to ensure they were age-appropriate. Initially, characters and scenarios were designed in 2D, which were then animated to children.

4.2 Content Validity of the ToMT

This study focuses on a mixed-methods design, combining qualitative and quantitative expert panel content analysis to verify the clarity, relevance, and pertinence of the ToMT tasks.

4.2.1 Qualitative content analysis of the ToMT

4.2.1.1 Participants

The panel included two academics renowned for their contributions to the theoretical framework of social cognition and decision-making studies in childhood. They were chosen for their substantial scholarly work published in high-impact journals and their significant influence on recent literature in the domain.

The panel comprised Expert 1, a Harvard cognitive psychology professor with a Ph.D., renowned for her groundbreaking research in infants' cognitive abilities. Her influential work, recognized in top-tier journals, has earned her prestigious awards for pioneering contributions to the field.

Expert 2, a Yale Associate Professor with a Ph.D. in Brain and Cognitive Sciences from MIT, is known for his groundbreaking research in children's reasoning and Decision-making. His contributions to cognitive development and social cognition have been honored with significant awards for innovation and excellence in science.

4.2.1.2 Procedures

Experts in Developmental Psychology, Cognitive Science, and Social Cognition were selected for their notable contributions to the literature. Invitations to participate were sent via email, and individual video calls were conducted through the Google Meet platform to introduce the ToMT. During these calls, academics provided qualitative feedback and discussed their experiences with applying similar tasks in experimental research settings. Their feedback was recorded and documented. Based on their insights and shared experiences, the tasks were updated and refined to create an improved version.

4.2.2 Quantitative content analysis of the ToMT

4.2.2.1 Participants

For the content validity analysis, we engaged a group of three Brazilian academics, each holding a Ph.D., with publications and research related to social cognition. These individuals were selected based on their substantial contributions to the field at the national level. Judge 1 is an Associate Professor with a Ph.D. in Psychology, focusing on neurodevelopmental disorders and the development and adaptation of neuropsychological assessments. Judge 2 is an Associate Professor with a Ph.D. in Psychology, concentrating on cognitive neuropsychology, social cognition, and neuropsychological assessment. Judge 3 is an Associate Professor with a Ph.D. in Developmental Disorders, specializing in neuropsychological assessment and neurodevelopmental disorders.

4.2.2.2 Instruments

Content Validity Sheet for the ToMT (Appendix A): The Content Validity Sheet for the ToMT, uses a 1 to 5 scale to calculate the Content Validity Coefficient (CVC) for tasks, to ensure their alignment with the test's objectives in assessing children aged 3 to 8. Experts evaluate each task on clarity (ease of understanding), pertinence (alignment with ToM and Decision-making constructs for assessment), and relevance (accuracy in reflecting specific mental state understandings). They rated the items on a scale of 1 to 5 across three dimensions—clarity, pertinence, and relevance defined as follows:

- Clarity: Ranging from "Not clear at all" to "Totally clear."
- Pertinence: Ranging from "Not pertinent at all" to "Totally pertinent."
- Relevance: Ranging from "Not relevant at all" to "Totally relevant."

4.2.2.3 Study procedures

For the quantitative panel analysis, five Brazilian experts in the field were invited to evaluate the ToMT, but only three accepted. During individual video calls, the animations were presented, and panelists rated each item for clarity, pertinence, and relevance on a scale from 1 to 5 using a Content Validity Spreadsheet. Their qualitative feedback and suggestions were documented. Although the instrument was revised based on this feedback, not all suggested updates were included in the version used for the pilot study. The instrument is currently being further updated.

4.2.2.4 Statistical Analysis

The statistical analysis process refers to the calculation of the Content Validity Coefficient (CVC) for evaluating the content validity of items in a research instrument. This process is based on the method developed by Hernandez-Nieto (2002) for calculating CVC. The analysis is conducted in the following steps:

Average Score Calculation. For each criterion of an item, the average score is computed. This is the sum of all the judges' scores for that particular criterion divided by the total number of judges.

$$\text{Average}_{\text{criterion}} = \frac{\sum \text{Judges' Scores for the criterion}}{\text{Number of Judges}}$$

CVC Calculation for Each Criterion of Each Item. The individual CVC for each criterion is determined by subtracting the calculated bias from the average score for that criterion and then dividing by the maximum possible score of the scale. Here bias is calculated as follows.

$$\text{bias} = \left(\frac{1}{\text{number of judges}}\right)^{\text{number of judges}}$$

Total CVC for Each Criterion. The Total CVC for a given criterion is the arithmetic mean of the individual CVCs for all items for that criterion. This is calculated by summing the individual CVCs for all items and dividing by the number of items.

$$\text{CVC Total}_{\text{criterion}} = \frac{\sum \text{CVC}_{\text{criterion}} \text{ for all items}}{\text{Number of items}}$$

The CVC provides a numerical value from 0 a 1 that indicates the degree to which a panel of experts agrees on the validity of the items, taking into account possible biases due to the number of judges involved in the assessment process (Hernandez-Nieto, 2002).

4.3 Pilot with Target Population

The study is characterized by a quantitative approach and employs a cross-sectional design. It aims to describe and explore the relationships between variables, without attempting to establish cause and effect (Campos, 2008).

4.3.1 Participants

The pilot study initially received a total of 15 applications. However, only 10 children with typical development, aged between 3 and 8 years, were included in the study. Five children were excluded as they did not meet the inclusion criteria. The criteria for inclusion required children to be within the age range of 3 to 8 years and to have typical development. The exclusion criteria included global neurodevelopmental disorders, neurological dysfunctions and injuries, as well as motor, auditory, and visual deficits. The excluded children had conditions such as ADHD, ASD, oppositional defiant disorder, and other neurodevelopmental disorders.

4.3.2 Instruments

4.3.2.1 Participant Identification Questionnaire (PIQ) (Appendix B)

This questionnaire includes personal identification, information about the parents' and children's education levels, profession, health, clinical conditions, and development. It also covers questions about the family's socioeconomic status. The completion time is approximately 15 minutes, and it is filled out using Google Forms.

4.3.2.2 ToMT: Theory of Mind Task (ToMT) (Souza et al., 2024). For more information, see page 37.

4.3.2.3 Developmental Neuropsychological Assessment Battery (Social Perception Domain Subtests) 2nd Edition (NEPSY-II) (Argollo et al., 2009).

The NEPSY-II is a comprehensive neuropsychological assessment specifically designed for children aged 3 to 16. It evaluates a wide range of cognitive skills and functions, including language, memory, sensorimotor abilities, social perception, and visuospatial skills across six domains (Argollo et al., 2009).

The NEPSY-II ToM subtest assesses children's comprehension of characters' beliefs, intentions, or feelings through verbal and contextual tasks involving stories and social scenarios. Verbal tasks focus on the child's ability to process and infer mental states from narrative information. In contextual tasks, children are shown images depicting various social interactions or situations, assessing the child's ability to interpret social cues and apply them to understand others' perspectives and emotions.

The Emotion Recognition subtest of the NEPSY-II assesses children's ability to recognize and distinguish emotions using photographs depicting various emotional expressions such as happiness, sadness, anger, fear, disgust, and neutrality. Tasks include matching emotions between photographs, selecting images that reflect a specific emotion, and recalling emotions from briefly viewed photographs.

4.3.2.4 Target Population Feedback Form (Appendix C)

A feedback form in the google forms containing the following questions: "How much did you like the video?" with responses ranging from "did not like" with a sad emoji, "neutral" with a neutral emoji, and "liked" with a happy emoji. "Would you recommend this video to a friend?" with responses being a thumbs down for "would not recommend" and a thumbs up for "would recommend." Finally, "What do you think could be improved in the video? Any suggestions?" and "What were your main difficulties in answering the video's questions?

4.3.3 Study Procedures

The study was submitted to the ethics committee through Plataforma Brasil and received approval from the ethics committee with Approval Number: 6.159.009. Children were recruited via social media to participate in the research at the Psychosocial Care Service (SAPSI), the psychology clinic school of the Federal University of Santa Catarina (UFSC).

The evaluation process was organized as follows: Appointments were scheduled with the families, during which they were sent an online questionnaire. Evaluations took place in a quiet room within the SAPSI, upon their arrival, guardians were requested to sign the informed consent form. Following, a video explaining the Informed Assent, was shown to the participants. With their assent to participate, the children proceeded to the next stage. They first completed the ToMT on a computer, which took about 15 minutes. After completing the ToMT children had a break of 5 minutes and then completed the NEPSY II RE and NEPSY II ToM subtests, average time 20 minutes. After the session, children provided their feedback related to the ToMT tasks via a simple form.

4.3.4 Statistical Procedures

The analysis of the pilot study data utilized both descriptive and inferential statistical methods to uncover patterns and relationships among the study variables. Given the quantitative nature of the study and the distribution of the data, appropriate statistical tests were selected. All statistical analyses were conducted using STATA software.

Descriptive statistics provided a summary of the basic features of the data in the study. For each variable—comprising demographic information such as age, sex and scores from the ToMT and NEPSY subtests (Emotion Recognition and Theory of Mind)—the mean, median, standard deviation (SD), minimum, and maximum values were calculated.

To determine the suitable statistical tests for data analysis, an assessment of normality was performed using the Shapiro-Francia test. This test was chosen for its efficacy in assessing the normality of distributions in small sample sizes, as is typical in pilot studies

(Royston, 1983; Shapiro, 1972).The Shapiro-Francia test indicated that the variables age, NEPSY II subtests ER and ToM, and the NUC subtests within the ToMT total conform to a normal distribution ($p > .05$). However, the ToM subtests within the ToMT do not conform to a normal distribution ($p = .044$), affecting their suitability for parametric tests.

Spearman's rank-order correlation was employed to explore the relationships between the ages of the participants and their scores on the ToMT, as well as the relationships among the assessment scores themselves (NEPSY ER and NEPSY ToM). This non-parametric measure of correlation was selected due to its appropriateness for ordinal data and its capability to manage non-normally distributed variables.

The strength of the Spearman's rho correlation was interpreted using Cohen's (1988) categorization, which classifies correlation coefficients ranging from -1 to 1. According to Cohen, the absolute value of the Spearman's rho correlation coefficient denotes the strength of association between two variables: a coefficient near -1 or 1 indicates a strong correlation, while a coefficient close to 0 suggests a weak correlation. Correlations are considered very weak for $|rho| \leq 0.19$, weak for $0.20 \leq |rho| \leq 0.39$, moderate for $0.40 \leq |rho| \leq 0.59$, strong for $0.60 \leq |rho| \leq 0.79$, and very strong for $0.80 \leq |rho| \leq 1.00$. This framework aids in interpreting the correlation results within this study.

To investigate the developmental progression in ToMT scores, participants were divided into three age categories: 3-5 years, 6 years, and 7-8 years. Descriptive statistics for ToMT scores were calculated for each age group to observe age-related trends and differences in performance. These analyses are exploratory due to the limited sample size, focusing on the applicability of the instrument across different age groups rather than providing evidence of validity.

5. RESULTS

In this section, the findings are presented, organized into sequential stages. Initially presenting the results for task development, following the content validity analysis, which includes both qualitative and quantitative analysis from a panel of experts, and then moving on to the results from a pilot test.

