

INVESTIGATING CULTURAL DIFFERENCES IN THE USE OF AUTISM SCREENER MEASURE MCHAT-R/F USING COGNITIVE INTERVIEWING TECHNIQUES

by

JOY MITCHELL

(Under the Direction of Ashley Harrison)

ABSTRACT

The Modified Checklist for Autism in Toddlers, Revised with Follow-Up (MCHAT-R/F) is one of the most widely used early screening tools for autism spectrum disorder (ASD); however, it was developed within a Western framework that may not align with the cultural norms, values, and experiences of all families. This study utilized cognitive interviewing techniques grounded in a phenomenological approach, to explore how caregivers from diverse cultural and linguistic backgrounds cognitively process MCHAT-R/F items. Fourteen caregivers of children aged 16–48 months participated in semi-structured cognitive interviews using the MCHAT-R/F. Interviews were analyzed using a blended coding framework that drew from a four-stage cognitive model that examined participant comprehension, retrieval, judgement, and response along with dimensions of cultural relevance. Analysis revealed notable variation in caregiver cognitive processing, particularly in the retrieval and response domains. Findings also highlighted notably increased difficulties for participants that were identified as bicultural on an acculturation measuring, indicating greater interpretation difficulties when navigating both their culture of origin and dominant Western culture. Items that assessed subtle or low-frequency

behaviors (e.g., unusual finger movements, social referencing), child response to social bids and joint attention (e.g., following a point, responding to name or verbal directions), and behaviors with culturally variable salience (e.g., upset by noise, emotional referencing) elicited the greatest number of interpretation errors. These results underscore the need for more culturally responsive adaptations of existing autism screeners and reinforce the value of cognitive interviewing as a tool for examining cross-cultural differences. Clinicians and researchers should be aware of how sociocultural frameworks influence item interpretation and avoid misattributing response variability to parental inaccuracy or disengagement.

INDEX WORDS: autism, culture, acculturation, screening, cognitive interviewing

INVESTIGATING CULTURAL DIFFERENCES IN THE USE OF AUTISM
SCREENER MEASURE MCHAT-R/F USING COGNITIVE INTERVIEWING TECHNIQUES

by

JOY MITCHELL

B.A., Emory University, 2014

M.A., Columbia University, 2020

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2025

© 2025

Joy Mitchell

All Rights Reserved

INVESTIGATING CULTURAL DIFFERENCES IN THE USE OF AUTISM SCREENER
MEASURE MCHAT-R/F USING COGNITIVE INTERVIEWING TECHNIQUES

by

JOY MITCHELL

Major Professor:	Ashley Harrison
Committee:	Sycarah Fisher
	Chitra Pidaparti
	Matthew Madison

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2025

DEDICATION

To my mother, my grandmother (Gam), and my big sister Serenity—my very first teachers, whose love, faith, and strength taught me how to endure and persevere. To my little sister, Love, my greatest reminder of joy, light, and possibility. And to my friends who stayed, encouraged, and held me up through every triumph and challenge, thank you for believing in me.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Ashley Harrison. Thank you for your unwavering support, invaluable insights, and continued belief in me. Your mentorship has shaped not only this dissertation but the scholar I am becoming. To the CAARE Lab, thank you for standing with me. From long meetings to late nights, your collaboration, encouragement, and tireless effort carried me through some of the most challenging and rewarding moments of this journey. I'm forever grateful to be surrounded by a team that pours so much heart and excellence into their work. To my dissertation committee, thank you for your thoughtful feedback and guidance throughout this process. Your contributions helped strengthen this project in ways that reflect both rigor and care. To the families who participated in this study, thank you for your honesty, your vulnerability, and your time. Your voices sit at the heart of this dissertation. This work is for and because of you.

Table of Contents

Chapter 1: Introduction.....	1
ASD Diagnosis Overview	1
ASD Screening	2
ASD Diagnostic Practices	9
Global ASD Screening	12
Global Screening Barriers	13
Current Progress in Global Screening: The MCHAT-R/F	17
The purpose of the Current Study.....	25
Chapter 2: Method	26
Participants	26
Participant Demographics.....	27
Procedure Overview	29
Measures	29
Cognitive Interviewing	31
Data Analysis.....	36
Items Analyzed	44
Chapter 3: Results	44
Total Cognitive Domain Results Across All Participants	44
Domain-specific Interpretation Patterns	47
Judgement.....	53
Response.....	55
Cultural Relevance	56
Person-specific Interpretation Difficulties.....	58
Chapter 4: Discussion.....	64
Cognitive Domain Interpretation Patterns	68
Implications for Practice.....	76
Limitations.....	80
Future Directions	82
Conclusion	83
References.....	84
Appendix: Tables & Figures.....	116
Tables.....	116
Figures	123

Chapter 1: Introduction

ASD Diagnosis Overview

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder that is known for its heterogeneous presentation across individuals. ASD is diagnosed using criteria from the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) and is characterized by deficits in social communication and interactions, as well as restrictive and repetitive behaviors (APA, 2013). The autism diagnostic criteria is placed into two categories: Category A is defined as social communicative and interactive deficits across multiple contexts, which consists of deficits in social-emotional reciprocity, deficits in nonverbal communicative behaviors, and deficits in developing, maintaining and understanding relationships. Category B is defined as restricted, repetitive patterns of behavior, consisting of stereotyped or repetitive motor movements or speech, insistence on sameness or rigidity, highly fixated interests, and sensory sensitivity.

There are two levels of identification in the process of making an autism diagnosis (Hyman et al., 2020). The first level consists of screening, which is typically a brief questionnaire given to all parents for developmental monitoring by generalist service providers (e.g., pediatrician or nurse practitioner; Hyman et al., 2020). If screeners indicate increased probability of ASD, the child would ideally be moved to the next level of identification which involves a comprehensive diagnostic assessment. Given the short period of time service providers are allotted for primary care visits with patients, they cannot always be certain of the presence of abnormality thus, comprehensive evaluations are an important next step (Hyman et

al., 2020). A comprehensive evaluation to diagnose ASD is a process that typically involves assessment across multiple domains including autism specific symptoms, cognition, and adaptive skills (Ozonoff et al., 2005).

ASD Screening

ASD Screening Importance

Amongst the areas with the highest prevalence rates of developmental disabilities, a system of regional centers called the Autism and Developmental Disabilities Monitoring Network has been put in place to conduct assessments, determine diagnosis, and inform the need for disability services. Within these sites, the CDC recently reported an increase in autism prevalence from 1 in 36 to 1 in 31, with diverse populations continuing to show higher prevalence than White counterparts (Maenner et al., 2023; Shaw et al., 2025). With these consistent increases, researchers have continued to investigate ways in which to improve early detection of ASD. Early identification of children with ASD is vital in order to intervene during critical periods of neuroplasticity (0-3 years old), as the brain is most responsive to experiences within the environment during this time (Huttenlocher, 1979; Pierce et al., 2016). Early intervention services including speech, behavior therapy, physical therapy, occupational therapy, and others specifically tailored to meet individualized family needs have been shown to improve developmental outcomes (e.g., gains in IQ, improved adaptive skills, reduction in core ASD symptoms) for children with ASD, especially when they are received early (Anderson et al., 2014; Bradshaw et al., 2017; Dawson et al., 2010; Oono et al., 2021; Reichow, 2012; Rogers et al., 2012). Research found that those previously diagnosed with autism who no longer meet criteria over time are more likely to have been diagnosed and provided interventions prior to the age of 3 (Anderson et al., 2014).

Because of the significant impacts that autism has on cognitive development, social relationships, and adaptive functioning, the American Academy of Pediatrics (AAP) recommends all children be screened at 18 months and 24 months of age in pediatric offices to provide intervention and support as early as possible (Johnson & Myers, 2007). The AAP Board of Directors recently published an advocacy letter calling for payers to shift away from specific evaluation requirements for ASD diagnosis and to allow general pediatricians to diagnose autism, as changing these requirements could remove many barriers that limit access to ASD diagnoses and/or services (American Academy of Pediatrics, 2024; Shaw, 2025; State of Georgia, 2022). This policy shift could fundamentally change how children are identified with ASD. Pediatricians have been identified as being optimal administrators for screening, as they are seen as the first line of defense given the continuity of care with well-child visits (Al-Qabandi, 2011; Crais et al., 2014). They play a major role in the early detection of autism, as they see children for 11 well-child visits prior to the age of 3 and are typically the first point of contact for caregivers. While caregivers are usually aware of concerns with their children, they may not share them unless the provider specifically asks about development; therefore, providers are suggested to formulate a standard of care that includes standardized developmental screening (Johnson & Myers, 2007). Additionally, continuity is helpful in screening, as 20-40% of autistic children show some type of developmental regression (Volkmar et al., 2005; Werner et al., 2005), and repeated screening at 24 months may help identify autistic children who may otherwise be missed (Barton et al., 2011; Earls & Hays, 2006; Pinto-Martin et al., 2005).

Autism Screening Practices

These early and routine developmental screenings in pediatric primary care are a highly recommended strategy to support timely autism evaluations, diagnoses, and treatment resources,

as elevations in screening often lead to referrals for a comprehensive evaluation and subsequent identification and intervention (Hyman et al., 2020). The process of screening alerts primary care providers and other healthcare professionals to children that may need further clinical examination and support. Screening is low-cost in that it requires little time, money, and resources (Robins, 2008). There are two types of screening measures. Level 1 screening measures are to be used among a widespread general pediatric population (Robins, 2008). Thus, they must be brief due to time constraints of pediatric appointments, and low cost due to frequent use among those who are both at risk and not at risk (Robins, 2008). Level 2 screening measures require more time and clinical expertise, as these are used on children who have an increased likelihood of having ASD. These measures are often used as a part of a comprehensive evaluation (Johnson & Myers, 2007). In hopes of identifying as many children as possible on the population level, Level 1 screening through primary care providers is vital (Robins, 2008). Without Level 1 autism screening, children with concerns are less likely to be seen by specialists for comprehensive evaluations (Robins, 2008).

The endorsement of screening as a standard of practice by the AAP has been considered one of the most successful public health policies for children with ASD (Pierce et al., 2011; Wetherby et al., 2008; Zwaigenbaum et al., 2015), and pediatric screening has been found to be objectively successful (Zwaigenbaum et al., 2015). Specifically, the use of screening tools has been found to contribute to more successful identification of risk as compared to provider judgement and monitoring (Gabrielsen et al., 2015; Miller et al., 2011). The development of screener usage has been found to diagnostically identify ASD in children 2 to 3 years earlier (Chlebowski et al., 2013, Gabrielsen et al., 2015; Miller et al., 2011; Pierce et al., 2011; Robins et al., 2014; Watson et al., 2007; Wetherby et al., 2008) than the national average diagnosis age

of 4 years old (Baio, 2014). This process serves an even greater utility for individuals from minoritized groups, specifically those from a lower socioeconomic status and diverse racial groups, who are significantly under-detected, often receive a later diagnosis, and subsequently experience delayed access to services (Fountain et al., 2011).

Despite AAP recommendations for universal screening in primary care, implementation research has found varied rates of autism screening with standardized measures via pediatricians' reports during child visits at 18 and 24 months (Arunyanart et al., 2012; Dosreis et al., 2006; Gillis, 2009; King et al., 2010; Pierce, 2011). Pediatricians have the role of monitoring children's development at a young age through well-child visits, making them uniquely positioned to contribute to early identification of autism (James, 2020). Unfortunately, primary care providers do not always follow the guidelines; some may screen at 18 months but not at the 24-month visit, and others have been found to screen with clinical judgement at both visits instead of using a validated measure. This clinical judgement can vary significantly and is less accurate than standardized measures (Gabrielsen, 2015; Miller, 2011). Despite the variation in autism screener utilization and adherence to specific AAP guidelines, research shows autism screening rates by pediatricians have notably increased in recent years (James & Smith, 2020). This increase is attributed to intentional efforts made to educate pediatricians on the importance of autism screening, which led to the incorporation of screening measures into primary care visits (Hyman & Johnson, 2012; Radecki et al., 2011). The AAP found pediatricians increased the utilization of screening tests from 23% in 2002 to 63% in 2016 (Lipkin et al., 2020).

Broad Autism Screening Measures

Screening measures are not used to provide an in-depth diagnosis, but serve to indicate if a detailed evaluation is required (Robins, 2008). General developmental screening measures can

identify language, cognitive, and motor delays. They are widely used but are not always sensitive to deficits uniquely associated with autism (Zwaigenbaum et al., 2015). There are both broad and autism-specific screening tool measures recommended by the CDC and AAP websites that should be considered for usage in pediatric or school settings (Hyman et al., 2020; Johnson & Myers, 2007; Rotholz et al., 2017).

The Ages and Stages Questionnaires (ASQ) is a general developmental screening tool completed by caregivers of children ages 1 month to 66 months in approximately 10-20 minutes. The ASQ psychometric properties are as follows: sensitivity = 0.75, specificity = 0.81. It includes a series of 19 age-specific questionnaires that screen across multiple developmental domains. These domains include communication, gross motor, fine motor, problem-solving, and adaptive skills. These questionnaires result in a pass/fail score by domain. The ASQ reliably identifies delays associated with autism, including delays in the behavior and communication domains (Squires et al., 1995).

The Communication and Symbolic Behavior Scales Developmental Profile - Infant Toddler Checklist (CSBS-ITC) is a caregiver screening tool of 24 items for children ages 6 months to 24 months. It takes 5 to 10 minutes to complete and assesses communication and symbolic abilities of children. The CSBS considers emotion, eye gaze, communication, gestures, understanding, and object use. The measure creates three composites: social, speech, and symbolic. The social composite score includes emotion, eye gaze, communication, and gestures. The speech composite score comprises the scores for sounds and words. The symbolic composite score consists of understanding and object use. The psychometric properties are as follows: sensitivity = 0.78, specificity = 0.84 for a communication delay (Limosani et al., 2021; Wetherby, 2003).

The Parents' Evaluation of Developmental Status (PEDS) is a parent interview form that screens for developmental and behavioral problems in children from birth to 7 years, 11 months. It asks 10 questions across 10 categories, including fine motor, gross motor, expressive language, receptive language, school, cognitive, health, behavioral, adaptive, and socio-emotional skills. The psychometric properties include sensitivity and specificity with values ranging from .70-.96 across all ages (Glascoe, 1998; Glascoe, 2003).

Autism Specific Screening Measures

Autism screening tools are designed to probe caregivers to identify and report symptoms related to autism. The screeners are typically based on early manifestations of autism, including behavioral symptoms and core deficits in social communication, play, and repetitive behaviors (Hyman et al., 2020). These manifestations are often referred to as red flags that cue the provider to the likelihood of autism. Some red flags include children who are unresponsive to their names at 12 months, children who are not pointing to show interest at 14 months, and/or children who are not displaying pretend play at 18 months. In order to identify red flags of ASD, there are many options for screening tools used in pediatric settings (Steiner et al., 2012).

The Social Communication Questionnaire (SCQ) is a screening instrument that consists of 40 yes or no questions to be answered by a caregiver. It takes about 10 minutes to complete and assesses children 4 years and older. It probes for information about repetitive body movements, verbal and nonverbal communication, and social interaction abilities. The psychometric properties are as follows: sensitivity = 0.85 and specificity = 0.75 (Berument et al., 2003).

The Screening Tool for Autism in Toddlers and Young Children (STAT) is considered a level 2 screener in that it is typically used for children already suspected to have a high risk of

ASD. It is an interactive screening tool for children aged 24 to 35 months. It consists of 12 activities that assess s play, communication, and imitation skills. It takes 20 to 30 minutes to complete and observations must be scored by a trained provider. The psychometric properties are as follows: sensitivity = 0.83, specificity = 0.86, positive predictive value = 0.77, negative predictive value = 0.90 for autism amongst children greater than 24 months (Stone et al., 2000).

The Rapid Interactive Screening Test for Autism in Toddlers (RITA-T) is an interactive and observational screening measure that must be administered by a professional. It is a 9-item measure for children aged between 18 to 36 months old. It is also considered a level 2 screener for children already suspected of having autism. The psychometric properties are as follows: sensitivity = 1, specificity = 0.84, positive predictive value = 0.88, and negative predictive value = 0.94 (Choueiri & Wagner, 2015).

The MCHAT-R/F is a two-part screener. The first part is a level 1 screener that consists of 20 yes or no questions that are designed for caregivers of children ages 16 to 30 months. It probes for common difficulties associated with individuals diagnosed with ASD (e.g., response to joint attention, initiation of joint attention, social interest, social responsiveness, etc.). The second part is a structured interview to be completed by a professional if a moderate risk of ASD is identified (Robins et al., 2014). Research has found that children screened using the MCHAT-R/F are identified with autism earlier than the national average (Lipkin et al., 2020). The psychometric properties are as follows: sensitivity = .91, specificity = .95, positive predictive value = 0.45 for low-risk 18 to 24-month-old children with the follow-up interview (Robins et al., 2014). The MCHAT-R/F is the most widely studied and implemented screener tool for autism (Aishworiya et al., 2023; Campbell et al., 2017; Marlow et al., 2019). It is also recommended by national guidelines and many autism experts as a screening tool for children

due to its sensitivity and ease of usage (Fuentes et al., 2021; Hyman et al., 2020; Zwaigenbaum et al., 2019).

ASD Diagnostic Practices

Following a positive screen indicating a child has a higher likelihood of autism, there are multiple options available to consider that can occur simultaneously in hopes of finding the option that results in the earliest evaluation. The family of the child will likely be given a referral for a comprehensive evaluation. The waitlists for these evaluations are often lengthy and can be time-consuming, happening over multiple days and/or visits (Gabrielsen et al., 2020).

The AAP algorithm recommends that children under 3 years of age with a positive screener be referred to Early Intervention (EI) services (Johnson & Myers, 2007). This is because EI services have been identified as the fastest path to an evaluation (Gabrielsen et al., 2020). Additionally, EI services are symptom-dependent opposed to diagnosis-dependent, meaning if delays are present, services are made available even without a diagnosis. These services are free of charge under the Individuals with Disabilities Education Act (IDEA), Part C for children from birth to 36 months of age (Lipkin et al., 2015). Prior to the initiation of EI, an Individual Family Service Plan (IFSP) is created by a team that includes caregivers and the child's providers. The team describes the current level of development, the family's strengths and weaknesses, specific services required, and a plan to transition into the public-school setting. EI assists children with learning new skills and can include family training, counseling, or home visits. It can also offer occupational, physical, or speech therapy. Additional services may include health, nutrition, social work, and coordination of care support, along with assistive technology devices and transportation (Lipkin et al., 2015).

Those between 3 and 21 years of age are accommodated through the IDEA, Part B in which services for school-aged children with developmental disabilities are provided at no cost through the public school system if evaluated and deemed eligible (Lipkin et al., 2015). The school's evaluation may or may not include autism specific measures and the level of therapeutic intensity wanted by the family may be beyond the school's scope (Irvin et al., 2012). With this, families may supplement school evaluations and interventions with private assessments and therapies. Following a determination of eligibility, an Individualized Education Plan (IEP) is developed by the necessary entities within the school district. The IEP, in contrast with the IFSP, focuses solely on the child's goal. If the state-regulated eligibility criteria are met, intervention support can include special education and related services, including physical, occupational, and speech therapy. It can also include aids, adaptive equipment, and communication devices (Hyman et al., 2020; Lord et al., 2006). Health insurance mandates have been put in place to support autistic children throughout the U.S. This includes coverage for the most commonly prescribed treatments for ASD, including speech therapy, occupational therapy, and Applied Behavior Analysis (ABA) (Bitterman et al., 2008). For health insurance mandates to be enforced, medical diagnoses are required; thus, comprehensive evaluations outside of school systems may be necessary (Irvin et al., 2012).

Parent-report Measures

Comprehensive evaluations include assessing for ASD specific symptoms that can be measured through direct observation and self-report measures from parents, teachers, or other primary caregivers. These measures rely on an operational definition of typical behavior and note autism symptoms as deviations from these (Hyman et al., 2020). Often, measures used as screeners can also be used as part of the comprehensive evaluation process, as data used to

diagnose ASD must include information from multiple sources across contexts. This is because symptoms of ASD may be present based on environmental changes and interactions with various people based on familiarity. Additionally, a short observation period may not elicit all difficulties an individual has experienced. With this in mind, self-report questionnaires can be additional tools given to caregivers (Ozonoff et al., 2005).

Some gold standard measures include the Social Responsiveness Scale, 2nd Edition (SRS-2), Autism Spectrum Rating Scales (ASRS), and The Autism Diagnostic Interview-Revised (ADI-R). The SRS-2 is a screening instrument with 65 questions that probes for information about restricted interests and repetitive behavior, as well as social communication and interaction, including social awareness, cognition, communication, and motivation (Constantino & Gruber, 2012). The ASRS is an instrument with two forms for age ranges 2-5 and 6-18 with 70 questions that probe for information about social/communication and unusual behaviors (Goldstein & Naglieiri, 2009). The Autism Diagnostic Interview-Revised (ADI-R) is a standardized interview used to diagnose autism and distinguish it from other developmental disorders (Rutter et al., 2003).

Observation-based Instruments

However, the gold standard combination of measures for a comprehensive autism evaluation following a positive initial screening includes a symptom observation component known as the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) (Lord et al., 2012) along with the ADI-R parent interview (Rutter et al., 2003). The ADOS-2 is semi-structured and designed as an interactive assessment specifically for identifying social and communication abnormalities. It consists of five modules that can be selected based on the age and language level of the individual being assessed. During the ADOS-2, the examiner offers

many opportunities for an individual to respond to ‘presses’ or bids for social interaction and communication, both verbal and nonverbal. This is done through various activities, play scenarios, and conversational prompts that are setup by the examiner. For individuals of a younger chronological or mental age, the following areas are typically examined: social awareness, motivation, joint attention, nonverbal communication, symbolic play, sensory sensitivity, highly fixated interests, and repetitive behaviors (Lord et al., 2012). For individuals of an older age and those with limited communication abilities, the following areas are also examined: reciprocity in social communication, insight into emotions, insight into social relationships, and unique interests (Lord et al., 2012). After an individual is offered several opportunities to display typical social behaviors and functional communication skills and they are missed, the examiner can more reliably deduce the presence of ASD symptoms (Ozonoff et al., 2005).

Global ASD Screening

Global autism rates were previously estimated to be about 1% worldwide (Elsabbagh et al., 2012) and have recently been estimated to have increased to approximately 1.5- 2% (Lyll et al., 2017; Zeidan et al., 2022). This makes autism one of the most common developmental disorders (Malcolm-Smith et al., 2013). Despite this increase, it can be difficult to bring global screening to the attention of the public and policymakers in low- and middle-income countries (LMICs), as survival related to physical health is often prioritized (Hahler & Elsabbagh, 2015; Scherzer et al., 2012).

Although rates of ASD are thought to occur worldwide with similar frequency (Elsabbagh et al., 2012), there is significant variability in screening rates and subsequent age of diagnosis (Zeidan et al., 2022). In epidemiological studies used to estimate the global prevalence

of ASD, 86.5% of cases identified were in North America, Europe, and Japan, where only 10% of the children across the world live. With this, the prevalence of ASD across LMICs are thought to be largely underestimated. (Matos et al., 2022). Globally there have not been consistently identified socio-demographic factors that account for prevalence variation, but it is likely related to regional differences in availability, accessibility, and utilization of tools. Specifically, global prevalence variability is often thought to be related to the utilization of standardized assessment tools (Matos et al., 2022; Zeidan et al., 2022). The usage of assessment tools are typically based on the knowledge and expertise of the administrators within the respected regions (Zeidan et al., 2022).

As awareness of autism and its barriers to identification (e.g., high cost of supplies, increased training) arise within underrepresented regions of the world, epidemiological studies have responded to the noticeable lack of contextually appropriate tools in three ways (Durkin et al., 2015; Elsabbagh et al., 2014). One is by translating and/or culturally adapting tools from their original form in English, which allows for comparable estimates across different world regions (Aldosari et al., 2019; Chaaya et al., 2016). Another way is by using originally developed screening tools for a specific country or region, which limits comparability with other autism research around the globe (Arora et al., 2018; Kakooza-Mwesige et al., 2014; Raina et al., 2017). Others have responded by using unvalidated tools that were originally designed for use in a non-autism specific context (e.g., teachers identifying students using a form based on informal descriptions of signs and symptoms) (Narsizi et al., 2020).

Global Screening Barriers

Assessment Access

Despite best screening practices advised by the AAP, there continues to be an imbalance between knowledge and screening services, especially in LMICs (Choueiri et al., 2023). One challenge is that professionals in low-resource areas often face difficulties accessing screening tools due to issues with accessibility, knowledge, and cost (Marlow et al., 2019). Additionally, administration of some screener measures can be resource intensive (i.e., time, personnel training) and local expertise is often limited (Durkin et al., 2015). There are also licensing concerns as it relates to adapting, translating, and validating tools in culturally and linguistically diverse settings, which is necessary because cultural influence can impact the interpretation of symptom presentation and development (Wallace et al., 2012). Additionally, in low-resource settings, lack of access to preventative healthcare, poverty, and psychosocial adversities have been identified as barriers to access to early autism screening. With this, children in these settings are often identified at a later age than their same-aged peers in resource rich environments (Durkin et al., 2015; Khowaja et al., 2015).

No Universal Screening Approach.

Early identification of ASD has been labeled a major problem for clinicians as well as researchers. Awareness continues to be a large barrier to early identification, which in turn increases screening disparities in low-resource environments. Along with this, deficits in qualified professionals who can recognize symptoms of ASD has also been identified as a significant concern (Durkin et al., 2015; Khowaja et al., 2015). Although there has been agreement across the globe to support the importance of early diagnosis and comprehensive management of autism, via the World Health Organization, a universal screening tool has yet to

be established (Jullien, 2021). Deficits in literacy skills are common amongst LMICs but it also affects vulnerable populations amongst high-income countries. Given this, screening instruments developed and validated with primarily highly educated families may not be as psychometrically effective for groups with lower education (Hoekstra, 2022; Khowaja et al., 2015). Furthermore, many well-validated and frequently used screening tools have been developed in countries using samples of high socioeconomic status that lack cultural diversity. These screening measures have been tested and found to be reliable and valid when assessing for ASD (Perera et al., 2017; Varma & Iskandar, 2014). However, as is common with much behavioral research, most measures used to assess autism are developed in Western countries using White samples and often skewed toward the upper SES (Heinrich et al., 2010; Kleinman et al., 2007). Typically, convenience samples like these consist of highly educated and affluent populations within the Western world, creating a global barrier to early screening. Many early screening tools have yet to be validated in environments with cultural values that differ from western society, including child-rearing practices. These barriers have created large socioeconomic and geographic disparities in the development, validation, and accessibility of early screening tools (Heinrich et al., 2010).

Cultural Measurement Invariance

When measures are translated into the appropriate language, this is often simply a word-for-word translation on a linguistic level without cultural consideration (Al Maskari et al., 2018; El-Behadli et al., 2015; Soto et al., 2015). The process of changing a measure based on both linguistic differences as well as cultural nuance is termed transadaptation (Zucker et al., 2005). Translation may not be sufficient to use screening tools with new populations, as word-level translation errors and cultural differences complicate translations (Gjersing et al., 2010;

McDermott & Palchanes, 1994). Culture creates distinctive lenses through which people view the world and impacts parenting and child development. For instance, it influences when children learn or are reinforced for particular skills or behaviors (Kleeck, 1994). Culture also impacts the way in which people understand and answer questions (Beaton et al., 2000). Culture's influence on parent practices will subsequently affect a parent's understanding of questions and their response patterns (Soto et al., 2015). To help ensure screening measures are valid for cross-country utility, research suggests transadaptation as a best practice (DuBay & Watson, 2019). Without appropriate cultural consideration in translation methods, the new version of the screening measure may differ, leading to an over or underidentification compared to the original instrument (Soto et al., 2015). Additionally, without representation across social classes and cultures in the development of new tools and modification of existing tools, sociocultural disparities in early ASD screening is inevitable (Durkin et al., 2015).

The majority of screening tools have been developed in high-income countries, which yields variable utility when used outside of these settings. This subsequently creates challenges in the utilization of instruments in LMICs due to barriers related to financial cost and the need for cultural adaptation and validation (Choueiri et al., 2023). The majority of screening tools were created and validated with English-speaking individuals in Western countries, but identifying children with disorders like autism is not specific to the western world (Aishworiya et al., 2023; Sangare et al., 2019). With this, the World Health Organization has launched significant efforts to increase the global ability to meet the needs of autistic children (World Health Organization, 2014). In order to do so, screening tools that are valid among culturally and linguistically diverse populations are necessary (DuBay & Watson, 2019). Until a universal tool is identified, efforts are needed to ensure that existing measures are transadapted in the most

culturally sensitive manner. The questions remain, what is the ideal screener and what are the best approaches for adapting these for other cultures.

Current Progress in Global Screening: The MCHAT-R/F

MCHAT-R/F Strengths

The resources required to formulate and validate new tools are costly; thus, the modification of existing tools is preferable, as it is more feasible and cost-effective (Guillemin et al., 1993; Ware et al., 1996). The MCHAT-R/F is a well-validated screener developed in western society by researchers in the United States. It assesses common constructs to evaluate the risk level for diagnosis (Robins et al., 2001; Robins et al., 2014) and is a good choice for transadaptation in the absence of a universal screener. The MCHAT-R/F is widely used on an international scale due to its brevity, affordability, and ease of access (Albores-Gallo et al., 2012; Barbaro & Dissanayake, 2009; DuBay & Watson, 2019; Kara et al., 2014; Kleinman et al., 2008; Levy et al., 2020). Its original validation also yields a high sensitivity, positive predictive value, and negative predictive value (Aishworiya et al., 2023). To date, the MCHAT-R/F has been translated into 76 different versions, including various languages and regional dialects with 16 more translations pending for use on the MCHAT-R/F website. This makes it accessible to many linguistically diverse populations (e.g., Canal-Bedia et al., 2011; Guo et al., 2019; Kamio et al., 2014; Vorster et al., 2022).

MCHAT-R/F Global Limitations.

