Title
Rethinking the Validity of Autism Assessment

Author
Choi, Sunghee

Publication Date
2022

UNIVERSITY OF CALIFORNIA

Santa Barbara

Rethinking the Validity of Autism Assessment

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Education

by

Sunghee Choi

Committee in charge:
Professor Diana Arya, Co-Chair
Professor Andrew Maul, Co-Chair
Professor Rachel Lambert
Professor Nancy Collins

June, 2022

The dissertation of Sunghee Choi is approved.

Rachel Lambert

Nancy Collins

Diana Arya, Committee Co-Chair

Andrew Maul, Committee Co-Chair

May, 2022

VITA OF SUNGHEE CHOI
May 2022

EUDCATION

Expected June, 2022	Ph.D in Education, UC Santa Barbara
2019	M.A. in Education, UC Santa Barbara
2001-2002	Graduate School, Art History, Northern Illinois University
2000	B.A .in English and German Literature and Language, Yonsei University, Seoul, Korea

EXPERIENCE

Teaching assistant at UCSB

2022	Asian American Migration (ASAM2)
2021	Introduction to Communication (COMM1)
2021	Grassroots Transnational Feminist Movements (FEMST50)
2021	Asian American Literature (ASAM5)
2020	Communication Research Method (COMM88)
2020	Introduction to Art (ARTHI1)
2019	Asian American History (ASAM1)
2018	Language, Power, and Learning (Ling 187)

Professional Experience

2017-2018	Case coordinator in UCSB OISS

CONFERENCE PRESENTATIONS

2022 Rachel Lambert, Kara Imm, Sunghee Choi, Rachel Schuck, & Avery McnNiff. “Empathy Is the Anchor”; Universal Design Learning as Design Thinking.” Roundtable at American Educational Research Association (AREA), San Diego, CA

2022 Rachel Schuck & Sunghee Choi. “Developing a Measure of Teachers Knowledge of Neurodiversity.” Roundtable at American Educational Research Association (AREA), San Diego, CA

2020 Sunghee Choi. “Textual Silence in Alternate Assessment Resources.” Paper at California Chapter of The Association of Severely Handicapped (CAL-TASH), Sacramento, CA

2020 Sunghee Choi. “Textual Silence in Alternate Assessment Resources.” Roundtable at American Educational Research Association (AREA), Accepted.

PUBLICATIONS

2022 Choi, S., Schuck, R. K., & Imm, K. (Redesigning Deficit-Laden Assessments for Neurodivergent Students. In *Handbook of Research on Challenging Deficit Thinking for Exceptional Education Improvement* (pp. 179-206). IGI Global.

2021 Lambert, R., Imm, K., Mcniff, A., Schuck, R., Choi, S. “UDL is the What, Design Thinking is the How:” Designing for Differentiation in Mathematics. *Mathematics Teacher Education and Development*

REFERENCES
Dr. Andrew Maul
Associate Professor, Gevirtz Graduate School of Education
University of California, Santa Barbara
(805) 893-7770
amaul@education.ucsb.edu

Dr. Diana Arya
Associate Professor, Gevirtz Graduate School of Education
University of California, Santa Barbara
(510) 363-1265
darya@education.ucsb.edu

Dr. Rachel Lambert
Associate Professor, Gevirtz Graduate School of Education
University of California, Santa Barbara
rlambert@education.ucsb.edu

Dr. Nancy Collins
Professor, School of Psychological & Brain Sciences
University of California, Santa Barbara
nancy.collins@psych.ucsb.edu

ABSTRACT

Rethinking the Validity of Autism Assessment

by

Sunghee Choi

Traditionally, most autism assessment instruments are based on medical models and designed to identify social communication deficits and behavioral abnormality in an individual. However, as more autistic narratives reveal the insider views of autists, some scholars and autistic activists support the neurodiversity model and assert the acceptance of autism as difference and diversity instead of deficits or impairment. With this conceptual shift, it is appropriate to rethink the validity of current autism assessment practice. Adopting contemporary validity theories, which emphasizes the ontological definition of an attribute, response process, and ethical consequences of measurement, this study examined how autism has been defined, how the definitions of autism affect the way autism is understood and assessed, and how autism assessment affects the way people define autism. Also, as a case study, the current form of Social Communication Question was revalidated following Wilson's (2004) Four Building Block approach. The interview conducted with autistic adults and autistic people's caregiver to examine the response process revealed how these people interpret the items in SCQ and how autism assessment tools like SCQ have affected the way they conceptualize autism and establish their self-identities.

Rethinking the Validity of Autism Assessment

Recently, autism[1] assessment validity studies are conducted by more and more advanced psychometric models and fit statistics to calculate the discriminant validity in order to distinguish autism from other developmental disabilities. However, these validity studies do not reflect the progress of contemporary validity theories, and hardly address the theoretical aspects of the attributes that are measured. Although each contemporary validity theory has slightly different emphases, these theories do not confine the role of validation to calculate the correlation between the test score and the criteria or identify the most precise cut score to distinguish different constructs. Instead, some scholars call for attention to the ontology of the attribute they measure and the causality between the attribute and the test score to make validity claims (Borsboom et al., 2004; Slaney, 2017). Also, other scholars insist that the response process of actual users of instruments and the social and ethical consequences of measurement should be considered in validation of an instrument (Kane, 2006; Messick, 1995).

As contemporary validity theories pay more attention to the ontological definition of the attribute they measure and the response of actual users of instruments, scholars in the disability studies and autistic activists also call for attention to similar three issues in autism research. As more and more first person narratives from autistic people are accumulated and respected, scholars and autistic self-advocates have emphasized the insider view to define autism and autistic people's agency in the decision making process (Dinishak, 2021; Hacking, 2009a, 2009b; Kapp et al., 2013; Nicolaidis, 2012; Nicolaidis et al., 2020).

---

[1] In this study, as I examined the history of autism diagnosis and the use of labels in diverse contexts, I use the term 'autism' instead of 'autism spectrum disorder (ASD)'. Only when indicating the change of the official label did I use the term ASD.

Traditionally, in most autism assessment instruments used by clinicians, autism has been defined by the medical model. In the medical model, autism is defined as deficits in social communication and abnormalities in their behavior due to neurobiological defects in individuals. For this reason, autism assessment instruments under the medical model are designed to identify deficits in social communication and abnormalities in the behaviors. With the advent of the neurodiversity model, however, more and more autistic people and scholars challenge the definition of autism under the medical model and contend that autism should be regarded as neurological diversity as we acknowledge human diversity in races or gender (Chapman2019; Nicolaidis, 2012). Even though the neurodiversity model is supported by autistic people and some scholars, widely used autism assessment instruments are still rooted in the medical model and autism assessment instruments based on the neurodiversity model have not yet to be developed. If a new autism assessment is developed based on the neurodiversity model, it might be more focused on the strengths of autistic people or the accommodations that can reduce the difficulties they encounter in the neurotypical world instead of focusing on identifying defects or abnormalities.

The second major issue in autism assessment is that most assessment tools are not designed to capture autistic people's voice directly. In autism assessment, children or adults under assessment are observed by caregivers or clinicians and this observation is rated by clinicians. In this process, the intention of autistic people's communication or behaviors can be misinterpreted by the third person who observes their behaviors. For example, the two gold standard tools in autism research and clinical practice, diagnostic assessment tools such as the Autism Diagnostic Interview – Revised (ADI-R; Lord et al., 1999) and the Autism Diagnostic Observation Schedule Second Edition (ADOS-2) (Lord et al., 1994) typically involve observing the child or adult under assessment and rating their behavior or ‘symptoms’ of autism by trained clinicians. Even though caregivers fill out some questionnaires and give accounts for the developmental history of a child, non-clinicians, such as caregivers and autistic adults, rarely have access to how ratings are decided or how diagnostic decisions are determined (Timmi et al., 2019). This exclusion of autistic people and their caregivers in the rating and diagnosis process has been justified with the rationale that only clinicians are specially trained for this process. It is quite ironic that on the one hand, the behavior reported by "non-clinically-trained" caregivers does not require direct observation by assessors and is accepted as valid data, while on the other hand their access to rating decisions are simultaneously denied because they are not trained (Timmi et al., 2016 b).

However, the more crucial issue is that direct collection of autistic people’s responses is very rare. As autistic individuals have been passively observed and rated by caregivers or clinicians, autistic people only exist in the description of the third person and it is hardly studied whether this description is valid without any bias or misunderstanding (Timmi et al., 2016 b; Timmi et al., 2019). Moreover, as Dinishak (2021) pointed out, this practice has deprived autistic people of their language to describe their firsthand experience. They do not have a chance to express their own language to describe autistic thoughts, feelings, sayings and doings. For example, some autism researchers use the term “visual acuity” to describe autistic people’s visual hypersensitivity (Baron-Cohen et al., 2009), but according to Hacking (2009 a), the term “acuity” does not capture what autistic people actually experience. Autistic people’s narrative revealed that too much sensation actually causes pain to them, but as “acuity” sounds neutral, this term is not subtle enough to convey the real experience of autistic people. For this reason, items in autism assessments are written by neurotypical clinicians or researchers without fully representing autistic people's experiences, and these items also involve some value-laden or subjective languages from non-autisic people.

The third major issue in autism assessment is the academic callousness of ethics in the consequence of autism assessment. As more and more scholars challenge the medical model and its natural kind approach to consider autism as a static entity present in an individual and recognize the social effect of labeling and categorizing on autistic individuals, they insist that the effect of life-long consequence of autism diagnosis on individuals should be more rigorously studied (Hacking,1995, 2007; Haslam, 2014; Teo, 2018; Timmi et al., 2019). Still, not much research has been conducted to identify longitudinal effects of autism assessment and diagnosis on autistic people's well-being in a community. In addition, considering that autism diagnosis is sometimes viewed as a means to gain financial or educational advantage, which is beyond the initial purpose of autism assessment, scholars assert that more attention should be paid to the 'partly non-purposeful' misuse of diagnosis as a solution to human problems (Latif, 2016).

As autistic people's autobiographies and first person narratives are acculturated, autism came to be understood more with the autistic people's perspective. Departing from the traditional medical model, the neurodiversity model which conceptualizes autism as human diversity instead of inherent neurobiological deficits gain more currency among autistic activists and disability studies scholars. With this conceptual shift, scholars have begun to pursue partnership with neurodivergent people to work on modifying existing assessment tools and to collect responses directly from neurodivergent people (Nicoladis et al., 2011; Nicolaidis et al., 2020). Other scholars have also challenged the deficit based languages in the assessment tools and the possible subjectivity of the raters (Dinishak, 2021; Timmi et al., 2019). In addition, some studies are questioning the effects of using labels on people and asking to consider the long-term effects of a diagnosis for autistic people (Hacking, 2019 b; Latif, 2016; Timmi et al, 2019).

The conceptual changes and questions correspond to what contemporary validity studies are aiming for: more attention to the ontological definition of an attribute, participants' response process, and social/ethical consequences of the measurement (Borsboom et al., 2004; Kane, 2006; Markus & Borsboom, 2013; Messick, 1995; Slaney, 2017). Therefore, it would be meaningful to examine the issues in autism research such as the shift of autism definitions, the response process of autistic people and their caregivers, and the effects of autism assessment. In this study, to address such issues in autism assessment, I adopted the framework of contemporary validity theories and examined the definition of autism and one autism screening instrument (Social Communication Questionnaire) with the following four research questions; 1) How do we define autism in autism assessment? 2) How do the definitions of autism affect the way autism is understood and assessed? 3) How can an autism assessment instrument be revalidated with contemporary validity theories? and 4) How does an autism assessment affect the way people define autism?

This study consists of four chapters to address these research questions. In the first chapter, a brief history of validity theories is introduced and issues that compromise autism assessment validity are identified. In the second chapter, different models to define autism and the problems of inconsistent definitions of autism are examined. In the third chapter, as a case study of autism validation, the Social Communication Questionnaire (SCQ)'s current form was revalidated based on Wilson's (2004) Four Building Block Approach. In the fourth chapter, based on the interview for SCQ validation, it is examined how the autism assessment instruments and their use affect the way autistic people and their family define autism.