5.1 Development of ToMT

The results from the development of the ToMT items are presented below. The narrative for the ToMT was developed adapting from experiments and tasks in existing literature, including works by Arslan et al. (2017), Masangkay et al. (1974), Flavell et al. (1981), Repacholi and Gopnik (1997), Wellman and Liu (2004) and Jara-Ettinger et al. (2016b).

Table 5 provides an overview of tasks aimed at assessing children's ToM. Through scenarios involving characters such as Ana, Pati, Pedro, and Caio, children are prompted to explore various mental states. Each narrative is crafted to test specific sub-components within the ToM framework, utilizing targeted questions to gauge children's ability to infer explicit ToM states.

The NUC tasks within the ToMT framework are designed around principles that explore the naive utility calculus underlying commonsense psychology (Jara-Ettinger et al., 2016b). These tasks aim to assess children's ability to infer intentions and preferences through the naive utility calculus. This approach evaluates how children understand and predict others' behaviors by considering the goals, desires, and the potential rewards or costs associated with the actions they observe.

Table 5

Theory of Mind Tasks (ToMT) - ToM Framework and Narratives

Sub-components	Narrative	References
Visual perspective-taking I and II	Which animal is Ana seeing? Which animal is Pati seeing? Which Ana is seeing? Which animal is Ana seeing? Which animal is Pati seeing? Which Ana is seeing? Are Ana and Pati seeing the same animal? Is Ana seeing the dog standing up or upside down? In other words, is the dog head up or head down?	Masangkay et al. (1974); Flavell et al. (1981)
Visual perspective-taking I and II	"Pati and Pedro were playing in the park, Pati was behind a wall and there was a dog close to her. On the other side of the wall there was a cat. Pedro, on the other hand, climbed a tree and sat on a high branch. Now I will ask some questions, okay? Can Paty see the cat? And Pedro, can he see the cat?	Masangkay et al. (1974); Flavell et al. (1981)

	Can Paty see the dog? And Pedro, can he see the dog?"	
Diverse desires	Which of these fruits is your favorite? Strawberry, banana, or pear? Ian likes.... It's snack time. Which of these fruits will Ian choose?	Repacholi & Gopnik (1997); Wellman & Liu (2004)
Knowledge access	Look, a bag! What do you think is inside? Let's open it and see? Look, there are shoes inside. Ana arrived, she is looking for clothes to go to a party. What do you think Ana is thinking is inside the bag? Let's open it and see? Look, Ana, there are shoes here inside! Let's play a trick on Pedro! Let's close it again and show him! Pedro arrived! Does Ana think Pedro is thinking there are shoes or clothes inside the bag?	Wimmer & Perner (1983); Baron-Cohen et al. (1985); Onishi & Baillargeon (2005); Wellman & Liu (2004)
First-order False belief	This is Ian. He has a backpack. This is Pati, she has a Box. Ian has a ball. He puts the ball in his backpack. Ian goes out to play outside. Pati then takes the ball out of the backpack and puts it in the box. Ian is now back, he wants to play with the ball. Where will Ian look for the ball? In the backpack or in the box? Where does Pati think Ian will look for the ball? In the backpack or in the Box?	Hogrefe et al. (1986); Perner et al. (1987); Ruffman & Olson (1989)
Second-order false belief inference	Ana and Pedro are in the kitchen. Ana has a jar of sweets and she doesn't want Pedro to see it, so she puts the jar in the yellow box, closes it, and leaves. Pedro wants to see what Ana put in the box. Ana sneaks up on the other side and sees Pedro going to open the box. Then he opens the box and finds a jar of sweets. Pedro then puts the jar of sweets in the drawer of the cabinet. Ana saw everything, but Pedro doesn't know that. Where did Ana first put the jar of sweets? Where did Pedro put the jar of sweets after? Does Pedro know that Ana was watching him? Ana arrives and will look for her jar of sweets. What is Pedro thinking where Ana will look for the jar of sweets, in the yellow box or in the cabinet?	Arslan et al. (2017); Perner & Wimmer (1985)
Emotion comprehension based on	This is Pedro, and it's his birthday! Pedro really wants a dog as a birthday gift. Pedro's father thinks he wants a remote control car as a gift.	Wellman & Liu (2004)

second-order false belief	Pedro's father bought a remote control car for him. Pedro doesn't know about the car and thinks his father bought a dog as a gift. What does Pedro want to receive as a gift? What does Pedro think his father bought for him? What does Pedro's father think he wanted? How does the father think Pedro will feel when he receives the car? Look, Pedro's father gave him a remote control car! How did Pedro feel when he received the car?
Emotion comprehension	Ana and Pedro were watching a movie about the sea. Suddenly a huge shark appeared. Ana got very scared and said she thinks the beach is very dangerous. Pedro thought it was really cool and really wants to meet a shark. Let's call the friends to go to the beach! Who do you think will accept? Pedro or Ana?

Note. The ToMT narrative tasks were designed by the author (Souza, 2024), adapted from Wellman and Liu (2004) and other existing literature as cited in the table.

Table 6 details the narrative tasks within the NUC framework of the ToMT. Each narrative is designed to assess a specific sub-component of the naive utility calculus, including inference of preference, predicting behavior, judgments of social behavior, and distinguishing between knowledgeable and ignorant agents. These tasks present scenarios where children must consider factors such as personal preferences, physical and social constraints, and past experiences to predict the decisions of others.

Table 6

Theory of Mind Tasks (ToMT) - NUC Framework and Tasks Narratives

Sub-components	Narrative	References

Inference of preference based on utility calculus	"This is Yellow. He is hungry and wants to eat a fruit. At his home, he has two fruits. A strawberry that is far away from him and another fruit that is inside a box, which is very close to him. He knows there is a fruit inside the box, but does not know which one. Look! Now he went to get the fruit he wants to eat! He chose the strawberry! Why do you think Yellow wanted to go further to get the strawberry? On another day, Yellow was hungry again and wanted to eat a fruit. He has two fruits, a banana that is very far and the fruit that is inside the box, very close. He knows there is a fruit in the box, but does not know which one. Now Yellow went to see what was inside the box. And then he chose the banana. Why do you think he chose to walk further to get the banana after seeing the fruit that was inside the box? If Yellow had a banana and a strawberry very close, which do you think Yellow would prefer?"	Jara-Ettinger et al. (2016b)
Predict Behavior based on utility calculus	"Bruno is hungry and decides to grab a fruit. He likes all fruits, but the fruit he likes the most is the orange. He does not like heights and would not climb a high place because he is very afraid. Now, you can see there is a wall before Bruno can reach the fruit shelf? So he still cannot know what is behind the wall before opening the door. I will ask some questions, okay? Let's remember a little bit of what we know about Bruno, his favorite fruit is the orange and he cannot see what is on the shelf before opening the door. Oh, do not forget also that he is very afraid of heights! Before opening the door and seeing the fruits on the shelf, which fruit do you think the boy most wants to eat? When Bruno opens the door and sees the shelf with the two fruits, which fruit do you think he will eat, the orange or the apple? If both fruits were on the lower part, which do you think Bruno would pick, the orange or the apple?"	Jara-Ettinger et al. (2016b)

Judgments of social behavior	"Pedro is in the kitchen with two friends, Ana and Caio. He really wants to eat a banana, but it is on the last shelf, where he cannot reach. Then, Pedro asks his friends, Ana and Caio to help him get the banana. Both friends answer “no,” and do not help him. Which fruit do you think Pedro prefers: the banana that is up there or the orange that is down below? Why do you think Ana did not want to help Pedro? Why do you think Caio did not want to help Pedro? Who do you think could have helped Pedro get the banana: Ian or Ana? Why?"	Jara-Ettinger et al. (2016b)
Knowledge and Past experience	"Paty and João are going to have a snack. In the kitchen, they see a fruit option on the table, which is the green fruit. Both friends grab the fruit and eat. Paty says “Ew! That's horrible”, and João says “Hmm! Tasty”. Which of the two do you think was eating the green fruit for the first time, the boy or the girl? Why?"	Jara-Ettinger et al. (2016b)

Note. The ToMT narrative tasks were designed by the author (Souza, 2024), adapted from Jara-Ettinger et al. (2016b).

The ToMT narrative tasks were designed to match the development of children aged 3 to 8, using simple stories and objects familiar to them. Subsequently, these narratives were brought to life through engaging visual content. The development of the visual content began with a PowerPoint presentation featuring characters in simple scenarios as show in figure 2, the characters were designed considering diversity and scenarios considering cultural and familiar objects. The narratives for each task were adapted into various settings and are presented to children through a PowerPoint presentation. The examiner displays the scenarios on a computer and narrates the story as it progresses.

Figure 2

PowerPoint Version Predict Behavior Task

Note. The image presents an example of a task designed to assess NUC abilities in the ToMT. It illustrates a boy with brown skin and black hair standing outside a room in front of a closed door. Inside the room, there is a tall shelf with an orange on the top shelf and an apple on a lower shelf, with a ladder leaning against the shelves.

The PowerPoint has been converted into an updated version based on animated videos, featuring 12 short stories, each with an accompanying voiceover and subtitles. These stories are easily adaptable to different languages through modifications to the voiceover and subtitles. Subtle animations of characters and scenes, along with a soft background soundtrack, have been introduced to engage children. Additionally, certain task designs have been refined to add depth and improve visual comprehension as seen in figure 3.

Figure 3

Updated Second-Order Belief Task

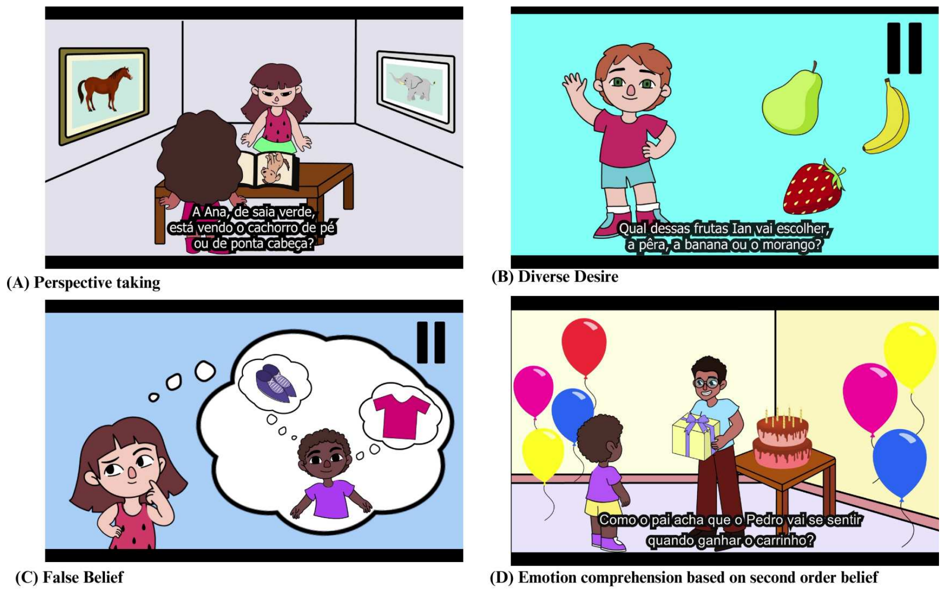
Note. The image shows the evolution of a second-order belief task, illustrating the changes made to improve visual comprehension and engagement.

During assessments, the examiner plays these videos on a computer and captures the children's responses digitally on a virtual form. If needed, the examiner may repeat the question or offer further clarification to ensure the children's understanding before documenting their answers. Target questions are displayed in the stories, and the examiner pauses the videos to record the children's answers in the digital form.