Although the internal consistency measured using Cronbach's alpha has been found to be reasonable across translations of the MCHAT-R/F (Brennan et al., 2016; Carakovac et al., 2016; Cuesta-Gómez et al., 2016; Divya et al., 2020; Oner & Munir, 2020; Tabril et al., 2023), some researchers have identified problems with other psychometric properties of the screening

instrument in diverse contexts. The metric positive predictive value (PPV) is the probability that the individual who screens positive actually has ASD (Parikh et al., 2008). It is considered more practical for clinicians as they investigate psychometrics to choose a measure based on probability of accurately identifying those with a disorder (Akobeng, 2006). The initial validation of MCHAT-R/F showed a high PPV (>0.90) for a development delay identification, but showed low PPV (<0.50) for identifying ASD; however, when the MCHAT-R/F early screener tool was adapted for use in diverse settings, the PPV showed wide variability. Likewise, some translated versions of the MCHAT-R/F also showed low PPV for ASD (e.g., Brennan et al., 2016; Canal-Bedia et al., 2011; Samadi & McConkey, 2015). Of note, some data shows PPV can be increased when the MCHAT-R/F follow-up interviews are conducted to clarify selected responses, though this is not typical practice during screening administration (Wallis et al., 2020).

Importantly, PPV has shown to vary even among populations for which a translated version is not required. For example, further related to PPV, a study within sociopolitically diverse populations showed lower PPV to be associated with decreased maternal education (Khowaja et al., 2015). One study investigating the real world accuracy of the MCHAT-R/F in a large pediatric setting found that specificity and PPV were higher for white children (97.9%; 24.0%) compared with black children (91.7%; 11.7%), Asian children (90.4%; 10.8%), and other/multiple racial groups (93.8%; 13.4%). Higher specificity and PPV were observed in children with English-only exposure compared to children with non-English exposure (95.2% vs 86.9%; 15.3% vs 8.5%), as well as children from higher- versus lower-income families (97.0% vs 92.3%; 20.4% vs 11.8%) (Guthrie et al., 2019).

In considering further psychometric properties, specificity refers to the ability to eliminate individuals who do not have autism. The initial MCHAT-R/F screener has solid specificity, which is > 0.90 . Some translations show similarly well-developed specificity (Canal-Bedia et al., 2015; Oner & Munir, 2020; Samadi & McConkey, 2015), while others show lower rates of specificity (Babaro & Halder, 2016; Jensen et al., 2021). With regard to sensitivity, high sensitivity was found in China (Guo et al., 2018) and Turkey (Oner & Munir, 2020), while low sensitivity was found in Peru (Jensen et al., 2021), Japan (Babaro & Halder, 2016), and Mali (Sangare et al., 2019). This variability in psychometric quality can have negative consequences. Specifically, psychometric variability in early screening can lead to inadequate results like high rates of false positives. When this occurs in low-resource communities, it likely floods the referrals for full evaluations in communities that often already have long wait times. This contributes to disparities in these communities of disproportionately minoritized individuals. Solutions are needed to ensure that, within these specific communities, the diagnostic process of accurate identification is streamlined despite or due to the limitation of resources (Zeidan et al., 2022).

MCHAT-R/F: Remaining Issues to Solve

Given these identified psychometric concerns with the MCHAT-R/F across cultural contexts, it is important to get to the root of the problem to move forward with identifying solutions. Currently, the MCHAT website has limited recommendations for how to best adapt new versions for use in different countries. The current standard for new version development is primarily translation without regard for cultural adaptation. Developers suggest having a bilingual individual translate all the original English components, and then having a second bilingual individual, who has not been exposed to the MCHAT-R/F, back translate the new

version. It is suggested this process be repeated until language discrepancies have been identified and fixed but have no recommendation for content issues that might arise (mchatscreen.com). Although this translation method provides linguistic access for various countries, it does not consider the cultural differences that may affect item-level responses on the questionnaire, likely resulting in a misrepresentation of symptomatology.

Research has shown that children from racially diverse backgrounds experience longer delays between initial concerns and autism diagnosis. International research supports this finding, suggesting that later age of diagnosis is tied to parental report. Some research suggests this may be related to varied cultural interpretation of early signs of autism (Grinker et al., 2012; Tek & Landa. 2012). Many early signs of autism (i.e., delayed developmental milestones) often go unidentified if they lack cultural significance. Specific delays in milestones may be culturally insignificant, and therefore go unreported by diverse groups (Coonrod et al., 2004). Given the age delays in ASD diagnosis for culturally diverse populations, the age range for the MCHAT-R/F (16-30 months) may be problematic for these populations, as subtle signs are not identified as early on as Western counterparts (Coonrod et al., 2004).

Transadapting the MCHAT-R/F

The use of translated versions of MCHAT-R/F across cultures is criticized, as one-to-one linguistic translations without cultural adaptation have been seen as problematic. There are multiple ways to address the lack of culturally relevant tools but adapting a validated measure for diverse populations has been suggested as the most effective and time efficient. Cultural adaptation is complex because it must go beyond language translation. It must examine values and traditions to increase congruence between the cultural views and beliefs of the population assessed and the social expectations described in the Western-developed measure. Literal

language translations are often minor, but content adaptation requires deeper examination into cultural values, structural beliefs, and social norms (Al Maskari et al., 2018). Within the fields of autism and developmental disabilities, simple linguistic adaptation methods are considered standard practice (Al Maskari, Melville, & Willis, 2018; Soto et al., 2015); however, this process often occurs with minimal methodological rigor and is poorly described within final peer-reviewed publications, which is a common issue in the field more broadly. Two meta-analyses of translations and cultural adaptations of screening measures in research found minimal descriptions of the translation processes and a lack of rigor in translation and adaptation (Al Maskari et al., 2018; Soto et al., 2015).

Transadaptation is necessary in order to obtain reliable and valid data when using evaluative measures (Zucker et al., 2005). Transadaptation should be considered for measures used in diverse contexts, as a poorly modified instrument can yield over or underestimation of prevalence and subsequent diagnosis rates (Soto et al., 2015). Without consideration through a cultural lens, risk and impairment across social communication and restrictive and repetitive behaviors can be falsely reported, further delaying access to intervention for optimal outcomes (Dubay & Watson, 2019). To date, very little has been published on the transadaptation of the MCHAT-R/F.

Considering how the sociocultural differences described above are embedded throughout the MCHAT-R/F items, it is likely that cultural differences on some items influence the overall properties of the measure (Harrison et al., 2017; de Leeuw et al., 2020; Norbury & Sparks, 2013). Relatedly, one meta-analysis found that while some items on the MCHAT-R/F performed consistently across cultural contexts, others did not (Stevanovic et al., 2022). For example, six items were found to show similar rating endorsements across eight countries that differ by

resource level, suggesting that these items may be universal indicators of ASD. Specifically, across seven countries in this study, there were consistently high endorsements for items such as: points to get help, points to show, brings things to show, follows a point, follows a gaze, and understands what is said. In contrast, items like makes eye contact, responds to name, hearing concerns, and reciprocal smile were endorsed highly in some countries but low in others, suggesting potential cultural differences. Additionally, these cultural endorsement differences appeared to vary between high-risk or low-risk screener populations. For instance, items like plays pretend and social referencing appeared more culturally distinctive amongst those who screened as high risk, while items like interest in other children, imitates actions, gets parent to watch, unusual finger movements, and upset by everyday noises differed most among low-risk samples (Stevanovic et al., 2022). Some cultural differences have begun to be estimated at an item level, and this research seeks to explore the manner in which cultural interpretation manifests at this level through cognitive interviewing. This would offer vital information to support the ways in which screeners like the MCHAT-R/F should be adapted to ensure cross-cultural validity.

Cultural Variability in Social Constructs at an item level

Research finds that culture impacts social norms and interactions, as each culture has implicit and explicit rules that guide a person's actions and perceptions (Clark, 1992; Clark & Robboy, 1992; Freeth et al., 2014). Specifically, research has uncovered cultural variability in play with partners (Lancy, 2007) as well as in types of play (Farver & Howes, 1993). For instance, in some cultures, it is not considered appropriate for parents to engage in pretend play with their children. Thus, the valuation of the pretend play construct may be inaccurate (Pufall & Pufall, 2008). Constructs of protodeclarative pointing and reciprocal social smiling have also

been found to vary based on cultural values divergent from western culture (Golson et al., 2022; Wormann et al., 2012). In considering the construct of eye contact, there are significant differences in its use and absence in societies outside of the west (Akechi et al., 2013; McCarthy et al., 2006; Uono & Hietanen, 2015). Cross-cultural differences are also seen in the expression of non-verbal facial cues (Marsh et al., 2003) and in how facial expressions are processed and interpreted (Yuki et al., 2007). Research has shown differences in social engagement among children in cultures that promote autonomy compared to those that value socialization, as well as differences in peer engagement (Wormann et al., 2012). Cross cultural differences have also been identified in sensory behaviors (Caron et al., 2012), challenging behaviors (Chung et al., 2012), social skills (Sipes et al., 2012), and overall variability in ASD symptom presentation across cultures (Freeth et al., 2013; Matson et al., 2017, 2011).

Taken together, social constructs like pretend play (Pufall & Pufall, 2008), eye contact (Akechi et al., 2013; Uono & Hietanen 2015, Yuki et al., 2007), peer social engagement (Golson et al., 2021; Sicorello et al., 2019), reciprocal smiling (Wormann et al., 2012), and protodeclarative pointing (Golson et al., 2022) have been identified to vary across cultures. This may lead to pathologizing behaviors that vary due to normative cultural rules as deficits (Harrison et al., 2017; Mandell et al., 2007). Sociocultural interpretations of symptomatology are important to explore, as culturally diverse individuals often experience greater delays in receiving an ASD diagnosis (Sansosti et al., 2012). For example, due to cultural differences in gender roles, girls are often identified later than boys in some cultures (Al-Salehi & Al-Hifthy, 2009). In India, parental reports indicate that impairments in children are identified months later than children in the U.S. due to cultural differences (Daley, 2004). It is not clear if ASD presentation truly differs across cultural groups or if the utilized measures are unable to capture

symptom presentation due to cultural variability. Although many items within the MCHAT-R/F examine social behaviors with cross-cultural variability, little research has examined the cognitive processing of diverse individuals that may impact screener outcomes. With this, the goal of this research is to examine cognitive thinking on an item-level to examine potential bias within the MCHAT-R/F.

MCHAT-R/F: Next Steps for cross cultural insight

Cognitive interviewing is an evidence-based method used to empirically investigate how individuals process and respond to questionnaires or surveys to examine if these measures fulfill their intended purpose (Willis & Artino, 2013). Questionnaires used to investigate a myriad of medical topics may seem straightforward, but many items believed to be clear are often significantly misinterpreted, resulting in failure to measure their intended purpose (Tourangeau et al., 2000; Willis et al., 2005; Willis & Artino, 2013). Cognitive interviewing has been conceptualized to not only explore individuals' mental processing but to also consider the sociocultural context that may influence a participant's ability to respond meaningfully and accurately (Gerber, 1999). It focuses on individualistic mental processing of items as well as the cultural context that may influence how questions capture the participant's experience (Gerber, 1999). It relies on a 4-stage cognitive model that includes comprehension, retrieval of information, judgement, and selection of a response (Tourangeau, 2000). Specifically, Cross cultural cognitive interviewing (CCCI) has been implemented across languages, cultures, and countries in order to uncover difficulties in answering questions. It highlights areas of inequivalence between original measures and adapted ones (Willis, 2015). By overtly exploring culture in this manner, sociological and anthropological perspectives are included to support cultural nuance (Gerber, 1999; Miller et al., 2014; Willis, 2015). Cognitive interviewing

methodology suggests a data-driven approach in that sample size is determined based on empirical evidence of saturation rather than a predetermined number to prioritize the identification of cognitive processing over all else (Bogdan & Bilken, 2003).

The purpose of the Current Study

Given the widespread use of the MCHAT-R/F, it is imperative to investigate this measure on an item level to ensure it has utility across cultures. Social values are shown to differ across cultural contexts and can result in item bias for certain cultures if not addressed (Matson et al., 2017). Simply translating the MCHAT without culturally adapting it for the population it will serve increases the risk of measurement bias (DuBay & Watson, 2019), thus, it may be necessary to undertake the transadaptation of the MCHAT-R/F in order to obtain more reliable and valid data across cultures (Zucker et al., 2005). Cognitive interviewing is recommended for the purpose of identifying measurement cultural adaptation needs (Andersen et al., 2014; Maillefert et al., 2009; Willis et al., 2005). This methodology offers insight into the responder's thought process, including their decision-making process at an item level, revealing possibilities of response challenges (Willis et al., 2005).

Qualitative research like this is used to gain an in-depth understanding of how individuals interpret and respond based on their unique experiences within a complex sociocultural context. Unlike quantitative research that emphasizes statistically significant measurement for generalizability, qualitative methods allow researchers to examine more nuanced perspectives to uncover challenges and highlight voices of often underrepresented populations. This method is particularly well suited to identify issues that may not be fully captured through traditional quantitative techniques. To gauge the need for adaptation of the MCHAT-R/F, the current study uses cognitive interviewing to evaluate response differences and comprehension across

participants from different cultures. Specifically, this study examines how a group of parents—distinct in culture, language, socioeconomic status, and acculturation—interpret the items on the MCHAT-R/F, using a four-stage cognitive processing model as a lens to pinpoint cultural influences on participant responses. Differences in interpretation are analyzed through a qualitative interviewing process.

Chapter 2: Method

Participants

Inclusion Criteria

Initially, parents of children ages 16–30 months were the target demographic for recruitment, as this is the age range for screening using the MCHAT-R/F. However, the specificity of this range limited recruitment feasibility. Thus, the upper limit was extended to 48 months of age. The child age at the time of the interview ranged from 16 to 48 months ($M = 30.6$, $SD = 10.1$). This approach aligns with the reality of clinical practice, where the MCHAT-R/F is sometimes administered outside of the recommended age range, particularly in pediatric settings where screening opportunities may have been delayed (Wieckowski, 2023). For children older than the MCHAT-R/F target range of 16–30 months ($n = 7$), caregivers were instructed to reflect on their child’s behavior during that developmental window. Importantly, consistent with best practices in cognitive interviewing and questionnaire validation, participants were not required to have a child with a diagnosed developmental disability. This approach is common in the field, as the goal of cognitive interviewing is to examine how items are interpreted and not to confirm diagnostic accuracy (Willis, 2005). Inclusion criteria instead prioritized cultural and/or linguistic diversity. Culture was defined as the values, norms, and traditions that influence thoughts, interactions, behaviors, and judgements within a given group (Chamberlain, 2005). Participants of various familial countries of origin were prioritized, as well as participants whose primary language(s) for communication was not English.

Recruitment

Caregivers were recruited through emails sent to local childcare centers, parenting community groups, and graduate student groups, as well as printed fliers posted around town (e.g., childcare settings, libraries, churches), and targeted Facebook advertisements. To specifically reach diverse participants, recruitment targeted diverse graduate student listservs, international student listservs, and preschools/daycares with diverse families. Participants were initially reimbursed \$20 (later increased to \$30 to increase the feasibility of recruitment) for their participation in an approximately one-hour-long interview.

Participant Demographics

The final sample included 14 caregivers of children between 16-48 months of age. Cognitive interviewing methodology suggests a data-driven approach to sample size selection, determined by empirical evidence of saturation rather than a predetermined number. That is, interviews stopped once no new ideas or problems emerged, referred to as reaching saturation (Bogdan & Biklen, 2003). In alignment with quantitative benchmarks for survey or questionnaire investigation, a sample size of 30 participants was estimated for interviews (DePaulo, 2000; Malterud, 2016; Perneger et al., 2015). However, the final sample size selection prioritized the gold standard approach of determining sample size based on saturation using cognitive interviewing methodology (Hennink et al., 2017, Schultz & Whitney, 2005; Willis, 2005). Saturation was assessed iteratively during the data analysis processes. Prior research in cognitive interviewing and qualitative methodologies suggest that saturation is often met within 5-15 interviews (Guest et al., 2006). Thus, this study did not rely on a predetermined sample size based on statistical power calculation like in quantitative research, as qualitative research should

be more flexible and rely on the richness of data to identify cognitive difficulties (Morse, 1995; Willis, 2005). Instead, data collection was stopped after no new ideas (codes) were introduced, indicating saturation was met. University Institutional Review Board (IRB) approval of the study protocol and electronic informed consent were obtained prior to the completion of any study procedures.

The majority of respondents identified as female mothers (71.4%) and were in the age range of 30-39 (71.4%). Participants represented a range of racial and ethnic backgrounds, including Asian (50%), White (35.7%), Black or African American (7.1%), and Latine (7.1%). The majority of participants (64.3%) were categorized bicultural based on their scores on the Stephenson Multigroup Acculturation Scale (SMAS), with smaller proportions identified as traditional (21.4%) or assimilated (21.4%). Participant worldview orientation, as measured by the Auckland Individualism and Collectivism Scale (AICS), was primarily collectivistic (64.3%). The majority of participants were first generation immigrants (57.10%). Participants spoke multiple languages which included: Urdu (n=2), Bengali (n=2), Spanish (n=2), Korean (n=1), French (n=1), Turkish (n=1), Congolese (n=1), Russian (n=1), Hindi (n=1), Marathi (n=1), German (n=1), and Hebrew (n=1). Countries of origin across the sample included: USA (n=3), Pakistan (n=2), Bangladesh (n=2), India (n=1), Russia (n=1), Germany (n=1), France (n=1), South Korea (n=1), China (n=1), and Democratic Republic of the Congo (n=1). The time living in the U.S. varied widely, ranging from 2 years to a lifetime (M=15.43 years, SD=11.62 years). Socioeconomic status (SES) was also assessed via self-reported household income using predefined income brackets. Participants represented a wide range of income levels, from under \$25,000 to over \$200,000 annually. Specifically, 1 participant (7.1%) reported a household

income between \$15,000–24,999; 1 (7.1%) between \$25,000–34,999; and 2 (14.3%) between \$35,000–49,999. Three participants (21.4%) reported incomes between \$50,000–74,999, while 4 participants (28.6%) reported income between \$75,000–99,999. Three participants (21.4%) reported higher income brackets, with 3 (21.4%) between \$150,000–199,999 and 1 (7.1%) over \$200,000. This diversity in household income allowed for representation across low, middle, and high SES groups (See Appendix Table 2).

Procedure Overview

This qualitative study consisted of interviews with individuals from culturally and linguistically diverse backgrounds and was designed based on the FDA PRO Guidance for cognitive interviewing (USDHHS, 2006). The study began with informed consent, followed by the completion of a sociodemographic measure and two additional questionnaires online through a Qualtrics platform, which took approximately 20 minutes. The primary component involved the qualitative interview, which is analyzed in detail below. The interviews lasted approximately 60 minutes.

Measures

Each participant completed demographic information and two questionnaires. The first questionnaire assessed the level of acculturation, and the second assessed participant worldview.

Sociodemographic Questionnaire

The first, sociodemographic questionnaire gathered information about the participant's age, gender, and generational status. Generational status was collected because all individuals, regardless of race, undergo a process of change to adapt to a society different from their society of origin (Berry & Sam, 1997). Thus, the societal immersion of a first-generation immigrant

differs from that of a fourth-generation immigrant (Suzuki et al., 1996). Additionally, information was collected about the participant and their parents, including country of origin, spoken language, years in the U.S., race, ethnicity, country of origin, and sociodemographic status. Additionally, parents provided their child's age. These questions help identify trends related to this study's thematic findings and identify manners in which to characterize the sample.

Acculturation Questionnaire

The Stephenson Multigroup Acculturation Scale (SMAS) looks at acculturation as an individual-level phenomenon and provides an index of the degree of cultural immersion (Stephenson, 2000). Specifically, it calculates acculturation as the degree of immersion in dominant and ethnic societies. It assessed acculturation across several domains, including language, interaction, media, and food, with each domain reflecting knowledge, behavior, and attitudes. The SMAS uses a 4-point Likert scale ranging from 1 (false) to 4 (true) across 32 questions to examine dominant society immersion (DSI) with 15 items and ethnic society immersion (ESI) with 17 items. These scores identify individuals as traditional, bicultural, or assimilated. The SMAS has confirmation of the two-factor model and has been found to have solid reliability (Huynh et al., 2009; Stephenson, 2000).

Collectivism Individualism Questionnaire

The Auckland Individual Collectivism Scale (AICS) is a self-report questionnaire that containing 26 items specifically designed to measure individualism and collectivism domains. These domains are made up of five subdomains. The collectivism domain consists of harmony and advice seeking subdomains. The individualism subdomains consist of responsibility,

competitiveness, and uniqueness. Each question is answered using a Likert scale from 1 (Strongly Disagree) to 5 (Strongly Agree). The scoring of AICS produces both individualistic domain and collectivist domain scores based on the means of their corresponding subdomains. A unitary individualist and collectivist score is used to reflect the continuum of each participant's individualism and collectivism. This score is calculated by subtracting the collectivist score from the individualist score, creating a scale ranging from -4 to 4. Scores above 0 equate to higher levels of collectivism and numbers below 0 equate to higher levels of individualism. Participants with scores above 0 will be categorized as having an individualistic point of view while those below 0 will be categorized as having a collectivistic point of view. While it is technically possible to receive an extreme score such as 0, such scores are uncommon and may indicate issues such as missing data, uniform response patterns, or participant disengagement (Shulruf et al., 2007).

Cognitive Interviewing

Measure Investigated: MCHAT-R/F

The Modified Checklist for Autism in Toddlers (M-CHAT) is a 23-item screener questionnaire given to caregivers of children aged 16-30 months to assess autism risk. It was originally adapted from the Checklist for Autism in Toddlers (CHAT), with modifications that removed the observation and parent report sections. The screener looks at constructs typically associated with autism, including: response to joint attention, initiation of joint attention, social orienting, protodeclarative and protoimperative pointing, social interest in peers, pretend play, behavioral requesting via nonverbal gestures, responsive social smiling, sensory sensitivity, eye contact, imitation, and receptive language (Robins et al., 2014). The scale also includes four

questions about general development designed to help reduce the face validity of the screener. The screener was designed to be at a 6th grade reading level and does not require formal training to administer. These factors improve ease of administration, making the screener quick and convenient for physicians managing many patients. The scoring considers 0-2 a low risk or negative screen, 3-7 a medium risk, and scores 8-20 are considered to be high risk for ASD (Robins, 2001).

The newest version of the M-CHAT or Modified Checklist for Autism in Toddlers Revised/Follow up (MCHAT-R/F) is shorter with 20 questions (See Appendix Table 1). This reduction in questions reduces affirmative response bias by excluding 3 of the original questionnaire items and rearranging the question order (Khowaja, 2015). The MCHAT-R/F also provides brief examples and uses simpler wording. These modifications were made to decrease the rate of false positives. Research has found that children who had a positive screen score are 114 times more likely to receive a later ASD diagnosis than those who screened negative. The screener was found to identify higher rates of autism while simultaneously reducing unnecessary follow-ups. The main goal of the screener is to maximize ASD detection through high sensitivity, therefore false positives continue to occur and many children with positive screens may not meet criteria for an ASD diagnosis. Consequently, children who fall into the medium risk category are strongly encouraged to complete an MCHAT-specific Follow-Up interview component (Robins et al., 2014).

Cognitive Interviewing Procedure

Cognitive interviewing is recommended for the purpose of identifying measurement cultural adaptation needs (Andersen et al., 2014; Maillefert et al., 2009; Willis et al., 2005). This

methodology offers insight into the responder's thought process, including their decision-making process at the item level, which can reveal potential response difficulties (Willis et al., 2005). To gauge the need for adaptation of the MCHAT-R/F, the current study uses cognitive interviewing to explore differences in response and understanding across participants from different cultures on each item of the MCHAT-R/F. The cognitive interviews and data analyses were conducted by researchers familiar with autism and were presented using a semi-structured interview guide. The guide included probes to promote consistency across interviews and to facilitate discussion surrounding the participant's cognitive processing of items.

The interview portion used a hybrid of two cognitive interviewing approaches: thinking aloud and verbal probing called concurrent probing. The thinking aloud approach encourages respondents to verbalize their thought process while completing the screener. Additionally, during a verbal probing component, respondents were asked questions about their understanding of the items and thoughts that would contribute to their decision to select an answer (Andersen et al., 2014). The study team, including individuals with clinical assessment and autism experience, reviewed and discussed interview transcriptions to determine whether the saturation point had been reached. Interviews were conducted through UGA's Zoom platform with audio and visual exchange. The audio was recorded for transcription and data gathering purposes using a voice memo application provided by Apple, but video was not recorded to provide additional protections to confidentiality.

Cognitive Interview Guide

An interview guide was created to establish consistency across interviews and facilitate discussion with the participants. The guide includes questions and prompts that are related to the

cognitive process when selecting an answer. Questions were developed based on guidelines by the US Department of Health and Human Services Food and Drug Administration (2006) formulated from sources of research (Beatty & Willis, 2007; Willis, 1999; Willis et al, 2005).

The interviews began with a brief introduction and warm-up as follows:

Thank you for participating in this cognitive interview. Your feedback will help us learn more about the utility of screening measures for children. The purpose of this interview is to find out how you think during the screening process and the screening material itself. One approach we will be using will be ‘Thinking aloud’ which may be new and unfamiliar to you, but please know there are no wrong answers. I am only interested in knowing what is going through your mind. Any information you provide is helpful. I did not design any of the materials and you will not hurt my feelings. Feel free to say anything you’re thinking. Before we begin the actual interview, I’d like to ask you a ‘warm-up’ question to introduce you to the think aloud process. Try to visualize the place where you live and think about how many windows there are in that place. As you count the windows, tell me what you are seeing and thinking about. (Willis, 1994).

Corrective feedback was provided as necessary to ensure the participant understood the task.

Once the participant demonstrated understanding of Thinking Aloud the interview continued with using the MCHAT-R/F.

Please answer these questions about your child. If your child is over 30 months old think about their behavior between 16 and 30 months of age. Keep in mind how your child usually behaves. If you have seen your child do the behavior a few times, but he or she

does not usually do it, then please answer no. Please respond yes or no for every question and tell me what is going through your mind. If you point at something across the room, does your child look at it? (For Example: if you point at a toy or an animal, does your child look at the toy or animal?) Yes or No

After thinking aloud, verbal probing questions were asked on an item level. There are six categories of verbal probing questions: Comprehension/interpretation, Paraphrasing, Confidence judgement, Recall, Specific, and General (Willis & Artino, 2015). Examples for each category include, “What does the term mean to you?” (comprehension/interpretation), “Can you repeat the question I just asked in your own words?” (paraphrasing), “How sure are you that your child looks?” (confidence judgement), “How did you come up with your answer?” (recall), “Is pointing something you would expect your child to do?” (specific), “How did you arrive at that answer?” (general).

Research also suggests noting participant reactions (e.g., facial cues, hesitancy to respond, nervous laughter) as these reactions could indicate inappropriate question functioning (Bracken & Barona, 1991). Transcripts captured any possible acquiescence response bias and/or socially desirable responding on this parent-report instrument by denoting the timing of response, perceived level of confidence, and reported evoked emotions. Additionally, data related to stigma surrounding symptoms related to child developmental disabilities within the culture (Cook et al., 1997; Reichenheim & Moraes, 2007; Sousa & Rojjanasrirat, 2011) was also gathered from interviews.

Data Analysis

Researcher Reflexivity Statement

This dissertation represents more than academic inquiry, it reflects a deep commitment to justice in early identification of autism and to honoring the diverse voices of caregivers navigating developmental systems. As a Black woman, researcher, clinician, and advocate, I entered this work not merely to examine cultural difference, but to amplify it. My personal and professional experiences have shaped my awareness of the ways in which culturally divergent perspectives are often overlooked or mischaracterized in research and practice. I sought to create space for the questions caregivers hesitate to ask, for the culturally incongruent moments that often go unseen, and for the nuanced responses that do not fit neatly into binary constructs. Throughout this project and my clinical internship, I have witnessed the profound emotional and logistical burdens families carry when engaging with systems that were not designed with them in mind. These experiences underscore the necessity of research that centers sociocultural context and caregiver voice, particularly within early autism identification. I remain grounded in the belief that seeking understanding and honoring lived experience is not radical, but it is essential. Research on diverse voices is a responsibility we bear as scientists and advocates for human-centered systems of care.

In analyzing the data, I was joined by a team of coders from varied cultural, linguistic, and academic backgrounds. We engaged in ongoing dialogue around how our personal identities and lived experiences informed our interpretations of participant responses. Regular meetings were held to discuss emergent codes, ensure consistency in applying the coding framework, and reflect on any assumptions we brought into the analytic process. This reflexive practice was

central to maintaining the integrity of the qualitative analysis and to honoring the phenomenological approach of centering participants' meaning-making processes.

Data Preparation

Transcripts were converted from audio to text using Otter Artificial Intelligence. An initial individual with transcription experience and knowledge of autism reviewed the quality of these transcriptions by listening to the interviews, comparing them to the produced transcription, and making adjustments as necessary. A second individual cleaned 100% of the transcriptions and served as a quality assurance check. The cleaner ensured that personally identifiable information within transcriptions was removed and that transcription errors were corrected. A third individual reconciled any discrepancies in initial transcription and cleaning data.

Qualitative Analysis

Thematic coding with a blended approach was used to allow for both existing codes to be used and novel codes to emerge. An inductive approach was also used in that four coders worked to develop labels to apply to participant quotes obtained from the interviews. The inductive codes were established by four coders who reviewed several initial interviews separately, made judgements about the meanings of the participants' cognitive process, assigned a possible label, and offered a preliminary definition of the label. The coders convened to review their codes until a consensus of labels were identified (Ryan & Bernard, 2000). Additional codes were added iteratively throughout the process until saturation was met, meaning no novel themes emerged (Fonteyn, 2008). This categorization and interpretation was guided by a phenomenological approach to data analysis in that the participant's lived experiences in relation to survey questions and constructs are applied to the interpretation of their response (van Manen, 1997).