Chapter 1: Development of validity theory

In this chapter, I traced the development of validity theory to examine how contemporary validity theories are established. Based on these contemporary theories, I identify four critical issues that autism assessment validity studies have not been addressed sufficiently.

Basic concepts in validity before 1950s

Basically, the validity is defined as the degree to which a test measures what it is purposed to measure (Kane, 2001). However, over the last 50 years validity theories have become more complex and expansive. Before the 1950s, the concept of validity reflected the dominant 'behavior' view in social sciences and the tests were primarily considered predictive devices. During this time, validity was understood in terms of the correlation between test scores and a criterion (i.e. the future or current behavior). For example, the validity of a job placement test might be expressed in terms of its correlation with measures of job performance, and if a test taker scores high on a job placement test, it was predicted that the test taker would show a better performance in the job. These test-criterion correlations were sometimes referred to as validity coefficients and this gave rise to the notion of "criterion validity." Angoff (1988) describes this early perspective on validity as follows;

> Consistent with other writers at that time, Bingham defined validity in purely operational terms, as simply the correlation of scores on a test with "some other objective measure of

> that which the test is used to measure” (Bingham 1937, p. 214). Guilford defined validity similarly: “In a very general sense, a test is valid for anything with which it correlates” (Guilford 1946, p. 429). (p. 20)

In other contexts such as educational tests, where external criterion is regarded as less important, tests were developed from sets of content specifications to ensure adequate coverage of the domain of interest (Maul, 2018). In these contexts, the goal is to develop a test as a representative sample from the domain of interest and evaluate the mastery within that domain. This content validity was primarily established by documenting the test-construction procedures and expert review.

Construct Validity

Although criterion validity and content validity seem to be appropriate for many tests, psychological attributes such as personality characteristics (e.g., aggression, contentiousness) and broadly defined cognitive abilities (e.g., general intelligence) are difficult to operationalize in terms of relations with specific external criteria. In 1955, Cronbach and Meehl introduced construct validity to provide a new approach to explain these attributes that are not ‘operationally defined.’ Construct validity was understood primarily in terms of how observable properties or quantities (i.e., item or scale scores) are related to theoretical constructs (i.e., psychological attributes) within an “interlocking system of laws” known as “nomological network.” For example, suppose that a theory of depression states that depression should be positively associated with anxiety, but not with social competence. Finding that scores on a depression test correlate positively with scores on an anxiety test but not with a social competence test would provide corroborative support for the theory that the depression test is a valid measure of depression. As validation of constructs is understood as an ongoing process, in which empirical findings lead to either confirmation of or adjustment to the theory for the construct, multiple sources of evidence might be required for the empirical evaluation of the theory.

However, this correlational evidence can be interpreted as evidence of validity only when the theory presupposes the association (Maul, 2018). For example, if the theory of depression does not specify whether depression is expected to be associated with introversion, finding that scores on the depression test correlate with scores on an introversion test (whether positively or negatively) could be theory generating, and it may be possible to generate an ad hoc explanation for the association, but the finding could not be interpreted as evidence either for or against the validity of the instrument. Furthermore, the finding cannot even be interpreted as evidence of an association between depression and introversion, unless it is presupposed that the depression test is in fact a valid measure of depression.

One contribution of construct validity was the distinction between a test and the psychological attribute, or construct, measured by the test, which is contrasted with the earlier, operationalist view that a test simply defined a construct (Slaney, 2017). This opened the door to the possibility that multiple tests could measure the same construct, and gave rise to two concepts; 1) convergent validity, reflecting the idea that multiple measures of a common construct should exhibit high levels of agreement with one another, and 2) discriminant validity, reflecting the idea that measures of distinct constructs should not be too highly correlated with one another, even if they used the same method of observation. To continue the previous example , suppose that depression and anxiety are both assessed via self-report and the reports of one's family members. Evidence for convergent validity could take the form of showing that self-reports and family reports of depression (or anxiety) are highly correlated while evidence for discriminant validity is showing that self-reports of depression and self-reports of anxiety are not so highly associated.

With the distinction between a test and a construct, there was a shift of focus from the test itself to the interpretation of the test score in the validation process. Although Cronbach and Meehl's (1955) description of construct validity is not easily distinguished as either a definition of validity or a process of validation, they clearly articulated how one might go about gathering evidence during the process of validation (Slaney, 2017). Along with this came an emphasis that validity and validation were about evaluating proposed interpretations of test scores, rather than a test itself. This remains a fundamental tenet of modern validity theory (Sireci, 2009).

The Unified Theory of Validity

Messick played a fundamental role in shaping validity theory in the latter decades of the twentieth century. Messick (1989) offered a unified view on validity that reflected a significant shift from previous viewpoints in three aspects. First, Messick's view combined separate categories of validity and validation under the generalized concept of construct validity (Markus & Borsboom, 2013). Thus, the idea of distinct types of validity (e.g., criterion, content, construct) was replaced with the notion of there being distinct types of evidence that could be brought to bear on the validity of a given test, depending on the intended purposes of the test.

Second, Messick pointed out the importance of the study on response processes (Zumbo & Chan, 2014). According to him, the evidence based on response process and consequences verifies that the test assesses as much as possible of what it should assess and as little as possible of what it should not: In Messick's language, this involves minimizing both construct underrepresentation and construct-irrelevant variance. To avoid these two threats to validity, Messick suggested systematic study of the response processes used by test takers or the consequences of test interpretation and use. The lack of evidence based on response processes and consequences raises concern for routine interpretations of test scores.

Third, Messick's view also called greater attention to the intended purposes of tests—including both interpretations of test scores and actions taken on the basis of such scores—and to the idea that quite different types of evidence could be necessary depending on these purposes (Maul, 2018). Messick (1989) states: "Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment" (p. 13, emphasis original). Therefore, he significantly raised consciousness around the importance of the ethical and social dimensions of testing, and he provided a framework for validity in respect to the use of tests.

Although there is controversy on how to consider consequences of test interpretation and use in validation, his unitary view of validity has remained influential since its introduction and is still the dominant conception of validity in the literature on educational assessment and measurement. Messick's view influenced the definition of validity presented by the AERA, APA, and NCME (1999) Standards for Educational and Psychological Testing: "Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests. . .The process of validation involves accumulating evidence to provide a sound scientific basis for the proposed score interpretations" (AERA et al. 1999, p. 9). This definition of validity shared three important tenets that Messick supported; (a) numerous sources of evidence can contribute to a judgment of validity, (b) validity is a matter of degree rather than all or none and, (c) validation of particular uses and interpretations of test scores, rather than a test itself.

Kane's argument based approach to validation

Based on Messick's definition of validity, Kane (2001, 2006) suggested the validation of pragmatic, context specific arguments tailored for specific audiences and circumstances. Kane's argument-based approach requires that validation should be based on an argument aimed at defending the appropriateness of a test for a particular use and the collection of evidence supporting the argument. Although Kane's argument-based approach is not a new theory of validity itself, it calls more attention to a clear statement of the proposed uses and interpretations of a test (Zumbo & Chan, 2014). If tests are used for purposes other than their original intentions, this will require a reexamination of the validity argument or the development of an entirely new argument. Also, Kane emphasized the importance of social consequences in a validity argument for a given test because the proposed use of the test implies an intention for certain consequences as a result (Markus & Borsboom, 2013).

A Causal Perspective on Validity

These mainstream validity theories weigh more on the epistemic character of validation, which requires evidence to justify the interpretation of the test score. On the other hand, there is other recent scholarship that has strongly emphasized understanding the ontological aspect of validity by claiming that validity is a property of measurement instruments. In particular, Borsboom and colleagues (2004) have developed an account of validity that traces back to the basic definition of the term, “validity is whether a test measures what it claims to measure.” According to Borsboom and colleagues (2004), a test is a valid measure of an attribute if (a) the attribute exists and (b) variation in the attribute causes variation in the outcomes of the test. This causal perspective on validity emphasizes that validity is a realistic attribute, independent of the evidence available at any given time, or the extent to which that evidence is found to be persuasive by any given community of observers (Slaney, 2017). These scholars argued that (a)the attribute being measured should have reference in the real world, not the meaning in the nomological network, (b) the variations in outcome should be explained by the causality, not correlation, and (c) validity is all-or-none ontological property, not the degree to which evidence justifies the interpretation of the test scores (Borsboom et al., 2004).

This causal perspective on validity merits scrutiny because it emphasizes the measurement aspect of a test, before we obtain test scores to be interpreted and used. Borsboom and his colleagues claim that the mainstream validity theories imply a “catch-all” approach to validation in which every test-related issue is integrated under a single header and treated as relevant to how validity should be conceptualized and approached (Borsboom, 2006; Borsboom et al., 2004; Slaney, 2017). Therefore, shifting focus on the ontology of measurement separately from the test score interpretations broadly calls attention to the scientific assumptions that have been taken for granted for the test development and requires articulated theories explaining the relationship between test scores and the attribute being measured.

Validity theories and autism assessment

In development of educational and psychological validity theory, Borsboom and colleagues' causal perspective contributes semantic and ontological components while Messick and Kane provides methodological tools for validation (Zumbo & Chan, 2014). The purpose of this brief review of the validity theory is not adopting one approach as opposed to the other, but taking the best from all to examine the practice of autism assessment and identify what is missing in the autism assessment validity and its validation process.

The causal perspective calls attention to issues of the inconsistent and incomplete definitions of autism and the lack of causal path between autism and the responsive behaviors that are assessed. Messick's unified theory and Kane's argument based approach point out that the response process of autism assessment participants and the consideration of social consequences of the test use are missing in the autism assessment validity studies. From the causal perspective, the ontological definition of autism, which has an independent reference in the real world and explains the essential features of autism, has not been achieved yet.

Broadly, there are three models that explain autism in academic research; medical model, social model, and neurodiversity model. It is very predictable that the ways these three models explain autism vary, but there are two different philosophical approaches in these three models. The medical model is rooted in the "natural kind" approach, in which autism is described as an independent and static entity that can be discovered and identified with systematic biological and neuroscience investigation like a physical disease. Alternatively, the social model and neurodiversity model, with some degree of variation, are rooted in the "social kind" approach, in which autism is classified by the sociocultural norm and this classification also affects the behavior of those who fall under that classification

(Chapman, 2020; Hacking, 1995). Most autism assessment has been developed on the medical model, but, even within the medical model, the diagnosis criteria and the label of autism have changed continuously whenever the Diagnostic and Statistical Manual of Mental Disorders (DSM) is revised.

In addition to the incomplete definitions of autism, the causal path of autism has not yet been established. The medical model of autism assumes that autism is caused by neuro-biological defects in individuals, but currently the cause(s) of autism has not been confirmed (Coleman, 2005; Tsai, 2004; Waterhouse, 2013). On the other hand, the social model explains that disabilities are constructed by social norms while the neurodiversity model contends that autism is caused by neurological diversity, not deficits, and struggles of autistic people are products of the neurotypical social norms and lack of accommodations for neurodivergent people (Chapman, 2020). Whichever model we take to define autism, the causal path from the cause of autism and the behaviors of autism cannot be established. In the medical model, as the cause of autism is still under investigation, it's improbable to confirm the causal path. Also, as the social model explains all kinds of disabilities in relation to social oppression, this model is not appropriate to establish the causal path from the cause of autism to the responsive behaviors. Lastly,the neurodiversity model is helpful to explain individual differences of autistic people and shift the focus from deficits and fixing these deficits to diversity and accommodating differences, this model is not appropriate for pinpointing the causal relationship of autistic attributes and the responsive behaviors.

Therefore, according to the causal perspective, autism assessments have been developed on the conceptual fallacy that behaviors that match the operationalized criteria are regarded as autistic attributes without making a distinction between the autistic attributes and the autistic behaviors that are caused by autistic attributes. This conceptual conflation between the operational criteria and the autistic attributes that should exist in the real world independently leads to inconsistent diagnosis of autism depending on the changes of autism criteria or the purposes of autism assessment such as educational assessment or clinical assessment.