Responses to questions are assessed using two types of scoring scales, depending on the nature of the task. Simple, binary questions responses are scored on a dichotomous scale from 0 to 1, where '0' indicates an incorrect or inadequate response, and '1' signifies a correct or satisfactory response. This scale is used for questions that have clear, definitive answers. Figure 4 presents the tasks for the ToM framework within the instrument and the target questions.

Figure 4

Examples of ToM Tasks within the ToMT



Note. Examples of target questions that have dichotomous answers. (A) Perspective-taking: Is Ana seeing the dog up or down? (B) Diverse Desire: During snack time, which of these 3 fruits will Ian choose? (C) False belief task: Does Ana think Pedro is thinking there are shoes or clothes inside the bag? (D) How does the father think Pedro will feel when he receives the car?

On the other hand, some tasks are designed to allow for more nuanced responses or require the respondent to demonstrate a degree of understanding or skill. For these tasks, responses are evaluated on a three-point scale, ranging from 0 to 2. A score of '0' is given for responses that fail to meet the basic requirements in the interpretation of the situation. A score of '1' indicates a response that meets the basic criteria but lacks depth or fullness. A score of '2' is awarded to responses that not only meet the criteria but also exhibit a higher level of understanding. Examples of the tasks and target questions are shown in figure 5.

Figure 5

Examples of NUC tasks within the ToMT

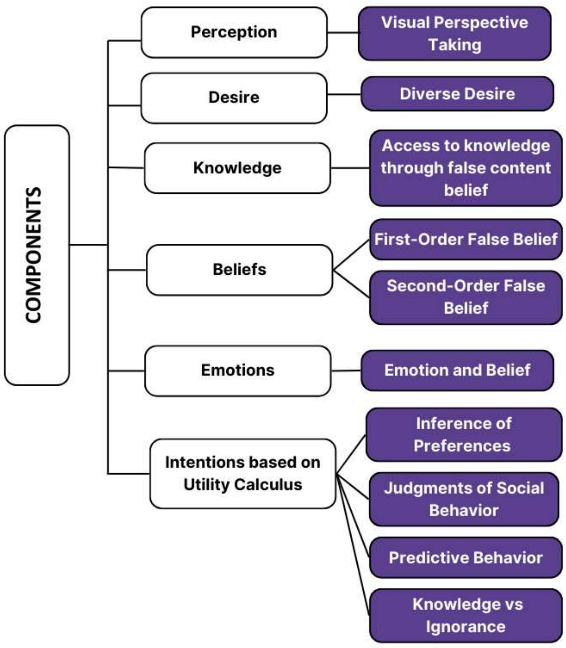
Note. (E) Preferences based on Utility calculus: Why did Yellow choose to walk more to get the strawberries? (F) Judgements of Social Behavior: Why do you think Caio did not want to help Pedro? (G) Predict Behavior: Before opening the door and seeing the fruits on the shelf, which fruit do you think the boy most wants to eat? (H) Inference of Knowledge vs Ignorance: Which of the two do you think was eating the green fruit for the first time, the boy or the girl? Why?

The ToMT is a computer-based tool for assessing ToM and NUC in 3 to 8-year-old children, inspired by the work of Wellman and Liu, (2004) and Jara-Ettinger et al. (2016b). It includes 12 video-animated tasks that present characters and narratives related to mental states across different sub-abilities, aiming to engage the target age group. The tasks have a total duration of approximately 15 minutes.

The ToMT comprises six main components. For the ToM framework, the components include Perception, Desire, Knowledge, Belief, and Emotions. The NUC tasks fall within the component of intentions and incorporate elements of the utility calculus framework, based on the experiments of Jara-Ettinger (2016a). Figure 5 presents the distribution of the components and subcomponents.

Figure 6

Components of the Theory of Mind Tasks (ToMT)



The ToM sub-components include Visual Perspective-taking I and II, Diverse Desire, Access to Knowledge, False Content Belief, First and Second-order False Belief, and Emotion and Belief. The NUC sub-components encompass Inference of Preference, Judgment of Social Behavior, Predictive Behavior, Knowledge vs Ignorance.

5.2 Content Validity Analysis of ToMT

5.2.1 Qualitative Experts Panel Analysis

Results from the qualitative expert's analysis are presented in the following. All items were recognized as relevant, suggesting they align well with the intended concepts. Additionally, experts proposed the inclusion of a new item within the Visual Perspective Taking sub-category. For item clarity, it was recommended to refine the animation design to aid children's understanding of costs and rewards. The table 7 provides a summary of the feedback received from the experts' qualitative review.

Table 7

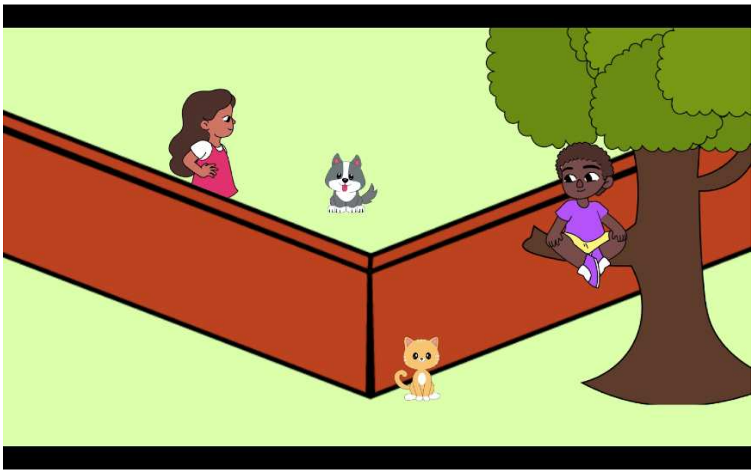
Expert Recommendations for ToMT Modifications

Tasks	Suggested Modifications	Purpose
Visual Perspective-Taking I and II	Add a new item to the category.	To more accurately capture this sub-component.
Inference of Preference Based on Utility Calculus	Increase the effort required to obtain the fruit.	To make the inferences of costs and rewards more pronounced.
Predict Behavior Based on Utility Calculus	Increase the effort required to reach the preferred object.	To underline the challenge posed by the physical effort needed to obtain the preferred object.
Judgments on Social Behavior	Omit color mentions; alter the characters' clothing.	To divert focus from visual descriptions to the core messages and choices of the narrative.

Based on the experts' suggestions, a new perspective-taking task was added as presented in Figure 7. Initially, we had one task covering Perspective Taking I and II. After the experts' review, they recommended adding another task on Perspective Taking I.

Figure 7

New Visual Perspective Task (ToMT)

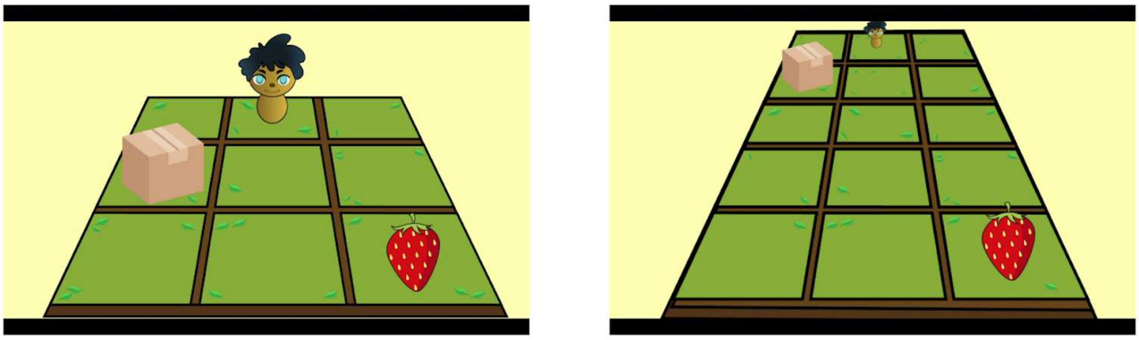


Note. The narrative following the target questions: Pati and Pedro were playing in the park. Pati was behind a wall, and close to her, there was a dog. On the other side of the wall, there was a cat. Pedro, on the other hand, climbed a tree and sat on a high branch. Now, I'm going to ask some questions, okay? Can Pati see the cat? And Pedro, can he see the cat? Can Pati see the dog? And Pedro, can he see the dog?

In the revised scenario design as shown in Figure 8, emphasis was placed on elucidating the concept of utility calculus to children by precisely delineating the efforts required to achieve specific objectives. The narrative was therefore refined to more clearly underscore this concept: "This is Yulu. He is hungry and desires to eat a fruit. At his home, he has two options: a strawberry that is far from him and another fruit that is nearby, encased within a box. Although he knows there is a fruit inside the box, its identity is unknown to him. Look! He now decides to go for the fruit he wants to eat. He selects the strawberry, despite it being further away. Why do you think Yulu preferred to undertake the longer journey to get the strawberry? Yulu's choice to pursue the distant strawberry over an accessible, unidentified fruit exemplifies his willingness to expend additional effort.

Figure 8

Inference of Preference task Based on NUC (ToMT)

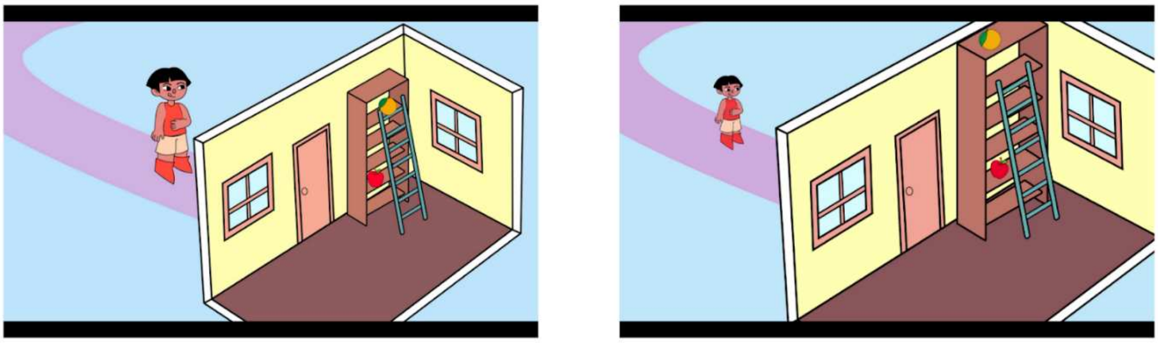


Note. The first figure shows a grid of 3x3, while the second figure shows a modified version of 5x3 to emphasize the distance and the effort required to reach the preferred object.

For the item focusing on predicting social behavior, the proportions of characters and obstacles were adjusted to underscore the challenges and costs within the narrative more effectively. These alterations make the animation more illustrative, highlighting the effort needed to achieve a goal, as shown in Figure 9.

Figure 9

Predict Behavior Task based on NUC (ToMT)



Note. Comparative visualization of scenario adjustments. The left panel illustrates the original animation with a proportionally sized character, making the orange within easy reach. The right panel shows the updated animation, with a comparatively smaller character and a higher-placed orange, indicating a greater effort needed to achieve the goal.

The initial portrayal did not effectively illustrate the challenge of obtaining the orange, considering the character's discomfort with heights. This inconsistency was addressed in the revised animation by resizing the character to be smaller and placing the orange higher. These changes enhance the visual representation to more accurately depict the narrative's portrayal of the character's apprehension of heights.