The identification of themes provides insight into differences in-item level interpretations of the screener questions on the MCHAT-R/F. Transcriptions were examined using deductive strategies and categorized based on the four steps of the cognitive process: comprehension, retrieval, judgement, and response (Tourangeau et al., 2000). Item-level cognitive processing differences across cultures were examined based on questions developed by cognitive interviewing researchers (Beatty & Willis, 2007; Gerber, 1999; Willis, 1999). In order to examine differences related to comprehension, the interviewer considered question intent (e.g., What does the participant believe the question to be asking?) and meaning of terms (e.g., What do specific words and phrases in the question mean to the participant?) (Willis, 2005). In order to examine differences related to retrieval from memory, the interviewer considered the recallability of information (e.g., What types of information did the participant use to recall in order to answer the question?) and recall strategy (e.g., What type of strategies are used to retrieve information? Does the participant tend to recall events through specific and individual, or do they use broad and generalized strategies?) (Beatty & Willis, 2007; Willis & Artino, 2013). In order to examine differences related to response processes, the interviewer considered how the participant may map the response (e.g., Can the participant match their internally generated answer to the response categories given by the survey question?; Tourangeau et al., 2000; Willis, 2005). Additionally, to examine differences related to response processes, the interviewer considered participant perception of response sufficiency given the yes or no categories, response confidence, and response latency (e.g., Does the participant take a long time to answer the question? Does the participant respond yes or no, or offer a contextualized alternative?) (Beatty & Willis, 2007; Miller et al., 2014; Willis & Artino, 2013). This research aims to delve

further into the cognitive processing model by investigating the role culture plays in how individuals cognitively process a social value-based survey. In order to examine this, specific questions related to culture were also asked to ascertain the influence on participants' responses. Cultural influences on interpretation were examined through specific probes about whether the behavioral construct was culturally typical, valued, or influenced by particular sociocultural expectations (Gerber, 1999; de Leeuw, 2011; Miller et al., 2014).

In order to maintain the integrity within a phenomenological approach, the focus of this study is not primarily on the frequency in which particular responses were elicited, but on the depth and significance of each participants' lived experience (Creswell & Poth, 2018; Moustakas, 1994). Thus, terms like some, most, or all will not be used to imply statistical generalizability or representativeness, but instead the analysis offers a sense of the range of experiences shared. Each experience is interpreted as meaningful in its own right and contributes to a broader understanding of the phenomenon of culture's influence on the cognitive processing of an autism screener (van Manen, 2016). Pseudonyms and gender-neutral pronouns (i.e., they/them/their) will be used to personalize participant narratives while maintaining anonymity. Responses were coded and categorized into themes and subthemes (Braun & Clarke, 2006) of recurring concepts within a codebook used in NVIVO 15 data analysis software.

Overview of Coding Framework

This study examined the interpretative difficulties by cognitive domains (i.e., comprehension, retrieval, judgement, response) across all 20 MCHAT-R/F items, which allowed for the identification of where culture influenced the interpretation process. A subset of the most relevant items was included in the analysis (see the section below titled, Items Analyzed) to gain

conceptual clarity about how caregivers engaged with and understood the key screener questions. The qualitative data extracted from the semistructured cognitive interviews were analyzed using a comprehensive codebook that was structured to align with Tourangeau's four-stage cognitive response model. It included codes for all four domains (i.e., Comprehension, Retrieval, Judgement, and Response) and a fifth domain (i.e., Cultural Relevance) to uniquely examine sociocultural influences on cognition. These code domains and subcodes were formulated based on prior empirical and theoretical literature in cognitive interviewing and cross-cultural survey methodology (Beatty & Willis, 2007; Gerber, 1999; Tourangeau et al., 2000; Willis, 2005), which provided guidance for structuring questions and categorizing interpretation challenges across comprehension, recall, judgment, response, and cultural relevance dimensions.

The initial codebook was developed with 26 a priori codes grounded in cognitive interviewing theory and previous research on cross-cultural survey response. As coding progressed, iterative team discussions led to the addition of 9 inductively derived subcodes to capture emerging patterns of emotional nuance, context-specific reasoning, and cultural interpretation. These included refinements in the judgment domain (e.g., distinguishing emotional reactions about the child vs. other sources, and adding parallel positive valence codes), and the response domain (e.g., specifying development-specific and general forms of ambiguous "sometimes" answers). In the cultural relevance domain, new subcodes were added to reflect gender-specific valuation (V.B.3.iii), uncertainty about cultural relevance (V.C.1), and stigma (V.D.1), the latter used when caregivers expressed community shame or fear of judgment surrounding a developmental behavior. These refinements brought the final code count to 35 and allowed for greater analytic precision while preserving the phenomenological intent to center

participant meaning making in context. The domains included distinct features and ways in which each caregiver understood, recalled, judged, and responded to MCHAT-R/F items. Saturation was considered met as no new themes or codes emerged from subsequent transcripts and all qualitative data fit meaningfully within the existing framework (Guest et al., 2006; Hennink et al., 2017).

Within the comprehension domain, codes were used to examine accurate understanding of item intent. Subcodes in this domain (i.e., full, partial) were used to capture response nuance, alongside methods through which understanding was expressed (i.e., examples, definitions, question repetitions). Codes to note inaccuracies and requests for clarification were also used. The retrieval domain codes were used to examine methods participants used to access information relevant to the social constructs in each MCHAT-R/F item (i.e., direct observation, indirect observation/reports, specific/personal memories or examples, broad examples or definitions). Codes were also used to denote inaccurate retrievals, timing in retrieving a relevant example (i.e., immediate, delayed, assisted), and participant confidence in their choice. This highlighted difficulty in recalling behavior. The judgement domain codes were used to examine participants' perceived difficulty (i.e., easy, moderate, hard) and emotional responses (i.e., negative, neutral, positive). The response domain codes were used to examine participants' ability to appropriately use binary yes or no response options. Codes captured the nuance within participant reasoning when binary options were deemed unfeasible (i.e., context specific, development specific, generalized sometimes yes/sometimes no) to reflect the complexity of their cognitive process.

The cultural relevance domain codes were used to examine how cultural background intersected in item interpretation. Codes examined participant ability to denote cultural and/or personal congruence or incongruence (i.e., personal incongruence, culturally congruent, culturally incongruent, partially incongruent, not sure). Participants reported instances that coders felt were stronger than incongruence of cultural values that was captured with the additional code culturally stigmatizing. For example, if a participant reported parent-child play interaction is not valued in the household, this suggests incongruence, whereas worry of judgement or community shame related to a behavior suggests stigma. These codes helped to operationalize a cultural incongruence when responses (explicit or inferred from context) suggested a notable divergence between the Western-based construct of the MCHAT-R/F item and participant cultural norms, values, or expectations. This was most clearly seen in the incongruence codes while the personal incongruence code was used for additional response nuance. This code was used when a participant's household values clashed with their cultural origin values and were thus forced to negotiate between the value of the item. For example, if a participant reported their culture does not value the item but they do. These judgements were based on coder consensus guided by participant statements and provided a framework for disaggregating results by cultural incongruence, as detailed in the following sections. This coding framework allowed for an analysis of complex caregiver responses across Tourangeau's four cognitive domains and highlighted cultural intersection within interpretations. It enabled coders to easily identify areas participants easily understood or misunderstood and offered details related to those misunderstandings.

Data Organization

In alignment with the phenomenological approach to qualitative data analysis, the intent was not to quantify prevalence but to highlight patterns of interpretive complexities that may influence caregiver responses on autism screeners. To further center participants' lived experiences and to explore the role cultural frameworks may have on interpretation, we evaluated responses in two ways. First, we examined themes across the entire group. Second, to hone in on those participants whose responses were disaggregated based on whether they explicitly endorsed a cultural incongruence or difference of social values with a survey item. This distinction provided a more nuanced understanding of how sociocultural norms can shape each cognitive processing domain. Participants who specifically endorsed a cultural incongruence often provided distinct examples of how developmental constructs conflicted with culturally informed family values and social norms. In contrast, participants who did not identify an incongruence still offered meaningful insights. The responses of participants who did not endorse an incongruence suggest more implicit ways in which culture can influence cognitive interpretation. While not always consciously acknowledged, cultural frameworks can shape how individuals evaluate developmental behaviors by functioning outside of one's awareness yet remaining aligned with norms within their sociocultural background (Geertz, 1973; Greenfield, 1997; Rogoff, 2003). These layered patterns of interpretation could uniquely contextualize response differences within both explicit cultural values and more deeply embedded beliefs and practices.

Items Analyzed

The data obtained from the cognitive interviews were coded using a structured qualitative framework aligned with Tourangeau's cognitive response model that included comprehension, retrieval, judgement, and response (Tourangeau, 2000). To capture additional nuance relevant to this study's focus on sociocultural variability, a fifth domain of cultural relevance was added and inductively developed by the research team. This domain was added to identify responses that reflected cultural norms, values, and/or expectations that may have shaped caregiver interpretation of screener items. Though not a part of Tourangeau's original model, the addition of this domain allowed for systematic documentation of culturally specific interpretations that may otherwise go overlooked. Consistent with prior research, specific MCHAT-R/F items were classified as filler or developmental milestone questions, which were incorporated during the measurement development process with the intention to balance the screener's focus and reduce face validity, thereby minimizing caregiver awareness of autism specific screening goals (Robins et al., 2001). From the literature, the following items were identified as fillers and excluded from analysis in the current study: Item 2 ("Have you ever wondered if your child might be deaf?"), Item 4 ("Does your child like climbing on things?"), Item 13 ("Does your child walk?"), and Item 20 ("Does your child like movement activities?"). Following the removal of the filler items, 16 MCHAT-R/F items were retained for analysis.

Chapter 3: Results

Total Cognitive Domain Results Across All Participants

Item-level Cross-Domain Difficulties

Across the 16 MCHAT-R/F items analyzed in this study, all items were associated with at least one instance of interpretive difficulty, with the number of reported difficulties ranging from 1 to 12 per item. While this study does not attempt to assign weighted value to these frequencies, examining the distribution of interpretive complexity across items provided insight into which questions elicited the broadest variation in cognitive response (See Appendix D1).

Low Concern Items. Items with the lowest concentration of difficulties, defined in this study as items flagged by four or less participants, included Question 15 (“Does your child try to copy what you do?”) with 1 participant, Question 11 (“When you smile at your child, does he or she smile back at you?”) and Question 9 (“Does your child show you things by bringing them to you or holding them up for you to see – not to get help, but just to share?”) with 3 participants each, as well as Question 7 (“Does your child point with one finger to show you something interesting?”) with 4 participants.

Moderate Concern Items. There were also a cluster of items with more moderate interpretive difficulties, defined in this study as items flagged by 5-8 participants. These included Question 3 (“Does your child play pretend or make-believe?”), Question 6 (“Does your child point with one finger to ask for something or to get help?”), Question 8 (“Is your child interested in other children?”), Question 10 (“Does your child respond when you call his or her name?”),

and Question 14 (“Does your child look you in the eye when you are talking to him or her, playing with him or her, or dressing him or her?”), each with 5–6 participants highlighting cognitive challenges. Question 17 (“Does your child try to get you to watch him or her?”) also fell into this midrange group with difficulties uncovered among 7 participants.

High Concern Items. Items with the highest concentration of interpretive difficulty were defined as items flagged by 10 to 12 participants. These included Question 1 (“If you point at something across the room, does your child look at it?”) and Question 5 (“Does your child make unusual finger movements near his or her eyes?”), both flagged by 11 participants; Question 12 (“Does your child get upset by everyday noises?”), and Question 18 (“Does your child understand when you tell him or her to do something?”), both with noted difficulties among 12 participants; as well as Question 16 (“If you turn your head to look at something, does your child look at your face to see how you feel about it?”) and Question 19 (“If something new happens, does your child look at your face to see how you feel about it?”) both flagged by 10 participants. Hereafter, to maintain brevity while centering the lived experiences reflected in each item, MCHAT-R/F items will be referred to by their question number and associated developmental construct.

Types of Domain Errors Per Item

Many of the MCHAT-R/F items were associated with interpretation difficulties across multiple cognitive domains. Items with the broadest distribution of domain-level errors, defined here as those flagged in three or more domains, included Q16 (response to joint attention) and Q18 (response to verbal directions), both of which showed elevated difficulties in comprehension, retrieval, judgment, and response domains. Items demonstrating difficulties

across two domains, most commonly in retrieval and response, included Q1 (follow a point), Q5 (unusual finger movement near eyes), Q6 (point to request), Q10 (response to name), and Q12 (upset by everyday noises). In contrast, a number of items exhibited domain-specific difficulty, with elevated errors concentrated within a single cognitive domain. This subset with more isolated interpretation concerns included Q3 (pretend play), Q11 (responsive social smile), Q14 (eye contact with interactions), Q15 (imitation), Q17 (social initiation), and Q19 (social referencing; See Tables 3-4).

A majority of the items emerged as having interpretation difficulties in isolated cognitive processing instances, suggesting potential interpretive strain that may be culture context specific by domain. As such, these patterns are noted here for completeness and further contextualized in the discussion (See Table 4, Figure 1)

Domain-specific Interpretation Patterns

Comprehension

Accurate Comprehension. Within the comprehension domain, responses indicating accurate comprehension fell into three categories: participants correctly repeating questions in their own words, offering accurate definitions of the construct, or providing relevant examples that demonstrated comprehension. Participant examples of accurate comprehension included, “If you're trying to get your child to look at something, or to notice something that you're noticing, and you point at it across the room, would they look in or at the object you're pointing at? (repetition via Q1 - follow a point, Interview 2), “Like whatever I do, she does the same.” (definition via Q15 - imitation, Interview 3), and “upset by common noises like blender, juicer,

household appliances, or animals, pets, construction work” (relevant examples via Q12 - upset by everyday noises, Interview 1).

Inaccurate Comprehension. There were also responses within this domain that conveyed inaccurate understanding, which fell into three common themes: i) reconceptualizing the behavioral construct as something slightly different from the question content, ii) a noted lack of understanding of the question, and iii) tangential/unrelated responses.

In terms of reconceptualizing constructs, one example is that participants redefined spontaneous child-parent dyadic social behaviors as prompted instructional acts or task-oriented exchanges. Participant examples of these inaccuracies include: “Does my child smile to me when I’m calling her name” (reconceptualizing via Q11 - responsive social smile, Interview 1), “How does your child respond to clear communication (reconceptualizing via Q11 - responsive social smile, Interview 9), and “How does your child react to something funny you do” (reconceptualizing via Q19 - social referencing, Interview 9). In these quotes, caregivers have shifted away from the casual social intent and instead reconceptualized the item to relate to social reactions prompted by parents, rather than spontaneous child behaviors. This reinterpretation suggests a possible incongruence between the intended construct of an unstructured social engagement and caregivers’ understanding, which appears shaped by more structured, adult-led interactions. Such reframing may reflect differences in caregiver expectations of appropriate social behavior or child-led initiation across cultural or parenting norms.

Within the theme of lack of understanding the question, participants exhibited confusion or uncertainty about the referenced construct and appeared to lack a frame of reference.

Participant response examples include, “He’s my first. So it um- all seems normal to me. I don’t know. I can’t really picture unusual.” (lack of understanding via Q5 - unusual finger movements near eyes, Interview 6) and “ Yeah uh you know sometimes he points two fingers to his eyes...like you see me- You know [from] watching wrestling and he does pointing like that.” (lack of understanding via Q7 - point to share interest, Interview 14). In both cases, caregivers appeared to lack grounding in the developmental expectations behind the item. The first quote reflects uncertainty stemming from having no reference point for what is considered to fit in the term “unusual,” while the second shows an attempt to interpret behavior through a familiar but unrelated reference—a media gesture— rather than the intended developmental cue of protodeclarative pointing. Together, these responses highlight how the absence of a shared developmental or behavioral framework can limit caregivers’ ability to access the intended meaning of the screener item.

Lastly in tangential or unrelated responses, participants responded in ways that were loosely connected or entirely tangential to the item construct. Participant response examples include, “So like, a lot of times, like [if] something [is] falling, I don’t even look. So I feel like that’s like a certain level of like caring about loud noises. I feel like that’s why I use the word cue. I think it’s something where like, I was raised to not snoop and not look” (tangential/unrelated via Q12 - upset by everyday noises, Interview 8), and “She just spends a lot of time with phone. She has developed a some sort of screen addiction but she loves- she loves human beings.” (tangential/unrelated via Q3 - pretend play, Interview 9). In both cases, participants diverged from the developmental behavior being assessed and instead introduced broader reflections on values, philosophies, or unrelated concerns. The first quote conflates the

screeners' focus on the child's sensory reactivity to common environmental sounds, instead centering on the caregiver's upbringing and beliefs about attentional restraint. The second response shifts attention away from pretend play behaviors and toward concerns about media exposure and general sociability. These responses suggest a lack of alignment with the item's intended construct.

Retrieval

This domain provided the highest concentration of difficulties. Here, participants exhibited accurate or inaccurate information retrievals by providing broad memories or examples, specific memories or examples, direct observations of the behavior happening during the interview, or indirect behaviors that other caregivers saw and shared with the participant. Retrieval was evaluated using multiple different criteria, including accuracy and timing of response.

Retrieval Confidence. While retrieving information, participants firmly and succinctly provided responses of aligned examples/memories (high confidence), expressed verbal uncertainty of their examples/memories (low confidence), or displayed nonverbal signs of uncertainty by pausing or waffling over examples/memories (low confidence). Lack of confidence was most prominent on Q12 (upset by everyday noises). A lack of confidence was most evident in the language used, hesitations, qualifiers, and expressions of doubt which were often coupled with difficulties anchoring the response in memory. One example included "Ah... (sigh, pauses) you know, I don't recall. They might not use a lot of finger pointing, but they talk about it. They always said, 'Mom, look, look,' you know. My three-year-old, she already learned this word 'look' so I don't - I don't know if she used a lot of finger pointing, to be honest."

(delayed response and low confidence via Q7 - point to share interest, Interview 12). The uncertainty and hesitation expressed through both verbal and non verbal cues, including pausing, suggests low retrieval confidence.

Retrieval Timing. Furthermore, while the majority of participant responses were retrieved immediately (within 3 seconds), others were delayed (over 3 seconds) and included some hesitancy or waffling regarding their examples. Most participants responded immediately to items, usually by reflecting on familiar contexts within 3 seconds. Delayed timing was most prominent on Q7 (point to share interest). These immediate responses were typically associated with high confidence. One example included, “Oh, when I was putting him to bed last night, yeah, he was looking [me] in the eye and we were singing and stuff. I think he gets excited by things, yeah he does it.” (immediate response and high confidence via Q14 - eye contact with interactions, Interview 6).

Retrieval Accuracy. Responses were deemed inaccurate by the research team if the provided memory, definition, or observation were irrelevant to the item construct. Accurate responses reflected the caregiver’s ability to recall appropriate behavioral examples, definitions, or observations that aligned with the intent of the screener item.

Participant examples of accurate retrievals included: “...just thinking how I always communicate with my baby like this all the time.” (broad memory/example via Q6 - point to request, Interview 9), “This morning, when we were driving into daycare ... As we're driving, I always point to things and go, look, there's a plane, look, there's a fire truck and he turns his head and looks at what I'm talking about.” (personal memory/example via Q1- follow a point, Interview 6), “Mmhmm yeah, I just tried to do that right now. (direct observation via Q16 -

response to joint attention, Interview 1), and “Yeah they said at daycare she like had a bottle and was feeding it to a baby doll.” (indirect observation via Q3 - pretend play, Interview 2). In each of these responses, participants successfully retrieved examples that aligned with the behavioral aspects of the items from broad experiences, personal memories, observations, or reports from others, demonstrating clear connection to the intended developmental construct.

There were also responses within this domain that conveyed inaccurate retrieval that fell into two common themes: surface level recall or intersecting misunderstanding and retrieval. Inaccurate retrieval was most prominent on Q19 (social referencing). Within the theme of surface level recall, caregivers focused on what the child did in terms of the observable features or topography of the behavior, rather than the function or social intent the question was designed to assess. One participant response example included: “She saw the laundry pods and opened the cabinet to get them. That was something new, she didn’t look at me but she went straight for it and kept looking at it.” (surface level recall via Q19 - social referencing, Interview 1). In this case, the caregiver described the child’s attention toward a novel object, but the behavior lacked the referential, social-sharing component the item was intended to measure specifically, whether the child looks to the caregiver to share interest or gain information. Another participant, when asked about social smiling, responded: “When we take her to like grocery stores um and she randomly smiles.” (surface level recall via Q11 - responsive social smile, Interview 3). While the caregiver recalled a specific smiling behavior, the context described involved a more self directed enjoyment in a preferred environment rather than a socially contingent smile directed at another person, their caregiver. In both cases, the recalled behaviors matched the form of what the item referenced (e.g., looking, smiling), but not the function resulting in misalignment between the caregiver's memory and the item’s construct.

Another theme was intersecting misunderstanding and retrieval. Here, participants appeared to conceptually understand the item behavior in isolation but offered examples that still reflected misinterpretation of the intended construct. In these cases, caregivers used language that suggested partial comprehension of the construct (e.g., joint attention or shared focus), but the behavioral examples provided were misaligned with the screener's intent. One participant, when asked about their child's ability to respond to joint attention, responded: "Mmm yeah, just when we are like go[ing] camping, or you notice like an animal and turn around and say, 'hey guys look,' and they turn around follow me, you know." (intersecting misunderstanding and retrieval via Q16 - response to joint attention, Interview 13). This response demonstrates an understanding of a shift to joint attention, but the example describes a caregiver verbally initiating a cue and the child's compliance, rather than the child's ability to spontaneously follow a nonverbal social cue which is a distinct feature of this item. Another participant responded to Q19 (social referencing, Interview 5) with: "I don't like the word 'look,' in this case, because um I think of examples. And I think that is similar to what they're uh soliciting here. Um, like, if there's a garbage truck going by the house, um and she's standing looking at it, she might run and get me. Like, she's not looking at me ... but she's trying to attract my attention to it. Um, it might not be exactly the same." (intersecting misunderstanding and retrieval via Q19 - social referencing, Interview 5). While the caregiver is clearly attempting to align their example with an important construct, the example described does not reflect social referencing, as it lacks the critical element of the child spontaneously checking with their caregiver for emotional or informational input. As with the prior example, this reflects an effortful but ultimately inaccurate retrieval, due to misalignment between the recalled behavior and the construct measured by the item. These responses illustrate how the misunderstanding of a construct may lead caregivers to

retrieve adjacent, but developmentally distinct, behaviors, ultimately resulting in misclassification due to subtle but important differences in the function and intent of the item's behavior.

Judgement

Within the judgement domain, participants shared their perceived level of difficulty along with any emotions the questions evoked.

Perceived Difficulty. Participants were asked to reflect on how hard or easy it was to answer each question after providing their response selection. Most participants described the majority of questions as relatively easy to answer. One example included, “Pretty easy. I think like I-I've been paying attention to the pretend play thing. He um he he's been doing more and more recently.” (easy via Q3 - pretend play, Interview 6), indicating a straightforward judgment process when the construct was familiar and salient. In contrast, a subset of responses expressed difficulty making a clear judgment tied to two themes: the need for additional information and lack of clarity in item wording. Participant examples included: “Um, difficult. I mean, I don't know if difficult is the right word - more like, I can't answer fully without adding to it.” (additional information needed via Q19 - social referencing, Interview 5) and “It was a lot harder, because the question wasn't quite as clear to me as the others.” (lack of clarity via Q5 - unusual finger movements near eyes, Interview 7). Perceived difficulty was most prominent on Q5 (unusual movements near eyes). These responses suggest that ease of judgment was often tied to the familiarity and observability of a behavior. Difficulty emerged when participants encountered vague or abstract wording, restricting their ability to evaluate more nuanced interpretations of their child's behavior.

Emotional Reaction. Within the judgment domain, participants also expressed a range of emotional reactions. Many participant responses indicated positive or even neutral emotions when recalling their child's behaviors. Participant examples included, "Um, glee, excitement thinking about when he plays." (positive emotions via Q8 interest in peers, Interview 6) and "Nothing in particular came up, no." (neutral emotions via Q15 imitation, Interview 5).

However, other participant responses exhibited clear emotional strain or negative emotions related to concerns about their child's development. For example, one participant stated "It made me a little, like, worried about whether my daughter behaves in the normal way." (negative emotions related to the child via Q1 follow a point, Interview 7). This was suggestive of anxiety surrounding developmental knowledge and whether their child was meeting expected developmental milestones. In another example, emotional reactions reflected feelings of personal inadequacy. For example, "That was surprisingly hard because I thought I'm not doing good paying attention to this, now [that] I can't actually think of an example." (negative emotions related to others via Q7 point to share interest, Interview 6). This suggests that caregivers may internalize their difficulty recalling specific examples as a personal failing, reflecting broader pressures as caregivers to be highly attuned to their child's behavior and development.

Response

In the response domain, the participants were asked if they felt the yes or no options were sufficient or if they needed to qualify their answers. Most participants indicated that mapping a response selection onto a binary yes or no was sufficient. For example, one participant's response example included: "Uh yes and no was just fine." (binary sufficiency via Q1 follow a point, Interview 8).

Other responses suggested inadequacy of a binary fit in which participants need to provide qualifying information. The need for qualifying information was tied to three themes that included: i) context specifiers, ii) development specifiers, or iii) that behavior happened sometimes in a more general sense without context or development specific patterns. As an example of context specificity, one participant said, “Like, sometimes, yes, sometimes no ... When she’s like playing and she’s paying too much attention to something.” (context specific via Q10 - response to name, Interview 4). This indicates that the caregiver perceived the behavior as dependent on the child’s level of engagement in an activity, and therefore felt the yes/no format did not adequately capture this situational variability. Similarly, in terms of developmental specificity, another participant said “...I said not anymore. Um because that’s something that used to be true for sure when he was younger. It’s just not right now.” (development specific via Q12 upset by everyday noises, Interview 5). This response reflects a discrepancy between the child’s current developmental abilities and earlier periods that are still within the window the question intended to capture, leading the caregiver to feel the need to qualify their answer. In a final example, a participant described a general preference for flexibility in response options, stating, “Exactly what I said right, that it’s kind of a yes and no question for it just can’t say, you know, 100%, yes, or 100% no. It’s in-between.” (general sometimes via Q14 eye contact with interactions, Interview 5).” This response illustrates how some caregivers resisted binary labeling when their experiences with the behavior fell along a spectrum, making it difficult to fully commit to either response option.

Together, these responses underscore that even when participants understand the item content, the structure of the response format itself can introduce interpretive friction, especially when the behavior is perceived as inconsistent, context-dependent, or developmentally shifting.

Cultural Relevance

Within the cultural relevance domain, participant responses reflected clear variation in how culturally salient or appropriate each screener item was perceived to be. Across responses, there was considerable cultural variability in the valence of items. That is, perceptions of behaviors viewed as desirable/expected, neutral, or irrelevant differed significantly across participants' cultural frameworks. Participant responses within this domain were tied to three themes: i) explicit and implicit cultural incongruence, ii) gendered interpretations of development, and iii) religious or stigmatized beliefs about particular behaviors.

Explicit and Implicit Cultural Incongruence. In terms of explicit and implicit cultural incongruence, some participant responses directly expressed that the constructs targeted by screener items were misaligned with their cultural values and norms. In these responses, the behavior in question was seen not only as uncommon, but as culturally inappropriate. For example, multiple participants indicated that eye contact with adults is not reinforced and is in some cases discouraged. Participant quotes included, "In our culture, especially around elders, you are not supposed to look them directly in the eyes. It's a sign of respect to look down." (explicit incongruence via Q14 eye contact with interactions, Interview 4). Similarly, another participant said, "Mmm... no, because culturally looking in the eyes is considered disrespect. [Those] who did not look in the eye [were considered] very good, obedient, well-behaved." (explicit incongruence via Q14 eye contact with interactions, Interview 3). These participants spoke about how eye contact was broadly culturally unacceptable, but another participant

provided a more personal stance by saying, “Honestly, I think I’d feel weird if I were looking at people straight into their eyes.” (explicit incongruence via Q14 eye contact with interactions, Interview 5). Even when not explicitly named as culturally incongruent, many responses suggested a disconnect between the behavior in question and the family's typical experience. This implicit incongruence often appeared as confusion, hesitancy, or the framing of certain behaviors as irrelevant or unnecessary. Participant examples included: “Um, I don’t think there’s necessarily anything [to think about]. I don’t think that’s um something that we would necessarily think of. Or even notice, probably.” (implicit incongruence via Q16 response to joint attention, Interview 6) and “Like pointing with the finger is not important, I never thought about it.” (implicit incongruence via Q7 point to share interest, Interview 4).

Gendered Interpretations of Development. The next theme, gendered interpretations of development, also influenced how participants interpreted behavior constructs. Several participants described certain behaviors as more appropriate for one gender than another. Participant examples included: “It is valued when we want very young girls to pretend to cook and help around. But not valued as a life skill.” (gendered interpretations of development via Q3 - pretend play, Interview 1), “If the kid is a boy this is okay. But if the kid is a girl, this is not encouraged. Girls are expected to be more reserved.” (gendered interpretations of development via Q8 interest in peers, Interview 9), and “Copying a parent? Yeah. we expect that of just young girls to try to imitate their mom’s cooking and helping in the house... something to look forward to.” (gendered interpretations of development via Q15 imitation, Interview 9).

Religious or Stigmatized Beliefs. A final theme involved religious or stigmatized beliefs about particular behaviors that shaped how participants responded to items. For some caregivers, screener items assessed behaviors that were considered taboo or were interpreted through

religious frameworks. Participant examples included: “Teddy bears and stuffed toys are considered not very holy. Something which could trap spirits.” (stigma via Q3 pretend play, Interview 3), and “If the abnormal body language is more severe, people tend to believe this is activity from the JIN... another creation of God. Not human. Not ghost.” (stigma via Q5 unusual finger movements near eyes, Interview 9)

Person-specific Interpretation Difficulties

To better understand which groups of participants would be at highest risk for challenges in interpreting MCHAT-R/F items, we examined patterns across several participant characteristics. First, we examined whether interpretation differences varied based on participant worldview and acculturation patterns. Participants were categorized into one of six profiles based on their acculturation and world view scores including: assimilated collectivistic, assimilated individualistic, bicultural collectivistic, bicultural individualistic, traditional collectivistic, and traditional individualistic. We then examined how interpretation challenges emerged across the six different profiles by cognitive domains.