In terms of the validation process contended by Messick and Kane, the response process of autism assessment participants has hardly been studied. For a long time, as clinicians have doubted that autistic people have enough intelligence and/or insights to self-report their conditions, thus most autism assessments are designed to be filled out by their caregivers. For this reason, autistic people's response process has hardly been studied. Moreover, the response process of caregivers has not been examined either, because it is a common practice in educational assessments that validity studies hardly include the evidence based on the response process (Zumbo & Chan, 2014).

Lastly, the consequence of autism assessment has hardly been studied. Messick (1989, 1995) and Kane (2006) insisted that the social and ethical dimension of test use should be included in the validation process. However, not enough studies for ethical and social consequences of autism assessment have been conducted (Zumbo & Chan, 2014). Considering that autism assessment results are used beyond its initial purpose of diagnoses such as providing social or educational services, validity studies should have paid more attention to the consequences of the autism assessment process.

In the following chapter, based on the causal perspective on validity, I examined the changes of meaning and definition of autism in the medical model, different models to define autism, and the practical problems that the lack of ontological definition of autism has brought to the diagnosis process in clinical psychology and education to address my first two research questions.

Chapter 2: Autism Definitions and Autism Assessment

The ontological definition of autism which articulates essential attributes of autism and distinguishes autism from other disabilities is still being studied and has not been achieved yet. It might be too strict to say that the validity of autism assessment cannot be established without the ontological definition of autism as the causal perspective on validity insists, but the lack of ontological definition of autism does cause confusion and inconsistency in autism diagnoses. In this chapter, I examined the history of the way the medical model has defined autism, the limitations of the medical model which is foundational for most autism assessments, the alternative models for defining autism, and the practical problems that are caused by the lack of ontological definition of autism.

Changes of Meaning and Definition of Autism in the Medical Model

In the history of autism research, the labels and meanings of autism have changed significantly. In 1911, psychiatrist Eugene Bleuler first used the term 'autistic' to describe the state of mind of a group of individuals who were socially withdrawn (Bleuler, 1911). The term autism introduced by Bleuler referred to a basic disturbance found in schizophrenia (another term introduced by Bleuler) characterized by an extreme withdrawal of oneself from social life due to hallucinations and delusions. Also, in 1922, Piaget described the pre-verbal stages of children's as 'autistic' or 'symbolic,' in which children could not follow ' logical' rules, did not think conceptually and there was a predominance of visual imagery in their minds (Piaget, 1923; 273-304). When autism first entered medical terminology in the 1930s, it described a particular mental state such as hallucination or visual fantasy in infants rather than an enduring set of behavioral traits (Evans, 2013; Timmi et al., 2016).

Autism as a diagnosis was proposed by Leo Kanner in 1943 in the case of eleven children who were characterized principally by “autistic aloneness” and “obsessive desire for sameness.” Since then, the conception of autism as “withdrawal into fantasy” was steadily eradicated and transformed into the concept of deficits or abnormalities in behavioral, communicative, and cognitive domains calculated through a sum of behavioral measures (Evans, 2013). This change was reflected in successive revisions of diagnostic manuals. Initially, autism was thought to be an early form of schizophrenia, in which a child attempts to engage with reality through his or her relationships with people and objects, but with the development of statistical methods in epidemiological studies and strong emphasis on purely behavioristic criteria by autism researchers like Rutter and Lotter in the 1960s, autism was conceptualized as abnormalities in behavioral variables within a total population that represented the norm (Evans, 2013; Lotter, 1996). The term “Infantile Autism” was introduced in DSM III (APA 1980) as a separate diagnosis under the category of childhood schizophrenia, but in 1987 when DSM III was revised, the term “infantile” was removed and the first diagnosis criteria of autism were established with the description of the developmental disorder of children who lack interests in people, have impairments in communication, and show abnormal responses to the environment. The most recent version DSM V (APA 2013), made another move from DSM IV (APA 1994) by combining previously separate disorders such as autistic disorder, Asperger Syndrome, Rett Syndrome, childhood disintegrative disorder (CDD) , and pervasive developmental disorder not otherwise specified (PDD-NOS) under the umbrella term of autism spectrum disorder (ASD) (Silberman, 2015).

Limitations of the Medical Model

Although there have been significant changes in meaning and label of autism in psychiatric and psychological autism research since the concept of autism was introduced, autism diagnosis research remains deeply rooted in the medical model, which explains autism as deficits in social communication and abnormalities in behavior due to neurobiological inherent defects. This medical model contributed to establishing two major directions in autism research. The first direction seeks out neurobiological defects that cause autistic symptoms because the medical model assumes that autism can be eradicated by identifying and removing the cause (Kapp et al., 2013). However, researchers have not confirmed any neurobiological causes for autism despite the continuous efforts to identify them (Coleman, 2005; Tsai, 2004; Waterhouse, 2013). As the cause of autism has not been discovered, researchers have not been able to eradicate autism by removing the cause. Instead, researchers chose to treat or eradicate behavioral marks of autism, and committed to developing behavioral interventions to extinguish abnormal behaviors and to make autistic people look like non-autistic people, which is the second major direction of autism research (Kapp et al., 2013).

However, more and more scholars challenge the logical and philosophical limitation of the assumptions in the medical model (Chapman, 2020; Hacking, 2007; Timmi et al., 2016). From the perspective of causal validity theory (Borsboom et al., 2004), the medical model fails to prove the first requirement; the existence of autistic attributes. Under the medical model, autistic attributes are replaced by the behavioral criteria in DSM V without ontological definition. These behavioral criteria cannot be regarded as the essential attributes of autism that distinguish autism from other disabilities because these criteria are also used for diagnosing other disabilities such as attention deficit hyperactivity disorder (ADHD) (fidgets with feet/hands, interrupts or intrudes into conversations and activities of others) or obsessive compulsive disorder (OCD) (repetitive behavior). Also, as these behaviors are symptoms caused by autistic attributes, if autistic attributes exist, these behaviors should not be confused with the attributes themselves. The failure to provide the ontological definition of autism attributes and their existence posits another challenge which is to prove the causality between changes in autism attributes and changes in their behaviors. Due to the conflation between the symptoms or the representation of the attribute and the attribute itself, there is no room for establishing the causal relationship. Even if it is assumed that there exist autistic attributes, theoretical explanations for the relationship between the not-yet-discovered inherent defects, autistic attributes, and autistic behaviors have not been articulated yet. Similar to the concept of general intelligence, the autistic trait is understood as an abstract concept which does not have a neurobiological entity and is only represented by the sum score of operationalized assessments.

Some scholars also challenge the natural kind approach, which is the logic underscoring the medical model, to explain autism (Chapman, 2020; Hacking, 2007; Timmi et al., 2016). Due to the tendency in psychology that explains human mind through hard science, autism has been explained as a natural kind as a stable and objective entity that can be discovered and identified with systematic biological and neuroscientific investigation like a physical disease (Chapman, 2020). A natural kind (classic examples being water and gold) refers to a class of entities, which exists independent of the efforts of investigators and has some inherent resemblance among the phenomena which dictate its belonging to a kind (Danzinger, 1997). Many physical diseases are regarded as natural kinds, and this natural kind concept is useful in providing the basis for explanation of salient phenomena of a disease and support reliable predictions about future phenomena in which the diseases may feature. One example is malaria, which is defined by the presence of the malaria parasite in the patient's body. Knowing that the patient is suffering from malaria implies a clear explanation for the infection, the likely course of the disease, and specific treatment for the disease. With the assumption that autism constitutes a natural kind like malaria, it is predicted that autism possesses a unitary and essential causal property that determines all observable characteristics of autism. Consequently, the majority of psychological and medical research of autism strives to identify neurobiological causes that determine all observable behavioral impairments (Happe et al., 2006; Haslam, 2014).

However, some scholars assert that autism is instead a *social* kind or *psychological* kind, for two reasons (Chapman, 2020; Hacking, 2007). First, as the label of autism spectrum disorder signifies, autistic people show such a wide range of differences in their sensory processing, cognitive abilities, and behavioral patterns. As one current slogan says, "if you know about one autistic person, you know about one autistic person" (Lord, 2011, p.166), it is almost impossible to identify the inherent unitary resemblance that applies for the whole group of autistic people. Second, these scholars explain that a label or classification of an individual is a social reality constructed by language and culture in a social system. Hacking (1995) called it a looping effect between individuals and concepts, meaning that the meaning of a scientific classification (autism, schizophrenia, or learning disability) affects the behavior of those who fall under that classification. Hacking explained that people tend to understand themselves through the concepts established by social institutions and thus reinforce the classification by the study of the concepts in academia.

To be more specific, until the middle of the twentieth century, autistic people were understood as eccentric people or feeble minded. But once autism was accepted as an academic psychological category, autism became a way to be a person, and autism became a natural kind to be assessed, measured, and quantified, an object that exists beyond history and geography, independent of constructions and culture (Hacking, 2007; Teo, 2018). The scholars who view autism as a social kind instead of a natural kind warn the natural kind approach strips off the human diversity of autistic people and reinforces the essentialism leading to harmful stereotypes that dehumanize autistic people. They assert that by respecting the first person view of autistic people and including the history, language and culture in the autism research, which are usually excluded in the clinical psychology research, we can understand essential features of autism and social force that interacts with autistic people more precisely (Danzinger,1997; Hacking, 1995, 2007; Teo, 2018).

Last but not least, some scholars criticize the deficit views in the medical model (Dinishak, 2016; Kapp et al., 2013). Due to the deficit views embedded in the medical model, autism research has focused mainly on deficits, overlooking alternative explanations, and researchers unconsciously assume that they offer a valid explanation of the relevant phenomena merely pointing to a lack or an absence (Dinishak, 2016). Also, the deficit views have yielded moral complacency in how deficit attributions are measured (Dinishak, 2016). Deficit attributions require a standard by which to measure deficiency or inadequacy, but in autism research, value judgments are made automatically and unthinkingly without presenting an appropriate standard. Furthermore, it is difficult to find answers from questions such as what legitimates the move from attributions of mere lack to attributions of deficits of autistic people in these domains or what makes it the case that autistic people' amygdala ought to be some determined size, or how long autistic people ought to make eye contact when greeting others (Humphrey et al., 2008; Nacewicz et al., 2006). Deficit views may be so entrenched in autism research that those making deficit determinations do not even recognize that they are making a choice to assign negative value to the deviation.

The Social Model

Reacting to the medical model and challenging its deficit views in the discipline of disability studies, scholars have suggested instead the social model (Shakespeare, 2006). Traditionally, proponents of this model make a crucial distinction between impairment and disability, arguing that disability is caused not by impairment, but rather by how society fails to accommodate and accept impaired individuals (Oliver, 1990). For instance, a paraplegic person who uses a wheelchair is always impaired (this is counted as an objective fact), but they are only considered disabled when their impairment is not accommodated for, such as when there are only steps instead of ramps. The social model contributed to shifting the way disability is framed away from being seen as an individual medical issue towards it being a political issue (Chapman, 2020). Under this model, the primary cause of disability is the way ableist societies are organized, rather than disability being framed as an individual matter (Shakespeare, 2006).

The social model played a crucial role in evolving disability civil rights movement and neurodiversity movements and increasing awareness around social accommodation, but still has two major limitations. First, the very concept of “impairment” in the social model still involves the concept of deviation from the norm because if anything is considered to be impaired, there must be something that is considered unimpaired (Chapman, 2020). Therefore, even though the social model changed the causal claim of disability from an individual to a society, it still relates the disability to impairment, rather than difference or diversity. Second, the social model is criticized because it overlooks daily struggles that people with disabilities go through due to their physical or mental traits, which cannot be easily reduced to a matter of marginalization and oppression (Chapman, 2020). This 'hardship denying' aspect of the social model is pointed out by people with disabilities.