In the item judgment of social behavior, the narrative's focus was shifted from character descriptions based on attire to avoid potential confusion and cater to the reader's ability to perceive color or their familiarity with specific visual details.

In Figure 10, the image on left, to enhance comprehension of the characters, the narrative describes the color of the clothes they are wearing: Pedro, who is wearing a white shirt, is at home with two friends. He is hungry and decides to go to the kitchen to grab a fruit. He really wants to eat a banana, but it is on the top shelf, out of his reach. So, Pedro asks his friends, Ian, who is in a blue shirt, and Caio, who is in a red one, to get the banana for him. Both friends respond with a 'no' and do not help him. Why do you think Ian, who is dressed in blue, said no? Do you know why Caio, in the red shirt, said no? Which friend do you think should have helped to get the banana? Why? On the right image, the revised version presents the altered characters' genders to avoid using color names in the scene description to identify the characters.

Figure 10

Judgment of Social Behavior Task (ToMT)

Note. Comparative visualization of scenario adjustments.

Expert feedback led to task modifications, such as adding visual perspective-taking items and adjusting narratives for utility calculus clarity. These revisions aim to better align tasks with developmental goals, pending validation through future research.

5.2.2 Quantitative Experts Panel Analysis

In the content validation process of the ToMT, a panel of three experts was convened to evaluate each item for clarity, relevance, and pertinence. The Content Validity Coefficient (CVC) was employed in the item analysis to assess each task's clarity, pertinence, relevance, and potential bias (Hernandez-nieto, 2002). According to predefined criteria, a CVC score above 0.80 is considered acceptable, indicating adequate content validity. The results of this analysis are presented in Table 8.

Table 8

Content Validity Coefficient (CVC) Scores for the ToMT

	Item	Bias	Clarit	Pertin	Relev
ToM Tasks	Visual perspective-taking I	0.04	0.896	0.963	0.963
	Visual perspective-taking II	0.04	0.896	0.963	0.963
	Diverse desires	0.04	**0.763**	0.963	0.830
	Knowledge & False Content Belief	0.04	**0.763**	0.963	0.963
	First-Order and Second-Order False Belief	0.04	0.896	0.963	0.963
	Second-Order False Belief	0.04	0.963	0.963	0.963
	Understanding emotions based on second-order false belief	0.04	0.830	0.963	0.963
	Emotion comprehension	0.04	0.896	0.963	0.963
NUC Tasks	Inference of Preference Based on Utility Calculus	0.04	0.896	0.963	0.963
	Inference of Preference Based on Utility Calculus	0.04	0.830	0.963	0.963
	Judgments on Social Behavior	0.04	**0.763**	0.963	0.963
	Knowledge vs Ignorance	0.04	**0.696**	0.896	0.963
CVC Total			**0.84**	**0.96**	**0.95**

The CVC scores for ToM and NUC tasks showed variations across categories. All tasks had a consistent bias score of 0.04. Clarity scores ranged from 0.696 to 0.963, indicating a need for improvement in how some tasks are communicated. Pertinence and relevance scores were high across the board, with most tasks scoring 0.963, demonstrating a strong alignment with theoretical constructs. The cumulative CVC totals were 0.84 for clarity, 0.96 for pertinence, and 0.95 for relevance, reflecting the overall content validity of the tasks.

5.3 Pilot with Target Population

5.3.1 Descriptive Analysis of Pilot Study: Demographic Characteristics and Assessment Scores

The pilot study comprised 10 children, with a mean age of 6.3 years (SD = 1.57), ranging from 3 to 8 years old. The median age of participants was 6 years. Of these participants, the majority were male (n = 8), with a smaller representation of female participants (n = 2), as presented in table 9. Children were assessed with the ToMT and Nepsy subtests two core areas of cognitive and emotional assessment were reported: Theory Emotional Recognition (ER) and ToM.

Table 9

Pilot Study Descriptive Statistics: Demographic Characteristics and Assessment Scores

	Mean	Median	SD	Min	Max
Age	6.3	6	1.57	3	8
ToMT	27.1	27.5	4.23	17	31
NEPSY II ER (raw score)	19.5	20	5.7	8	28
NEPSY II ToM (raw score)	17.9	17	4.09	10	25

For the ToMT assessment, the mean score achieved by participants was 27.1 (SD = 4.23), with scores ranging from 17 to 31. In the Nepsy ER component, the children displayed a mean score of 19.5 (SD = 5.7), with the lowest and highest scores being 8 and 28, respectively. The Nepsy ToM assessment yielded a mean score of 17.9 (SD = 4.09), with participant scores spanning from 10 to 25.

5.3.2 Exploratory Analysis of ToMT

This study employed Spearman correlation analysis to explore the relationship between ToMT and the NEPSY subtests for ER and ToM. Table 10 presents the matrix correlation for Age, ToMT, NEPSY II subtest of ToM and NEPSY II subtest of ER.

Table 10

Spearman Correlation Matrix for Age, ToMT, ToM, and ER

		Age	ToMT	Nepsy ToM	Nepsy ER
Age	Spearman's rho	—			
	p-value	—			
ToMT	Spearman's rho	0.888***	—		
	p-value	< .001	—		
Nepsy ToM	Spearman's rho	0.699*	0.522	—	
	p-value	0.025	0.122	—	
Nepsy ER	Spearman's rho	0.757*	0.671*	0.621	—
	p-value	0.011	0.034	0.055	—

Note. The use of asterisks (*) denotes the level of significance, with a single asterisk (*) indicating a p-value less than 0.05, and three asterisks (***) indicating a p-value less than 0.001.

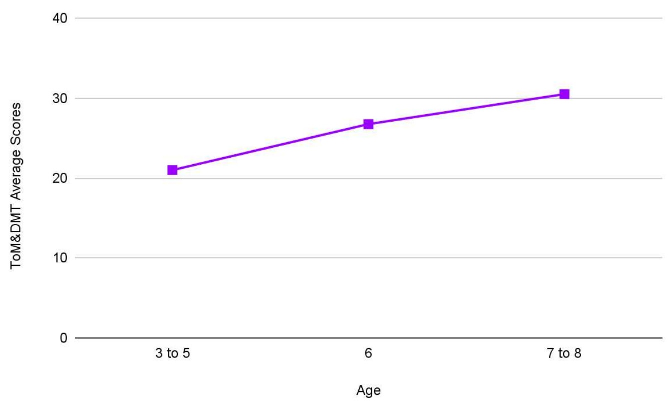
In analyzing the correlations among ToMT with Nepsy subtests ER and ToM, Spearman's rank-order correlation was utilized. A significant positive correlation was observed between the children's age and their ToMT scores (Spearman's rho = 0.888, $p < .001$), indicating a strong association where older children tended to have higher ToMT scores.

Between the assessment scores, the ToMT scores were moderately correlated with NEPSY ER scores (Spearman's rho = 0.671, $p = 0.034$). However, the correlation between ToMT scores and NEPSY ToM scores was found to be not statistically significant (Spearman's rho = 0.522, $p = 0.122$).

Figure 11 presents the scores analysis across age groups: 3-5 years, 6 years, and 7-8 years. The Children aged 3 to 5 years exhibited an average ToMT score of 21. This average score increased to 26.75 for children aged 6 years, and further to 30.5 for children within the 7 to 8-year age bracket.

Figure 11

ToMT Scores Across Age Groups



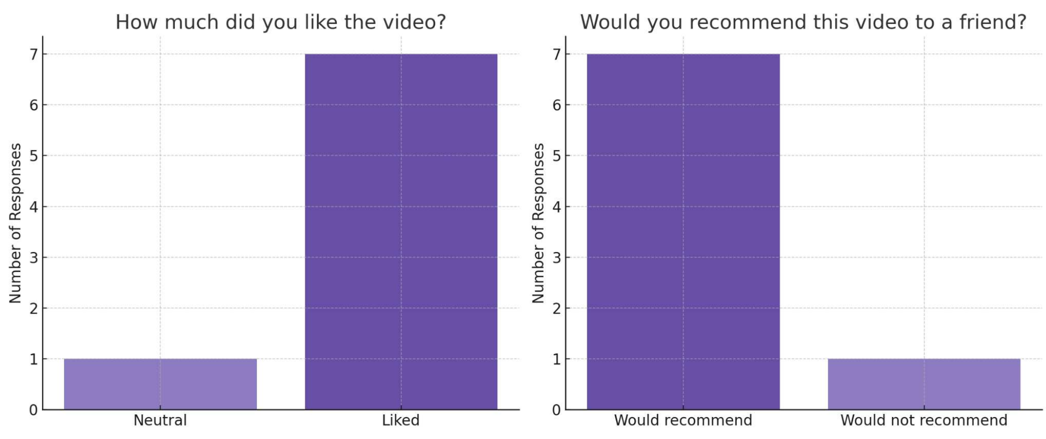
This analysis suggests a trend of increasing cognitive abilities with age, as indicated by ToMT scores.

5.3.3 Feedback on ToMT from Target Population

In the analysis of feedback from the target population regarding ToMT, a feedback form was utilized to obtain insights into children's satisfaction. Eight children completed the form. Figure 12 presents the results, indicating a predominantly positive reception of the video animation, with the majority of respondents expressing enjoyment and a willingness to recommend it to friends. Only one participant responded neutrally regarding enjoyment, and one expressed a reluctance to recommend the video.

Figure 12

Children's Feedback on ToMT



In qualitative responses, children generally found the video animation engaging and had minimal challenges in recalling specific details. No suggestions for enhancement were detected, indicating overall satisfaction with the content and presentation.

6. DISCUSSION

This study aimed to develop a computer-based, video-animated tool for assessing ToM abilities in children aged 3 to 8 years. The process involved a literature review, development of the narrative, creation of cartoons, and animated stories. Furthermore, a content analysis was conducted by judges, and a pilot study was carried out with the target audience.

6.1 Development of ToMT

Based on principles of Developmental Psychology and Cognitive Science, this assessment tool aims to provide a comprehensive measure to assess a broad spectrum of sub-components related to the developmental trajectory of ToM. Recent literature review identified gaps in ToM assessment tools, traditional single-task measures fail to capture ToM's varied cognitive and affective aspects, necessitating more comprehensive evaluation methods for understanding and supporting children's ToM development (Beaudoin et al., 2020; Fu et al., 2023; Wellman, 2018).

The development of the ToMT tool employs a multidimensional approach integrating 6 components: intentions, perceptions, desires, knowledge, beliefs, emotions, within 10 sub-components. Moreover, ToMT assesses the development of ToM across various developmental stages in childhood. The ToM framework within the test based its tasks in Wellman and Liu scale (2004), and also draws on studies involving second-order false belief tasks (Arslan et al., 2017) and visual perspective levels I and II (Flavell, 1991).

In the literature review by Beaudoin et al. (2020), the Wellman and Liu ToM scale is presented as a comprehensive measure that assesses multiple aspects of ToM abilities through direct testing. It is associated with 183 articles, amounting to 22.0% of the total number of studies (830) included in the review. This extensive utilization underscores its relevance and importance in research related to cognitive development and social cognition, evidencing its significant role in the field.