Second, in order to further uncover ways in which cultural values and a misalignment with values in the screener might uniquely shape interpretation, we divided the full sample into two broad groups: participants who had responses that explicitly endorsed cultural incongruence (i.e., directly endorsed that the items or behaviors were not aligned with their sociocultural norms), and participants whose responses did not. We then took this analysis one step further and examined how interpretation patterns varied across the six acculturation/worldview profiles (bicultural collectivistic, bicultural individualistic, assimilated collectivistic, assimilated individualistic, traditional collectivistic, traditional individualistic) *within* the explicit endorsed

incongruence response group compared to the group responses that did not endorse incongruence. This allowed us to assess whether certain profiles were more prone to interpretation difficulties only when paired with overt cultural incongruence, or whether profile related challenges emerged more broadly, even in the absence of explicit incongruence (See Appendix Figures 2-3). This multilayered analytic approach enabled a more nuanced understanding of how both explicit incongruence and acculturation/worldview profiles can influence cognitive interpretation patterns on the MCHAT-R/F. (See Appendix Figures 4–15.)

Cognitive Domain Patterns by Acculturation & Worldview Profiles

The examination of cognitive interpretation by the presence or absence of cultural incongruence provided meaningful insight into how specific items on the MCHAT-R/F may conflict with diverse caregiver norms and expectations across many domains. To expand this understanding, participants' broader cultural orientations were examined. Cognitive interpretation patterns were further investigated in relation to participants' acculturation profiles, measured by the Stephenson Multigroup Acculturation Scale (SMAS), and their worldview orientation, measured by the Auckland Individualism and Collectivism Scale (AICS). This approach provided a deeper understanding of how these dimensions of cultural identity may shape caregiver interpretation of this autism screener across cognitive domains.

When examining cognitive interpretation difficulties across acculturation and worldview profiles, distinct patterns emerged in frequency and distribution. Participants with traditional or assimilated profiles, whether collectivistic or individualistic, tended to exhibit more isolated difficulties, usually amongst the response or comprehension domains, but not across multiple domains simultaneously. Participants identified as bicultural individualistic and bicultural

collectivistic demonstrated the highest overall concentration of cognitive interpretation difficulties across domains. (See Appendix Figures 4-7)

For bicultural participants, items Q5 (unusual finger movements near eyes), Q6 (point to request), Q12 (upset by everyday noise), Q14 (eye contact with interactions), and Q18 (response to verbal directions) exhibited consistent difficulties across at least three cognitive domains, regardless of worldview. Bicultural participants showed the highest concentration of difficulties within the retrieval and response domains. Simultaneous elevations were seen in both retrieval and response domains on Q12 (upset by everyday noises). Higher concentrations of retrieval difficulties were also seen independently on item Q19 (social referencing). Additionally, bicultural participants exhibited higher concentrations of cognitive difficulties in the comprehension domain, most notably on items: Q3 (pretend play), Q5 (unusual finger movements near eyes), Q16 (response to joint attention), Q18 (response to verbal directions) and Q19 (social referencing). Though to a significantly lesser degree, bicultural participants also showed higher concentrations of cognitive difficulties in the judgement domain, primarily for collectivistic, and most notably on Q14 (eye contact with interactions).

Differences Between Participants With and Without Explicit Cultural Incongruence

To further explore how sociocultural context may shape interpretive difficulty, participant responses that explicitly endorsed cultural values misaligned with screener items were examined as a distinct group. This allowed for a specific comparison to responses without overt cultural incongruence, in order to assess whether explicit incongruence was associated with more consistent or pronounced interpretation difficulties, as opposed to the more variable or subtle challenges observed across all response groups (See Appendix Figures 8-15).

Participant responses with endorsed explicit cultural incongruence tended to exhibit more frequent and consistent interpretation difficulties, particularly within the retrieval and response domains. For this response group, retrieval difficulties were frequently accompanied by response challenges, especially on items such as Q1 (follow a point), Q5 (unusual finger movements near eyes), Q6 (point to request), Q12 (upset by everyday noise), and Q14 (eye contact with interactions). Notably, the response domain showed elevated difficulty independent of retrieval on Q6 (point to request) and Q7 (point to share interest), suggesting a consistent pattern of needing to qualify binary answers or experiencing uncertainty with interpreting culturally unfamiliar behaviors. In this explicit incongruence response group, comprehension challenges also emerged more prominently on Q5 (usual finger movements near eyes), Q16 (response to joint attention), and Q18 (response to verbal directions), while judgment difficulties were concentrated on Q5 (unusual finger movements near eyes) and Q14 (eye contact with interactions), often reflecting discomfort with item clarity or sociocultural discomfort.

In contrast, participant responses without explicit cultural incongruence also demonstrated the most difficulty in retrieval and response, but these challenges were less frequent and more variable across items. Retrieval difficulties in this response group appeared on a wider array of items, including Q1 (follow a point), Q10 (response to name), Q12 (upset by everyday noises), Q16 (response to joint attention), Q18 (response to verbal directions), and Q19 (social referencing). Response challenges were similarly distributed and appeared independently of retrieval on items such as Q11 (responsive social smile), Q12 (upset by everyday noises), and Q19 (social referencing), indicating a more item-specific, nuanced interpretation strain rather than a broad misalignment. In this response group, comprehension and judgment difficulties

were noticeably less frequent and dispersed, with minor elevations on Q3 (pretend play), Q18 (response to verbal directions), and Q19 (social referencing) within the comprehension domain as well as Q5 (unusual finger movements near eyes), Q14 (eye contact with interactions), Q16 (response to joint attention), Q18 (response to verbal directions), and Q19 (social referencing) within the judgment domain. Of note, some items only elicited interpretive difficulty for participant responses with endorsed cultural incongruence, including Q6 (point to request), Q7 (point to share interest), and Q15 (imitation). Conversely, Q18 (response to verbal directions) showed a higher concentration of cognitive challenges among participant responses without cultural incongruence.

Taken together, these findings suggest that explicit cultural incongruence is associated with more pervasive and predictable interpretation challenges, particularly in retrieving examples and mapping responses for behaviors that are culturally inappropriate or less frequently encouraged. Meanwhile, those without overt incongruence still encountered difficulties, but these appeared to be more specific.

Cognitive Domain Difficulties by Profile and Cultural Incongruence

When comparing cognitive interpretation difficulties across acculturation profiles, the most concentrated difficulties were again observed among participants classified as bicultural, regardless of cultural incongruence reporting. However, within this subdomain of participants, interpretive challenges were even more prevalent and spanned more MCHAT-R/F items in participants who endorsed cultural incongruence. The retrieval and response domains consistently accounted for the greatest concentration of difficulties. Retrieval domain difficulties for Q1 (follow a point) and Q16 (response to joint attention) were uniquely elevated among

participants who reported cultural incongruence, suggesting challenges specific to culturally influenced understandings of joint attention. Similarly, response domain difficulties for Q1 (follow a point), Q6 (point to request), and Q7 (point to share interest) were more prominent in participants who reported cultural incongruence as well, indicating hesitancy or contextual variability in how caregivers framed their responses to these behavioral probes. In contrast, some items demonstrated domain-specific difficulties only among participants who did not endorse cultural incongruence. For example, Q18 (response to verbal directions) showed elevated response difficulty in participants who did not report incongruence, but not for those who did. Similarly, Q12 (upset by everyday noise) showed more retrieval difficulty amongst participants who did not report incongruence as well (See Appendix Figures: 8-15).

Chapter 4: Discussion

The purpose of this study was to explore how culturally, and linguistically diverse caregivers interpret items on a widely used autism screener, MCHAT-R/F, using cognitive interviewing methodology. Interpretation variability was examined using Tourangeau's model of cognitive processing for surveys, which examines four core domains (comprehension, retrieval, judgement, response). Given the sociocultural variability, in many of the behaviors assessed on the MCHAT-R/F we specifically examined an additional cultural domain. Results revealed that while all items showed some susceptibility to misinterpretation, certain questions consistently elicited higher frequencies of cognitive challenges across participants. We also examined how participant characteristics related to distinct item interpretation and generally found that bicultural caregivers, whether collectivistic or individualistic, exhibited the most interpretive difficulty with the screener. This may reflect the unique cognitive and cultural negotiation processes involved in reconciling multiple worldviews when interpreting items rooted in Western developmental expectations (Aishworiya et al., 2023; Zuckerman et al., 2014).

First, and potentially the most clinically relevant, this study identified which MCHAT-R/F items elicited the most interpretation difficulties across participants and cognitive domains. While all analyzed items showed some level of interpretive strain, questions such as Q1 (follow a point), Q5 (unusual finger movement near eyes), Q12 (upset by everyday noises), Q16 (response to joint attention), Q18 (response to verbal directions), and Q19 (social referencing) exhibited the highest concentrations of participant-level difficulties. Prior research on measurement suggests

that items with consistent interpretation difficulty across respondents should be examined for potential revision, particularly if the construct is essential for diagnostic accuracy (Glascoe, 2005; Squires et al., 1996). These frequently misinterpreted items share some common features: they involve low-frequency or subtle behaviors (e.g., Q5 unusual finger movements near eyes, Q19 social referencing), rely on inferences of a child's social intent (e.g., Q1 follow a point, Q16 response to joint attention, Q19 social referencing), and assess constructs that are culturally variable in salience (e.g., Q1 follow a point, Q12 upset by everyday noises, Q16 response to joint attention, Q18 response to verbal directions, Q19 social referencing). Of particular interest to this study, several high-difficulty items may not be equally emphasized or expected across cultures. For example, Q1 (follow a point) and Q16 (response to joint attention) rely on the concept of joint attention, which is a key focus in Western models of social development. However, in some cultures where children are expected to listen and observe more than actively engage with adults, behaviors like pointing or following someone's gaze may not be encouraged or closely monitored (Rogoff, 2003; Vinden, 1999). These more respect-based or adult-led communications are often referred to as hierarchical communication styles, which can influence whether caregivers notice or value such behaviors (Chavajay & Rogoff, 2002). Q18 (respond to verbal directions) also reflects cultural expectations regarding obedience and early verbal compliance, which can vary depending on whether families emphasize independence and autonomy or expect young children to consistently follow directions (Chavajay & Rogoff, 2002; Harkness & Super, 2002). Similarly, Q12 (upset by everyday noises) may be viewed through the lens of personality or temperament rather than as a potential developmental concern, especially in families where sensory sensitivity is not commonly discussed (de Leeuw et al., 2020; Zuckerman et al., 2014).

For Q19 (social referencing), a child checking an adult's face for emotional cues may also be culturally variable. In some communities, children are taught to be emotionally self-reliant or avoid overt facial expressions, making this behavior less visible or expected (Grinker et al., 2012; Keller, 2007). Finally, Q5 (unusual finger movements near eyes), while not explicitly social, may trigger discomfort or avoidance due to cultural stigma around repetitive or atypical behaviors. In some cultures, such behaviors are associated with shame or disability and may be minimized or reinterpreted by caregivers (Burke et al., 2020; Mandell & Novak, 2005), making it harder to interpret the item accurately. These characteristics may compound interpretive difficulties, particularly among caregivers unfamiliar with Western developmental frameworks or who can lack clear behavioral examples to draw from (Aishworiya et al., 2023; Grinker et al., 2012; Zuckerman et al., 2014).

Other items like Q7 (point to share interest) and Q8 (interest in peers) also showed fewer overall difficulties across domains but exhibited more domain specific challenges. These items may not warrant full revisions, but documented challenges may highlight the need for clinicians to conduct targeted follow-up, particularly during the structured follow-up interview phase of the MCHAT-R/F, which is strongly recommended to improve diagnostic accuracy and reduce false positives (Aishworiya et al., 2023; Robins et al., 2009). In contrast, item Q9 (social sharing) was rarely flagged, suggesting improved clarity or greater consistency in interpretation among diverse groups. One possible reason is that Q9 (social sharing), slightly similar to Q7 (point to share interest) and Q8 (interest in peers), references a distinct observable behavior. All three items involve visible, externally directed social actions that are relatively clearer for caregivers to notice or describe (Grinker et al., 2012; Zuckerman et al., 2014). However, Q9 (social sharing)

may have stood out as particularly clear because it involves a recognized act of sharing with others, which may carry less cultural variation in expectation compared to pointing (Q7), which is considered rude and is discouraged in some cultures, or social interest with peers (Q8), which may be less emphasized in early childhood depending on cultural norms around family structure, age-mixing, and autonomy (Gernsbacher et al., 2005; Keller, 2007; Vinden, 1999).

Taken together, these findings suggest that items grounded in tangible, observable behaviors tend to produce fewer interpretive difficulties across diverse populations. However, even when error rates are low overall, specific cognitive domains may still present challenges, especially when caregivers attempt to map nuanced experiences onto binary response formats. This reinforces the importance of structured follow-up interviews to clarify meaning when ambiguity arises. The occasional appearance of difficulty, even on seemingly straightforward items, underscores the potential for conceptual ambiguity in assessing early social-communicative functioning. These findings align with prior research showing that culturally shaped developmental expectations influence how early behaviors are noticed, interpreted, and evaluated, and emphasize the need for cultural framing, clarification strategies, and targeted probing in clinical screening practice to support more accurate and equitable assessment (Aishworiya et al., 2023; Grinker et al., 2012; Zuckerman et al., 2014).

Additionally, this study explored differences based on participant characteristics, particularly their acculturation and worldview profiles. Participants identified as bicultural (navigating both their ethnic cultural and more dominant Western culture), regardless of collectivist or individualist orientation, demonstrated the highest frequency of interpretive challenges spread across domains, particularly within retrieval and response domains. In line

with existing research, these findings suggest that navigating between two cultural frameworks may introduce unique cognitive strain, even if language-level comprehension is intact (Benet-Martínez & Haritatos, 2005; LaFromboise et al., 1993; Schwartz et al., 2010). Further disaggregation of this data underscored how explicit and implicit cultural incongruences can influence not just whether an item makes sense to a caregiver, but whether it is perceived as relevant or appropriate for their child, particularly for behaviors embedded within the social norms and values of Western culture. Caregivers with greater awareness of cultural incongruence may make deliberate adjustments or provide useful qualifications to their responses, reflecting metacognitive insight into the contrast between their own values and those embedded in the screener. In contrast, for caregivers with lower awareness cultural discord, interpretation challenges may be less clearly articulated, requiring clinicians to monitor for signs of dissonance or disengagement (Schwartz et al., 2010; Sue & Sue, 2012; Zayas & Solari, 1994).

Cognitive Domain Interpretation Patterns

First, among each of Tourangeau's four cognitive domains (comprehension, retrieval, judgment, and response) and a fifth cultural relevance domain, retrieval and response were the most frequently impacted domains. This suggests that many caregivers experienced cognitive strain not in understanding the items themselves, but in retrieving relevant examples and/or mapping those examples onto a binary response format. This finding supports the idea that sociocultural context shapes how caregivers' access, evaluate, and report behavioral information in developmental screeners (Bornstein, 2012; Rogoff, 2003). Consistent with prior research, this pattern was anticipated. Previous cross-cultural research has shown that retrieval and response processes are particularly vulnerable to cultural variability, as these processes rely on memory

salience, contextual relevance, and communication norms, all of which are embedded in cultural models of child development and socialization (Chasiotis et al., 2006; Harkness & Super, 2002; de Leeuw et al., 2020). For instance, caregivers from cultures where certain behaviors are not emphasized or expected may have fewer salient memories to retrieve or may be unsure whether their examples fit in the context of a screener with a Western lens. Furthermore, the act of fitting nuanced experiences into binary yes/no options may conflict with cultural norms that value explanation or situational flexibility over rigid categorization (Sue & Sue, 2012). Overall, these findings align with prior research suggesting the cognitive interpretation steps of retrieval and response are socially and culturally mediated processes (de Leeuw et al., 2020; Rogoff, 2003).

Comprehension Domain

Within the comprehension domain, participants commonly demonstrated understanding by repeating item content in their own words, offering clear definitions of key constructs, or providing relevant behavioral examples. However, several forms of inaccurate comprehension also emerged, including themes of reconceptualizing item intent, confusion about the construct, or tangential and unrelated responses. The reconceptualization findings suggest that while surface-level comprehension may appear intact, nuanced misalignments in meaning can occur, often linked to the reframing of developmental constructs to fit more salient cultural expectations and norms. For example, some participants reconceptualized behaviors such as responsive smiling or social referencing as responses to adult prompting rather than naturally occurring social cues. This suggests a culturally grounded interpretation of early interactions, where parent-child exchanges are framed more as opportunities for teaching than as mutual social engagement. This aligns with prior research, which highlights how in many non-Western or immigrant

communities, adult-child exchanges are often oriented toward teaching and social responsibility rather than casual or expressive interaction (Farver, 1999; Gaskins, 2006; Leyendecker et al., 2020; Super & Harkness 2022). It also aligns with prior literature suggesting that cultural variation in social communication expectations can subtly shift how caregivers conceptualize normative behavior (Buriel, 1993; Rogoff, 2003). More explicit misunderstandings may also be shaped by cultural norms around communication with perceived authority figures. In some cultures, questioning or seeking clarification from professionals may be seen as disrespectful or inappropriate, potentially limiting participants' willingness to indicate confusion or request clarification during assessment (Sue & Sue, 2012; Zayas & Solari, 1994). These dynamics can further obscure areas of misunderstanding, particularly when clinician and caregiver cultural frameworks are not aligned.

Retrieval Domain

The retrieval domain revealed the highest overall concentration of interpretive difficulties. Participants reflected on a range of memory retrieval sources, including broad examples, personal memories, direct observations, and reports from others. Consistent with previous literature, retrieval successes appeared strongly linked to the salience and routinization of behaviors, which helps facilitate encoding and memory access (Bornstein, 2012; de Leeuw et al., 2020; Tourangeau et al., 2000). For example, caregivers who reported consistent routines were able to retrieve examples rapidly and with high confidence. In contrast, behaviors that were less culturally emphasized or less frequently observed resulted in delayed, hesitant, and/or inaccurate responses. Two notable retrieval difficulty themes emerged: surface level recall and the intersection of misunderstanding, in which partial comprehension led to misaligned

examples. These difficulties reinforce how culturally shaped expectations can affect the accessibility and accuracy of behavioral recall (Benet-Martínez et al., 2002; LaFromboise et al., 1993). When particular behaviors (e.g., pointing, social referencing, eye contact) are not emphasized or socially reinforced within a caregiver's cultural context, parents may struggle to retrieve relevant examples—even if they comprehend the question's wording—resulting in continued cognitive strain. This pattern is consistent with prior research indicating that the salience and meaning of child behaviors are shaped by sociocultural expectations, which in turn affect parental recall and reporting (Bornstein, 2012; de León, 2011; Leyendecker et al., 2020; Miller et al., 2015).

Judgement Domain

In the judgment domain, we evaluated both participants' expressed cognitive and emotional responses to the MCHAT-R/F items. Some participants found questions straightforward and easy to answer, while others reported difficulty due to unclear item wording or the need for contextual clarification. Emotional reactions varied from neutral or positive reflections to strain that included anxiety about developmental expectations or guilt for not having noticed certain behaviors. These emotions were particularly salient for items that appeared to carry different meanings across cultures or assessed behaviors more likely to be stigmatized in some communities (Grinker et al., 2012; Mandell & Novak, 2005). The emotional component of the judgment domain demonstrates that caregivers' interpretive responses are not only cognitive, but also affectively charged, especially when items touch on socially valued or culturally irrelevant behaviors unexpected of diverse families. These findings build on calls from

prior research to consider both the cognitive strain and emotional load that developmental screeners may elicit in diverse populations (Aishworiya et al., 2023; Zuckerman et al., 2014).

Response Domain

In the response domain, some participants found the yes/no response format sufficient, while others reported difficulties with the binary constraint. Caregivers frequently sought to qualify their answers, offering context-specific, development-specific, or general sometimes yes/sometimes no clarifications. These qualifiers reflect how binary formats may be insufficient in capturing a non-Western interpretation of behavioral constructs. Prior research measurement has similarly questioned the cultural appropriateness of forced choice or dichotomous formats, noting that respondents from diverse backgrounds may prefer or require elaborative or contextualized answers to accurately reflect their experiences (de Leeuw et al., 2020; Maxwell, 1996; van de Vijver & Leung, 1997). Such difficulty is particularly pronounced in diverse cultural contexts, where expectations around child behavior may be more fluid, relational, or context dependent. For example, in some cultures, behaviors like social referencing or eye contact are not consistently expected across settings and may develop differently based on family interaction patterns or socialization practices and expectations (Miller et al., 2015; Rogoff, 2003), consistent with participant responses in this study. As a result, caregivers may struggle to give a definitive yes or no. This difficulty cannot simply be labeled as confusion related to the behavior itself, but rather stems from lived experience resisting dichotomous categorization imposed by a Western lens. The discord between nuanced lived experiences and binary response demands reflects a structural limitation of standardized screeners. In subjecting participants to forced-choice formats, there is increased risk of obscuring cultural variation, leading to potential misclassification or reduced measurement validity (Betancourt et al., 2003; van de Vijver &

Leung, 1997). Prior research highlights that this issue may disproportionately impact caregivers from non-Western backgrounds, whose behavioral expectations are not well-aligned with the norms embedded in screening tools like the MCHAT-R/F (Tek & Landa, 2012; Zuckerman et al., 2013).

Furthermore, response mapping difficulties were found to compound the interpretive strain already introduced in earlier cognitive stages, such as retrieval, where those who hesitated, expressed uncertainty, or used inaccurate memories when retrieving an example often displayed similar uncertainty in selecting a definitive response. This pattern suggests that cultural misalignment may produce a ripple effect across cognitive domains, ultimately manifesting in how caregivers report behavior, regardless of their overall understanding of the question's wording (Tourangeau, 2000; van de Vijver, 2003; Willis, 2005)..

Cultural Relevance

Cultural Variability in Item Valence. The cultural relevance domain, perhaps the most significant domain, illuminated how the perceived value of screener items varied across participants' cultural frameworks. This cultural variability in valence showed that constructs evaluated by the screener can be deemed valuable, valuable in specific contexts only, or irrelevant depending on cultural norms. Cultural relevance responses fell into key themes which included: i) explicit cultural incongruence, where participants directly stated that an item conflicted with their values (Hong et al., 2000; van de Vijver & Leung, 1997)), ii) implicit incongruence, where misalignment was evident through confusion or irrelevance (Tek and Landa, 2012), iii) gendered cultural norms, where behaviors were seen as valuable only for one gender (Chavajay & Rogoff, 2002; Soto et al., 2015); and iv) religious or stigmatized

interpretations, where behaviors were viewed through spiritual or culturally taboo lenses (Gernsbacher et al., 2005; Zayas & Solari, 1994). These patterns are consistent with existing research on how cultural models of child development shape behavioral expectations (Hong et al., 2000; Soto et al., 2015; Tek & Landa, 2012). Additionally, this observed variability in cultural valence underscores the importance of distinguishing between conceptual and cultural understanding (Bernal et al., 2009; Rogoff, 2003; Zuckerman et al., 2014). Caregivers may understand what a question is asking, yet still view the behavior as culturally irrelevant or inappropriate for their child, and therefore have difficulty cognitively processing it to produce a valid response selection. This has implications for both item construction and clinical interpretation, especially in cross-cultural contexts where assumptions about child behavior may not be universally shared (van de Vijver & Leung, 1997).

Influence of Participant Characteristics. Additional analyses by acculturation and worldview profiles revealed that participants with traditional or assimilated profiles showed more domain-specific challenges. Consistent with the literature (Benet-Martínez & Haritatos, 2005; Schwartz et al., 2010), this suggests that stability in cultural orientation, whether in ethnic or dominant culture, may provide clearer interpretive anchors. However, bicultural participants, regardless of whether their worldview leaned collectivist or individualist, demonstrated the most consistent cognitive strain. Specifically, items Q5 (unusual finger movements near eyes), Q6 (point to request), Q12 (upset by everyday noises), Q14 (eye contact with interaction), Q18 (response to verbal directions), and Q19 (social referencing) showed elevated difficulty for bicultural participants across multiple domains. Retrieval and response domains again exhibited the most consistent challenges among these participants, with comprehension and judgment

difficulties also emerging higher for bicultural individuals, most notably on Q3 (pretend play), Q5 (unusual finger movements near eyes), Q16 (response to joint attention), Q18 (response to verbal directions), and Q19 (social referencing). This aligns with literature suggesting that navigating dual-cultural frameworks can create internal dissonance or increased interpretive effort during assessments that push for a binary response (Benet-Martinez et al., 2002; Hong et al., 2000; LaFromboise et al., 1993; Schwartz et al., 2010).

Further within these profiles, participants who explicitly endorsed cultural incongruence were more likely to show multi-domain difficulties, while those without explicit incongruence still demonstrated nuanced, subtle retrieval or response misalignments. For example, while items were still elevated in the full group of congruent and incongruent responses (e.g., Q3 pretend play, Q14 eye contact with interactions), they became more prominent among the group of participant responses that endorsed cultural incongruence when the data were disaggregated. Additionally, some items (e.g., Q1 follow a point, Q6 point to request, Q7 point to share interest) exhibited response domain difficulties only among bicultural participants that reported cultural incongruence. Similarly, Q16 (response to joint attention) and Q19 (social referencing) had retrieval difficulties unique to the reported cultural incongruence group. These findings support the notion that cultural incongruence can still affect cognitive processing of screener items, even among caregivers who do not perceive or report it (Benet-Martínez et al., 2002; Grinker et al., 2012). Cultural incongruence, whether overt or covert, can shape how caregivers retrieve, evaluate, and respond to screening questions independent of language fluency background (Beaton et al., 2000; Grinker et al., 2012; Mandell & Novak, 2005; Zuckerman et al., 2013).

Taken together, these findings are consistent with research highlighting that although screeners like the MCHAT-R/F are normed on majority populations, they may not align with lived experiences of families from immigrant, multilingual, or racially minoritized backgrounds (Magana et al., 2015; Tek & Landa, 2012; Zuckerman et al., 2013;). Though there is a growing number of studies addressing linguistic translation of autism screeners, relatively few have focused on cultural adaptation that seeks to modify item content, constructs, or administration procedures to better reflect the social norms tied to caregiver values and expectations for diverse populations (DuBay et al., 2021; Levante et al., 2023; Soto et al., 2015). This lack of cultural calibration may contribute to persistent disparities in interpretation, screening outcomes, and access to diagnostic services.

Implications for Practice

The phenomenological approach to this study aligns with calls to prioritize the credibility of difficulties with a lower frequency but high contextual value, recognizing that not every cultural difficulty will reveal itself through quantifiable or statistically significant frequencies (O'Dell et al., 2016; Smith & Osborn, 2007). This study values the meaning of each individual response, not just the frequency of cognitive interpretation difficulties (Smith & Osborn, 2017). Although particular items emerged with higher concentrations of difficulties by domains, it is equally as important to consider the interpretive richness of lower concentration difficulties (O'Dell et al., 2016; Pinder, 1998). Participant responses that flagged one domain on items deemed to have higher clarity amongst participants (e.g., Q9 social sharing, Q15 imitation) still offered valuable insight and reflected how cultural norms complicate the interpretation of

construct and exhibit how culture is embedded within seemingly straightforward child development expectations (Buriel, 1993; Rogoff, 2003).

Specific findings within this study suggest that retrieval and response components of the cognitive interpretation process are particularly vulnerable domains across diverse populations, especially those navigating a bicultural framework (Benet-Martínez et al., 2002; Schwartz et al., 2010). Families navigating both ethnic and dominant cultural norms, values, and expectations may experience heightened cognitive strain during screening, especially when asked to compress contextually rich behaviors into binary yes-or-no answers (Hong et al., 2000; van de Vijver & Leung, 1997). Prior research has reported that many binary formats obscure nuanced cultural perspectives, often resulting in difficulties with observed behaviors and screening response (Grinker et al., 2012; Mandell & Novak, 2005). Clinicians should be cautious not to interpret inconsistencies as lack of insight or concern, but rather as a sign of possible cultural friction embedded in the question. It would be helpful for clinicians to be familiar with cognitive processing stages while using the MCHAT-R/F to support caregiver interpretation. Additionally, clinicians should be aware of the impacts of both explicit and implicit incongruence. While some caregivers were able to report the overt ways in which an item construct misaligned with their values, signs of incongruence emerged more subtly for others, yet both caregiver groups produced cognitive strain (Benet-Martínez et al., 2002; LaFromboise et al., 1993). This reinforces the need for clinician awareness and reflective probing during screening encounters (DuBay et al., 2021; Levante et al., 2023), as misinterpretation of screener items have the ability to contribute to false negatives and/or false positives, particularly for families whose social

values deviate from Western culture norms (Magaña et al., 2015; Soto et al., 2015; Tek & Landa, 2012; Zuckerman et al., 2014).

This study supports important implications for early autism screening in diverse populations. Clinicians should consider incorporating cognitive probes to assess how caregivers are interpreting key items, as they are reliant on constructs influenced by Western specific culture. This could include asking for examples from caregivers and clarifying understanding as necessary or asking if the behavioral construct within the item is a culturally valid expectation for their child. This could also include asking about the construct and the unique ways in which their child exhibits the skill in a more culturally familiar form as a more open-ended probe option (Beaton et al., 2000; van de Vijver & Leung, 1997). Additionally, providers should work to notice cultural incongruences, both implicit and explicit, that may improve accuracy of screener response by offering scaffolded support in these scenarios. This could include asking questions about hesitancy or delayed responses, as they may be indicators of cultural misalignment. Overall, these findings suggest that without deeper attention to cultural variability, screeners can place undue burden on families to be able to easily shift between ethnic and dominant cultural frameworks while using cognitive processing skills.