The Neurodiversity Model

Influenced by social model, the disability rights movement strived to change the definition of disabilities based on individual deficits and the corresponding discrimination, and to frame the disability issues as political issue based on the shared identity, calling for accessibility and equal opportunities in independent living, education, and employment. Following this tradition, the neurodiversity concept and neurodiversity movement began to emerge as a ground swell movement during the late 1990s among autistic self-advocates (Ortega, 2009; Orsini & Smith, 2010). The autistic activists insisted that autism is not an impairment to be cured by medical treatments or behavioral intervention, but should be appreciated as neurological variation among the human population (Kapp, 2019; Kapp et al., 2013; Nicolaidis, 2012). Also, they call for the equal opportunities and social acceptance of autistic people and stress their participation in political decisions and scientific research agenda to reduce challenges of autistic people in daily living and improve the quality of life (Kapp, 2019; Kapp et al., 2013; Nicolaidis, 2012).

Based on their lived experience and embodied knowledge, neurodiversity self-advocates

challenge people to rethink autism through the lens of human diversity, not through deficits or impairment (Nicolaidis, 2012; Silverman, 2015). They view autism as a result of natural variations and genetic legacy in the process of evolution and value diversity in neurobiological development as people would value diversity in gender, race, ethnicity, religion, or sexual orientation. As opposed to only focusing on impairments, the neurodiversity model sees autistic individuals as possessing a complex combination of cognitive strengths and challenges. For example, difficulties in understanding social nuances, filtering competing sensory stimuli, and planning the tasks of daily living may be coupled with strengths in detailed thinking, memory, and complex pattern analysis.

Unlike the social model which underestimate the daily struggle that people with disabilities go through, neurodiversity proponents recognize neurobiological differences, and autistic self-advocates often vividly describe the challenges they have experienced in their daily life (Hacking 2009a; Nicolaidis, 2012). However, they also maintain that difficulties experienced by neurodivergent people are contextual and that living in a society designed for non-autistic people which systematically fails to accommodate autistic people's needs exacerbates the challenges experienced by autistic individuals (Chapman, 2019; Chapman, 2020). In line with the academic field of disability studies, neurodiversity advocates explain how the social norms established by neurotypical people pathologize and disable autistic people (Nadesan, 2013). For example, an increase in open office plans and the overuse of bright lights in working environments might tend to disable autistic people by making them experience “sensory overload” or “sensory fatigue,” which can, in turn, hinder social understanding and participation (Booth, 2016, pp. 43–44).

Also, many autistic self-advocates maintain that being autistic cannot be separated from who they are. Like gender, race, or sexual orientation, one’s neurobiology is only part of a person's legitimate identity and certainly not the sole defining factor of who he or she is. Therefore, neurodiversity advocates opposed interventions that eliminate unusual but harmless behaviors, like avoiding eye contact or repetitive body movements, across all contexts without regard for the coping mechanism they serve (Bascom, 2012; Kapp et al., 2013; Silberman, 2015). A number of autistic writers have criticized behavioral interventions based on applied behavior analysis (ABA), sharing that these programs for normalization increase their anxiety of being found out for their “weirdness”, suppress their unique ways to experience the world and soothe themselves, and force them to internalize a self-blaming view (Bascom, 2012; Milton, 2012; Silberman 2015).

Meanwhile, some have argued that the concept of neurodiversity may make sense for the “high-functioning” end of the autism spectrum, but not the “low-functioning” end (Jaarsma & Welin, 2012). Many autistic self-advocates and researchers, however, maintain that the use of concepts such as “high- and low-functioning” are inaccurate, demeaning, and potentially harmful (Hacking, 2009a; Nicolaidis, 2012). A linear autism spectrum on which researchers or clinicians can place individuals based on their functioning is misguided. For example, there is no way on the linear autism spectrum to categorize an individual with minimal spoken language and very little ability to perform activities of daily living but excellent written communication skills and the ability to analyze complex patterns or an individual whose functioning varies tremendously from day to day or in different environments. Autistic self-advocates assert that rather than categorizing people as high-or low functioning, the diversity of autistic people should be respected because categorizing autistic people as "low-functioning" may deprive them of their agency and opportunities to reach their potential (Ne'eman, 2010; Nicolaidis, 2012). Similarly, categorizing autistic people as "high-functioning" can deprive them of necessary support and services. Therefore, an individual's complex combinations of strengths and challenges, as well as the potential for wide variations in functioning should be understood to promote self-determination and increase the effectiveness of care and services.

Currently most autism assessment instruments are designed to identify deficits defined by medical models based on observing autistic people's behaviors. As the assessments are focused on the operationalized behaviors, the assessment results are also used for planning behavioral interventions to remove these abnormal behaviors to look more like neurotypical people. However, if new instruments were designed with different conceptual models of autism such as social kinds or the neurodiversity model, the instrument would reflect a more first person view of autistic people and identify the challenges caused by the interaction between society and autistic people, not within the autistic individuals.

Inconsistent Autism Definitions in Practice

The lack of ontological definition of autism is not only the issue of validity theories or philosophical approach but also the issue of the diagnosis process in practice and daily life of autistic people. In the US, one of the biggest issues of the inconsistent definition of autism is that the diagnosis criteria of clinical institutions and those of educational institutions are not identical and the assessment results for the same individual can be different depending on the context of the assessment (MacFarlane & Kanaya, 2009; Safer-Lichtenstein et al., 2021).

In the clinical institutions, the autism construct under DSM IV consisted of three criteria; deficits in language communication, social interaction, and repetitive behaviors. Under this construct, there were three different categories under autism, which are Asperger's disorder, autistic disorder, and PDD-NOS. With the advent of DSM V, the autism construct has changed with two diagnosis criteria (deficits in social communication and repetitive behaviors) and the three subgroups of autism are combined into one big group called autism spectrum disorder (ASD). With this change, a person who used to be diagnosed as PDD-NOS might be diagnosed as ASD without any personal trait changes. Also, a person who used to be diagnosed with asperger might be diagnosed either as ASD or as non-austistic without any personal trait changes. These changes of labels without any changes in the people's attributes are due to the lack of solid ontological definition of autism. For researchers or clinicians, these changes could be interpreted as scientific development or updated knowledge, but for the autistic people who go through this diagnosis change, it could result in a crisis of identity or the threats of losing essential services or income that support their lives.

In the educational institutions in the US, the autism construct follows the Individuals with Disabilities Education Act (IDEA). The diagnosis criteria of IDEA are more similar to DSM IV criteria, but IDEA has additional criterion, which is "adversely affecting children's educational performance." Due to this requirement, a person who is clinically diagnosed as autistic might not be qualified for an autism diagnosis in the school system if the person shows 'decent' academic performance. In 2007, there were 256,809 students between the ages of 6–21 with an educational classification of autism (OSEP 2008) in US public schools. The OSEP data indicated that this number of educational classification did not include approximately one-third of all students who are clinically diagnosed with autism because they are not qualified for the autism category in the school system due to their decent academic performance (MacFarlane & Kanaya 2009). This data reveals that some autistic people who have grade level academic skills have different diagnoses depending on the different assessment contexts.

The more serious problem in educational autism assessment is that each state has the flexibility to create their own eligibility criteria as long as it meets or exceeds the minimal requirements set forth by the Code of Federal Regulations. Since guidelines for autism assessment and eligibility differ by state, a student that qualifies for autism services in one state may qualify for a different special education category or fail to qualify for any services altogether in a different state (Kurth, 2015; MacFarlane & Kanaya; 2009; Safer-Lichtenstein & McIntyre, 2020). This variability of practice guidelines among states make autistic students and their family vulnerable to exploitation of assessment results because the result interpretation and diagnosis may not be based on the scientific criteria or students needs but on the school's financial resources (Kurth, 2015; MacFarlane & Kanaya; 2009; Safer-Lichtenstein et al., 2021). As the per-student expenditures for autism is approximately $5,000 higher than expenditures for OHI or ED and students with autism tend to use more services (average 5.4 services compared with 3.5 services in other disability categories), even slight differences in the eligibility criteria and different eligibility decisions can have substantial financial implications for school districts and educational experiences for autistic students (MacFarlane & Kanaya; 2009). This inconsistency of autism diagnosis raises more concerns about whether scholars, clinicians, and educators reached consensus on the definition of autism and whether it is scientific to have multiple definitions for the same attribute.

Admittedly, the lack of ontological definition of autism or at least the lack of consistent definition of autism weakens the theoretical foundation of autism assessment as well as causes confusion and inconsistency in educational decisions. To examine this issue in a more specific case of an autism screening tool, I revalidated the social communication questionnaire adopting four steps for constructing measures suggested by Wilson (2004).

Chapter 3: A Case Study, Revalidation of Social Communication Questionnaire

As described in Chapter 2, the lack of ontological definition of autism has caused confusion in the autism assessment process, which also influences autistic people's quality of life. Even though this study cannot address all the confusion and inconsistency that current autism assessment practice has caused, I chose an autism assessment instrument, the Social Communication Questionnaire (SCQ) current form, as a case study to revalidate it adopting Wilson's (2004) "Four Building Blocks" approach. In this revalidation, more attention was paid to the definition of autism and the response process which have not commonly been included in most validity studies.

SCQ was developed to be an autism screening instrument to refer an individual at risk for autism to subsequent diagnostic assessment which is often expensive and time-consuming work completed by a multidisciplinary team of professionals (Brooks & Benson, 2013). For this case study, SCQ was selected because it is one of the most widely used screening instruments in both educational and clinical settings (Barnad-Brak et al., 2016). Considering that the two golden standard of autism diagnosis tools, Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999) and the Autism Diagnostic Interview Revised (ADI-R; Lord et al., 1994) are hardly used in educational settings due to the limited time and financial resources, SCQ is more appropriate for this study because it is used in different settings such as schools and clinical institutions which adopt slightly different criteria for autism diagnosis.

Methods

I examined the validity of SCQ following Wilson's (2004) "Four Building Blocks" approach. To guide the development of an instrument to measure a construct, Wilson suggested the Four Building Blocks: the construct map, the item design, the outcome space, and the measurement model. Wilson added that this approach is not a linear process, but a cycle that might be repeated to develop and refine an instrument. Although the current study's purpose is not to develop a new instrument to screen autistic attributes, I will follow these Four Building Blocks approach to review the original validity study of SCQ (Berument et al., 1999) and revalidate this instrument.

The Social Communication Questionnaire

SCQ was developed as a screener for autism based on the ADI-R (Berument et al,1999). Before requesting a full clinical evaluation, which requires more financial and human resources, the SCQ is employed as a screener that caregivers (e.g., parents, guardians, teachers) can fill out to help determine the likelihood of a child having ASD (Wei et al., 2015). SCQ is a 40-item, caregiver/teacher-report measure for screening behavioral symptoms associated with autism. All 40 items are administered in a dichotomous format (i.e., yes/no), with Item 1 simply documenting whether or not the child is able to speak with short phrases or sentences, and Items 2 through 40 used for scoring. There are two separate versions of the questionnaire: the SCQ lifetime form, which is completed with reference to the individual’s entire developmental history for a wider range of age, and the SCQ current form, which is completed with reference to the individual’s behavior during the last three months mostly at the age four to five (Wei et al., 2015). The cut score for the suspected autism diagnosis is 15 (Berument et al,1999).

The Four Building Blocks

The construct map

To clarify the essential features of the latent attribute to be measured, Wilson (2004) suggested the idea of a construct map. A construct map is a visual representation of a unidimensional latent attribute that is continuous but has distinguishable qualitative levels such as high to low or positive to negative. In the initial stage of instrument development, using a construct map helps the measurer to examine a coherent and substantive definition of the latent variable and the qualitative order of levels inherent in the attribute.

Based on this construct map, I examined how autistic attributes are represented in the SCQ. As SCQ was not developed with the idea of the construct map, SCQ did not pay attention to unidimensionality or linearity of the attributes it assesses. Instead, SCQ conceptualizes autistic attributes based on the operational autism diagnosis criteria of DSM IV, and the three autism diagnosis criteria are assumed as three domains of autistic attributes in SCQ. After the DSM V update, SCQ is still being used without modifications to reflect the changes in DSM V. In this section, I analyzed conceptual fallacies caused by the lack of clarification on autistic attributes in the evidence of the initial SCQ validity studies.