ToM and NUC are fundamental components that collectively provide a comprehensive understanding of social cognition (Gergely & Csibra , 2003; Jara-Ettinger et al., 2020). People inherently possess intuitive psychology, a capacity involving the comprehension of the invisible mental states underlying observable actions (Liu, 2020). A growing body of research suggests that children's Decision-making begins with grasping goals and efficiency, progressing to employing statistical reasoning for preferences. This progression entails learning about costs, rewards, and others' perspectives, ultimately enhancing their ToM (Baker et al., 2017; Biro et al., 2013; Jara-Ettinger et al., 2016b; Liu, 2020; Lucas et al., 2014). The integration of ToM and NUC tasks within an assessment tool aims to provide a comprehensive measure of children's social cognition. This innovative tool addresses existing research gaps (Beaudoin et al., 2020; Fu et al., 2023) by utilizing a diverse range of components to capture how children infer other agents' mental states and predict agents' actions.

Variability in ToM assessment methods complicates data comparison and undermines reliability (Fu et al., 2023). Computer-based tools in psychological evaluations enhance standardization and psychometric robustness, reducing variability (Casaletto & Heaton, 2017). These tools provide standardized procedures, real-time insights, cost-effectiveness, and quicker corrections (Casaletto & Heaton, 2017; Schatz & Browndyke, 2002). This approach is necessary for comparing results, understanding diverse developmental patterns, and identifying social cognition characteristics across cultures and individuals (Casaletto & Heaton, 2017; Galindo-Aldana et al., 2018).

Integrating technology into psychological practices demands balancing innovation with ethical responsibility. The introduction of technology has transformed psychological assessments, highlighting the need for strong ethical standards, professional integrity, and online individual rights protection (Muniz et al., 2021). Responsible digital tool usage ensures participant confidentiality and dignity. Given increased data privacy concerns in the digital realm, robust safeguards are essential to maintain the individual's safety and trust, crucial for responsibly adapting technology in psychological practices (Muniz et al., 2021).

The ToMT incorporates digital technology into psychological assessments, emphasizing adherence to ethical standards. Examiners are required to be psychologists or psychology graduate students with supervision trained in its theoretical constructs and ethical guidelines to ensure adherence to the ethical protocols. This tool is applicable in varied settings such as clinics, schools, and for remote use, facilitating a dynamic assessment process. By utilizing animated videos presented on electronic devices, it engages children interactively. Capable of operating offline, ToMT offers flexibility and accessibility, even in regions with limited internet connectivity, underscoring its utility and ethical consideration in diverse contexts.

Incorporating visual aids in ToM assessments benefits children with limited verbal abilities, fostering inclusivity. This approach aligns with research advocating tailored cognitive assessment methods (Fu et al., 2023). For children with ASD and ADHD, who often face challenges in social cognition and communication, visual aids are crucial. These aids provide concrete representations of abstract concepts, aiding comprehension for individuals with ASD or ADHD (Fletcher-Watson et al., 2014; Pineda-Alhucema et al., 2018).

Recent studies, supported by rigorous evidence, highlight the effectiveness of socio-emotional interventions in boosting children's social cognition skills (Hoffman et al., 2016). The ToMT assessment tool is promising for its ease of use in diverse settings, suggesting potential for facilitating data collection in large-scale experimental studies and longitudinal research. This tool can facilitate the measurement of the efficacy of social cognition interventions over different times, offering insights into efficient interventions that enhance children's social cognition abilities.

Its adaptability, achieved through the use of modifiable narratives and voiceovers, ensures broad applicability. This flexibility allows the tool to be tailored to meet the specific linguistic and cultural needs of diverse populations. Cross-cultural validation demands precision, statistical rigor, and cultural insight, ensuring tools maintain integrity and relevance, enabling accurate comparisons across cultures. This involves confirming universal applicability and uniform measurement of constructs, despite cultural differences (Casalleto et al., 2017; Gjersing et al., 2010).

In conclusion, the ToMT's innovative framework presents a promising direction in assessing children's social cognition. However, the assessment requires validation through psychometric tests for reliability and validity, plus transcultural longitudinal studies on diverse populations, to verify its adaptability in different assessment settings.

6.2 Content Validity of ToMT

The ToMT demonstrate potential and areas needing refinement, as revealed by qualitative and quantitative analyses. Feedback from the qualitative panel highlighted the necessity for task modifications to better simulate real-life scenarios.

A growing body of studies in Cognitive Science suggests an interplay between intuitive physics and ToM development (Baker et al., 2017; Gergely & Csibra, 2003; Jacobs et al., 2021; Pelz et al., 2020; Tenenbaum et al., 2006; Ullman et al.2017; Yildirim et al., 2019). The task modifications aim to clarify the conceptual connections between perceived effort and reward for children, enhancing the accuracy of measuring their reasoning of intentions within a utility maximization framework. Furthermore, the ToMT may offer additional evidence on the interaction between intuitive physics and psychology with ToM development.

Understanding the transition from a basic to a more nuanced grasp of perspectives in early childhood ToM development, Level I and II visual perspective-taking are identified as pivotal milestones (Rakoczy, 2022). Introducing an additional visual perspective-taking task can enrich the comprehension of ToM's developmental trajectory, illuminating the evolution of children's ability to understand others' viewpoints. Furthermore, this enhancement promises to aid future studies in evaluating the internal consistency of such tasks, thereby augmenting our understanding of the instrument's reliability by verifying that its components consistently assess the same construct across diverse populations, a critical factor in tool evaluation highlighted by Souza et al. (2014).

Following expert recommendations and Gibson et al. (2017)'s insights on the variability of color significance across cultures and developmental stages, this modification aims to reduce potential bias from children's variable color perception. By omitting color references in the Inference of Social Behavior task, the adjustment seeks to enhance the task's accessibility and relevance across diverse cultural and developmental contexts, ensuring it remains unaffected by perceptual differences.

The analysis of CVC scores for the ToM and NUC Tasks offers insights into both its conformity with the target constructs and opportunities for improvement. Notably, high scores in pertinence (0.96) and relevance (0.95) suggest the tool's strong alignment with the constructs it aims to measure, thereby demonstrating its relevance across evaluators (Hernandez-Nieto, 2002). However, clarity scores reveal variability, ranging from 0.696 to 0.963, signaling differences in the tasks' comprehensibility. Particularly, tasks like Diverse Desires, Knowledge & False Content Belief, Judgments of Social Behavior and knowledge vs ignorance based on Utility Calculus highlight a clarity gap, emphasizing the need for refinement to improve understanding. The comparison of the overall clarity CVC score (0.84) with those for pertinence (0.96) and relevance (0.95) identifies an opportunity for refinement, which is essential for enhancing the tool's overall utility.

Translating complex laboratory experiments into tools like ToMT poses significant challenges, notably in tasks such as "Judgment of Social Behavior" and "Diverse Desires," inspired by Jara-Ettinger et al. (2016a) and Repacholi and Gopnik (1997). These tasks, which exhibit lower scores on the CVC for clarity, underscore the difficulty of preserving the essence of original experiments while ensuring they are engaging and comprehensible. The "Inference of knowledge vs ignorance" task, receiving the lowest clarity score at 0.696, emphasizes the necessity for adjustments to enhance clarity and effectively communicate complex cognitive concepts to a young audience.

6.3 Pilot with Target Population

This study demonstrates a significant strong correlation between age and ToMT scores (rho = 0.888, $p < .001$), indicating that as children age, their ability to understand others' mental states and their proficiency in NUC concurrently improve.This finding aligns with Wellman and Liu's (2004) theory on the gradual maturation of ToM. Additionally, by kindergarten age, studies show that children exhibit advanced social reasoning abilities, recognizing that actions are taken to maximize outcomes by balancing benefits against costs (Jara-Ettinger, 2016a). This observation underscores the interconnected growth of cognitive skills critical for social interaction and Decision-making in early childhood.

The strong correlation between ToMT and NEPSY Emotional Recognition (ER) scores (Spearman's rho = 0.671, $p = 0.034$) suggests a linkage between children's emotional understanding and ToM abilities. Studies analyzed in the review by Berggren et al. (2017) highlight this relationship, underscoring the critical role of emotional recognition in developing ToM, which is essential for interpreting social cues and making informed decisions. Enhanced emotional skills are associated with improved empathy. Regarding the performance related to age of the children in the ToMT, these preliminary findings highlight the need for further research with a larger sample size to comprehensively understand the various factors influencing cognitive development across age groups.

7. STUDY LIMITATIONS

The ToMT tool proposes an innovative assessment compared to traditional psychological assessments. However, this assessment does not capture the nuances of a mentalistic understanding such as irony, sarcasm, various lies, humor, and faux pas recognition. These facets are key to ToM development, crucial for understanding social cognition complexities. Furthermore, the NUC framework within the tool does not encompass all components for a comprehensive measurement of this construct, highlighting the need for a more inclusive revision.

In these preliminary exploratory findings, no indication of an association was found between the scores obtained from the NEPSY ToM subtest and the ToMT. The small sample size of 10 children limits the robustness and generalizability of these findings, emphasizing the need for caution in interpretation. Moreover, the NEPSY ToM test features tasks that assess understanding of common sayings and social situations, designed for children up to 16 years old. This aspect might have resulted in a reduced ability to differentiate among the younger participants in our study.

Additionally, a limitation of the computerized tool is that the scores are computed manually through a digital form. The lack of an automated scoring system may introduce potential errors and inconsistencies. Implementing a software solution that can accurately compute these scores automatically would enhance the reliability and efficiency of the assessment process.

8. CONCLUSION

This study presents a comprehensive computer-based animated tool for evaluating ToM in children. By providing a detailed framework and engaging content, it aims to offer deeper insights into social cognitive development. This research lays the groundwork for future studies to further investigate social cognition and developmental patterns across diverse populations.

In light of the increasing integration of Artificial Intelligence (AI) in research methodologies, future studies should consider adapting the ToMT into a platform that utilizes AI for more precise and accurate measurement of social cognition constructs. Moreover, the use of Virtual Reality could enhance children's comprehension and engagement in a simulation of the real world. Additionally, the development of version B of the tool for follow-up assessments is crucial to facilitate longitudinal studies, enabling the comparison of baseline and follow-up data to track the developmental trajectory of ToM skills over time.

Collaborating with researchers experienced in laboratory experiments related to the construct ensures task consistency within the experimental setting. Continuous testing, feedback, and adjustments are recommended. Recognizing that initial attempts might not fully capture real-world complexities, this iterative process is crucial for refining and enhancing the tools to improve comprehension for the target population.

Additionally, incorporating missing components of ToM and NUC, especially those related to understanding non-literal communication and complex Decision-making scenarios, will enhance the tool's comprehensiveness. Future research should also focus on evaluating the psychometric properties of the ToMT to ensure its validity and reliability. Cross-cultural validation and adaptation of the ToMT will also be necessary to ensure its applicability and effectiveness in diverse cultural settings, paving the way for global research on cognitive development. These additions and modifications are aimed at addressing the current study's limitations while setting a robust foundation for future research endeavors.

9. REFERENCES

Aboulafia-Brakha, T., Christe, B., Martory, M. D., & Annoni, J. M. (2011). Theory of mind tasks and executive functions: A systematic review of group studies in neurology. *Journal of Neuropsychology, 5*(1), 39-55. https://doi.org/10.1348/174866410X500350

Ahmed, F. S., & Miller, S. L. (2011). Executive function mechanisms of theory of mind. *Journal of Autism and Developmental Disorders, 41*(5), 667-678.