This study contributes to a growing body of literature calling for screening tools that are not only linguistically translated, but also culturally adapted to ensure meaningful, equitable use across diverse communities via transadaptation (DuBay & Watson, 2019; Levante et al., 2023; Soto et al., 2015; Zucker, 2005). To improve the cultural responsiveness of ASD screening, providers using tools like the MCHAT-R/F should consider incorporating brief cognitive probes to assess caregiver understanding and alignment with behavioral constructs. For example,

clinicians might ask: “Is this something your child typically does?” “Can you think of a time when it happened?”, “Does this behavior show up in your family or culture in a different way?”, “Was this something you expected your child to do at that age?”. These small but meaningful adjustments can offer insight into whether the caregiver’s interpretation reflects their child’s abilities or a cultural mismatch between the screener’s assumptions and their cultural expectations (Beaton et al., 2000; van de Vijver & Leung, 1997).

This study also contributes to the transadaptation progress by identifying specific item-level challenges that could be addressed through measure/wording refinement. Items about pretend play could include culturally diverse examples, such as mimicking chores or imitating adult roles, rather than only referencing Western pretend play. Pointing behaviors may benefit from clarifying the function (e.g., to request vs. to share attention) or asking about appropriate ways children indicate needs or show things to others. Eye contact could be accompanied by phrasing of relevancy or supplemented with examples that distinguish respectful gaze norms across cultural groups. These revisions reflect the principle that transadaptation must go beyond literal translation to ensure conceptual fit and salience of behaviors across cultures (Soto et al., 2015; DuBay & Watson, 2019). Without modifications, screeners may unintentionally place cognitive and emotional burden on families to switch between cultural frames, ultimately increasing the risk of false positives or false negatives among marginalized populations (Magaña et al., 2015; Tek & Landa, 2012; Zuckerman et al., 2014).

This study highlights the critical role of the MCHAT-R/F clinical interview, which allows providers to clarify caregiver understanding, probe for examples, and mitigate cultural incongruence that may otherwise lead to misinterpretation. The follow-up offers a built-in space

to clarify caregiver understanding, elicit examples, and mitigate difficulties. For families navigating multiple cultural frameworks, this step can serve as a great component of support. (Aishworiya et al., 2023; Robins et al., 2009).

Limitations

Several limitations should be considered when interpreting the findings of this study. Though common for qualitative research, sample size was guided by data saturation opposed to statistical power (Creswell & Poth, 2018). Although this approach is appropriate for increasing the depth and variation of cognitive processing, it limits generalizability to broader populations (Creswell & Poth, 2018; Smith & Osborn, 2007). The emphasis on nuanced responses within a phenomenological approach often means that findings reflect complex lived experiences and interpretations of individuals rather than statistically significant trends (O'Dell et al., 2016; Smith & Osborn, 2007;).

Additionally, while the sample was diverse in linguistics, immigration, and race/ethnicity, all participants were required to speak English for feasibility purposes. This may have limited the ability to completely capture interpretation patterns of caregivers with more limited English proficiency (DuBay & Watson, 2019; Soto et al., 2015). Further, cognitive interviewing methodology relies on each participant's metacognitive awareness and ability to verbally articulate their thought processes (Willis, 2005). This may have excluded perspectives from caregivers who had difficulties verbalizing abstract reasoning in a second language (Folke et al., 2016; Nath, 2021). Additionally, in cross-cultural contexts, constructs of power distance, authority, and formality can create feelings of discomfort that can influence participant comfortability and willingness to share (Chun & Chelsea, 2011; Dai et al., 2022; Karnieli-Miller

et al., 2009). Another methodological limitation is that the study did not collect information on participants' education levels or whether they were first-time parents. Both factors may influence how caregivers interpret and respond to developmental screening questions. For example, education level could affect familiarity with certain terminology or confidence in answering questionnaires (Bethell et al., 2011; Glascoe, 2000), while parenting experience may shape knowledge of developmental milestones and expectations of child behavior (DeBaryshe & Binder, 1994; Petersen et al., 2013). Without these data, it is unclear how such factors may have contributed to participants' cognitive processing or introduced variability in interpretation patterns. Although caregivers of children outside the MCHAT-R/F target window (16–30 months) were asked to reflect retrospectively on their child's behavior during that age range, the increased temporal distance may have contributed to challenges in memory retrieval. While participants did not explicitly attribute retrieval difficulties to this, it remains a potential factor influencing recall accuracy. Finally, although the structured interview and coding process allowed for consistency for the categorization of interpretation difficulties across the cognitive processing domains, it is possible that some nuance was left uncovered in applying a more fixed framework to the interview (Jamshed, 2014; Ruslin et al., 2022).

Despite these limitations, this study offers several strengths. The use of a phenomenological approach allowed for rich exploration of participants' complex lived experiences, illuminating both subtle and explicit cultural influences on item interpretation. For example, one caregiver interpreted the imitation construct item (Q15 imitation) through a gender-specific cultural lens, explaining that copying behaviors such as pretending to cook or clean are expected primarily of young girls, while boys are not held to the same expectation.

Although this item elicited the fewest overall difficulties, this response revealed a culturally embedded expectation that might impact how behaviors are interpreted or valued across different families. Such insight would be missed without a qualitative, interpretive lens. The diverse, multilingual, and predominantly immigrant sample also enhanced the relevance of findings for a more applied use in global screening practices. Furthermore, the item level analyses provided unique insights that can inform future screener modifications and culturally adaptive practices.

Future Directions

Future research should continue to explore how culturally, and linguistically diverse families interpret autism screening items based on social values, using methods that prioritize the complexities of lived experiences. Further developing this cognitive interviewing approach in ways to include non-English interviews with interpreters and/or community based researchers could deepen insight into metacognitive experiences when interpreting screeners. Longitudinal qualitative research could provide information regarding the evolution of interpretation over time, especially after any diagnosis or intervention service engagement. This could examine how cultural frameworks adapt or resist following increased exposure to Western developmental norms. Further studies could also integrate qualitative findings with quantitative data on screener outcomes of false positives and false negatives. This could examine if specific cognitive processing difficulties are associated with delays in referral and/or diagnostic accuracy. Such findings could underscore the need for provider scaffolded support in screening procedures, particularly in under-resourced areas, given the consistent rise of developmental disabilities like autism. Further research into how caregiver demographics (e.g., immigrant generation, refugee background, gender expectations) interact with cognitive processing within a Western

sociocultural context could offer insight into additional layered experiences of screener interpretation. The findings of the current study suggest that even when caregivers endorse higher rates of individualism, more aligned with Western values, there is still strain in cognitive processing of screener items. This is possibly due to bicultural negotiation, but further research should explore this persistent discrepancy, as it may relate to stigma or acculturation pressures.

Conclusion

The current study highlights the complex manner in which caregivers of culturally and linguistically diverse backgrounds interpret items on the MCHAT-R/F screener. In using a cognitive interviewing approach grounded in Tourangeau's cognitive processing model, findings revealed interpretation is not simply a matter of word comprehension addressed through linguistically translating measures. Instead, it shows the influence social values, norms and expectation can have on various levels of cognitive processing through both explicitly and implicitly perceived differences amongst those navigating ethnic and western cultural contexts. These findings highlight the importance of centering variations in family-based social norms and expectations, while valuing unique perspectives. Efforts to increase retrieval and response accuracy amongst diverse populations in screening must move beyond translation for the purposes of demographic inclusion alone and instead focus on how cultural adaptations could better shape understanding. In elevating experiences of diverse families related to their strain in the cognitive interpretation of clinical tools, a more equitably informed approach to early autism identification can be pursued.

References

- Aishworiya, R., Ma, V. K., Stewart, S., Hagerman, R., & Feldman, H. M. (2023). Meta-analysis of the Modified Checklist for Autism in Toddlers, Revised/Follow-up for Screening. *Pediatrics*, 151(6), e2022059393.
- Akechi, H., Senju, A., Uibo, H., Kikuchi, Y., Hasegawa, T., & Hietanen, J. K. (2013). Attention to eye contact in the West and East: Autonomic responses and evaluative ratings. *PloS one*, 8(3), e59312.
- Akobeng, A. K. (2006). Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta paediatrica*, 96(3), 338-341.
- Albores-Gallo, L., Roldán-Ceballos, O., Villarreal-Valdes, G., Betanzos-Cruz, B. X., Santos-Sánchez, C., Martínez-Jaime, M. M., ... & Hilton, C. L. (2012). M-CHAT Mexican version validity and reliability and some cultural considerations. *International Scholarly Research Notices*, 2012.
- Aldosari, M., Fombonne, E., Aldhalaan, H., Ouda, M., Elhag, S., Alshammari, H., Ghazal, I., Alsaleh, A., Alqadoumi, T., Thomson, R., al Khasawneh, M., Tolefat, M., & Alshaban, F. (2019). Validation of the Arabic version of the social communication questionnaire. *Autism*, 23(7), 1655–1662.
- Al Maskari, T. S., Melville, C. A., & Willis, D. S. (2018). Systematic review: cultural adaptation and feasibility of screening for autism in non-English speaking countries. *International journal of mental health systems*, 12, 1-19.
- Al-Qabandi, M., Gorter, J. W., & Rosenbaum, P. (2011). Early autism detection: are we ready for routine screening?. *Pediatrics*, 128(1), e211-e217.

- Al-Salehi, S. M., & Al-Hifthy, E. H. (2009). Autism in Saudi Arabia: Presentation, clinical correlates and comorbidity. *Transcultural Psychiatry*, 46(2), 340–347.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorder* (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- American Academy of Pediatrics. (2024). *National payer advocacy letter*. American Academy of Pediatrics.<https://downloads.aap.org/AAP/PDF/AAP%20Autism%20Spectrum%20Disorder%20Payer%20Advocacy%20Letter.pdf>
- Anderson, D. K., Liang, J. W., & Lord, C. (2014). Predicting young adult outcome among more and less cognitively able individuals with autism spectrum disorders. *Journal of child psychology and psychiatry*, 55(5), 485-494.
- Arunyanart, W., Fenick, A., Ukritchon, S., Imjaijitt, W., Northrup, V., & Weitzman, C. (2012). Developmental and autism screening: A survey across six states. *Infants & Young Children*, 25(3), 175-187.
- Arora, N. K., Nair, M. K. C., Gulati, S., Deshmukh, V., Mohapatra, A., Mishra, D., ... & Vajaratkar, V. (2018). Neurodevelopmental disorders in children aged 2–9 years: Population-based burden estimates across five regions in India. *PLoS medicine*, 15(7), e1002615.
- Baio, J. (2014). Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, United States, 2010.
- Barbaro, J., & Dissanayake, C. (2009). Autism spectrum disorders in infancy and toddlerhood: a review of the evidence on early signs, early identification tools, and early diagnosis. *Journal of Developmental & Behavioral Pediatrics*, 30(5), 447-459.

- Barbaro, J., & Halder, S. (2016). Early identification of autism spectrum disorder: Current challenges and future global directions. *Current Developmental Disorders Reports*, 3, 67-74.
- Barton, M. L., Dumont-Mathieu, T., & Fein, D. (2012). Screening young children for autism spectrum disorders in primary practice. *Journal of autism and developmental disorders*, 42, 1165-1174.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186-3191.
- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public opinion quarterly*, 71(2), 287-311.
- Benet-Martínez, V., Leu, J., Lee, F., & Morris, M. W. (2002). *Negotiating biculturalism: Cultural frame switching in biculturals with oppositional versus compatible cultural identities*. *Journal of Cross-Cultural Psychology*, 33(5), 492–516.
<https://doi.org/10.1177/0022022102033005005>
- Bernal, G., Jiménez-Chafey, M. I., & Domenech Rodríguez, M. M. (2009). Cultural adaptation of treatments: A resource for considering culture in evidence-based practice. *Professional Psychology: Research and Practice*, 40(4), 361–368. <https://doi.org/10.1037/a0016401>
- Berry, J. W ., & Sam, D. (1997). Acculturation and adaptation. In J. W . Berry, M. H. Segall, & C. Kagitcibasi (Eds.), *Handbook of cross- cultural psychology. Social behavior and applications* (pp. 291-326). Boston: Allyn & Bacon.
- Bethell, C. D., Reuland, C., Schor, E. L., Abrahms, M., & Halfon, N. (2011). Rates of parent-centered developmental screening: Disparities and links to services access. *Pediatrics*, 128(1), 146–155. <https://doi.org/10.1542/peds.2010-0424>

- Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: diagnostic validity. *The British Journal of Psychiatry*, 175(5), 444-451.
- Bitterman, A., Daley, T. C., Misra, S., Carlson, E., & Markowitz, J. (2008). A national sample of preschoolers with autism spectrum disorders: Special education services and parent satisfaction. *Journal of autism and developmental disorders*, 38, 1509-1517.
- Bogdan, R. C., & Biklen, S. K. (2003). *Qualitative Research of Education: An Introductive to Theories and Methods* (4th ed.). Boston: Allyn and Bacon.
- Bornstein, M. H. (2012). Cultural approaches to parenting. *Parenting: Science and Practice*, 12(2-3), 212–221. <https://doi.org/10.1080/15295192.2012.683359>
- Bracken, B. A., & Barona, A. (1991). State of the art procedures for translating, validating and using psychoeducational tests in cross-cultural assessment. *School Psychology International*, 12(1-2), 119-132.
- Bradshaw, J., Koegel, L. K., & Koegel, R. L. (2017). Improving functional language and social motivation with a parent-mediated intervention for toddlers with autism spectrum disorder. *Journal of autism and developmental disorders*, 47, 2443-2458.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- Brennan, L., Fein, D., Como, A., Rathwell, I. C., & Chen, C. M. (2016). Use of the modified checklist for autism, revised with follow up-Albanian to screen for ASD in Albania. *Journal of autism and developmental disorders*, 46, 3392-3407.

- Buriel, R. (1993). Acculturation, respect for cultural differences, and biculturalism in school-age children. *The Elementary School Journal*, 93(5), 287–299.
<https://doi.org/10.1086/461730>
- Burke, M., Zuckerman, K. E., & Dovidio, J. F. (2020). Examining the role of cultural factors in the clinical encounter: Attitudes, attributions, and the use of screening tools among caregivers of children with autism spectrum disorder. *Cultural Diversity and Ethnic Minority Psychology*, 26(4), 512–523. <https://doi.org/10.1037/cdp0000327>
- Campbell, K., Carpenter, K. L., Espinosa, S., Hashemi, J., Qiu, Q., Tepper, M., ... & Dawson, G. (2017). Use of a digital modified checklist for autism in toddlers—revised with follow-up to improve quality of screening for autism. *The Journal of Pediatrics*, 183, 133-139.
- Canal-Bedia, R., García-Primo, P., Martín-Cilleros, M. V., Santos-Borbujo, J., Guisuraga-Fernández, Z., Herráez-García, L., ... & Posada-de La Paz, M. (2011). Modified checklist for autism in toddlers: cross-cultural adaptation and validation in Spain. *Journal of autism and developmental disorders*, 41, 1342-1351.
- Carakovac, M., Jovanovic, J., Kalanj, M., Rudic, N., Aleksic-Hil, O., Aleksic, B., ... & Pejovic-Milovancevic, M. (2016). Serbian language version of the modified checklist for autism in toddlers, revised, with follow-up: Cross-cultural adaptation and assessment of reliability. *Scientific reports*, 6(1), 1-5.
- Caron, K. G., Schaaf, R. C., Benevides, T. W., & Gal, E. (2012). Cross-cultural comparison of sensory behaviors in children with autism. *American Journal of Occupational Therapy*, 66(5), e77–e80.

- Chaaya, M., Saab, D., Maalouf, F. T., & Boustany, R.-M. (2016). Prevalence of autism spectrum disorder in nurseries in Lebanon: A cross sectional study. *Journal of Autism and Developmental Disorders*, 46(2), 514–522.
- Chamberlain, S. P. (2005). Recognizing and responding to cultural differences in the education of culturally and linguistically diverse learners. *Intervention in School and Clinic*, 40(4), 195-211.
- Chasiotis, A., Kiessling, F., Hofer, J., & Campos, D. (2006). Autonomy as a universal and culturally specific developmental goal: A multilevel comparison between Cameroon and Germany. *Journal of Cross-Cultural Psychology*, 37(6), 660–679.
<https://doi.org/10.1177/0022022106292076>
- Chavajay, P., & Rogoff, B. (2002). Schooling and traditional collaborative social organization of problem solving by Mayan mothers and children. *Developmental Psychology*, 38(1), 55–66. <https://doi.org/10.1037/0012-1649.38.1.55>
- Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics*, 131(4), e1121-e1127.
- Choueiri, R., & Wagner, S. (2015). A new interactive screening test for autism spectrum disorders in toddlers. *The Journal of Pediatrics*, 167(2), 460-466.
- Choueiri, R., Garrison, W. T., & Tokatli, V. (2023). Early identification of autism spectrum disorder (ASD): strategies for use in local communities. *Indian journal of pediatrics*, 90(4), 377-386.
- Chung, K.-M., Jung, W., Yang, J.-W., Ben-Itzhak, E., Zachor, D. A., Furniss, F., ... Barker, A. A. (2012). Cross cultural differences in challenging behaviors of children with autism spectrum disorders: An international examination between Israel, South Korea, the

- United Kingdom, and the United States of America. *Research in Autism Spectrum Disorders*, 6(2), 881–889.
- Chun, K. M., & Chesla, C. A. (2011). Cultural Norms Shaping Research Group Interviews with Chinese American Immigrants. *Journal of Advanced Nursing*, 67(5), 972–980.
<https://doi.org/10.1111/j.1365-2648.2010.05543.x>
- Clark, C., & Robboy, H. (1992). *Social interaction: Readings in sociology* (4th ed., pp. 107-111). New York: St. Martin's Press.
- Clark, C. (1992). Culture and Sympathy. In *Social interaction: Readings in sociology* (4th ed., pp. 37-54). New York, NY: St. Martin's Press.
- Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale: SRS-2* (p. 106). Torrance, CA: Western psychological services.
- Cook, P., Cook, M., Tran, L., & Tu, W. (1997). Children enabling change: A multicultural, participatory, community-based rehabilitation research project involving Chinese children with disabilities and their families. In *Child and Youth Care Forum* (Vol. 26, No. 3, pp. 205-219). New York: Kluwer Academic Publishers-Human Sciences Press.
- Coonrod, E. E., Turner, L. M., & Pozdol, S. L. (2004). Psychometric properties of the STAT for early autism screening. *Journal of autism and developmental disorders*, 34, 691-701.
- Crais, E. R., McComish, C. S., Humphreys, B. P., Watson, L. R., Baranek, G. T., Reznick, J. S., ... & Earls, M. (2014). Pediatric healthcare professionals' views on autism spectrum disorder screening at 12–18 months. *Journal of autism and developmental disorders*, 44, 2311-2328.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications

- Cuesta-Gómez, J. L., Andrea Manzone, L., & Posada-De-La-Paz, M. (2016). Modified checklist for autism in toddlers cross-cultural adaptation for Argentina. *International journal of Developmental Disabilities*, 62(2), 117-123.
- Dai, Y., Li, H., Xie, W., & Deng, T. (2022). Power Distance Belief and Workplace Communication: The Mediating Role of Fear of Authority. *International Journal of Environmental Research and Public Health*, 19(5), 2932.
<https://doi.org/10.3390/ijerph19052932>
- Daley, T. C. (2004). From symptom recognition to diagnosis: children with autism in urban India. *Social Science & Medicine*, 58(7), 1323–1335.
- Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., ... & Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics*, 125(1), e17-e23.
- DeBaryshe, B. D., & Binder, J. C. (1994). Developmental expectations and child-rearing values of first-time and experienced mothers. *Journal of Psychology*, 128(5), 485–494.
<https://doi.org/10.1080/00223980.1994.9712743>
- DePaulo, P. (2000). Sample size for qualitative research. *Quirks Marketing Research Review*, 1202.
- Divya, P. S., Selvaraj, J. L., & Priscilla, P. (2020). Translation and validation of modified checklist for autism in toddlers-revised version in Tamil (T-MCHAT-R). *Journal of Child Language Acquisition and Development-JCLAD*, 136-143.
- Dosreis, S., Weiner, C. L., Johnson, L., & Newschaffer, C. J. (2006). Autism spectrum disorder screening and management practices among general pediatric providers. *Journal of Developmental & Behavioral Pediatrics*, 27(2), S88-S94.

- DuBay, M., & Watson, L. R. (2019). Translation and cultural adaptation of parent-report developmental assessments: Improving rigor in methodology. *Research in Autism Spectrum Disorders*, 62, 55-65.
- DuBay, M., Watson, L. R., Baranek, G. T., Lee, H., Rojevic, C., Brinson, W., Smith, D., & Sideris, J. (2021). Rigorous translation and cultural adaptation of an autism screening tool: First Years Inventory as a case study. *Journal of Autism and Developmental Disorders*, 51(7), 2349–2362. <https://doi.org/10.1007/s10803-020-04837-1>
- Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., ... & Shih, A. (2015). Autism screening and diagnosis in low resource settings: Challenges and opportunities to enhance research and services worldwide. *Autism Research*, 8(5), 473-476.
- Earls, M. F., & Hay, S. S. (2006). Setting the stage for success: Implementation of developmental and behavioral screening and surveillance in primary care practice—The North Carolina Assuring Better Child Health and Development (ABCD) Project. *Pediatrics*, 118(1), e183–e188.
- El-Behadli, A., Neger, E., Perrin, E., & Sheldrick, R. (2015). Translations of developmental screening instruments: An evidence map of available research. *J Dev Behav Pediatr*, 36, 471–483.
- Elsabbagh, M., Divan, G., Koh, Y. J., Kim, Y. S., Kauchali, S., Marcín, C., ... & Fombonne, E. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism research*, 5(3), 160-179.
- Farver, J. M. (1999). Activity setting analysis: A model for examining the role of culture in development. *Human Development*, 42(4), 238–259. <https://doi.org/10.1159/000022644>

- Farver, J. M., & Howes, C. (1993). Cultural differences in American and Mexican mother-child pretend play. *Merrill-Palmer Quarterly*, 39, 344–358.
- Fonteyn, M. E., Vetteese, M., Lancaster, D. R., & Bauer-Wu, S. (2008). Developing a codebook to guide content analysis of expressive writing transcripts. *Applied Nursing Research*, 21(3), 165-168.
- Fountain, C., King, M. D., & Bearman, P. S. (2011). Age of diagnosis for autism: individual and community factors across 10 birth cohorts. *Journal of Epidemiology & Community Health*, 65(6), 503-510.
- Freeth, M., Sheppard, E., Ramachandran, R., & Milne, E. (2013). A cross-cultural comparison of autistic traits in the UK, India and Malaysia. *Journal of Autism and Developmental Disorders*, 43(11), 2569–2583.
- Fuentes, J., Hervás, A., Howlin, P., & (ESCAP ASD Working Party). (2021). ESCAP practice guidance for autism: a summary of evidence-based recommendations for diagnosis and treatment. *European child & adolescent psychiatry*, 30(6), 961-984.
- Gabrielsen, T. P., Farley, M., Speer, L., Villalobos, M., Baker, C. N., & Miller, J. (2015). Identifying autism in a brief observation. *Pediatrics*, 135(2), e330-e338.
- Gabrielsen, T. P., Manwaring, S. S., & Stuart, A. L. (2020). Screening and surveillance. *Interprofessional Care Coordination for Pediatric Autism Spectrum Disorder: Translating Research into Practice*, 15-36.
- Gaskins, S. (2006). Cultural perspectives on infant–caregiver interaction. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 279–298). Berg.
- Geertz, C. (1973). The interpretation of cultures. Basic Books.

- Gerber, E. R. (1999). The view from anthropology: Ethnography and the cognitive interview. *Cognition and survey research*, 217-234.
- Gernsbacher, M. A., Stevenson, J. L., & Dern, S. (2005). Specificity, contexts, and reference groups matter when assessing autism-related traits. *Journal of Child Psychology and Psychiatry*, 46(5), 500–505. <https://doi.org/10.1111/j.1469-7610.2005.00401.x>
- Gillis, J. M. (2009). Screening practices of family physicians and pediatricians in 2 southern states. *Infants & Young Children*, 22(4), 321-331.
- Gjersing, L., Caplehorn, J. R., & Clausen, T. (2010). Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC medical research methodology*, 10, 1-10.
- Glascoe, F. P. (1998). The Value of Parents' evaluations of developmental status' in detecting and addressing children's developmental and behavioral problems. *Diagnostic*, 23(4), 185-203.
- Glascoe, F. P. (2000). Evidence-based approach to developmental and behavioral surveillance using parents' concerns. *Child: Care, Health and Development*, 26(2), 137–149. <https://doi.org/10.1046/j.1365-2214.2000.00173.x>
- Glascoe, F. P. (2003). Parents' evaluation of developmental status: how well do parents' concerns identify children with behavioral and emotional problems?. *Clinical pediatrics*, 42(2), 133-138.
- Goldstein, S., & Naglieri, J. A. (2009). *Autism spectrum rating scales (ASRS)*. North Tonawanda, NY: Multi-Health System.

- Golson, M. E., Ficklin, E., Haverkamp, C. R., McClain, M. B., & Harris, B. (2022). Cultural differences in social communication and interaction: A gap in autism research. *Autism Research, 15*(2), 208-214.
- Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist, 52*(10), 1115–1124. <https://doi.org/10.1037/0003-066X.52.10.1115>
- Grinker, R. R., Chambers, N., Njongwe, N., Lagman, A. E., Guthrie, W., Stronach, S., ... & Wetherby, A. M. (2012). “Communities” in Community Engagement: Lessons Learned From Autism Research in S outh K orea and S outh A frica. *Autism Research, 5*(3), 201-210.
- Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of clinical epidemiology, 46*(12), 1417-1432.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59–82. <https://doi.org/10.1177/1525822X05279903>
- Guo, C., Luo, M., Wang, X., Huang, S., Meng, Z., Shao, J., ... & Jing, J. (2019). Reliability and validity of the Chinese version of modified checklist for autism in toddlers, revised, with follow-up (M-CHAT-R/F). *Journal of Autism and Developmental Disorders, 49*, 185-196.
- Guthrie, W., Wallis, K., Bennett, A., Brooks, E., Dudley, J., Gerdes, M., ... & Miller, J. S. (2019). Accuracy of autism screening in a large pediatric network. *Pediatrics, 144*(4)

- Hahler, E. M., & Elsabbagh, M. (2015). Autism: A global perspective. *Current Developmental Disorders Reports*, 2, 58-64.
- Harkness, S., & Super, C. M. (2002). Culture and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting: Vol. 2. Biology and ecology of parenting* (2nd ed., pp. 253–280). Mahwah, NJ: Lawrence Erlbaum Associates.
- Harrison, A. J., Long, K. A., Tommet, D. C., & Jones, R. N. (2017). Examining the role of race, ethnicity, and gender on social and behavioral ratings within the Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders*, 47(9), 2770–2782.
- Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). *Code saturation versus meaning saturation: How many interviews are enough?* *Qualitative Health Research*, 27(4), 591–608. <https://doi.org/10.1177/1049732316665344>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and brain sciences*, 33(2-3), 61-83.
- Hoekstra, R. A. (2022). Serving the underserved: How can we reach autism families who systemically miss out on support?. *Autism*, 26(6), 1315-1319.
- Hong, Y. Y., Morris, M. W., Chiu, C. Y., & Benet-Martínez, V. (2000). *Multicultural minds: A dynamic constructivist approach to culture and cognition*. *American Psychologist*, 55(7), 709–720. <https://doi.org/10.1037/0003-066X.55.7.709>
- Huttenlocher, P. R. (1979). Synaptic density in human frontal cortex-developmental changes and effects of aging. *Brain Res*, 163(2), 195-205.

- Huynh, Q. L., Howell, R. T., & Benet-Martínez, V. (2009). Reliability of bidimensional acculturation scores: A meta-analysis. *Journal of cross-cultural psychology*, 40(2), 256-274.
- Hyman, S. L., & Johnson, J. K. (2012). Autism and pediatric practice: Toward a medical home. *Journal of Autism and Developmental Disorders*, 42, 1156-1164
- Hyman, S. L., Levy, S. E., Myers, S. M., Kuo, D. Z., Apkon, S., Davidson, L. F., ... & Bridgemohan, C. (2020). Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics*, 145(1).
- Irvin, D. W., McBee, M., Boyd, B. A., Hume, K., & Odom, S. L. (2012). Child and family factors associated with the use of services for preschoolers with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 6(1), 565-572.
- James, S. N., & Smith, C. J. (2020, October). Early autism diagnosis in the primary care setting. In *Seminars in Pediatric Neurology* (Vol. 35, p. 100827). WB Saunders.
- Jamshed, S. (2014). Qualitative research method-interviewing and observation. *Journal of Basic and Clinical Pharmacy*, 5(4), 87–88. <https://doi.org/10.4103/0976-0105.141942>
- Jensen, K., Noazin, S., Bitterfeld, L., Carcelen, A., Vargas-Cuentas, N. I., Hidalgo, D., ... & Autism Working Group in Peru Vanessa Cavallera Ricardo Zavaleta Juan Flores Dennis Nuñez Alejandro Dioses Anna Smith. (2021). Autism detection in children by combined use of gaze preference and the M-CHAT-R in a resource-scarce setting. *Journal of autism and developmental disorders*, 51, 994-1006.
- Johnson, C. P., & Myers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, 120(5), 1183-1215.

- Jullien, S. (2021). Screening for autistic spectrum disorder in early childhood. *BMC pediatrics*, 21, 1-9.
- Kakooza-Mwesige, A., Ssebyala, K., Karamagi, C., Kiguli, S., Smith, K., Anderson, M. C., Croen, L. A., Trevathan, E., Hansen, R., Smith, D., & Grether, J. K. (2014). Adaptation of the “ten questions” to screen for autism and other neurodevelopmental disorders in Uganda. *Autism*, 18(4), 447–457.
- Kamio, Y., Inada, N., Koyama, T., Inokuchi, E., Tsuchiya, K., & Kuroda, M. (2014). Effectiveness of using the Modified Checklist for Autism in Toddlers in two-stage screening of autism spectrum disorder at the 18-month health check-up in Japan. *Journal of autism and developmental disorders*, 44, 194-203.
- Kara, B., Mukaddes, N. M., Altinkaya, I., Güntepe, D., Gökçay, G., & Özmen, M. (2014). Using the Modified Checklist for Autism in Toddlers in a well-child clinic in Turkey: Adapting the screening method based on culture and setting. *Autism*, 18(3), 331-338.
- Karnieli-Miller, O., Strier, R., & Pessach, L. (2009). Power Relations in Qualitative Research. *Qualitative Health Research*, 19(2), 279–289. <https://doi.org/10.1177/1049732308329306>
- Keller, H. (2007). *Cultures of infancy*. Lawrence Erlbaum Associates.
- Khowaja, M. K., Hazzard, A. P., & Robins, D. L. (2015). Sociodemographic barriers to early detection of autism: screening and evaluation using the M-CHAT, M-CHAT-R, and follow-up. *Journal of autism and developmental disorders*, 45, 1797-1808.
- King, T. M., Tandon, S. D., Macias, M. M., Healy, J. A., Duncan, P. M., Swigonski, N. L., ... & Lipkin, P. H. (2010). Implementing developmental screening and referrals: lessons learned from a national project. *Pediatrics*, 125(2), 350-360.