The items design

Wilson (2004) suggested items should be developed based on the construct map. However, items on SCQ constitute three dimensions, and these dimensions do not have distinct qualitative levels. The three dimensions are based on the three operationalized criteria of DSM IV, which assume that the inherent defects in an autistic individual cause these autistic behaviors. Therefore, the items under these dimensions are developed to identify the behaviors caused by these defects in the individual, reflecting the medical model on which DSM IV criteria are based. Through conducting a content review of individual items on the questionnaire, I examined how deficit views in the medical model are represented in the items. I also addressed the value-laden and subjective language issues. To verify the issues raised in the content review with the potential users of SCQ, I developed questions (Appendix) for the cognitive interview in the “outcome space” stage.

The outcome space

The outcome space is a set of categories to organize qualitatively distinctive responses from the respondents and score them according to the construct map (Wilson, 2004). To set up a well-defined, research-based, exhaustive, but finite outcome space, Wilson (2004) suggested the instrument should be implemented with potential users in order to find out whether the respondents understand the items as they are intended.

In this study, to investigate whether the users understand the items without any confusion, whether the language of the instrument offends the respondents, or whether the items include exhaustive aspects of characteristics of autistic people, I conducted cognitive interviews with 20 potential users of this instrument. The potential users are teachers, families of autistic children, and autistic people. Although SCQ is originally designed to be filled out by teachers and caregivers, I included autistic adults in the interview to capture the views of autistic people who were excluded from the process of self-reporting for a long time.

I interviewed 20 people; four teachers, six parents, and ten autistic people. Interviewees were recruited by snowball sampling through personal contacts and targeted recruitment through postings on social media (Facebook and Instagram). Fifteen interviews were conducted via Zoom, four autistic participants elected to do the email interview over the verbal interview, and one interview was conducted in-person. One out of four email interview participants answered the follow-up questions. Two mother-son pairs (Grace and Josh, Sarah and Luke) participated in the Zoom interviews together. The interviews lasted from about 20 to 60 minutes and the total interview time was 356.63 minutes. The interviews were recorded via Zoom or an audio recording app and transcribed. The transcriptions were analyzed to understand how the items are interpreted by the potential users and were coded to generate meaningful themes for future research to develop more valid instruments.

Table1

Demographic information of the participants

Name (Pseudonyms)	Description	Gender	Ethnicity	Interview Method
Zoe	Family of autistic children	Female	Asian/Asian American	Zoom
Stella	Family of autistic children	Female	Latinx/Hispanic	Zoom
Hazel	Family of autistic children	Female	White/Caucasian	Zoom
Simon	Family of autistic children	Male	Black/African American	Zoom
Grace	Family of autistic children	Female	White/Caucasian	Zoom interview with Josh
Sarah	Family of autistic children	Female	White/Caucasian	Zoom interview with Luke
Lucy	Former preK-12 educator	Female	Latinx/Hispanic	Zoom
Matthew	Current preK-12 educator	Male	Indigenous American/Native American	Zoom

Mia	Former preK-12 educator	Female	Asian/Asian American	Zoom
Liam	Current preK-12 educator	Male	White/Caucasian	In-person
Caleb	Autistic adults, diagnosed at the age of 18	Male	White/Caucasian	Email interview
Tyler	Autistic adults, Diagnosed at the age of 12	Male	Black/African American	Email interview
Vincent	Autistic adults, diagnosed at the age of 27	Male	Black/African American	Email interview
Brandon	Autistic adults, diagnosed at the age of 29	Male	White/Caucasian	Email interview
Myles	Autistic adults, diagnosed at the age of 29	Male	Black/African American	Zoom
Cole	Autistic adults, diagnosed at the age of 12	Male	White/Caucasian	Zoom
Lily	Autistic adults, diagnosed at the age of 7	Female	Black/African American	Zoom
Tristan	Autistic adults, diagnosed at the age of 29	Male	Black/African American	Zoom
Josh	Autistic adults, diagnosed at the age of 5	Male	White/Caucasian	Zoom interview with Grace

Luke	Autistic adults, diagnosed at the age of 4	Male	White/Caucasian	Zoom interview with Sarah

The Measurement Model

In this stage, Wilson (2004) proposed to use a measurement model or psychometric model to relate the score to the construct. To analyze whether the item functions as the developers of the SCQ intended and to address the issues that I raised in the content review of items, I conducted a secondary data analysis to examine the item fit, item characteristic curves, and the Wright map with the Rasch analysis. For this secondary data analysis, I originally acquired 1,040 responses for SCQ items from the National Database for Autism Research (NDAR), but after eliminating all the missing values, the total number of cases that I used was 635 for the Rasch analysis.

Result

The Construct Map: What SCQ Assesses

When Berument and his colleagues (1999) developed SCQ (originally it was named Autism Screening Questionnaire), they stated that the questionnaire was based on the revised version of the ADI algorithm (ADI-R; Lord et al., 1994) used for ICD 10 (World Health Organization, 1992) and DSM IV (APA, 1994) diagnosis of autism, which provided an operational diagnosis based on the behavioral item scores in three areas of functioning; reciprocal social interactions, language and communication, and repetitive and stereotyped of patterns of behaviors. They also indicated that SCQ was designed to be completed by caregivers of individuals who might have a pervasive developmental disorder (PDD) which was an umbrella term under which the autism related diagnoses were defined at that time.

Although PDD in DSM IV was replaced by a comparable umbrella term, autism spectrum disorder (ASD), in DSM V, and the criteria in three domains in DSM IV were changed into two criteria in DSM V, the items have never been modified to reflect the changes in criteria. Also, the lack of enough conceptualization of autism attributes in SCQ has never been paid enough attention. As it is a widespread practice in clinical psychology to regard operationalized criteria as the attribute itself, the validity studies of SCQ have been focused on reporting reliable psychometric results or high correlation with existing measures without considering the unidimensionality or linearity of autism attributes SCQ tries to assess (Barnad-Brak et al., 2016; Marvin et al., 2017; Wei et al., 2015). Given that essential features of autism are still under scientific research, the construct map approach which visually represents one dimension of a latent attribute at a time and examines the levels of qualitative characteristics of the attribute could have helped to clarify autistic attributes and increase the validity of this measure. Due to the lack of conceptual clarity of attributes that should have been addressed in the initial stage of SCQ development, the initial validity study of SCQ (Berument et al., 1999) involves a few logical fallacies. To be more specific, the lack of clarification on what is non-PDD group, incoherent language use, and the different assumptions between SCQ and ADI-R compromised the validity evidence suggested by the authors.

First, the authors did not explain what non-PDD group is and why PDD group should be distinguished from other emotional or developmental disorders, not from people without disorders. If DSM criteria on PDD assume the behavioral deficits as deviance from the norm, it is logical to assess participants behavioral deficits compared with the children without any disorders (or *normal* children, if we can define *normal* children). If the authors decided the reference group with the children with specific disabilities, not children without disorders, they should have provided rationale to select this reference group and explained why it is meaningful to distinguish PDD children from this reference group. In the study, the sample were 200 participants; 160 participants were diagnosed as PDD (comprising 83 with autism, 49 with atypical autism, 16 with Asperger syndrome, seven with fragile X but not autism, and five with Rett Syndrome) and 40 individuals with non-PDD diagnosis (comprising 10 with conduct disorder, seven with specific developmental language disorder, 15 with mental retardation and 8 with other psychiatric diagnoses such as anxiety disorders). Authors did not provide any reasons to include ‘fragile X but not autism’ in PDD participants and to constitute the non-PDD participants with conduct disorder, language disorder, intellectual disability, and anxiety disorder, and this lack of explanation adds more confusion about the purpose of SCQ. Considering the initial screening of PDD children is the purpose of SCQ, it is worth examining which is more meaningful and practical: discriminating PDD children from children with other developmental/emotional disorder is or discriminating PDD children from children without developmental disorders. Furthermore, if authors believe that discriminating PDD children from children with conduct disorder, developmental language disorder, intellectual disability, or anxiety disorder increases validity of this study, they should have provided rationale regarding their belief and explain why other developmental/emotional disorders are excluded in the non-PDD sample.

Second, the authors did not provide a clear definition of the factors that they found in the factor analysis. The factor analysis of this study revealed that SCQ consists of four factors; social interaction (20 items) , communication (6 items) , abnormal language (5 items) , and stereotyped behavior (8 items). Without articulating the definitions of these factors, the authors stated that the four factors correspond to the three key autistic domains in ADI-R, adding that "the communication domain in ADI-R divided between the *social* factors, mainly reflecting communicative deficits, half being in *abnormal language*, reflecting abnormal language features in SCQ" (Berument et al., 1999, p. 445, emphasis added). This explanation rouses confusion again because the meanings of "communication" or "social" in SCQ are different from the meanings of "communication" or "social" of autistic domains in ADI-R according to the authors' explanation. The confusion posits another question: if social factors reflect communicative deficits, what does communication factor mean? Moreover, as "social" or "communication" are daily expressions, used outside of this discipline, when these daily terms have different meanings in the different assessment instruments as academic terms, it is more equivocal to understand the attributes that these terms describe.

Third, the authors did not examine the assumptions between autistic attributes represented in SCQ and autistic traits in ADI-R when reporting the correlation between the two instruments. The study reported the high correlation (.71) between ADI-R and SCQ as validity evidence of SCQ (Berument et al., 1999), following the practice of psychology that the new instrument's validity is assessed by establishing its correlation with an existing, dominant concept (Slaney, 2017; Teo, 2018). As SCQ items are designed similar to the ADI-R's items, the high correlation might be the expected result. However, as some scholars pointed out, reporting the correlation between the two instruments is not enough for proving validity (Borsboom et al., 2004; Slaney, 2017; Teo, 2018). As ADI-R was designed for a thorough assessment for three to four year old children and was validated with small sample size of 50 autistic children and 30 "mentally handicapped" children, it is doubtful that the autistic attributes assessed by ADI-R are the identical attributes that are assessed by SCQ which was designed for an initial screening tool for four to five year old children. Because both instruments are designed with operationalized criteria with different purposes, even if they have similar items, if the attributes assumed in the two instruments were represented on a construct map, the two maps would look different. If the two construct maps based on each instrument would look different, the high correlation between SCQ and ADI-R cannot be regarded as validity evidence for SCQ.

The Items Design

According to Wilson(2004), items in an instrument should be developed based on the construct map to reflect the order of different qualitative levels of the attribute that the instrument tries to measure. In the case of SCQ, the items are developed based on the three criteria in DSM IV and items in ADI-R, not on a construct map with an order of distinct qualitative levels. Therefore, items in SCQ belong to three behavioral domains; reciprocal social interaction, language and communication, repeated and stereotyped patterns of behavior without consistent definitions of these three domains and qualitative continuum of each domain.

Even if it is tentatively assumed that the three behavioral domains are essential characteristics of autistic attributes, it is hard to determine whether the items in SCQ are valid or they are the representative description of behaviors in these domains because these three domains have never been defined clearly. As ‘social interaction’, ‘language’, and ‘communication’ are lay words that are also used outside of clinical psychology, these words should be defined to clarify their meanings used in the DSM criteria or SCQ validity studies. However, DSM criteria only provide examples of behaviors related to each domain without articulated definition. In the initial validity study of SCQ (Berument et al., 1999), the same words for three domains of ADI-R are also used to refer to the factors found in the factor analysis in SCQ with different meanings, which increases the confusion and reduces the clarity of these concepts. For example, item 3 in SCQ (Does she/he ever use odd phrases or say the same thing over and over in almost exactly the same way?) is from the communication domain in ADI-R, but belongs to the abnormal language factor in SCQ. Also, item 9 in SCQ (Does her/his facial expressions usually seem appropriate to the particular situation as far as you can tell?) is from social interaction domains in ADI-R, but is classified as a communication factor in SCQ. These examples suggest that the language used in three domains represent inconsistent concepts. Due to these inconsistent concepts, it becomes improbable to figure out whether the items in SCQ represent the qualitative features of three domains appropriately or not.