Alexandre, N. M. C., & Coluci, M. Z. O. (2011). Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. *Ciência & Saúde Coletiva, 16*, 3061-3068.

Andreou, M., & Skrimpa, V. (2020). Theory of mind deficits and neurophysiological operations in autism spectrum disorders: A review. *Brain Sciences, 10*(6), 393. https://doi.org/10.3390/brainsci10060393

Argollo, N., Bueno, O. F. A., Shayer, B., Godinho, K., Abreu, K., Durán, P., … Seabra, A. G. (2009). Adaptação transcultural da Bateria NEPSY - avaliação neuropsicológica do desenvolvimento: estudo-piloto. *Avaliação Psicológica, 8*(1), 59–75.

Arslan, B., Taatgen, N. A., & Verbrugge, R. (2017). Five-year-olds' systematic errors in second-order false belief tasks are due to first-order theory of mind strategy selection: A computational modeling study. *Frontiers in Psychology, 8*, 275. https://doi.org/10.3389/fpsyg.2017.00275

Aschersleben, G., Hofer, T., & Jovanovic, B. (2008). The link between infant attention to goal-directed action and later theory of mind abilities. *Developmental Science, 11*(6), 862–868. https://doi.org/10.1111/j.1467-7687.2008.00736.x

Astington, J. W. (2006). The developmental interdependence of theory of mind and language. In *Roots of Human Sociality* (1st ed., pp. 28). Routledge. https://doi.org/10.4324/9781003135517

Astington, J. W., Pelletier, J., & Homer, B. (2002). Theory of mind and epistemological development: The relation between children's second-order false-belief understanding and their ability to reason about evidence. *New Ideas in Psychology, 20*(2–3), 131-144. https://doi.org/10.1016/S0732-118X(02)00005-3

Astington, J., & Jenkins, J. (1999). A longitudinal study of the relation between language and theory of mind development. *Developmental Psychology, 35*(5), 1311-1320. https://doi.org/10.1037//0012-1649.35.5.1311

Baillargeon, R. (2004). Infants' Physical World. *Current Directions in Psychological Science, 13*(3), 89-94. https://doi.org/10.1111/j.0963-7214.2004.00281.x

Baillargeon, R. (1994). How do infants learn about the physical world? *Current Directions in Psychological Science, 3*(5), 133–140. https://doi.org/10.1111/1467-8721.ep10770614

Baillargeon, R., Spelke, E., & Wasserman, S. (1985). Object permanence in 5-month-old infants. *Cognition, 20*, 191-208.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition, 113*(3), 329-349. https://doi.org/10.1016/j.cognition.2009.07.005

Baker, C., Saxe, R., & Tenenbaum, J. (2011). Bayesian ToM: Modeling joint belief-desire attribution. In *Proceedings of the annual meeting of the Cognitive Science society* (Vol. 33, No. 33).

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires, and percepts in human mentalizing. *Nature Human Behaviour, 1*(4), Article 0064. https://doi.org/10.1038/s41562-017-0064

Banerjee, R., Watling, D., & Caputi, M. (2011). Peer relations and the understanding of faux pas: Longitudinal evidence for bidirectional associations. *Child Development, 82*, 1887–1905. https://doi.org/10.1111/j.1467-8624.2011.01669.x

Baron-Cohen, S. (1999). Evolution of a theory of mind. In M. Corballis & S. Lea (Eds.), *The Descent of Mind: Psychological Perspectives on Hominid Evolution* (pp. 1-31). Oxford University Press.

Baron-Cohen, S. (2000). Theory of mind and autism: A review. *International Review of Research in Mental Retardation, 23*, 169-184. https://doi.org/10.1016/S0074-7750(00)80010-5

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

Bartsch, K. (1996). Between desires and beliefs: Young children's action predictions. *Child Development, 67*(4), 1671-1685. https://doi.org/10.2307/1131719

Bartsch, K., & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60*, 946-964.

Beaudoin, C., Leblanc, É., Gagner, C., & Beauchamp, M. H. (2020). Systematic review and inventory of theory of mind measures for young children. *Frontiers in Psychology, 10*, 2905. https://doi.org/10.3389/fpsyg.2019.02905

Berggren, S., Fletcher-Watson, S., Milenkovic, N., Marschik, P. B., Bölte, S., & Jonsson, U. (2017). Emotion recognition training in autism spectrum disorder: A systematic review of challenges related to generalizability. *Developmental Neurorehabilitation, 20*(3), 141–154. https://doi.org/10.1080/17518423.2017.1305004

Biro, S., Yu, C., & Baillargeon, R. (2013). Infants' strategic use of information about personal preferences in social Decision-making. *Developmental Psychology, 49*(8), 1640. https://doi.org/10.1037/a0032289

Bora, E., & Pantelis, C. (2016). Meta-analysis of social cognition in attention-deficit/hyperactivity disorder (ADHD): Comparison with healthy controls and autistic spectrum disorder. *Psychological Medicine, 46*, 699–716. https://doi.org/10.1017/S0033291715002573

Campos, L. F. L. (2008). *Métodos e técnicas de pesquisa em psicologia*. Alínea.

Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: The role of prosocial behavior. *Developmental Psychology, 48*(1), 257-270. https://doi.org/10.1037/a0025309

Carey, S. (2009). The origin of concepts. *Oxford University Press*.

Carey, S., & Spelke, E. S. (1994). Domain-specific knowledge and conceptual change. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169–200). Cambridge University Press.

Casaletto, K. B., & Heaton, R. K. (2017). Neuropsychological assessment: Past and future. *Journal of the International Neuropsychological Society, 23*(9-10), 778-790. https://doi.org/10.1017/S1355617717001060

Cavadel, E. W., & Frye, D. A. (2017). Not just numeracy and literacy: Theory of mind development and school readiness among low-income children. *Developmental Psychology, 53*(12), 2290–2303. https://doi.org/10.1037/dev0000409

Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2014). *Testagem e avaliação psicológica: Introdução a testes e medidas* (8ª ed.). AMGH.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Coluci, M. Z. O., Alexandre, N. M. C., & Milani, D. (2015). Construção de instrumentos de medida na área da saúde. *Ciência & Saúde Coletiva, 20*, 925-936.

Dehaene, S., Izard, V., Pica, P., & Spelke, E. (2006). Core knowledge of geometry in an Amazonian indigenous group. *Science, 311*(5759), 381–384. https://doi.org/10.1126/science.1121739

Delgado-Rico, E., Carretero-Dios, H., & Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology Espana, 12*(3), 449-460.

Dennett, D. C. (1989). *The intentional stance*. MIT Press.

Derksen, D. G., Hunsche, M. C., Giroux, M. E., Connolly, D. A., & Bernstein, D. M. (2018). A systematic review of theory of mind's precursors and functions. *Zeitschrift für Psychologie, 226*(2). https://doi.org/10.1027/2151-2604/a000325

Dias, N. M., Batista, L. S., & Mecca, T. P. (2020). Evidências de validade do Teste de Teoria da Mente para Crianças: relação com funções executivas. *Neuropsicologia Latinoamericana, 12*(2).

Dixson, H. G. W., Komugabe-Dixson, A. F., Dixson, B. J., & Low, J. (2017). Scaling Theory of Mind in a small-scale society: A case study from Vanuatu. Child Development. https://doi.org/10.1111/cdev.12919

Ebert, S. (2020). Theory of mind, language, and reading: Developmental relations from early childhood to early adolescence. *Journal of Experimental Child Psychology, 191*, 104739. https://doi.org/10.1016/j.jecp.2019.104739

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307–314.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1-Level 2 distinction. *Developmental Psychology, 17*(1), 99-103.

Fletcher-Watson, S., McConnell, F., Manola, E., & McConachie, H. (2014). Interventions based on the theory of mind cognitive model for autism spectrum disorder (ASD). *Cochrane Database of Systematic Reviews*, (3), CD008785. https://doi.org/10.1002/14651858.CD008785.pub2

Fletcher, G. J. (1984). Psychology and common sense. *American Psychologist, 39*(3), 203–213. https://doi.org/10.1037/0003-066X.39.3.203

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. https://doi.org/10.1098/rstb.2002.1218

Fujita, N., Devine, R. T., & Hughes, C. (2022). Theory of mind and executive function in early childhood: A cross-cultural investigation. *Cognitive Development, 61*, 101150.

Fu, I.-N., Chen, K.-L., Liu, M.-R., Jiang, D.-R., Hsieh, C.-L., & Lee, S.-C. (2023). A systematic review of measures of theory of mind for children. *Developmental Review, 67*, 101061. https://doi.org/10.1016/j.dr.2022.101061

Galindo-Aldana, G., Meza-Kubo, V., Castillo-Medina, G., Ledesma-Amaya, I., Galarza-Del-Angel, J., Padilla-López, A., & Morán, A. L. (2018). Computer-based neuropsychological assessment: A validation of structured examination of executive functions and emotion. In D. Harris (Ed.), *EPCE 2018, LNAI 10906* (pp. 306–316). Springer International Publishing. https://doi.org/10.1007/978-3-319-91122-9_26

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences, 7*, 77–83. https://doi.org/10.1016/S1364-6613(02)00025-6

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences, 7*(7), 287-292. https://doi.org/10.1016/S1364-6613(03)00128-1

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. Cognition, 56(2), 165–193.

Gibson, E., Futrell, R., Jara-Ettinger, J., & Conway, B. R. (2017). Color naming across languages reflects color use. *Proceedings of the National Academy of Sciences, 114*(40), 10785-10790. https://doi.org/10.1073/pnas.1619666114

Gjersing, L., Caplehorn, J. R. M., & Clausen, T. (2010). Cross-cultural adaptation of research instruments: Language, setting, time and statistical considerations. *BMC Medical Research Methodology, 10*(13). https://doi.org/10.1186/1471-2288-10-13

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory.

Hamlin, J. K., Wynn, K., & Bloom, P. (2008). Social cognition in infancy: Evidence from human infants. *Cognition, 109*(2), 159-178.

Happé, F., & Frith, U. (2014). Annual research review: Towards a developmental neuroscience of atypical social cognition. *Journal of Child Psychology and Psychiatry, 55*(6), 553-577.

Hernández-Nieto, R. A. (2002). *Contributions to statistical analysis* (pp. 119). Universidad de Los Andes.

Heider, F. (1958). *The psychology of interpersonal relations.* John Wiley & Sons Inc. https://doi.org/10.1037/10628-000

Hofmann, S. G., Doan, S. N., Sprung, M., Wilson, A., Ebesutani, C., Andrews, L. A., ... & Harris, P. L. (2016). Training children's theory-of-mind: A meta-analysis of controlled studies. *Cognition, 150*, 200-212. https://doi.org/10.1016/j.cognition.2016.01.006

Imuta, K., Henry, J. D., Slaughter, V., Selcuk, B., & Ruffman, T. (2016). ToM and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology, 52*(8), 1192–1205. https://doi.org/10.1037/dev0000140

Jacobs, C., Lopez-Brau, M., & Jara-Ettinger, J. (2021). What happened here? Children integrate physical reasoning to infer actions from indirect evidence. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43, pp. 872).

Jara-Ettinger, J. J. (2016a). *The inner life of goals: Costs, rewards, and commonsense psychology* [Doctoral dissertation, Massachusetts Institute of Technology]. http://hdl.handle.net/1721.1/106433

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016b). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences, 20*(8), 589-604.