- Kleeck, A. V. (1994). Potential cultural bias in training parents as conversational partners with their children who have delays in language development. *American journal of speech-language pathology*, 3(1), 67-78.
- Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., ... & Fein, D. (2007). The modified checklist for autism in toddlers: a follow-up study investigating the early detection of autism spectrum disorders. *Journal of autism and developmental disorders*, 38, 827-839.
- Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., ... & Fein, D. (2008). The modified checklist for autism in toddlers: a follow-up study investigating the early detection of autism spectrum disorders. *Journal of autism and developmental disorders*, 38, 827-839.
- LaFromboise, T., Coleman, H. L. K., & Gerton, J. (1993). *Psychological impact of biculturalism: Evidence and theory*. *Psychological Bulletin*, 114(3), 395–412.
<https://doi.org/10.1037/0033-2909.114.3.395>
- de Leeuw, E. D. (2011). *Improving data quality when surveying children and adolescents: Cognitive and social development and its role in questionnaire construction and pretesting*. Report prepared for UNICEF.
- de Leeuw, A., Happe, F., & Hoekstra, R. A. (2020). A conceptual framework for understanding the cultural and contextual factors on autism across the globe. *Autism Research*.
<https://doi.org/10.1002/aur.2276>
- de Leeuw, E., Harkness, J., Harkness, J. A., & Mohler, P. P. (2020). *Surveying across cultures: Toward improved equivalence and comparability of data*. In D. L. Vannette & J. A.

- Krosnick (Eds.), *The Palgrave handbook of survey research* (pp. 111–132). Palgrave Macmillan.
- de Leeuw, E. D., Hox, J. J., & Boeijs, H. R. (2020). Surveying across cultures: Comparative survey research in multicultural, multiregional, and multilingual contexts. In P. Vannette & J. Krosnick (Eds.), *The Palgrave handbook of survey research* (pp. 101–122). Palgrave Macmillan.
- de León, L. (2011). Language socialization and the construction of knowledge in Zinacantec Maya child language. *Annual Review of Anthropology*, 40, 131–147.
<https://doi.org/10.1146/annurev-anthro-081309-145626>
- Levante, A., Petrocchi, S., & Lecciso, F. (2023). Systematic review of translation and cultural adaptations of autism spectrum disorder’s screening tool: The Modified Checklist for Autism in Toddlers, Revised (M-CHAT-R). *F1000Research*, 12, 471.
<https://doi.org/10.12688/f1000research.130426.1>
- Levy, S. E., Wolfe, A., Coury, D., Duby, J., Farmer, J., Schor, E., ... & Warren, Z. (2020). Screening tools for autism spectrum disorder in primary care: A systematic evidence review. *Pediatrics*, 145(Supplement_1), S47-S59.
- Leyendecker, B., Agache, A., & Dimitrova, R. (2020). Parenting in cultural context: A commentary on theoretical models and empirical studies. *New Directions for Child and Adolescent Development*, 2020(170), 125–134. <https://doi.org/10.1002/cad.20358>
- Limosani, C., Marchitto, M., & Panuzio, B. (2021). Communication and Symbolic Behavior Scales–Normed Edition (2003). *Encyclopedia of Autism Spectrum Disorders*, 1089-1093.

- Lipkin, P. H., Macias, M. M., Baer Chen, B., Coury, D., Gottschlich, E. A., Hyman, S. L., ... & Levy, S. E. (2020). Trends in pediatricians' developmental screening: 2002–2016. *Pediatrics*, 145(4).
- Lipkin, P. H., Okamoto, J., Council on Children with Disabilities and Council on School Health, Norwood Jr, K. W., Adams, R. C., Brei, T. J., ... & Young, T. (2015). The individuals with disabilities education act (IDEA) for children with special educational needs. *Pediatrics*, 136(6), e1650-e1662.
- Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of general psychiatry*, 63(6), 694-701.
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule, 2nd ed. (ADOS-2)*. Torrance, CA: Western Psychological Services.
- Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., ... & Newschaffer, C. (2017). The changing epidemiology of autism spectrum disorders. *Annual review of public health*, 38, 81-102.
- Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V., Bilder, D. A., ... & Shaw, K. A. (2023). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2020. *MMWR Surveillance Summaries*, 72(2),
- Magaña, S., Lopez, K., Aguinaga, A., & Morton, H. (2015). Access to diagnosis and treatment services among Latino children with autism spectrum disorders. *Intellectual and Developmental Disabilities*, 53(1), 13–23. <https://doi.org/10.1352/1934-9556-53.1.13>

- Maillefert, J. F., Kloppenburg, M., Fernandes, L., Punzi, L., Günther, K. P., Mola, E. M., ... & Hawker, G. A. (2009). Multi-language translation and cross-cultural adaptation of the OARSI/OMERACT measure of intermittent and constant osteoarthritis pain (ICOAP). *Osteoarthritis and cartilage*, 17(10), 1293-1296.
- Malcolm-Smith, S., Hoogenhout, M., Ing, N., Thomas, K. G., & de Vries, P. (2013). Autism spectrum disorders—Global challenges and local opportunities. *Journal of Child & Adolescent Mental Health*, 25(1), 1-5.
- Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: guided by information power. *Qualitative health research*, 26(13), 1753-1760.
- Mandell, D. S., & Novak, M. M. (2005). The role of culture in families' treatment decisions for children with autism spectrum disorders. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(2), 110–115. <https://doi.org/10.1002/mrdd.20061>
- van Manen, M. (1997). From meaning to method. *Qualitative health research*, 7(3), 345-369.
- van Manen, M (2016). *Researching lived experience: Human science for an action sensitive pedagogy* (2nd ed.). Routledge.
- Marlow, M., Servili, C., & Tomlinson, M. (2019). A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low-and middle-income countries. *Autism Research*, 12(2), 176-199.
- Marsh, A. A., Elfenbein, H. A., & Ambady, N. (2003). Nonverbal “accents” cultural differences in facial expressions of emotion. *Psychological Science*, 14(4), 373–376.
- Matos, M. B., Bara, T. S., & Cordeiro, M. L. (2022). Autism Spectrum Disorder Diagnoses: A Comparison of Countries with Different Income Levels. *Clinical Epidemiology*, 959-969.

- Matson, J. L., Matheis, M., Burns, C. O., Esposito, G., Venuti, P., Pisula, E., ... & Goldin, R. (2017). Examining cross-cultural differences in autism spectrum disorder: A multinational comparison from Greece, Italy, Japan, Poland, and the United States. *European Psychiatry, 42*, 70-76.
- Matson, J. L., Worley, J. A., Fodstad, J. C., Chung, K.-M., Suh, D., Jhin, H. K., ... Furniss, F. (2011). A multinational study examining the cross cultural differences in reported symptoms of autism spectrum disorders: Israel, South Korea, the United Kingdom, and the United States of America. *Research in Autism Spectrum Disorders, 5*(4), 1598–1604.
- Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage.
- McCarthy, A., Lee, K., Itakura, S., & Muir, D. W. (2006). Cultural display rules drive eye gaze during thinking. *Journal of Cross-Cultural Psychology, 37*(6), 717–722.
doi:[10.1177/0022022106292079](https://doi.org/10.1177/0022022106292079).
- McDermott, M. A. N., & Palchanes, K. (1994). A literature review of the critical elements in translation theory. *Image: The Journal of Nursing Scholarship, 26*(2), 113-118.
- Miller, J. S., Gabrielsen, T., Villalobos, M., Alleman, R., Wahmhoff, N., Carbone, P. S., & Segura, B. (2011). The each child study: systematic screening for autism spectrum disorders in a pediatric setting. *Pediatrics, 127*(5), 866-871.
- Miller, K., Chepp, V., Willson, S., & Padilla, J. L. (Eds.). (2014). *Cognitive interviewing methodology*. John Wiley & Sons.
- Miller, P. J., Wang, S., Sandel, T. L., & Cho, G. E. (2015). Self-construction through narrative practices: A Chinese and American comparison on parental socialization. *Child Development, 86*(6), 1571–1585. <https://doi.org/10.1111/cdev.12408>

- Mishra, D., Patel, V., Pandey, R. M., Das, B. C., Divan, G., Murthy, G. V. S., Sharma, T. D., Sapra, S., Aneja, S., Juneja, M., Reddy, S. K., Suman, P., Mukherjee, S. B., Dasgupta, R., ... Vajaratkar, V. (2018). Neurodevelopmental disorders in children aged 2–9 years: Population-based burden estimates across five regions in India. *PLoS Medicine*, 15(7), e1002615.
- Morse, J. M. (1995). The significance of saturation. *Qualitative Health Research*, 5(2), 147–149. <https://doi.org/10.1177/104973239500500201>
- Moustakas, C. (1994). *Phenomenological research methods*. SAGE Publications.
- Narzisi, A., Posada, M., Barbieri, F., Chericoni, N., Ciuffolini, D., Pinzino, M., Romano, R., Scattoni, M. L., Tancredi, R., Calderoni, S., & Muratori, F. (2020). Prevalence of autism spectrum disorder in a large Italian catchment area: A school-based population study within the ASDEU project. *Epidemiology and Psychiatric Sciences*, 29, e5.
- Norbury, C. F., & Sparks, A. (2013). Difference or disorder? Cultural issues in understanding neurodevelopmental disorders. *Developmental Psychology*, 49(1), 45–58.
- O'Dell, L., Brownlow, C., & Bertilsdotter Rosqvist, H. (2016). Critical autism studies: Exploring epistemic dialogues and intersections, challenging dominant understandings of autism. *Disability & Society*, 31(2), 166–179. <https://doi.org/10.1080/09687599.2016.1164026>
- Oner, O., & Munir, K. M. (2020). Modified checklist for autism in toddlers revised (MCHAT-R/F) in an urban metropolitan sample of young children in Turkey. *Journal of autism and developmental disorders*, 50, 3312–3319.
- Oono, I. P., Honey, E. J., & McConachie, H. (2016). Parent-mediated early intervention for young children with autism spectrum disorders (ASD). *BJPsych Advances*, 22(3), 146–146.

- Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 34(3), 523-540.
- Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian journal of ophthalmology*, 56(1), 45.
- Perera, H., Jeewandara, K. C., Seneviratne, S., & Guruge, C. (2017). Culturally adapted pictorial screening tool for autism spectrum disorder: A new approach. *World journal of clinical pediatrics*, 6(1), 45.
- Perneger, T. V., Courvoisier, D. S., Hudelson, P. M., & Gayet-Ageron, A. (2015). Sample size for pre-tests of questionnaires. *Quality of life Research*, 24, 147-151.
- Petersen, J. M., Kube, D. A., & Palmer, F. B. (2013). Infant development and parenting education for first-time mothers. *Infant Behavior and Development*, 36(3), 470–478.
<https://doi.org/10.1016/j.infbeh.2013.03.004>
- Pierce, K., Carter, C., Weinfeld, M., Desmond, J., Hazin, R., Bjork, R., & Gallagher, N. (2011). Detecting, studying, and treating autism early: the one-year well-baby check-up approach. *The Journal of pediatrics*, 159(3), 458-465.
- Pierce, K., Courchesne, E., & Bacon, E. (2016). To screen or not to screen universally for autism is not the question: Why the task force got it wrong. *The Journal of pediatrics*, 176, 182-194.
- Pierce, R. (2011). The expressive function of public health policy: the case of pandemic planning. *Public Health Ethics*, 4(1), 53-62.

- Pinto-Martin, J. A., Souders, M. C., Giarelli, E., & Levy, S. E. (2005). The role of nursing in screening for autistic spectrum disorder in pediatric primary care. *Journal of Pediatric Nursing*, 20(3), 163–169.
- Pufall, P. B., & Pufall, E. (2008). The relationship between play and culture.
- Radecki, L., Sand-Loud, N., O'Connor, K. G., Sharp, S., & Olson, L. M. (2011). Trends in the use of standardized tools for developmental screening in early childhood: 2002–2009. *Pediatrics*, 128(1), 14-19..
- Raina, S. K., Chander, V., Bhardwaj, A. K., Kumar, D., Sharma, S., Kashyap, V., Singh, M., & Bhardwaj, A. (2017). Prevalence of autism spectrum disorder among rural, urban, and tribal children (1–10 years of age). *Journal of Neurosciences in Rural Practice*, 8(03), 368–374.
- Reichenheim, M. E., & Moraes, C. L. (2007). Operationalizing the cross-cultural adaptation of epidemiological measurement instruments. *Revista de saúde pública*, 41, 665-673.
- Reichow, B. (2012). Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *Journal of autism and developmental disorders*, 42, 512-520.
- Robins, D. L., Casagrande, K., Barton, M., Chen, C. M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*, 133(1), 37-45.
- Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and

- pervasive developmental disorders. *Journal of autism and developmental disorders*, 31, 131-144.
- Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism*, 12(5), 537-556.
- Rogers, S. J., Estes, A., Lord, C., Vismara, L., Winter, J., Fitzpatrick, A., ... & Dawson, G. (2012). Effects of a brief Early Start Denver Model (ESDM)–based parent intervention on toddlers at risk for autism spectrum disorders: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(10), 1052-1065.
- Rogoff, B. (2003). *The cultural nature of human development*. Oxford University Press.
- Rotholz, D. A., Kinsman, A. M., Lacy, K. K., & Charles, J. (2017). Improving early identification and intervention for children at risk for autism spectrum disorder. *Pediatrics*, 139(2).
- Ruslin, R., Mashuri, S., Abdul Rasak, M. S., Alhabsyi, F., & Syam, H. (2022). Semi-structured interview: A methodological reflection on the development of a qualitative research instrument in educational studies. *IOSR Journal of Research & Method in Education (IOSR-JRME)*, 12(1), 22–29. <https://doi.org/10.9790/7388-1201052229>
- Rutter, M., Le Couteur, A., & Lord, C. (2003). ADI-R. *Autism diagnostic interview revised. Manual*. Los Angeles: Western Psychological Services.
- Ryan, G. W., & Bernard, H. R. (2000). Data management and analysis methods. *Handbook of qualitative research*, 2(1), 769-802.
- Samadi, S. A., & McConkey, R. (2015). Screening for autism in Iranian preschoolers: Contrasting M-CHAT and a scale developed in Iran. *Journal of autism and developmental disorders*, 45, 2908-2916.

- Sangare, M., Toure, H. B., Toure, A., Karembe, A., Dolo, H., Coulibaly, Y. I., ... & Geschwind, D. H. (2019). Validation of two parent-reported autism spectrum disorders screening tools M-CHAT-R and SCQ in Bamako, Mali. *Eneurologicalsci*, 15, 100188.
- Sansosti, F. J., Lavik, K. B., & Sansosti, J. M. (2012). Family experiences through the autism diagnostic process. *Focus on Autism and Other Developmental Disabilities*, 27(2), 81-92.
- Scherzer, A. L., Chhagan, M., Kauchali, S., & Susser, E. (2012). Global perspective on early diagnosis and intervention for children with developmental delays and disabilities. *Developmental Medicine & Child Neurology*, 54(12), 1079-1084.
- Schultz, K. S., & Whitney, D. J. (2005). *Measurement theory in action*. Thousand Oaks.
- Schwartz, S. J., Unger, J. B., Zamboanga, B. L., & Szapocznik, J. (2010). *Rethinking the concept of acculturation: Implications for theory and research*. *American Psychologist*, 65(4), 237–251. <https://doi.org/10.1037/a0019330>
- Shulruf, B., Hattie, J. and Dixon, R., 2007. Development of a new measurement tool for individualism and collectivism. *Journal of Psychoeducational Assessment*, 25(4), pp.385-401.
- Shaw, K. A., Williams, S., Patrick, M. E., Hughes, M. M., Ali, A., Bakian, A. V., ... & Maenner, M. J. (2025). Prevalence and early identification of autism spectrum disorder among children aged 4 and 8 years—Autism and Developmental Disabilities Monitoring Network, 16 sites, United States, 2022. *Morbidity and Mortality Weekly Report. Surveillance Summaries*, 74(SS-2), 1–22. <https://doi.org/10.15585/mmwr.ss7402a1>

- Schwartz, S. J., Unger, J. B., Zamboanga, B. L., & Szapocznik, J. (2010). Rethinking the concept of acculturation: Implications for theory and research. *American Psychologist*, 65(4), 237–251. <https://doi.org/10.1037/a0019330>
- Sicorello, M., Stevanov, J., Ashida, H., & Hecht, H. (2019). Effect of gaze on personal space: A japanese–german cross-cultural study. *Journal of Cross-Cultural Psychology*, 50(1), 8–21.
- Sipes, M., Furniss, F., Matson, J. L., & Hattier, M. (2012). A multinational study examining the cross cultural differences in social skills of children with autism spectrum disorders: A comparison between the United Kingdom and the United States of America. *Journal of Developmental and Physical Disabilities*, 24(2), 145–154.
- Smith, J. A., & Osborn, M. (2007). Interpretative phenomenological analysis. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods* (2nd ed., pp. 53–80). SAGE Publications.
- Soto, S., Linas, K., Jacobstein, D., Biel, M., Migdal, T., & Anthony, B. J. (2015). A review of cultural adaptations of screening tools for autism spectrum disorders. *Autism*, 19(6), 646–661.
- Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *Journal of evaluation in clinical practice*, 17(2), 268–274.
- Squires, J., Potter, L., & Bricker, D. (1995). The ASQ user's guide for the Ages & Stages Questionnaires: A parent-completed, child-monitoring system. Paul H Brookes Publishing.

State of Georgia.(2022). *Georgia Medicaid: TEFRA/Katie Beckett*.

<https://medicaid.georgia.gov/programs/all-programs/tefrakatie-beckett>

Steiner, A. M., Goldsmith, T. R., Snow, A. V., & Chawarska, K. (2012). Practitioner's guide to assessment of autism spectrum disorders in infants and toddlers. *Journal of autism and developmental disorders*, 42, 1183-1196.

Stephenson, M. (2000). Development and validation of the Stephenson Multigroup Acculturation Scale (SMAS). *Psychological assessment*, 12(1), 77.

Stevanovic, D., Robins, D. L., Costanzo, F., Fuca, E., Valeri, G., Vicari, S., ... & Knez, R. (2022). Cross-cultural similarities and differences in reporting autistic symptoms in toddlers: A study synthesizing M-CHAT (-R) data from ten countries. *Research in Autism Spectrum Disorders*, 95, 101984.

Stone, W. L., Coonrod, E. E., & Ousley, O. Y. (2000). Brief report: screening tool for autism in two-year-olds (STAT): development and preliminary data. *Journal of autism and developmental disorders*, 30(6), 607.

Sue, D. W., & Sue, D. (2012). *Counseling the culturally diverse: Theory and practice* (6th ed.). John Wiley & Sons.

Super, C. M., & Harkness, S. (2022). Culture and developmental niches in child development. *Annual Review of Psychology*, 73, 235–258. <https://doi.org/10.1146/annurev-psych-030221-010313>

Suzuki, L . A ., V raniak, D . A ., & K ugler, J. F . (1996). Intellectual assess- ment across cultures. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 141-177). Thousand Oaks, CA: Sage.

- Tabril, T., Chekira, A., Moukhless, S., Ouazzani Housni Touhami, Y., Kourissen, M., Semlali, C., Najid, K., Hamidou, A., Bout, A., Aarab, C., Boujraf, S., Rammouz, I., Berraho, M., El Fakir, S., & Aalouane, R. (2023). Cultural Adaptation and Validation of the Modified Checklist for Autism in Toddlers, Revised with Follow-up in Moroccan Arabic dialect. *L'Encephale*, 49(1), 15–20. <https://doi.org/10.1016/j.encep.2021.09.010>
- Tek, S., & Landa, R. J. (2012). Differences in autism symptoms between minority and non-minority toddlers. *Journal of autism and developmental disorders*, 42, 1967-1973.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.
- U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes. 2006 Oct 11;4:79. doi: 10.1186/1477-7525-4-79. PMID: 17034633; PMCID: PMC1629006.
- Uono, S., & Hietanen, J. K. (2015). Eye contact perception in the west and east: A cross-cultural study. *Plos one*, 10(2), e0118094.
- Varma, A., & Iskandar, J. W. (2014). Challenges in diagnosis of autism and the struggle of using western screening tools in different cultures. Psychiatrist's perspective. *Indian pediatrics*, 51(5), 356-357.
- van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. SAGE Publications.

- van de Vijver, F. J. R. (2003). Bias and equivalence: Cross-cultural perspectives. In J. A. Harkness, F. J. R. van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 143–155). Wiley.
- Vinden, P. G. (1999). Children's understanding of mind and emotion: A multi-culture study. *Cognition and Emotion*, 13(1), 19–48. <https://doi.org/10.1080/026999399379375>
- Volkmar, F., Chawarska, K., & Klin, A. (2005). Autism in infancy and early childhood. *Annual Review of Psychology*, 56, 315–336.
- Vorster, C., Kritzinger, A., Lekganyane, M., Taljard, E., & van der Linde, J. (2022). Cultural adaptation and Northern Sotho translation of the Modified Checklist for Autism in Toddlers. *South African Journal of Childhood Education*, 11(1), 968.
- Wallace, S., Fein, D., Rosanoff, M., Dawson, G., Hossain, S., Brennan, L., ... & Shih, A. (2012). A global public health strategy for autism spectrum disorders. *Autism Research*, 5(3), 211-217.
- Wallis, K. E., Guthrie, W., Bennett, A. E., Gerdes, M., Levy, S. E., Mandell, D. S., & Miller, J. S. (2020). Adherence to screening and referral guidelines for autism spectrum disorder in toddlers in pediatric primary care. *PloS one*, 15(5), e0232335.
- Ware Jr, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical care*, 220-233.
- Watson, L. R., Baranek, G. T., Crais, E. R., Steven Reznick, J., Dykstra, J., & Perryman, T. (2007). The first year inventory: retrospective parent responses to a questionnaire designed to identify one-year-olds at risk for autism. *Journal of autism and developmental disorders*, 37, 49-61.

- Werner, E., Dawson, G., Munson, J., & Osterling, J. (2005). Variation in early developmental course in autism and its relation with behavioral outcome at 3–4 years of age. *Journal of Autism and Developmental Disorders*, 35, 337–350.
- Wetherby, A. M., Brosnan-Maddox, S., Peace, V., & Newton, L. (2008). Validation of the Infant—Toddler Checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism*, 12(5), 487-511.
- Wetherby, A. M., Goldstein, H., Cleary, J., Allen, L., & Kublin, K. (2003). Early identification of children with communication disorders: Concurrent and predictive validity of the CSBS Developmental Profile. *Infants & Young Children*, 16(2), 161-174.
- Wieckowski, A. T., Williams, L. N., Rando, J., Lyall, K., & Robins, D. L. (2023). Sensitivity and specificity of the modified checklist for autism in toddlers (original and revised): A systematic review and meta-analysis. *JAMA Pediatrics*, 177(4), 373–383.
- Willis, G. B. (1999). Cognitive interviewing: A “how to” guide. *Research Triangle Park, NC: Research Triangle Institute*.
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications.
- Willis, G., Lawrence, D., Thompson, F., Kudela, M., Levin, K., & Miller, K. (2005). The use of cognitive interviewing to evaluate translated survey questions: lessons learned. In *Conference of the Federal Committee on Statistical Methodology, Arlington, VA*.
- Willis, G. B., & Artino Jr, A. R. (2013). What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *Journal of graduate medical education*, 5(3), 353-356.

- Willis, G. B. (2015). The practice of cross-cultural cognitive interviewing. *Public opinion quarterly*, 79(S1), 359-395.
- World Health Organization (2014). *WHA67.8: Comprehensive and coordinated efforts for the management of autism spectrum disorders. Sixty-Seventh World Health Assembly, Geneva, Switzerland*. Retrieved from http://www.who.int/mental_health/maternal-child/WHA67.8_resolution_autism.pdf.
- Wörmann, V., Holodynski, M., Kärtner, J., & Keller, H. (2012). A cross-cultural comparison of the development of the social smile: A longitudinal study of maternal and infant imitation in 6-and 12-week-old infants. *Infant Behavior and Development*, 35(3), 335-347.
- Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology*, 43(2), 303-311.
- Zayas, L. H., & Solari, F. (1994). Early childhood socialization in Hispanic families: Context, culture, and practice. *American Journal of Orthopsychiatry*, 64(3), 358–373.
<https://doi.org/10.1037/h0079535>
- Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., ... & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, 15(5), 778-790.
- Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of management review*, 12(2), 250-264.
- Zucker, S., Miska, M., Alaniz, L. G., & Guzmán, L. (2005). Transadaptation: Publishing assessments in world languages. *Assessment report. September*.

Zuckerman, K. E., Mattox, K., Donelan, K., Batbayar, O., Baghaee, A., & Bethell, C. (2013).

Pediatrician identification of Latino children at risk for autism spectrum disorder.

Pediatrics, 132(3), 445-453.

Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Kasari, C., Carter, A., Granpeesheh, D., ... &

Natowicz, M. R. (2015). Early intervention for children with autism spectrum disorder under 3 years of age: recommendations for practice and research. *Pediatrics*,

136(Supplement_1), S60-S81.

Zwaigenbaum, L., Brian, J. A., & Ip, A. (2019). Early detection for autism spectrum disorder in

young children. *Paediatrics & Child Health*, 24(7), 424-432.

Appendix: Tables & Figures

Tables

Table 1

M-CHAT-R™		
Please answer these questions about your child. Keep in mind how your child <u>usually</u> behaves. If you have seen your child do the behavior a few times, but he or she does not usually do it, then please answer no . Please circle yes or no for every question. Thank you very much.		
1. If you point at something across the room, does your child look at it? (FOR EXAMPLE , if you point at a toy or an animal, does your child look at the toy or animal?)	Yes	No
2. Have you ever wondered if your child might be deaf?	Yes	No
3. Does your child play pretend or make-believe? (FOR EXAMPLE , pretend to drink from an empty cup, pretend to talk on a phone, or pretend to feed a doll or stuffed animal?)	Yes	No
4. Does your child like climbing on things? (FOR EXAMPLE , furniture, playground equipment, or stairs)	Yes	No
5. Does your child make <u>unusual</u> finger movements near his or her eyes? (FOR EXAMPLE , does your child wiggle his or her fingers close to his or her eyes?)	Yes	No
6. Does your child point with one finger to ask for something or to get help? (FOR EXAMPLE , pointing to a snack or toy that is out of reach)	Yes	No
7. Does your child point with one finger to show you something interesting? (FOR EXAMPLE , pointing to an airplane in the sky or a big truck in the road)	Yes	No
8. Is your child interested in other children? (FOR EXAMPLE , does your child watch other children, smile at them, or go to them?)	Yes	No
9. Does your child show you things by bringing them to you or holding them up for you to see – not to get help, but just to share? (FOR EXAMPLE , showing you a flower, a stuffed animal, or a toy truck)	Yes	No
10. Does your child respond when you call his or her name? (FOR EXAMPLE , does he or she look up, talk or babble, or stop what he or she is doing when you call his or her name?)	Yes	No
11. When you smile at your child, does he or she smile back at you?	Yes	No
12. Does your child get upset by everyday noises? (FOR EXAMPLE , does your child scream or cry to noise such as a vacuum cleaner or loud music?)	Yes	No
13. Does your child walk?	Yes	No
14. Does your child look you in the eye when you are talking to him or her, playing with him or her, or dressing him or her?	Yes	No
15. Does your child try to copy what you do? (FOR EXAMPLE , wave bye-bye, clap, or make a funny noise when you do)	Yes	No
16. If you turn your head to look at something, does your child look around to see what you are looking at?	Yes	No
17. Does your child try to get you to watch him or her? (FOR EXAMPLE , does your child look at you for praise, or say "look" or "watch me"?)	Yes	No
18. Does your child understand when you tell him or her to do something? (FOR EXAMPLE , if you don't point, can your child understand "put the book on the chair" or "bring me the blanket"?)	Yes	No
19. If something new happens, does your child look at your face to see how you feel about it? (FOR EXAMPLE , if he or she hears a strange or funny noise, or sees a new toy, will he or she look at your face?)	Yes	No
20. Does your child like movement activities? (FOR EXAMPLE , being swung or bounced on your knee)	Yes	No

MCHAT-R/F

Note. This table shows the questions posed on the MCHAT-R/F.

Table 2

Demographics

Demographic Category	n	%
Gender		
Female	10	71.40%
Male	4	28.60%
Age Range		
30-39	10	71.40%
40-49	2	14.30%
20-29	2	14.30%
Race/Ethnicity		
Asian	7	50.00%
White	5	35.70%
Latinx	1	7.10%

Black or African American	1	7.10%
---------------------------	---	-------

Country of Origin

United States of America	3	21.40%
--------------------------	---	--------

Pakistan	2	14.30%
----------	---	--------

Bangladesh	2	14.30%
------------	---	--------

India	1	7.10%
-------	---	-------

Russian Federation	1	7.10%
--------------------	---	-------

Germany	1	7.10%
---------	---	-------

France	1	7.10%
--------	---	-------

South Korea	1	7.10%
-------------	---	-------

China	1	7.10%
-------	---	-------

Democratic Republic of the Congo	1	7.10%
-------------------------------------	---	-------

Language(s) Spoken in Household

Urdu	2	14.30%
------	---	--------

Bengali	2	14.30%
---------	---	--------

Hebrew	1	7.10%
Spanish	2	14.30%
German	1	7.10%
Congolese	1	7.10%
Hindi and Marathi	1	7.10%
French and Turkish	1	7.10%
Russian, Hebrew, English	1	7.10%
Unspecified	2	14.30%

Years in the U.S.