Even after the update of DSM V or ICD 11 (World Health Organization, 2019), which combined social interaction domain and communication domain, deemphasized abnormal language usage, and emphasized sensory sensitivity, these items are being used without any modifications. Some items, which are irrelevant to the autism criteria such as self-injury (item 17) or carrying objects other than soft toys or blanket (item 18) are still in SCQ without valid rationale. Also, SCQ does not address sensory sensitivity, which is regarded as another seminal characteristic of autistic people in DSM V or ICD 11, and still has five items in abnormal language factor even after DSM 5 removed abnormal language use from ASD criteria. These five items in the abnormal language factor are more problematic because three of them belong to the communication domain and two of them to the repetitive and stereotyped behavior domain. Therefore, it is questionable whether the items represent the current diagnosis criteria properly regardless of a construct map approach.

A more fundamental problem of items in SCQ is the frequent use of value-laden subjective language in the items. Eight items involve the value-laden terms such as “odd”, “(in)appropriate”, “unusual(ly)”. The distinction between “socially appropriate” and “socially inappropriate” depends not only on personal value but also on family or social culture. To address these subjective language which could be interpreted differently, Wilson highly recommended listening to the respondents’ thoughts about the items. However, in the validity study of autism assessment in clinical psychology, the response process of respondents was hardly investigated. Therefore, how these items with subjective terms are interpreted by the respondents is still under examined.

Also, there are some items that require the understanding of communicative intention of the child who is assessed. This is a common problem for most autism assessment instruments which are designed to be filled out by caregivers. As many instruments aim for early detection of an autistic individual, most autism assessments are designed to be filled out by the child’s caregivers. In the case of SCQ which is designed to be filled out based on observation of four to five year old children, the respondents are their caregivers or their teachers in schools. The problem is that to fill out some items in the social interaction domain, the caregivers should clearly understand the communicative intention of their child. However, it is doubtful that parents and teachers understand the communicative intention of their child or student who are referred to the assessment procedure due to their limited communication skills. Even if the parents and teachers think that they fully understand the communicative intention of the children they observe, it has never been tested whether the caregiver or teacher's interpretation or understanding is correct. Moreover, the lifetime form of SCQ, which is used for diagnosis of a much wider range of age, is also designed to be filled out by caregivers with the memory of their children, which does not allow the autistic adults to report their thoughts even if they are able to understand items and respond to the items appropriately. Considering the general limitations of self-report such as response bias, incorrect memory, or item misunderstanding (Morsbach & Prinz, 2006) it should be noted that these autism assessment instruments including SCQ might be more subject to measurement error when the caregivers might not fully understand the communicative intention of autistic people.

The Outcome Space

The outcome space is a set of categories that are well defined, finite and exhaustive, ordered, and context-specific to reflect the continuum of the construct map which is developed in the first stage (Wilson, 2004). However, SCQ was not developed with a construct map approach, and every item in SCQ is evenly scored as one when the response shows autistic behaviors without distinguishable quality levels. In this section, instead of creating outcome space for SCQ, I analyzed the cognitive interview with a family of autistic people, teachers, and autistic adults to examine how they interpret the items and how they answer value-laden questions.

Different interpretation

Item 4: What Is Socially Inappropriate Questions or Statements. Each interviewer has his/her own interpretation of "socially inappropriate questions or statements". It was interpreted as revealing personal information in front of the public, too straightforward statements which might offend others' feelings, out of context questions, or talking about sensitive topics such as politics. Also, several interviewees revealed their concerns about responding to this item. One teacher stated that it is not obvious to judge whether it is socially inappropriate or not as a teacher because each family has their own conversation cultures or patterns. The teacher added that some statements in the same context could be interpreted totally appropriate to one family but not to another family. Another teacher stated that she is very cautious when she judges whether it is socially appropriate or not because some questions or statements from students might sound awkward to her but it does not necessarily mean that the statements or questions are completely wrong or inappropriate in other cultures. Also, one parent said that it is a very common practice for a four or five year old child to ask embarrassing questions like “when will you die, mom?” or “Are you going to die at 27?” as she watched a TV show. She added that she cannot understand why these questions which children can ask out of their curiosity are regarded as autistic symptoms.

Item 9 and Item 33: Appropriate/Normal Facial Expression. Several interviewees showed their concerns about the subjectivity of “normal” or “appropriate” facial expression. One autistic adult said, “if the child had a big smile with their tongue stuck out at someone then that can be considered inappropriate. It really depends on the setting and the situation at hand”. Two interviewees pointed out the unnatural facial expression of autistic children depending on different contexts. A parent said that when her son is asked to smile while taking pictures, his facial expression is very unnatural because he learned “how to smile”, but when he really enjoys something, his smile is very natural and genuine. A teacher also pointed out unnatural facial expressions of autistic children due to their behavioral intervention. He said that autistic children’s facial expressions are somewhat robotic when they are trained to practice certain conversational routines, but they show natural facial expressions when they genuinely feel something. Moreover, one teacher mentioned how difficult it is to define “normal” facial expression due to the change of culture and gender identity among the students. The teacher, who has worked at a junior high school for more than twenty years as both a special education teacher and an instructional assistant, said that what students in his school regard as “normal”, not just facial expressions but other aspects as well, has drastically changed since he started to work with students. These comments suggest that it is not explicit how to answer the items that require the judgment of being appropriate/normal from being inappropriate/abnormal without considering certain contexts or changes of values.

Item 19 Particular Friends or the Best Friend. Teachers showed concerns about filling out item 19, “Does she/he have any particular friends or a best friend?”. Two teachers said that they watch a student playing with other students well in the class or in the playground, but it is hard for them to tell whether they are really close friends or not. Also, one parent mentioned that her son regularly plays with their neighbor’s children but she does not know whether they are faithful, ‘best’ friends or just nice people to spend time with. She questioned the level of intimacy to be a “friend” or a “best friend” especially when the child is young.

Communicative intention

Item 20 through Item 32 are asking the communicative intention of a child being observed. Some items like item 24 (nodding head to indicate yes) and item 25 (shaking head to indicate no) do not ask implicit intention of the child. However, there are some items such as item 20 (talk to you just to be friendly) or item 22 (spontaneously point things to show you things not because she/he wants them) require the caregiver’s interpretation of the implicit intention of the child. Interviewees indicated some hesitation to answer these questions. First of all, parents whose child was almost nonverbal at the age of four to five mentioned that they did not fully understand what their child tried to communicate until their child had certain levels of expressive language. Second, two teachers indicated that the definition of “being friendly” and expression of “being friendly” vary depending on the family culture or individual personality. Also, they added that as students tend to show more affection once they build up more rapport, the result of observing students in the beginning of the school year would be different from the observation result at the end of the school year. Third, one teacher said that she does not believe “purely spontaneous” behavior because there are some settings or social cues that might provoke certain behaviors.

> I think spontaneous is an interesting word, because I think that if we look at it from just a quick viewpoint, it can look spontaneous. I think that if we, if it's followed by, like, what are their thoughts or what was the setting okay, maybe this isn't spontaneous…I think copying isn't necessarily spontaneous. I think copying is like, being people being aware of a social, or, like maybe not even quite sure about the social cue, but like, following through on that. So I don't know about spontaneous in that sense as a definition, because I think there's always something before it that either in their mind, or my mind even that, like has prompted that a little bit.

These comments suggest that parents or teachers might not fully understand the implicit intention of the observed child, which might threaten the validity of the data collection for these items.

Overall concerns for SCQ items

Given that SCQ has a binary answer choice, some respondents mentioned that it would be better to have some space to make notes for additional information about individual differences or varied behavioral patterns depending on the contexts. Also, one parent expressed her concerns about interpreting the word “ever” in the majority of items. She said, if ‘ever’ means just one time behavior or verbal utterance, everyone could fall into ASD category because sometimes we all make mistakes about pronouns (item 5), say things over and over (item 7) not to forget something, or have special interest that are unusual in their intensity (item 13).

Teachers and parents also denoted different concerns for filling out SCQ in general. Teachers are more worried about their interpretation of the behaviors of students because they do not know individual students' diverse family culture and they spend limited time at school. One teacher stated that she wonders whether the student behaves differently when he or she is with more intimate family members like his or her parents or siblings. Meanwhile, parents are more interested in whether the items explain their children’s behavior correctly or not. They expressed concerns when they cannot decide to answer yes or no because they feel their answers fall somewhere in the middle. Finally, several autistic adults indicated that the questionnaire is surprisingly long and somewhat redundant. They also stated that there should be some items about the strengths of autistic people such as their long attention span or extraordinary memory and about their sensory issues.

The Measurement Model

The last step of the four building blocks Wilson suggested is using the measurement model to relate the scored outcomes from the designed item to the construct map that was the original inspiration of the items (Wilson, 2004). However, SCQ is designed to use the sum score of adding individual item scores, without relating the scores to the construct map. The problem of using the sum score focusing only on the instrument is that even though it is understandable that some sort of an aggregation across the items is needed, the means of aggregation is either left vague or assumed on the basis of historical precedent to add item scores to give a total score (Wilson, 2004). Also, in this summing procedure, all items are assumed to contribute equally to the total scaled score but all items do not discriminate equally well (Wilson, 2004). For these reasons, Wilson (2004) suggested using the Rasch model to focus more on the relationship between items from a construct map to the results of the model.

The Rasch model differs from the instrument-focused-model used to interpret SCQ in two ways. First, the Rasch model provides both item level information and instrument level information, not just the instrument level information. Second, as the Rasch model focuses attention on modeling the probability of the observed responses, rather than on modeling the responses, it provides the respondent location (ability) and item location (item difficulty) on the construct map. The relationship between location and the probability is visualized with a blue line in Figure 1-Figure 4, and the items that do not fit for the construct map are visualized with distance (black line) from the expected function (blue line). With the Wright Map (Figure 6), which is created by combining the construct map idea with the Rasch model, it is possible to determine the relationship of the construct to the probability of response.

The sample used in the analysis based on the Rasch model consisted of 635 individuals from the National Database for Autism Research (NDAR) who had SCQ current form item-level data . As discussed in Novikova et al. (2013), one of the clear strengths of NDAR for secondary analysis studies is the National Institutes of Health (NIH) peer review process that provides a level of quality control for recruitment, diagnosis, and experimental procedures that is often absent or omitted from other data repositories. The basic analysis of the sample indicated three difficult items; item 28, 18, and 10 (Table 2). As the difficult items in this case are the ones that participants hardly endorsed rather than the ones that require higher ability to endorse, it could be interpreted that these items may not be closely relevant to the attributes that this instrument tries to assess. Specifically, Berumet and his colleagues' (1999) validity study indicated that item 18 (Does she/he ever have any objects (other than soft toy or comfort blanket) that she/he has to carry around?) is not relevant to the autism criteria. Also, item 10 (Does she/he ever use your hand like a tool or as if it were part of his/her own body) only works for the autistic children with limited expressive language. In the interview, some parents stated that their children did not need to use their parents' hands as a tool because their children were able to communicate verbally when they were four to five years old. Considering that all the participants in this data marked their children as verbal, it is probable that most parents marked "no" for item 10, which makes item 10 as a difficult item. Item 28 (Does she/he ever show you things that interest her/him to engage your attention?) could be a difficult item because it could be rare that autistic children show things to engage parents' attention. However, this item needs a more thorough investigation with respondents because , as one teacher indicated in the interview, it is ambiguous to distinguish the "pure" intention of engaging parents' attention from other intentions such as wanting to have more similar items when the children are young and parents do not fully understand the communicative intention of their children.

The two easy items (Item 13:Does she/he ever have any special interests that are unusual in their intensity but otherwise appropriate for his/her age and peer group ? & Item 3:

Does she/he ever use odd phrases or say the same thing over and over in almost exactly the same way?, in Table 3) are ones that most participants endorse. These items could be the ones that describe autistic behaviors that most autistic children exhibit. However, at the same time, these items could describe behaviors that parents can easily observe without further interpretation. For example, unlike item 28 which needs more interpretation, easy items like item 3 or 13 are more descriptive ones that can be answered through simple observation without interpreting the implicit communicative intention.