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015a). Children's understanding of the costs and rewards underlying rational action. *Cognition, 140*, 14-23.

Jara-Ettinger, J., Tenenbaum, J. B., & Schulz, L. E. (2015b). Not so innocent: Toddlers' inferences about costs and culpability. *Psychological Science, 26*(5). https://doi.org/10.1177/0956797615572806

Jara-Ettinger, J., Schulz, L. E., & Tenenbaum, J. B. (2020). The Naïve Utility Calculus as a unified, quantitative framework for action understanding. *Cognitive Psychology, 123*, Article 101334. https://doi.org/10.1016/j.cogpsych.2020.101334

Jara-Ettinger, J. (2019). Theory of mind as inverse reinforcement learning. *Current Opinion in Behavioral Sciences, 29*, 105-110. https://doi.org/10.1016/j.cobeha.2019.04.010

Kara, B., & Selcuk, B. (2021). The role of socioeconomic adversity and armed conflict in executive function, theory of mind, and empathy in children. *Child Psychiatry & Human Development*. Advance online publication. https://doi.org/10.1007/s10578-021-01155-4

Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A developmental neuropsychological assessment*. The Psychological Corporation.

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, 21(8), 1134-1140. https://doi.org/10.1177/0956797610376652

Lecce, S., & Devine, R. T. (2021). Theory of mind at school: Academic outcomes and the influence of the school context. *Infant and Child Development, e2274*. https://doi.org/10.1002/icd.2274

Lecce, S., Caputi, M., & Hughes, C. (2011). Does sensitivity to criticism mediate the relationship between theory of mind and academic achievement? *Journal of Experimental Child Psychology, 110*(3), 313-331. https://doi.org/10.1016/j.jecp.2011.04.011

Liu, S. (2020). *Nature and Origins of Intuitive Psychology in Human Infants*. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences. https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37365698

Liu, S., & Spelke, E. S. (2017). Six-month-old infants expect agents to minimize the cost of their actions. *Cognition, 160*, 35-42. https://doi.org/10.1016/j.cognition.2016.12.007

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science, 358*(6366), 1038-1041. https://doi.org/10.1126/science.aag2132

Lucas, C. G., Griffiths, T. L., Xu, F., Fawcett, C., Gopnik, A., Kushnir, T., Markson, L., & Hu, J. (2014). The child as econometrician: A rational model of preference understanding in children. *PLoS ONE, 9*(3), e92160. https://doi.org/10.1371/journal.pone.0092160

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development, 45*(2), 357-366. http://www.jstor.org/stable/1127956

Mayer, A., & Träuble, B. (2015). The weird world of cross-cultural false-belief research: A true- and false-belief study among Samoan children based on commands. Journal of Cognition and Development, 16(4), 650-665. https://doi.org/10.1080/15248372.2014.926273

Mayes, L. C., Klin, A., Tercyak, K. P., Jr., Cicchetti, D. V., & Cohen, D. J. (1996). Test-retest reliability for false-belief tasks. *Journal of Child Psychology and Psychiatry, 37*(3), 313-319.

Menezes, A. B. de C., Cruz, M. M. da S., Corrêa, H. V. V., & Brito, R. (2014). Perception and development of Theory of Mind reasoning in Asurini children, from Amazonia, Brazil. Revista de Etologia, 13(2).

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology, 24*, 603–613. https://doi.org/10.1348/026151005X55370

Mecca, T. P., Dias, N. M., Oliveira, P. V. D., & Muniz, M. (2018). Teste de teoria da mente para crianças: Validade de conteúdo. *Psico-USF, 23*, 393-407.

Muniz, M., Cardoso, L. M., Rueda, F. J. M., & Noronha, A. P. P. (2021). Desafios da Avaliação Psicológica para a Prática diante da Atuação Profissional Mediada pela Tecnologia de Informação. *Psico-USF, Bragança Paulista, n/e*, 2021, 9-19. http://dx.doi.org/10.1590/1413-8271202126nesp03

Oliveira, P., & Mecca, T. P. (2016). Ferramentas para avaliação de teoria da mente. In T. P. Mecca, N. M. Dias, & A. A. Berberian (Eds.), *Cognição Social: Teoria, Pesquisa e Aplicação* (pp. 138-151). São Paulo: Memnon.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255-258. https://doi.org/10.1126/science.1107621

Parsey, C. M., & Schmitter-Edgecombe, M. (2013). Applications of technology in neuropsychological assessment. *Clinical Neuropsychologist, 27*(8). https://doi.org/10.1080/13854046.2013.834971

Pasquali, L. (2009). Psicometria. *Revista da Escola de Enfermagem da USP, 43*, 992-999. https://doi.org/10.1590/S0080-62342009000500002

Pelz, M., Schulz, L., & Jara-Ettinger, J. (2020). The signature of all things: Children infer knowledge states from static images. In *Proceedings of the Annual Conference of the Cognitive Science Society*.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology, 39*(3), 437-471.

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: Knowledge, belief, and communication. *Journal of Child Psychology and Psychiatry, 30*(5), 689-700.

Peterson, C.C., & Wellman, H.M. (2019). Longitudinal Theory of Mind (ToM) Development From Preschool to Adolescence With and Without ToM Delay. *Child development.*

Peterson, C. C., Slaughter, V. P., & Paynter, J. (2007). Social maturity and theory of mind in typically developing children and those on the autism spectrum. *Journal of Child Psychology and Psychiatry, 48*(12), 1243-1250.

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition, 69*(1), 1–34.

Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition, 98*(2), 137–155. https://doi.org/10.1016/j.cognition.2004.11.005

Pineda-Alhucema, W., Aristizabal, E., Escudero-Cabarcas, J., Acosta-López, J. E., & Vélez, J. I. (2018). Executive function and theory of mind in children with ADHD: A systematic review. *Neuropsychology Review, 28*(3), 341–358. https://doi.org/10.1007/s11065-018-9381-9

Pica, P., Lemer, C., Izard, W., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*(5695), 499–503.

Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. *Nature Reviews Psychology, 1*, 223. https://doi.org/10.1038/s44159-022-00037-z

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*(1), 12-21. https://doi.org/10.1037/0012-1649.33.1.12

Reppold, C. T., Gurgel, L. G., & Hutz, C. S. (2014). O processo de construção de escalas psicométricas. *Avaliação Psicológica, 13*(2), 307-310.

Royston, J. P. (1983). A simple method for evaluating the Shapiro-Francia W' test of non-normality. *Journal of the Royal Statistical Society. Series D (The Statistician), 32*(3), 297-300. https://doi.org/10.2307/2987935

Ruffman, T., Taumoepeau, M., Perkins, C., & Gibbs, V. (2015). A longitudinal study of mental state talk in low-socioeconomic status mother–child dyads. *Journal of Child Language, 42*(5), 1058-1071. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4698072/.

Rule, J. S., Tenenbaum, J. B., & Piantadosi, S. T. (2020). The child as hacker. *Trends in Cognitive Sciences, 24*(11), 900-915. https://doi.org/10.1016/j.tics.2020.07.005

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-analysis of theory of mind and peer popularity in the preschool and early school years. Child Development, 86(4), 1159–1174. https://doi.org/10.1111/cdev.12372.

Sarmento-Henrique, R., Quintanilla, L., & Giménez-Dasí, M. (2019). The longitudinal interplay of emotion understanding, theory of mind, and language in the preschool years. *International Society for the Study of Behavioural Development, 44*(3). https://doi.org/10.1177/0165025419866907.

Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P. (2021). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin, 147*(3), 293–327. https://doi.org/10.1037/bul0000303

Schatz, P., & Browndyke, J. (2002). Applications of computer-based neuropsychological assessment. *Journal of Head Trauma Rehabilitation, 17*(5), 395-410. https://doi.org/10.1097/00001199-200210000-00008

Schuwerk, T., Vuori, M., & Sodian, B. (2014). Implicit and explicit theory of mind reasoning in autism spectrum disorders: The impact of experience. *Autism, 19*(4). https://doi.org/10.1177/1362361314526004

Scott, R. M., & Baillargeon, R. (2013). Do infants really expect agents to act efficiently? A critical test of the rationality principle. *Psychological Science, 24*(4), 466–474. https://doi.org/10.1177/0956797612457395

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology, 47*(5), 1239.

Shapiro, S. S., & Francia, R. S. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association, 67*(215).

Skerry, A. E., Carey, S. E., & Spelke, E. S. (2013). First-person action experience reveals sensitivity to action efficiency in prereaching infants. *Proceedings of the National Academy of Sciences of the United States of America, 110*(46), 18728–18733. https://doi.org/10.1073/pnas.1312322110

Souza, A. C. de., Alexandre, N. M. C., & Guirardello, E. de B. (2017). Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. Epidemiologia e Serviços de Saúde, 26(3), 649–659. doi:10.5123/s1679- 49742017000300022.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science, 14*(1), 29-56. https://doi.org/10.1016/0364-0213(90)90025-R

Spelke, E. S., Katz, G., Purcell, S. E., Ehrlich, S. M., & Breinlinger, K. (1994). Early knowledge of object motion: Continuity and inertia. *Cognition, 51*(2), 131-176. https://doi.org/10.1016/0010-0277(94)90013-2

Spelke, E. S. (2022). *What Babies Know: Core Knowledge and Composition Volume 1*. Oxford University Press. https://doi.org/10.1093/oso/9780190618247.001.0001

Spelke, E. S., Bernier, E. P., & Skerry, A. E. (2013). Core social cognition. In *Navigating the Social World* (Vol. 6, pp. 11–16). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199890712.003.0003

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Spelke, E., Lee, S. A., & Izard, V. (2010). Beyond core knowledge: Natural geometry. *Cognitive Science, 34*(5), 1-25. https://doi.org/10.1111/j.1551-6709.2010.01110.x

Stojnić, G., Gandhi, K., Yasuda, S., Lake, B. M., & Dillon, M. R. (2023). Commonsense psychology in human infants and machines. *Cognition*, 105406. https://doi.org/10.1016/j.cognition.2023.105406

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences, 10*(7), 309-318.https://doi.org/10.1016/j.tics.2006.05.009

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2007). Intuitive psychology: A task for Cognitive Science. *Trends in Cognitive Sciences, 11*(4), 125-132. https://doi.org/10.1016/j.tics.2007.01.005

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675-735. https://doi.org/10.1017/S0140525X05000129

Uekermann, J., Kraemer, M., Abdel-Hamid, M., Schimmelmann, B. G., Hebebrand, J., Daum, I., Wiltfang, J., & Kis, B. (2010). Social cognition in attention-deficit hyperactivity disorder (ADHD). *Neuroscience and Biobehavioral Reviews, 34*(5), 734-743. https://doi.org/10.1016/j.neubiorev.2009.10.009

Ullman, T. D., & Tenenbaum, J. B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *Annual Review of Developmental Psychology, 2*, 533-558. https://doi.org/10.1146/annurev-devpsych-021420-101405

Ullman, T. D., Spelke, E., Battaglia, P., & Tenenbaum, J. B. (2017). Mind games: Game engines as an architecture for intuitive physics. *Trends in Cognitive Sciences*, 21(9), 649. https://doi.org/10.1016/j.tics.2017.05.012

Ullman, T. D. (2015). On the nature and origin of intuitive theories: Learning, physics and psychology (Publication No. 1721.1/97788) [Doctoral dissertation, Massachusetts Institute of Technology]. MIT DSpace Repository. http://hdl.handle.net/1721.1/97788

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1-B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition, 30*(3), 239-277.