2-9 years	6	42.85%
10-19 years	2	14.28%
20-29 years	3	21.42%
30+ years	3	21.42%

Immigrant/Refugee Status

1st generation	8	57.10%
not reported	3	21.40%

3rd generation	1	7.10%
2nd generation	1	7.10%
Refugee - 1st generation	1	7.10%

Annual Household Income

\$15,000 - \$24,999	1	7.10%
\$25,000 - \$34,999	1	7.10%
\$35,000 - \$49,999	2	14.30%
\$50,000 - \$74,999	3	21.40%
\$75,000 - \$99,999	4	28.60%
\$150,000 - \$199,999	3	21.40%
Over \$200,000	1	7.10%

Age of Child (in months)

16 -24 months	4	28.6%
25 - 30 months	3	21.4%
31 - 36 months	2	14.3%
37 - 48 months	5	35.7%

Acculturation

Bicultural	9	64.30%
Traditional (ethnic centered)	3	21.40%
Assimilated (dominant centered)	2	14.30%

Worldview

Collectivistic	7	50.00%
Individualistic	7	50.00%

Participant Demographics

Note. This table shows full sample demographics

Table 3

Question	Construct	N	Concern Level
Q12	Upset by everyday noises	12	High Concern
Q18	Response to verbal instructions	12	High Concern
Q1	Follow a point	11	High Concern
Q5	Unusual finger movement near eyes	11	High Concern
Q16	Response to joint attention	10	High Concern
Q19	Social referencing	10	High Concern
Q17	Social initiation	7	Moderate Concern
Q14	Eye contact during interactions	6	Moderate Concern
Q3	Pretend play	5	Moderate Concern
Q6	Point to request	5	Moderate Concern
Q8	Interest in other children	5	Moderate Concern
Q10	Response to name	5	Moderate Concern
Q7	Point to share interest	4	Low Concern
Q9	Shared showing	3	Low Concern
Q11	Responsive social smile	3	Low Concern
Q15	Imitation	1	Low Concern

Cognitive Domain Errors by Item

Note: This table shows how many participants of the full sample (n=14) endorsed concerns for each analyzed item.

Table 4

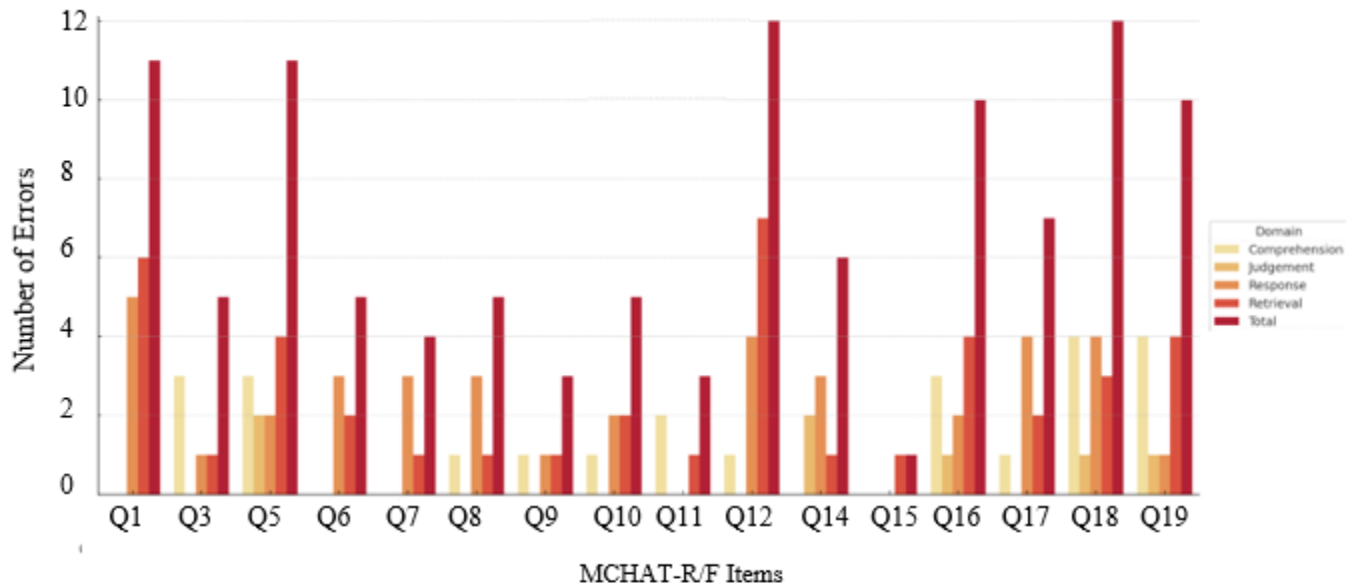
Item	Construct	Domains with Errors	Domain
Q16	Response to joint attention	Comprehension, Retrieval, Judgment, Response	4 domains
Q18	Response to verbal instructions	Comprehension, Retrieval, Judgment, Response	4 domains
Q1	Follow a point	Retrieval, Response	2 domains
Q5	Unusual finger movement near eyes	Retrieval, Response	2 domains
Q6	Point to request	Retrieval, Response	2 domains
Q10	Response to name	Retrieval, Response	2 domains
Q12	Upset by everyday noises	Retrieval, Response	2 domains
Q3	Pretend play	Judgment	1 domain
Q11	Responsive social smile	Judgment	1 domain
Q14	Eye contact with interactions	Judgment	1 domain
Q15	Imitation	Response	1 domain
Q17	Social initiation	Response	1 domain
Q19	Social referencing	Retrieval	1 domain

Interpretative Concern by Domain Spread

Note. This table shows which domain(s) were flagged with interpretation difficulties for each analyzed item

Figures

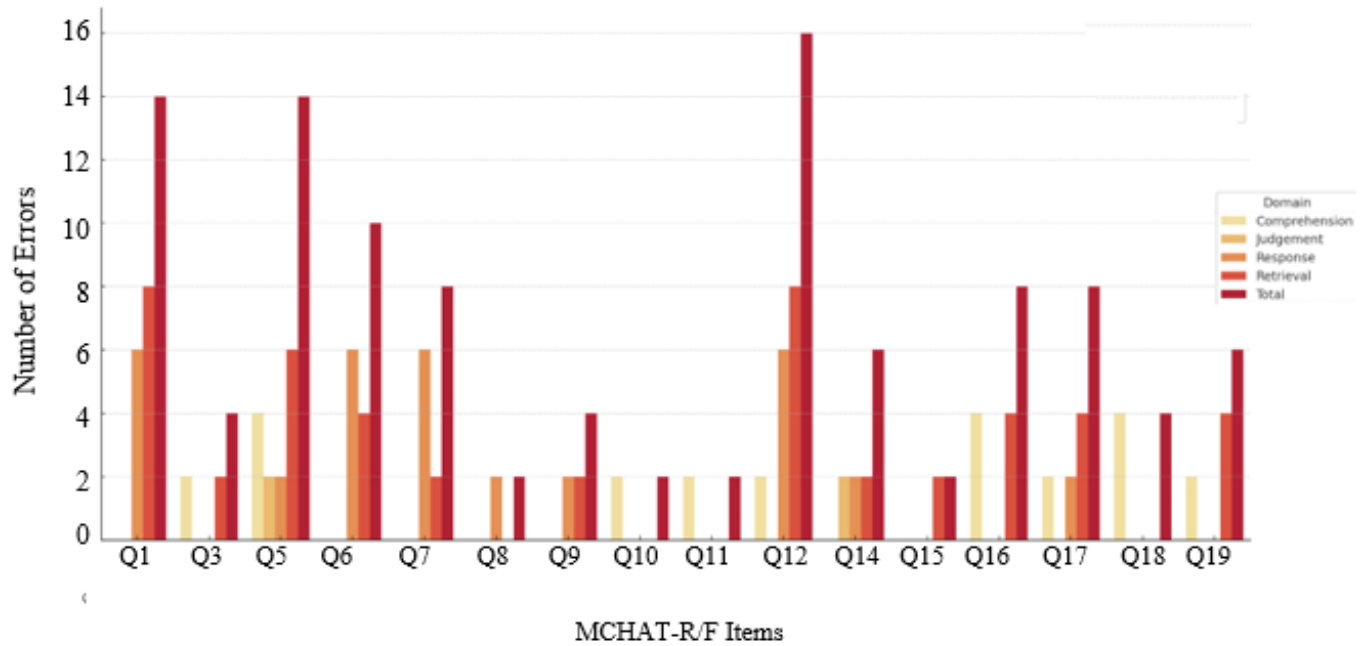
Figure 1



Total cognitive domain errors by MCHAT-R/F question

Note. This figure displays the total number of cognitive domain errors (comprehension, retrieval, judgement, response) flagged per MCHAT-R/F question.

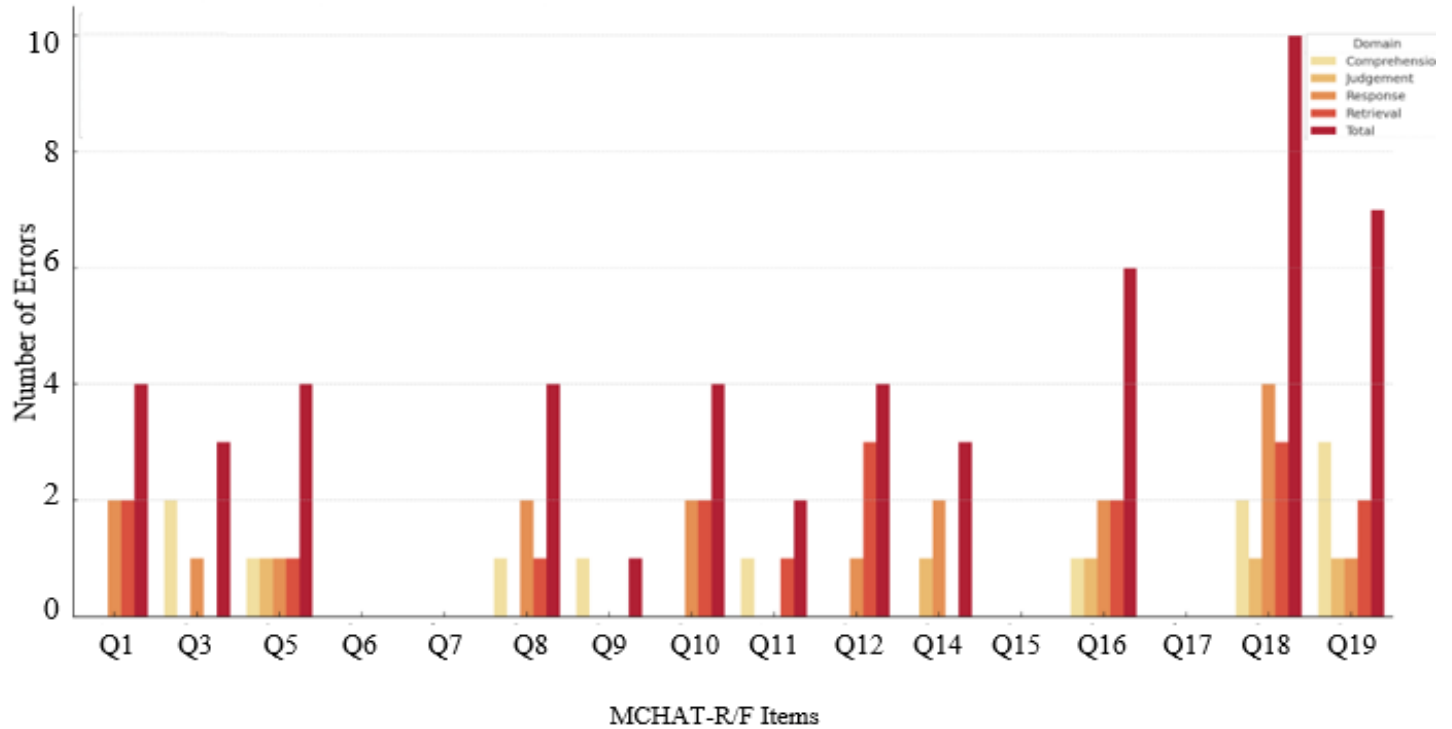
Figure 2



Total cognitive domain errors by MCHAT-R/F question – cultural incongruence

Note. This figure displays the number of cognitive domain errors (comprehension, retrieval, judgement, response) flagged by MCHAT-R/F question for participant responses with cultural incongruence

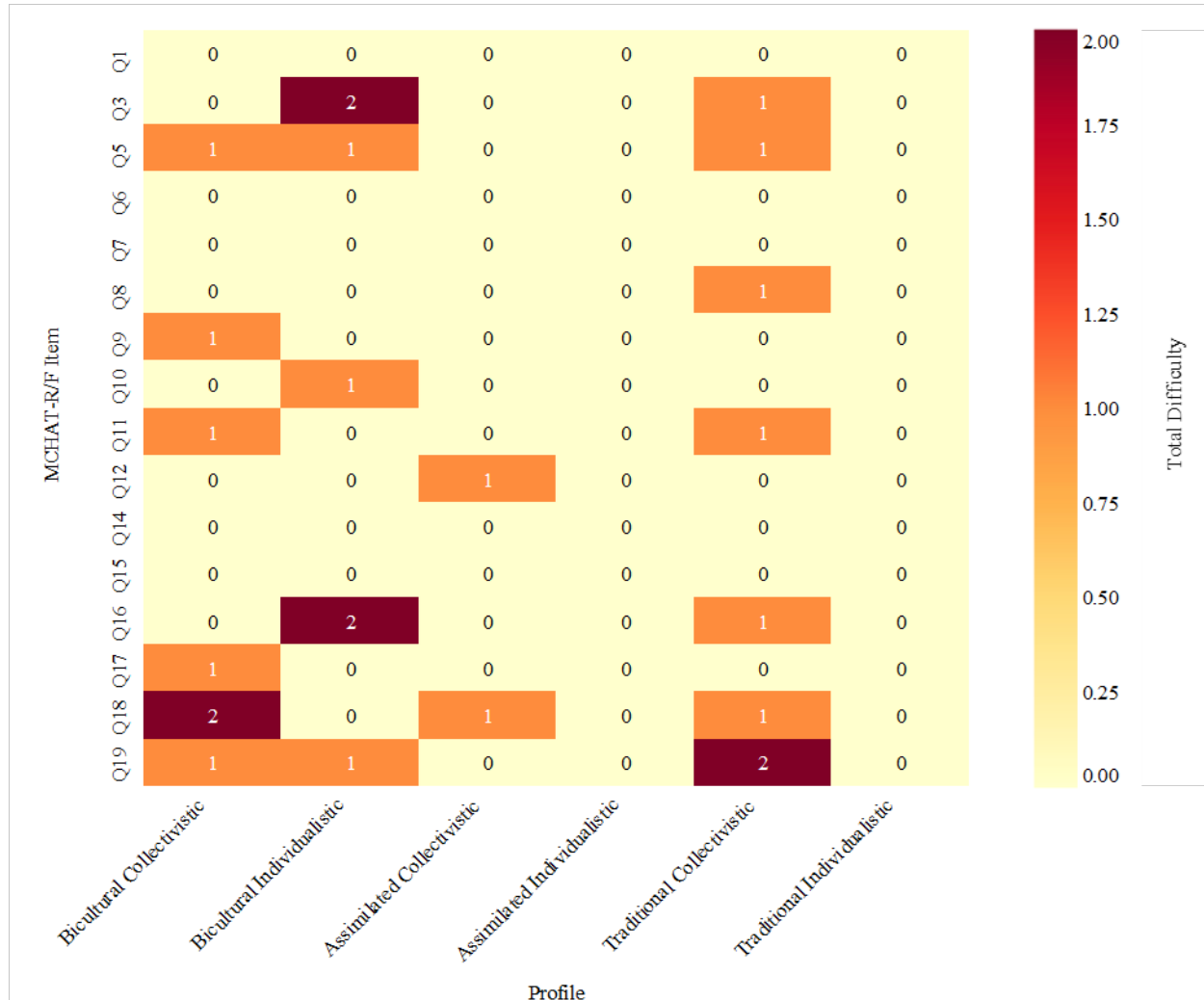
Figure 3



Total cognitive domain errors by MCHAT-R/F question – without cultural incongruence

Note. This figure displays the number of cognitive domain errors (comprehension, retrieval, judgement, response) flagged by MCHAT-R/F question for participant responses without cultural incongruence

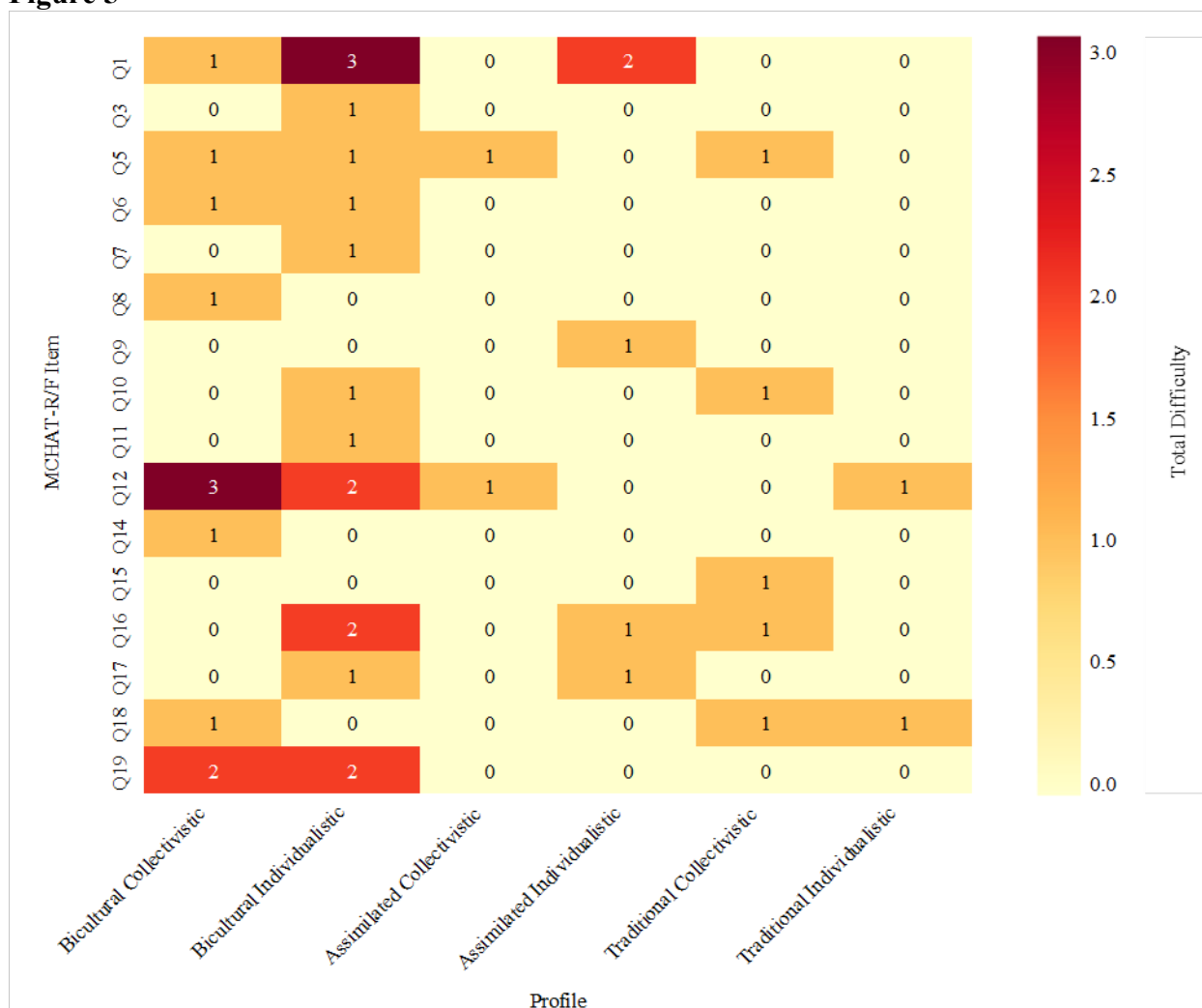
Figure 4



Comprehension domain errors by MCHAT-R/F question & profile - Total

Note. This figure displays the number of comprehension domain difficulties flagged by MCHAT-R/F question for all participant responses by acculturation and worldview profile

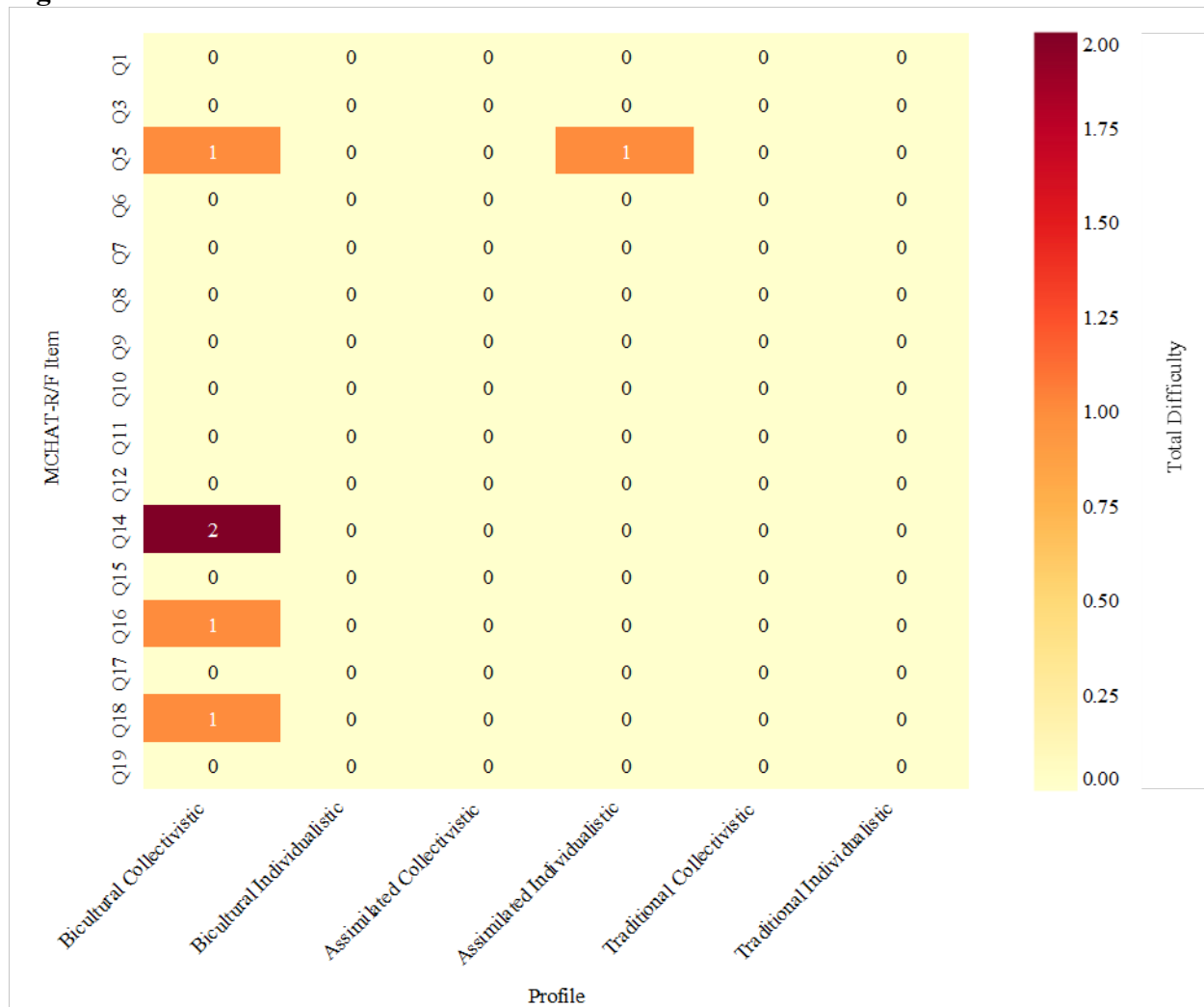
Figure 5



Retrieval domain errors by MCHAT-R/F question & profile - Total

Note. This figure displays the number of retrieval domain difficulties flagged by MCHAT-R/F question for all participant responses by acculturation and worldview profile

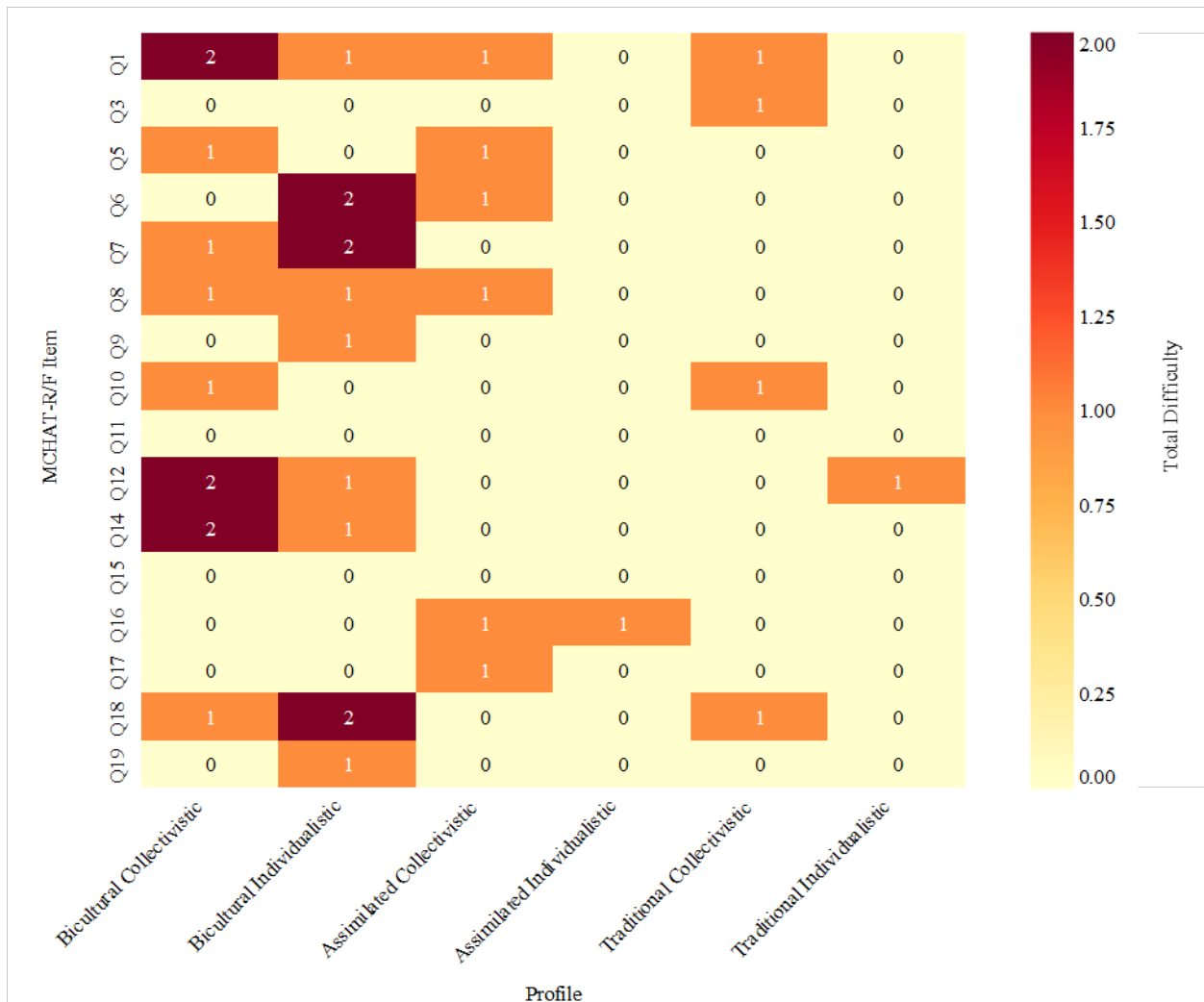
Figure 6



Judgement domain errors by MCHAT-R/F question & profile - Total

Note. This figure displays the number of judgement domain difficulties flagged by MCHAT-R/F question for all participant responses by acculturation and worldview profile