The fit statistics revealed four underfit items and two overfit items. The convention to determine the misfitting items is below .75 and over 1. 33 (Wu & Adams, 2013) . Given that the sample consists of 635 observations, a more stringent range, below .89 and over 1.11, was employed to determine underfit items and overfit items (Figure 5). Underfit items (Table 4) do not follow the expectation in the model as the black lines show in Figure 1-Figure 4. These underfit items need more attention from the developers of the instrument because respondents might misunderstand the items due to unclear or subjective wording or certain contexts that might influence the respondents' answer.

The interview gave some clues to interpret what problems these underfit items have. Item 4 was frequently pointed out by participants as a subjective item due to the diverse opinions regarding "socially inappropriate questions". For example, interviewees displayed varied interpretations of socially inappropriate questions such as too personal questions, too straightforward questions, out of context questions, or political questions. Item 21 and item 22 are questions requiring the understanding of "spontaneity" of their children. As some teachers during the interview pointed out, for caregivers it might not be clear whether the behavior is conducted spontaneously or not. Therefore, the answers might vary depending on whether the participant thinks more about this spontaneity or does not weigh too much about the meaning of spontaneity. Specifically, one interviewee, who was a father of autistic child, interpreted “spontaneously copy” in item 21 as following others’ behavior without thinking and shared that he does not want his child to be a person who just copy others. As this example indicated, this item might be understood differently among other respondents. Item 22 is a double barreled question, asking two points at the same time; spontaneity and the purpose of communication. Therefore, it entails another layer of interpretation, what the purpose of showing things is; just to show the things or to show the things because a child wants the things. For example, when a child shows a picture of a puppy to his or her parents, it looks like the child “just shows” the picture, but the child may actually “want to have” the puppy in his or her mind when the child shows the picture to the parents. Regarding item 23, parents whose children were verbal when they were four or five mostly said that their children did not need a range of gestures to communicate. Other parents or teachers hardly mentioned this item in the interview, but it is not clear what counts for “gestures” in this item. For example, if a child just picks up and brings a toy to let the parents know that the child wants to play with the toy, it might not be clear for the parents to understand whether this behavior is counted as “gestures” or simple behavior .

Overfit items (Table 4) are more discriminating ones than typical items and follow less random pattern than the model expects. Given that many interviewed parents indicated their concerns about the lack of social interaction with their peers, item 29 and item 40 might correctly represent autistic children’s characteristics perceived by their parents. However, it is doubtful that “not offering to sharing things” can be persistent autistic attributes throughout their life because quite a few autistic adults frequently mentioned their willingness to help their friends, families or neighbors in the interview, which signifies that autistic people are willing to share their “time” and “efforts” even though they hardly shared their “things” when they were young.

Finally, I also examined highly correlated items (Table 5). It is very reasonable that item 24 and 25 are highly correlated because they are two items but actually assess the paired behaviors; nodding the head for ‘yes’ and shaking the head for ‘no.’. Other than these two items, there are not particularly correlated items, but item 28 is somewhat correlated with item 2 and item 20 (.51). Considering that both item 2 and item 20 are addressing verbal ability of a child, it might be interpreted that the endorsement of item 28 is likely to depend on the child’s verbal communication ability. As autistic children at the age of four to five have varied verbal abilities, more information about the child’s verbal ability would be helpful to understand how these correlated items work or how the underfit items such as item 22 or 23 can be modified to capture the reality more precisely.

Table 2

Difficult Items of Social Communication Questionnaire Current Form

Item #	Description	Difficulty	SE
Item 28	Does she/he ever show you things that interest her/him to engage your attention?	1.67	.11
Item 18	Does she/he ever have any objects (other than soft toy or comfort blanket) that she/he has to carry around?	1.48	.10
Item 10	Does she/he ever use your hand like a tool or as if it were part of his/her own body (e.g. pointing with your finger or putting your hand on a doorknob to get you to open the door?)	1.48	.10

Table 3

Easy Items of Social Communication Questionnaire Current Form

Item #	Description	Difficulty	SE
Item 13	Does she/he ever have any special interests that are unusual in their intensity but otherwise appropriate for his/her age and peer group (e.g., trains or dinosaurs)?	-.83	.09

Item 3	Does she/he ever use odd phrases or say the same thing over and over in almost exactly the same way (either phrases that she/he hears other people use or ones that she/he makes up)?	-.24	.09

Table 4

Fit Statistics of Social Communication Questionnaire Current Form

Item #	Description	Outfit_ t	Infit_t
Item 21	Does she/he ever spontaneously copy you(or other people) or what you are doing (such as vacuuming, gardening, or mending things)?	13.53	10.47
Item 23	Does she/he ever use gestures, other than pointing or pulling your hand, to let you know what she/he wants?	11.89	9.20
Item 4	Does she/he ever use socially inappropriate questions or statements? For example, does she/he ever regularly ask personal questions or make personal comments at awkward times?	7.00	5.14
Item 22	Does she/he ever spontaneously point things around her/him just to show you things (not because she/he wants them)?	7.48	4.55
Item 29	Does she/he ever offer to share things other than food with you?	-7.37	-4.75
Item 40	Does she/he play cooperatively in games that need some form of joining in with a group of other children, such as hide-and seek or ball games?	-6.35	-4.01

Table 5

Highly Correlated Items in Social Communication Questionnaire Current Form

Item Correlation	Item #	Description	Item #	Description
.81	Item 24	Does she/he nod her/his head to indicate yes?	Item 25	Does she/he shake her/his head to indicate no?
.51	Item 28	Does she/he ever show you things that interest her/him to engage your attention?	Item 2	Do you have a to and from "conversation" with her/him that involves taking turns or building on what you have said?
.51	Item 28	Does she/he ever show you things that interest her/him to engage your attention?	Item 20	Does she/he ever talk with you just to be friendly (rather than to get something)?

Figure 1

Item Characteristic Curve, Item 21

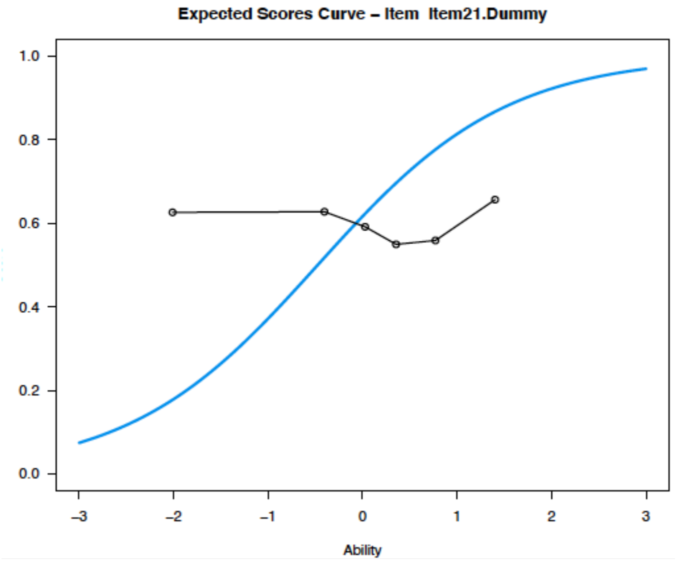


Figure 2

Item Characteristic Curve, Item 23

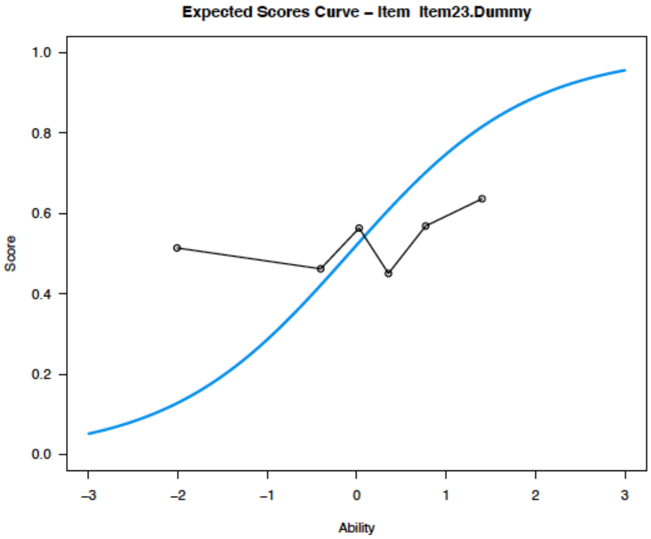


Figure 3

Item Characteristic Curve, Item 4

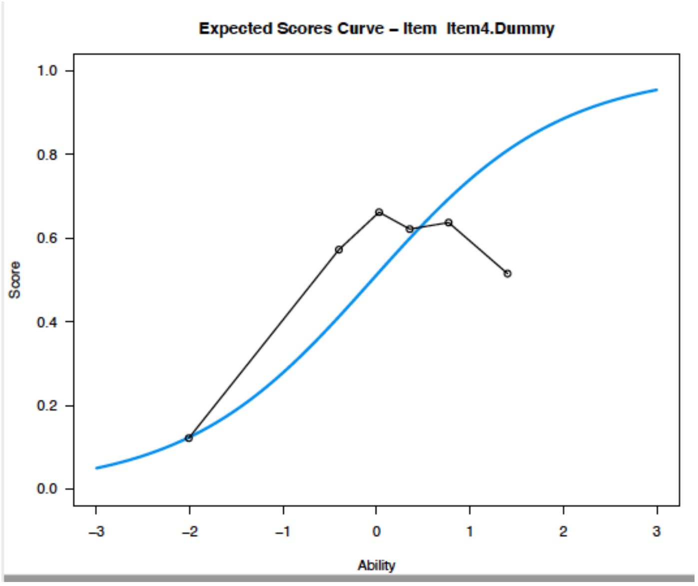


Figure 4

Item Characteristic Curve, Item 22

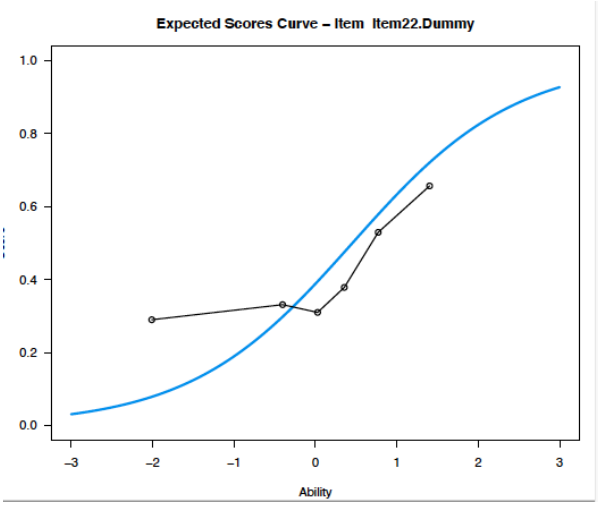


Figure 5

Fit Plot of Social Communication Current Form (below .89 and over 1.11)