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition, 35*(3), 245-275. https://doi.org/10.1016/0010-0277(90)90024-E

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655-684.

Wellman, H. M., & Liu, D. (2004). Tasks. *Child Development, 75*(2), 523-541.

Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford University Press.

Wellman, H. M., Kushnir, T., Xu, F., & Brink, K. A. (2016). Infants use statistical sampling to understand the psychological world. *Infancy, 21*(5), 668-676. https://doi.org/10.1111/infa.12131

Wellman, H. M. (2018). Theory of mind: The state of the art. European Journal of Developmental Psychology. https://doi.org/10.1080/17405629.2018.1435413

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

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor’s reach. *Cognition, 69*(1), 1–34.

Yildirim, I., Saeed, B., Bennett-Pierre, G., Gerstenberg, T., Tenenbaum, J., & Gweon, H. (2019). Explaining intuitive difficulty judgments by modeling physical effort and risk. arXiv preprint arXiv:1905.04445.

Yu, C.-L., Stanzione, C. M., Wellman, H. M., & Lederberg, A. R. (2021). Theory-of-Mind Development in Young Deaf Children With Early Hearing Provisions. Psychological Science, 32(1), 109-119. https://doi.org/10.1177/0956797620960389

10. APPENDIX

APPENDIX A – Content Validity Sheet for the ToMT

Instruções

O TTOMI tem como objetivo avaliar a habilidade das crianças de 4 a 8 anos em Teoria da Mente (ToM), é composto por 12 tarefas baseadas em narrativas de animações em vídeo. As tarefas foram desenvolvidas para avaliar uma ampla gama de componentes da ToM, dentro das respectivas categorias: percepções, desejos, emoções, intenções e crenças.

Considere três aspectos principais para avaliar cada tarefa do instrumento:
1) Clareza: Avalie o quão clara e compreensível está a tarefa.Isso inclui a facilidade de compreensão da tarefa para a faixa etária e se as instruções e perguntas são facilmente entendidas.
2) Pertinência: Avalie se a tarefa contribui para o objetivo do instrumento. Verifique se a tarefa realmente é pertinente para avalia o construto Teoria da Mente.
3) Relevância: Avalie se a tarefa representa o componente que quer medir.

Cada tarefa será avaliada de acordo com estas três categorias numa escala de 1 a 5. Sendo 1 Pouco e 5 Muito. Por favor, forneça um feedback detalhado e construtivo para cada uma, incluindo sugestões de alterações, se necessário.

Construto - Teoria da Mente	Categorias	Componentes	Tarefa	Clareza (1-5)	Pertinência (1-5)	Relevância (1-5)	Sugestões de alteração
A Teoria da Mente (ToM) é um dos sub-domínios da Cognição Social (CS). Corresponde a capacidade humana de atribuir estados mentais a si mesmo e aos outros, como emoções, crenças, desejos e intenções.	Percepção	**Tomada de perspectiva visual** - Refere-se ao reconhecimento de que outras pessoas podem ter percepções visuais diferentes e a capacidade de adotar a perspectiva visual de outra pessoa.	Acesse a tarefa neste link				
	Percepção	**Tomada de perspectiva visual** - Refere-se ao reconhecimento de que outras pessoas podem ter percepções visuais diferentes e a capacidade de adotar a perspectiva visual de outra pessoa.	Acesse a tarefa neste link				
	Desejo	**Desejo diverso** refere-se ao entendimento de que diferentes pessoas podem ter preferências ou desejos divergentes em circunstâncias idênticas.	Acesse a tarefa neste link Qual dessas frutas é a sua preferida? Morango, banana ou pera? Ian gosta de .... **(neste momento o avaliador tem que falar a fruta que a criança não indicou como preferida).** É hora do lanche. Qual dessas frutas o Ian vai escolher?				
	Crenças	**Crença Falsa de Conteúdo** refere-se à compreensão de que um indivíduo pode ter uma crença incorreta sobre o conteúdo de um recipiente familiar, sem nunca ter aberto ou verificado esse recipiente.	Acesse a tarefa neste link				
	Crenças	**Crença Falsa de Primeira Ordem** refere-se à compreensão de que um indivíduo pode ter uma crença incorreta sobre uma situação específica. **Crença Falsa de Segunda Ordem** envolve a compreensão de que um indivíduo pode estar errado sobre a crença de outra pessoa em relação a uma situação.	Acesse a tarefa neste link				
	Crenças	**Crença Falsa de Segunda Ordem** envolve a compreensão de que um indivíduo pode estar errado sobre a crença de outra pessoa em relação a uma situação..	Acesse a tarefa neste link				
	Emoções/Crenças	**Compreensão das emoções com base em crença falsa de segunda ordem** refere-se à capacidade de entender e interpretar as emoções de outra pessoa considerando não apenas o que eles acreditam, mas também o que acreditam que outra pessoa acredita, mesmo que essa crença seja equivocada.	Acesse a tarefa neste link				
	Intenções	**Inferência de Preferência Baseada na Maximização da Utilidade** refere-se habilidade de inferir as preferências de alguém considerando sua tendência em escolher opções que maximizem as recompensas e minimizem os custos. ENCONTRE O REFERENCIAL TEORICO AQUI PARA REFERÊNCIA.	Acesse a tarefa neste link Essa tarefa quer mensurar a capacidade das crianças de aplicar um cálculo de utilidade ingênuo, um raciocínio que considera custos e recompensas para prever comportamentos Na primeira situação, o "amarelinho" opta pelo morango, uma escolha segura, porque ele sabe o que é, evitando a incerteza ,da fruta na caixa e mesmo a caixa estando mais perto. Pode-se inferir que ele te/m preferencia por morangos. Na segunda situação, antes de decidir pela banana que está mais longe, ele verifica a caixa, indicando que está disposto a considerar a incerteza, e que banana não seria a sua fruta favorita, se isso puder reduzir o esforço necessário para obter a recompensa. No entanto, após verificar a caixa e descobrir qual fruta está lá, ele ainda escolhe a banana, indicando que a fruta da caixa não era tão boa quanto a banana, e portanto ele iria mais longe para conseguir algo que ele prefere. A pergunta target avalia o racicionio da criança em relação a preferencia com base na maximização de utilidade				
	Intenções	**Inferência de Preferência com Base na Maximização da Utilidade** é a capacidade de deduzir as preferências de um indivíduo considerando sua tendência em escolher opções que maximizem as recompensas e minimizem os custos	Acesse a tarefa neste link				
	Intenções	**Inferência de Preferência Baseada na Maximização da Utilidade** refere-se à habilidade de inferir as preferências de alguém considerando sua tendência em escolher opções que maximizem as recompensas e minimizem os custos. **Avaliação Social** refere-se à habilidade de fazer julgamentos sobre as características sociais de outras pessoas, como simpatia, amizade, confiabilidade e competência.	Acesse a tarefa neste link				
	Intenções/emoções	**Inferência de experiência prévia e preferência a partir de reações emocionais** é a capacidade de deduzir as vivências passadas e preferências de alguém com base em suas expressões emocionais ao experimentar algo novo.	Acesse a tarefa neste link				
	Emoções/intenções	**Compreensão social** refere-se ao processamento e interpretação de informações sociais, incluindo emoções, intenções e comportamentos de outras pessoas, juntamente com o contexto social envolvido.	Acesse a tarefa neste link				
Algum componente importante do construto não foi abordado nos itens? Se sim, qual?			**Sugestão de item?**				

APPENDIX B - Participant Identification Questionnaire (PIQ)

Questionário de Identificação do Participante

Olá! Este questionário tem como objetivo coletar algumas informações sobre a criança participante e sua família para fins de identificação e caracterização dos participantes da avaliação de Teoria da Mente em crianças de 3 a 8 anos. Essas informações nos ajudarão a entender melhor como as habilidades de Teoria da Mente se desenvolvem em diferentes grupos de crianças. Suas respostas serão mantidas em sigilo e não serão compartilhadas com terceiros. Agradecemos antecipadamente pela sua colaboração!

* Indicates required question

Identificação pessoal

1. Nome da Criança: *

2. Data de Nascimento: *

Example: January 7, 2019

3. Idade da criança *

Mark only one oval.

- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- Other: ______________

4. Sexo da criança *

Mark only one oval.

- Maculino
- Feminino

5. Quem está respondendo? *

Check all that apply.

- [ ] Pai
- [ ] Mãe
- [ ] Outro Cuidador (especificar abaixo)
- [ ] Other: ______________

6. Nome do Responsável: *

____

7. Idade do(a) Responsável *

8. Telefone: *

9. E-mail: *

10. Cidade: *

11. Estado: *

12. Estado civil do(a) responsável: *

Mark only one oval.

- Solteiro(a)
- União estável
- Divorciado(a)
- Casado(a)
- Viúvo(a)
- Other:

13. Qual é o nível de escolaridade do responsável? *

Mark only one oval.

- Ensino Fundamental Incompleto
- Ensino Fundamental Completo
- Ensino Médio Incompleto
- Ensino Médio Completo
- Ensino Superior Incompleto
- Ensino Superior Completo
- Pós-graduação Incompleta
- Pós-graduação Completa

14. Qual a profissão do responsável? *

____

15. Quantas pessoas moram na casa? *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5 ou mais

16. A criança tem irmãos ou irmãs? *

Mark only one oval.

- Sim
- Não

17. Se sim, quantos? *

____

18. Qual é a renda mensal familiar? *

Mark only one oval.

- Até 1 salário mínimo
- De 1 a 2 salários mínimos
- De 2 a 3 salários mínimos
- De 3 a 5 salários mínimos
- Mais de 5 salários mínimos

Condição Clínica

Esta seção tem como objetivo coletar informações sobre a saúde da criança e eventuais diagnósticos médicos, para fins de identificação de possíveis fatores que possam influenciar as habilidades de Teoria da Mente.

19. A criança tem algum problema de saúde ou já recebeu algum diagnóstico médico? *

Mark only one oval.

- Sim
- Não

20. A criança tem algum diagnóstico neuropsicologico?

Mark only one oval.

- Sim
- Não
- Other: _______________________

21. Se a resposta anterior for "Sim", por favor especifique o diagnóstico na pergunta abaixo. Caso contrário, escreva "Nenhum". *

____

22. Alguma observação importante sobre a criança que a equipe de avaliação deveria saber? *

____
____
____
____
____

23. Você declara que as informações acima são verdadeiras e que autoriza o uso desses dados para fins de pesquisa científica? *

Mark only one oval.

- Sim
- Não

APPENDIX C – Target Population Feedback Form

Feedback TTOMI

* Indicates required question

1. Nome da criança *

E então, vocês se divertiram?

Você acabou de finalizar o instrumento TTOMi e precisamos do seu feedback sobre essa experiência. As crianças respondem com a ajuda de um adulto, ok?

2. O quanto você gostou do vídeo?

Mark only one oval.

Não Gostei

Neutro

Gostei

3. Você indicaria esse vídeo a um amigo?

Mark only one oval.

◯ Não indicaria

◯ Indicaria

4. Para finalizar, o que você acredita que poderia melhorar no vídeo? Alguma sugestão?

____

____

____

____

____

5. Quais as suas principais dificuldades que você teve para responder as perguntas do vídeo?

____

____

____

____

____