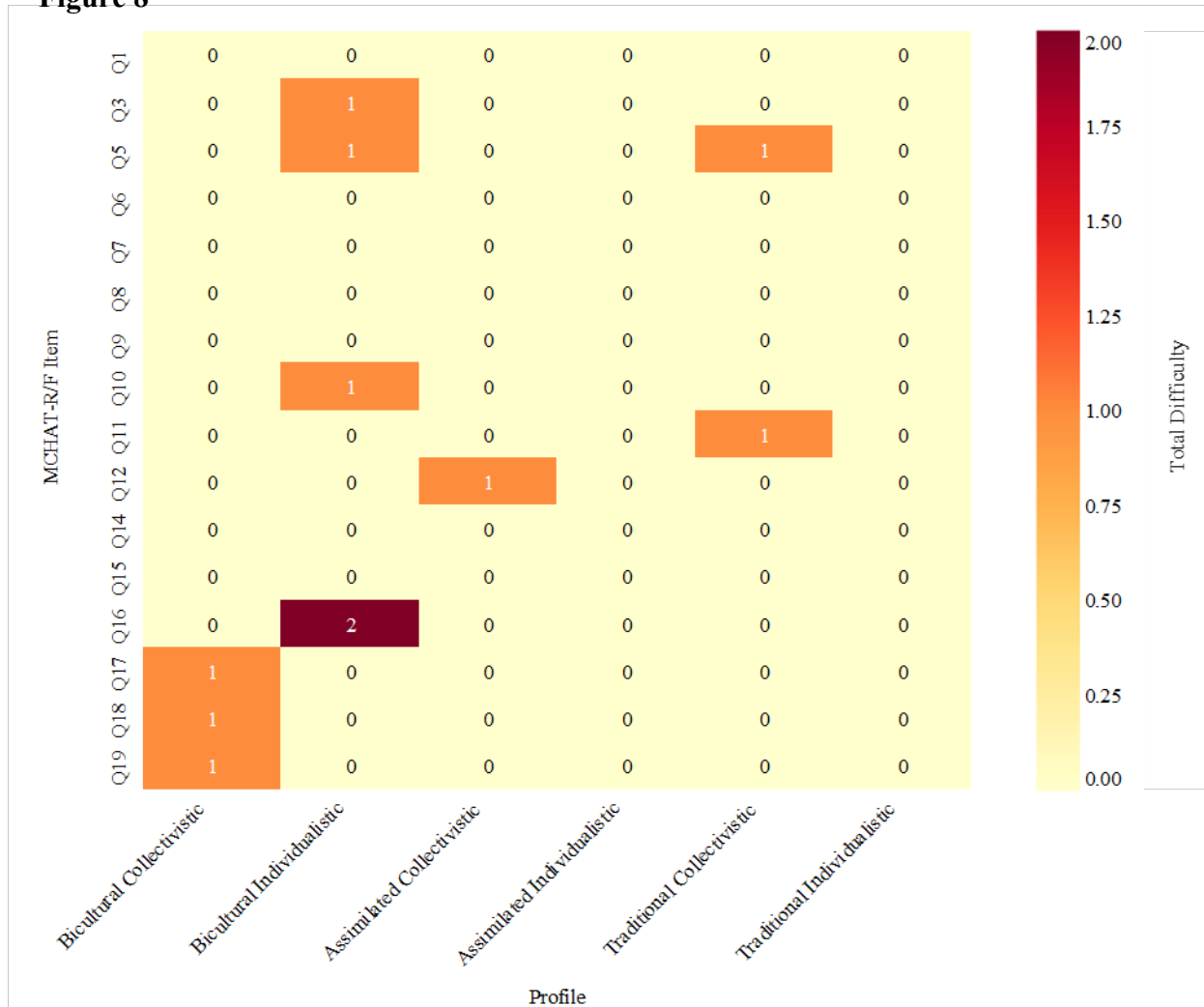
Figure 7



Response domain errors by MCHAT-R/F question & profile - Total

Note. This figure displays the number of response domain difficulties flagged by MCHAT-R/F question for all participant responses by acculturation and worldview profile

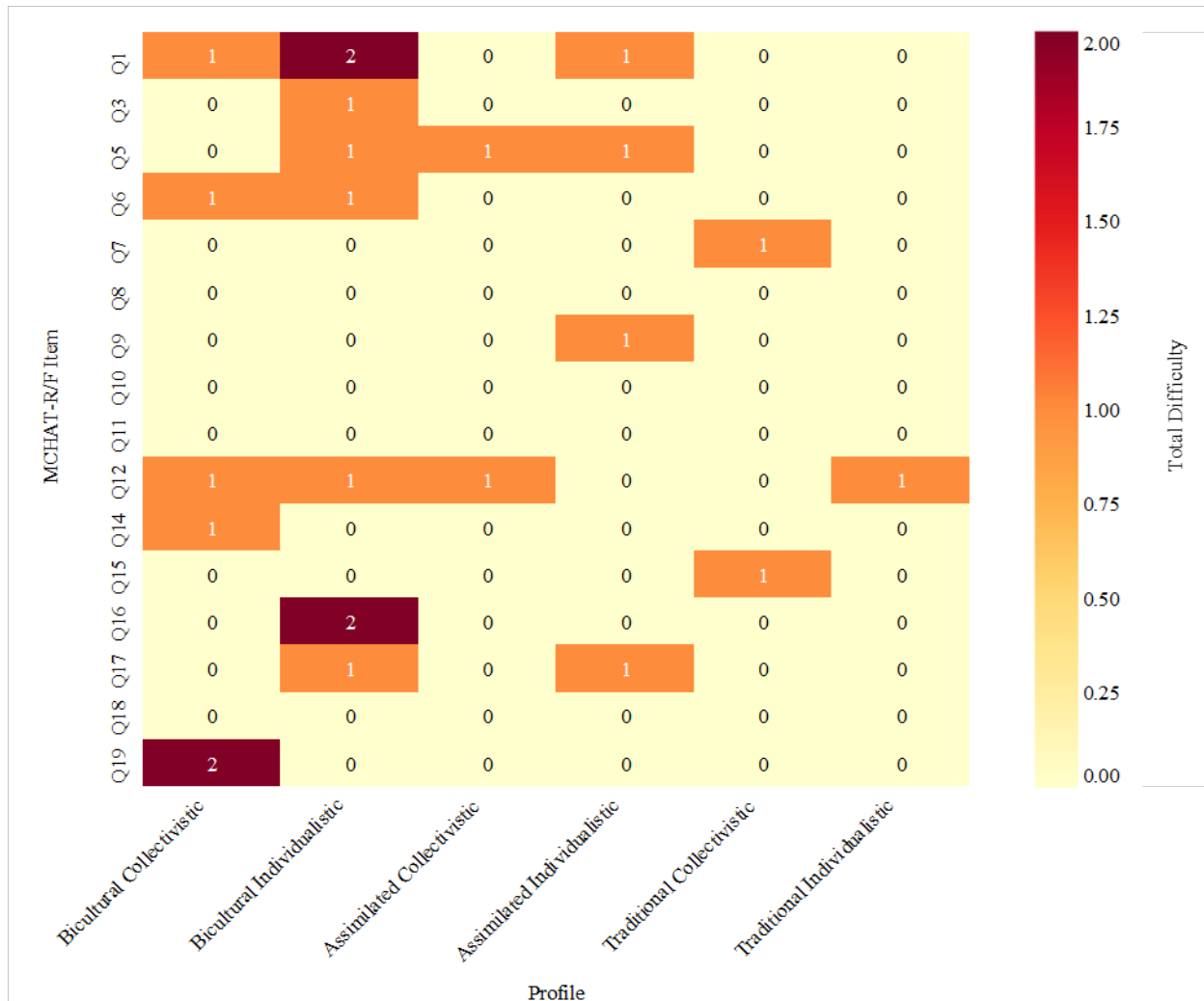
Figure 8



Comprehension domain errors by MCHAT-R/F question - Cultural Incongruence

Note. This figure displays the number of comprehension domain difficulties with cultural incongruence by MCHAT-R/F item and acculturation and worldview profile

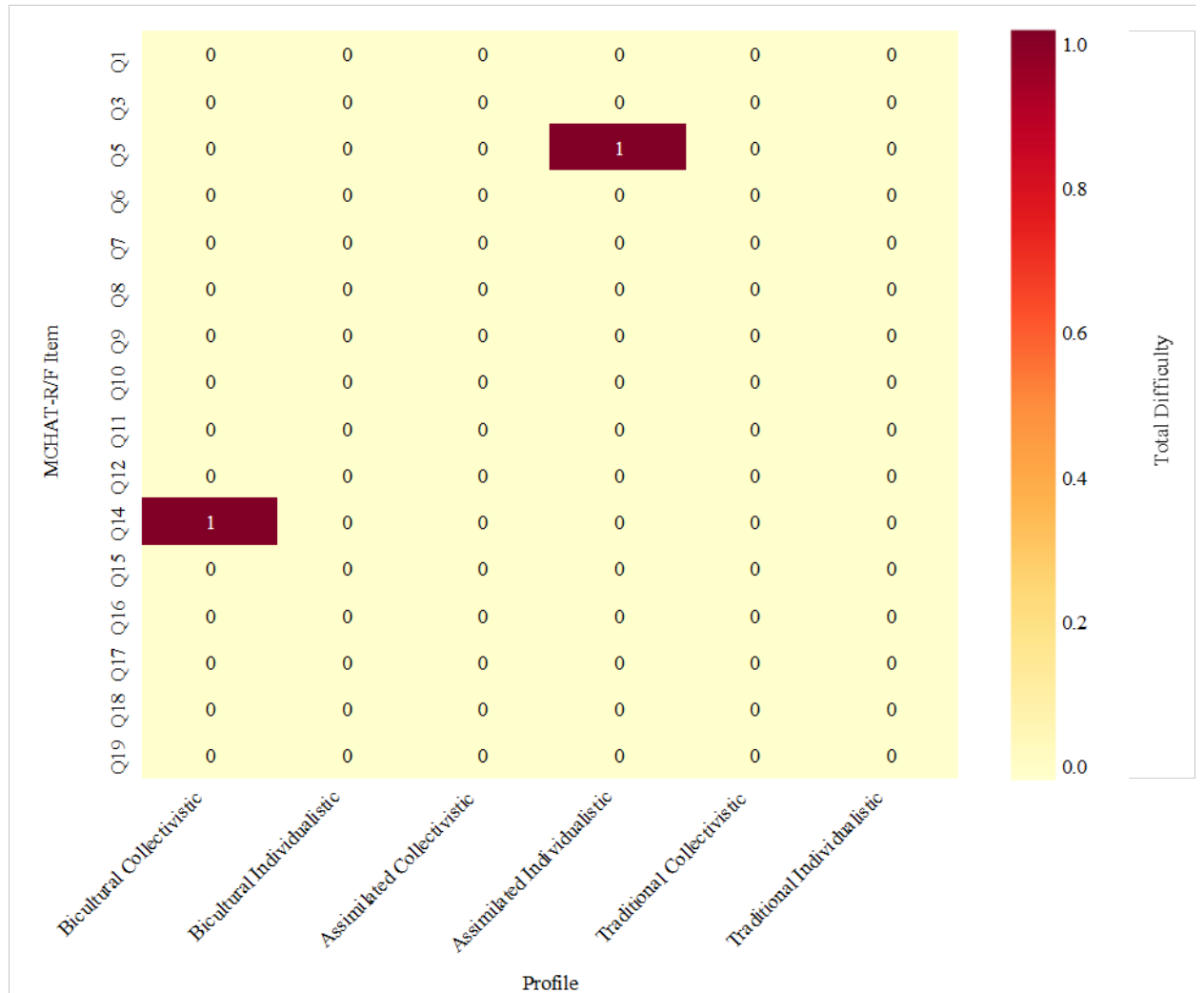
Figure 9



Retrieval domain errors by MCHAT-R/F question - Cultural Incongruence

Note. This figure displays the number of retrieval domain difficulties with cultural incongruence by MCHAT-R/F item and acculturation and worldview profile

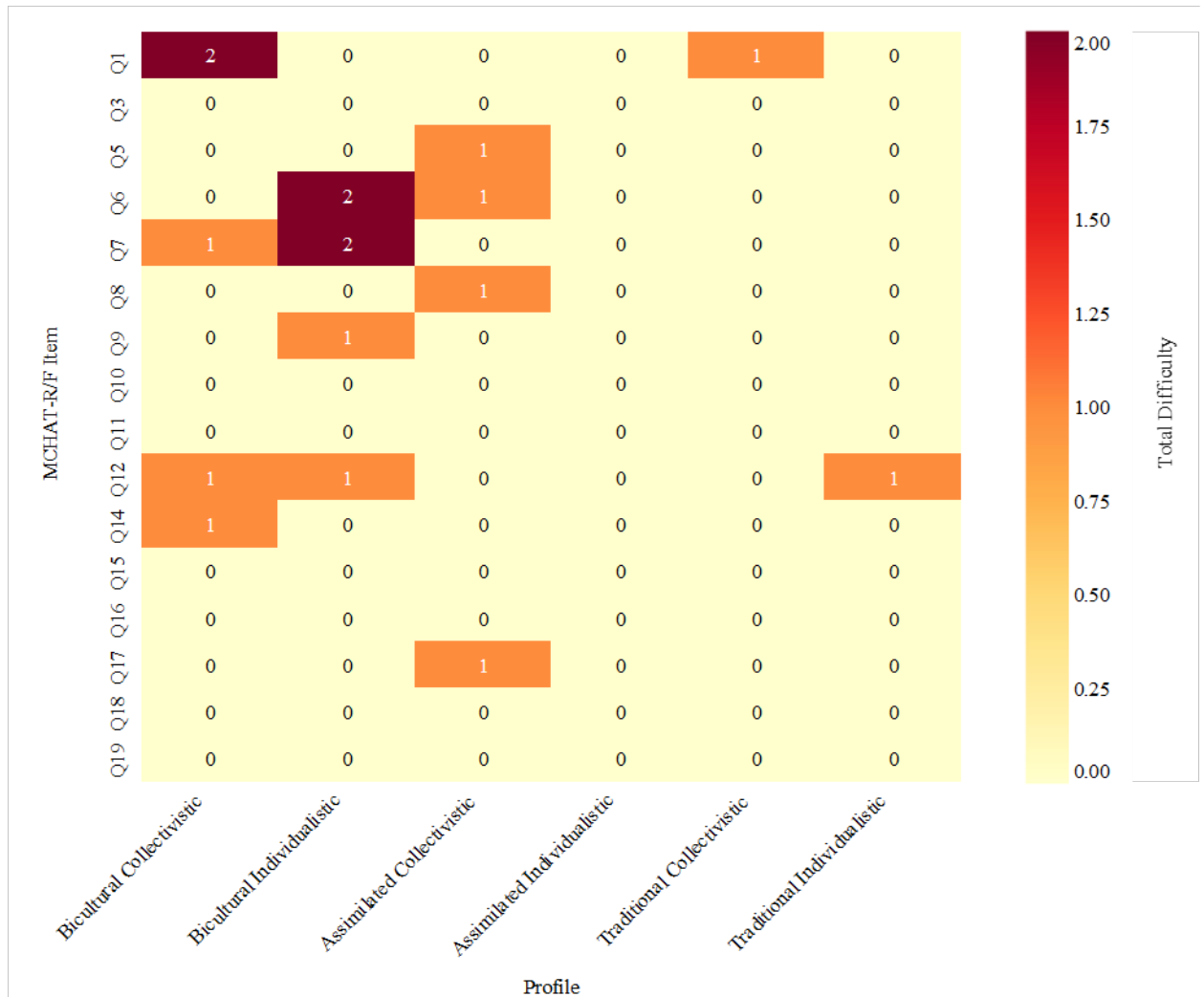
Figure 10



Judgement domain errors by MCHAT-R/F question - Cultural Incongruence

Note. This figure displays the number of judgement domain difficulties with cultural incongruence by MCHAT-R/F item and acculturation and worldview profile

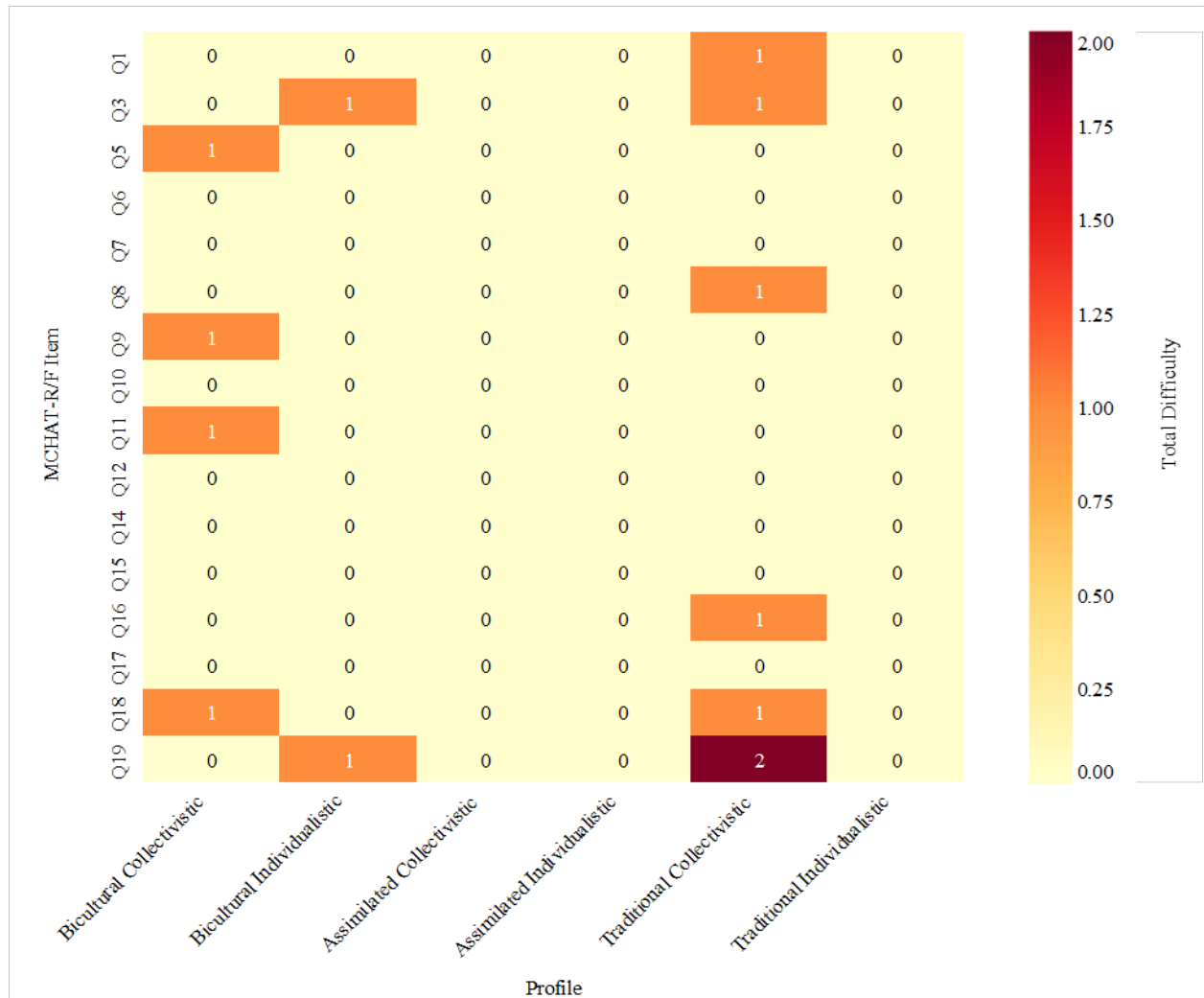
Figure 11



Response domain errors by MCHAT-R/F question - Cultural Incongruence

Note. This figure displays the number of response domain difficulties with cultural incongruence by MCHAT-R/F item and acculturation and worldview profile

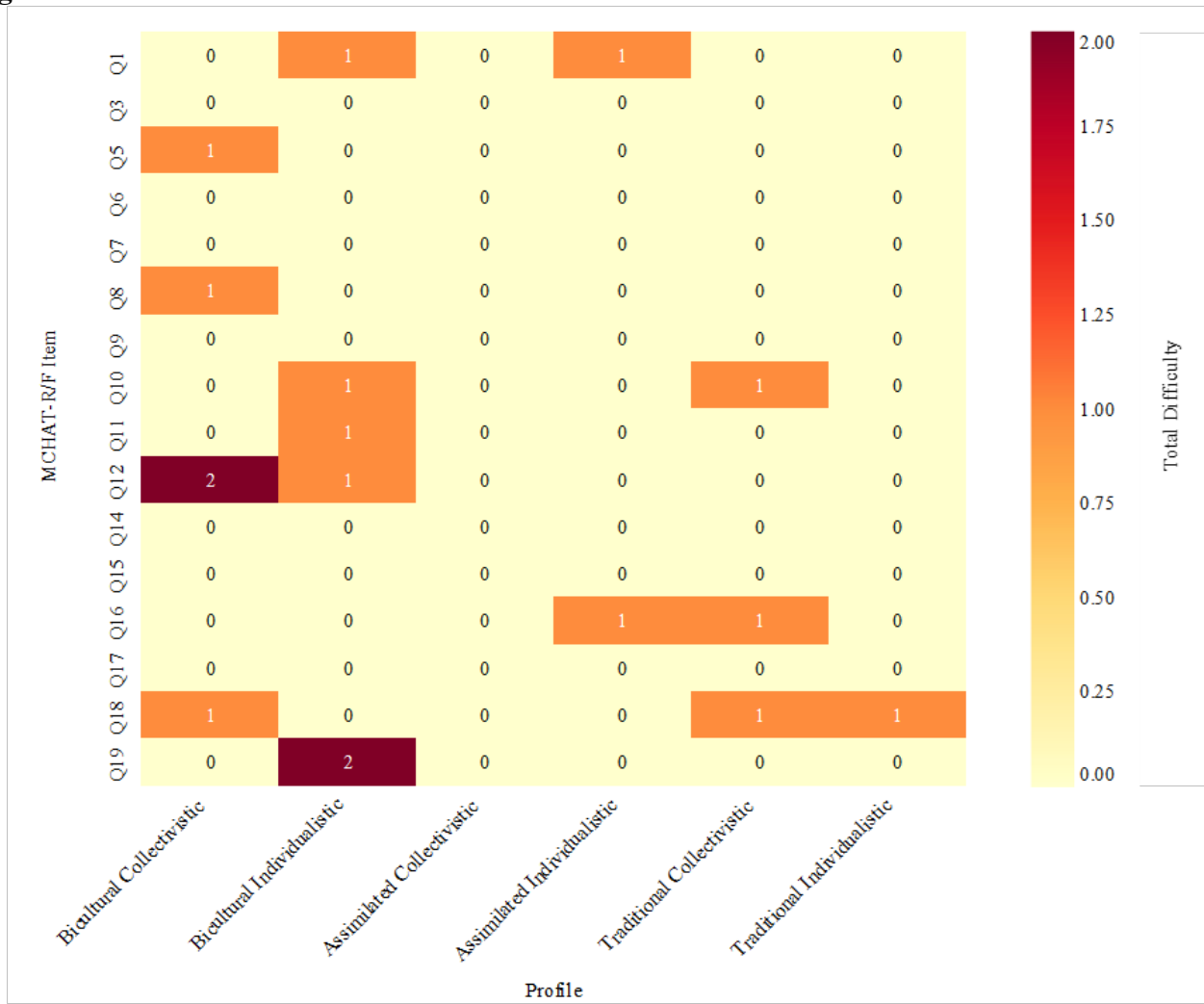
Figure 12



Comprehension domain errors by MCHAT-R/F question without Cultural Incongruence

Note. This figure displays the number of response domain difficulties without cultural incongruence by MCHAT-R/F item and acculturation and worldview profile

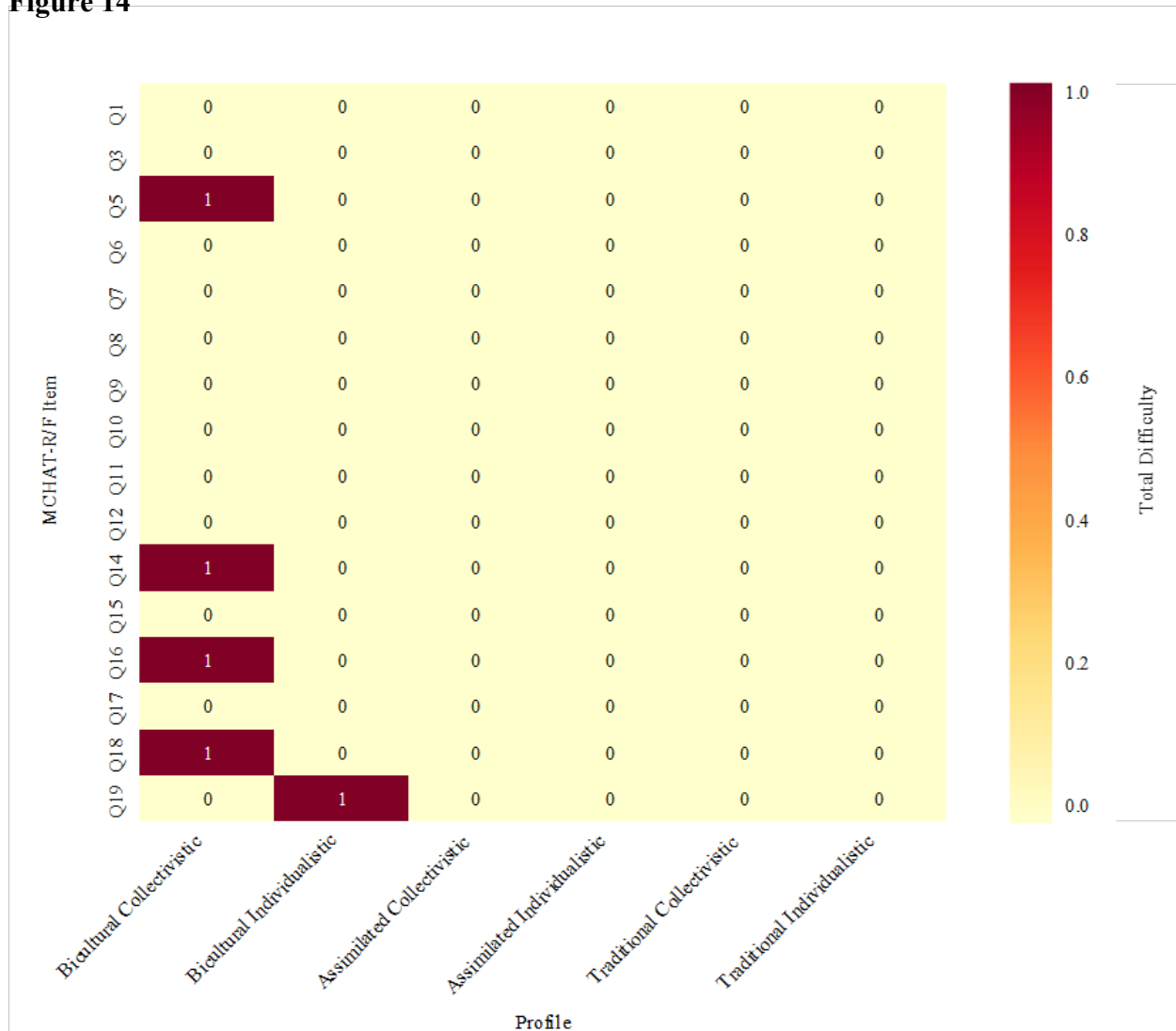
Figure 13



Retrieval domain errors by MCHAT-R/F question without Cultural Incongruence

Note. This figure displays the number of response domain difficulties without cultural incongruence by MCHAT-R/F item and acculturation and worldview profile

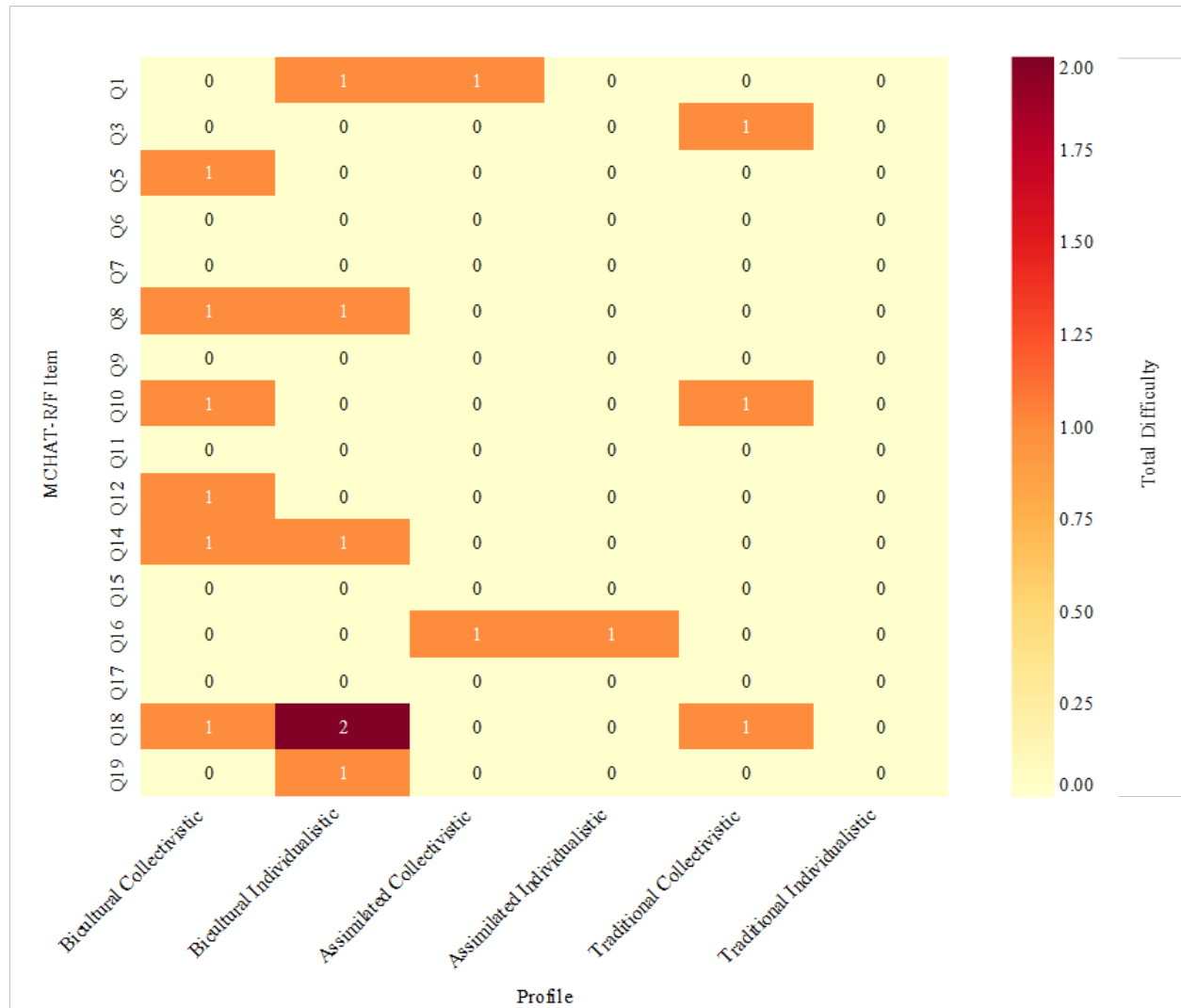
Figure 14



Judgement domain errors by MCHAT-R/F question without Cultural Incongruence

Note. This figure displays the number of judgement domain difficulties without cultural incongruence by MCHAT-R/F item and acculturation and worldview profile

Figure 15



Response domain errors by MCHAT-R/F question without Cultural Incongruence

Note. This figure displays the number of response domain difficulties without cultural incongruence by MCHAT-R/F item and acculturation and worldview profile