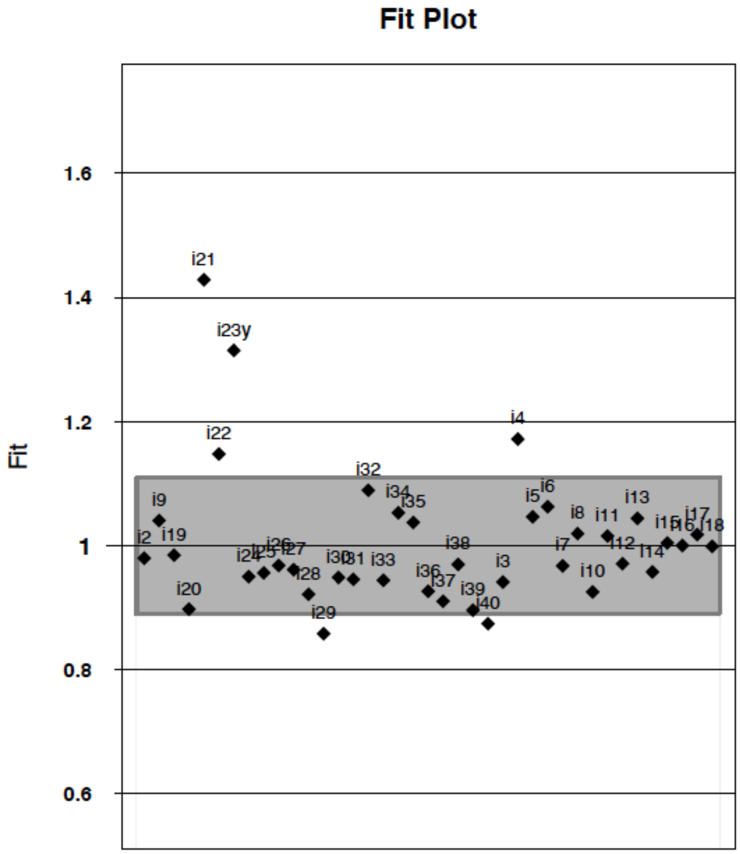
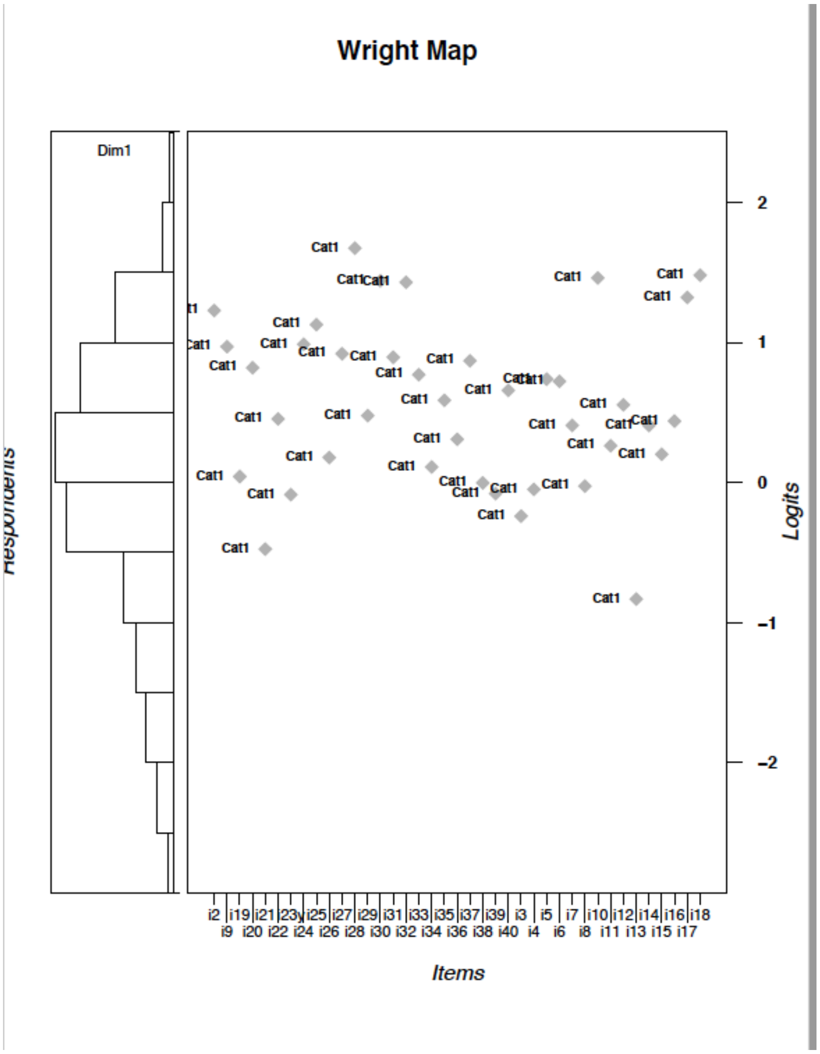


Figure 6

Wright Map of Social Communication Current Form (below .89 and over 1.11)



Discussion

The revalidation of SCQ following Wilson’s (2004) Four Building Blocks approach denoted four major problems regarding its validity. Three problems expected in the item content reviews were affirmed by the interview and the Rasch analysis. The three problems are the lack of definitions of the factors in SCQ validity study, the use of value-laden subjective language in items, and the items that require the caregiver's interpretation of the child’s communicative intentions. One additional problem was identified by the Rasch analysis, which is how to address the items that might function differently depending on the child's language ability.

First, the lack of definitions of the factors used in the initial validity study (Berument et al., 1999) weakens the conceptual foundation on which SCQ items are developed. Even though the developers of SCQ indicated that SCQ is founded on the three behavioral criteria of autism in DSM IV and ADI-R items, the factor analysis result does not correspond to the three categories of criteria in DSM IV and ADI. For example, three underfit items (item 21, 22, 23) are classified as "communication" related items in ADI-R, but these items belong to the "social interaction" factor in SCQ validity study. According to the SCQ developers this difference signifies the need to distinguish between "language deviance" and "language deficit" or the lack of validity of SCQ items. However, as SCQ developers never provided articulated definition of the four factors (communication, social interaction, abnormal language, and repetitive behavior) and did not explain how communication in the SCQ factor and communication in the ADI domain are different, it is very confusing to figure out the conceptual foundation of those underfit items. Moreover, as all of these terms (communication, social interaction, deviance or deficit) are lay terms used in everyday conversation, it is even more challenging to understand what phenomena or behaviors these terms signify. For this reason, their suggestion to distinguish between language deviance and language deficit is also ambiguous because they did not articulate the difference between "deviance" and "deficit" (Berument et al., 1999, p. 450). Another underfit item (item 4), which is from communication criteria but loaded on the abnormal language factor, affirms the inconsistent conceptual foundation issue in SCQ. Finally, as some interviewees indicated, SCQ does not have items that represent the sensory issues. Given that DSM V and autistic narratives emphasize the sensory issues of autistic people, it is doubtful whether SCQ's conceptual foundation is comprehensive enough to represent the major characteristics of autistic people.

Second, the value-laden items are subject to different interpretations, which might yield the misleading data collection. The interview revealed that the interviewees have different definitions of "(in)appropriate", "(ab)normal", or "odd" and they do not always have clear and consistent criteria to distinguish whether it is appropriate or not. Parents showed a tendency that they regard the conversation patterns or behaviors abnormal or inappropriate when those verbal and behavioral patterns add challenges in their daily life. On the other hand, teachers were cautious about making judgements because they do not know the students' family culture or conversation patterns. The interviewed teachers also tend to regard some behaviors more problematic when the behaviors disrupt the class severely. As seen from the result of the fit statistics, one of the underfit items (item4) is about this value-laden question, which put more weight on the need to revise this item. Therefore, it would be necessary to revise these items with value-laden questions into items with more descriptive and objective questions for the validity of this instrument.

Third, it is doubtful whether the items that require the caregivers' interpretation of the child's communicative intention reflects reality. Three underfit items are based on the caregiver's interpretation of the communicative intentions. The interviews indicated that when the children are nonverbal or have limited expressive language, the caregivers did not fully understand their intentions of certain behaviors. On the other hand, when their children had sufficient expressive language, parents said that "pointing" behaviors were rare because these children were able to convey their message verbally rather than with pointing gestures.

Both parents and teachers showed concerns about distinguishing “spontaneous” behaviors, “just to be friendly”, or “just to show things”. Considering that autism assessment instruments are mostly designed to be filled out the caregivers, not the autistic people themselves, it is crucial to pay enough attention to this indirect process of collecting responses, because autistic advocates have argued that parents do not always understand them and autistic adults in my interview also mentioned that their autistic friends or siblings understand them better than parents (Carey et al., 2019; Nicolaidis et al., 2011). Even though it is understandable that early detection is important and autistic children cannot fill out this instrument at the age of four or five, it should be noted that items for communicative intention are more subject to the observers’ personal interpretation which might be misleading or biased.

Last but not least, the Rasch analysis revealed that some difficult items and underfit items are items that can be influenced by the receptive and expressive language ability of a child. This result indicates that the interpretation of the assessment results will be more meaningful and valid if more detailed receptive and expressive language ability is provided. As the data set that I used for the secondary data analysis does not have data about different levels of language ability, I was not able to examine the different item functioning (DIF) of items (10, 21, 22, 23, 28) regarding different language abilities. If the DIF in language ability is examined in the future studies, it would provide more practical suggestions for refining these items.

Limitations and Future Directions

This revalidation study of SCQ is limited in two ways. First, the cognitive interviews were not thorough enough to cover the whole items. Except for the two participants who replied to emails with follow up questions, most interviews were conducted only once. Due to the time limit and the fatigue of interviewees, the interview was not conducted in a way to go over every single item. Also, as some interviewees were not used to expressing their opinions about the items, they tended to explain their answers rather than analyze items or verbalize their thought process. If more interviews could be conducted or a more direct way such as the response process evaluation (RPE) method would be used in the future, more detailed data about how people understand each item will be acquired. Second, both the interview and the Rasch analysis were limited in the number of the participants. The interviews were conducted only with 20 people, and the Rasch analysis was conducted with 635 observations. If more interviewees with diverse backgrounds were recruited and more cases for the secondary analysis were available, the result might reveal different pictures for the validity of SCQ.

Despite these limitations, this study is still meaningful because it highlights the importance of the theoretical foundation on which an assessment instrument is based. Currently, the theoretical foundation of autism is still in controversy as seen in diverse models that define autism in different ways. SCQ is designed to assess deficits defined by DSM IV, but the diagnostic criteria in medical model including DSM criteria only list examples of operationalized behaviors and have not confirmed the causes of autistic attributes or objective criteria regarding what norms are, what deficits are, or the criteria of the severity or intensity to determine whether a behavior is regarded as abnormal. Considering its lack of theoretical limitations, if an instrument is developed with more practical purposes to relate the results to help autistic people such as assessing challenges in the classroom or assessing the accommodation needs for autistic children in the community, the initial construct map would be more logical and the validation process would meaningful for developing and refining items. Also, if more instrument development involves the response process, not only caregivers' voice but also autistic people's voice, more valid assessment instruments not just aiming for presenting labels on people but aiming for benefits autistic people's education and thriving community living could be developed.

Chapter 4: How Autism Assessment Tools Shape the Way People Define Autism

Ian Hacking's (1995, 2007, 2009a, 20009b) social kind notion provides an alternative conceptual framework to understand autism, gearing away from the natural kind approach on which the medical model was founded. The social kind approach noted that the labels of human beings which are determined by seemingly objective and scientific assessment processes are actually influenced by social norms or the linguistic practice (Hacking 1995; Teo, 2018). Also, this label and categorization is not a static entity but a moving target because the label influences the way people identify themselves as well as the educational or life paths, and the people with the label also support or challenge the notion of the label (Hacking, 2007; 2009a). According to Hacking (2009a), the label of "autism" is quite modern conception, and after the society and academia set the category of autism, autistic people react to the notion created by non-autistic people by producing autistic narratives on various media and collecting their political power and presenting their presence by neurodiversity movement.

This social kind concept is supported by the Kapp and Ne'eman's (2020) explanation of how they influence the DSM V revision by exerting collective political power of autistic people. Before DSM V, the organized autistic community did not exist, but when DSM V draft was proposed, autistic people were able to organize a community with autistic researchers and self-advocates such as Autistic Self Advocacy Network (ASAN) to exert social, political, and scientific power on the DSM V process. There were two major concerns from autistic community for DSM V revision, which combined DSM IV's three main autism diagnosis, Autistic Disorder, Asperger's Syndrome, and PDD-NOS, narrowing the scope of autism diagnosis and the integration of the Asperger's diagnosis into the autism spectrum.

ASAN persuaded both autistic people who feared losing their benefits and services and the DSM committee to unify the autism diagnosis. ASAN provided research findings and documentation to demonstrate that the three autism diagnoses were applied inconsistently depending on the age and background of the person being diagnosed and separate diagnoses contributed to service eligibility gaps and limiting access to services. Also, ASAN addressed the concerns of Aspergers by advocating for a broad formulation of a unified diagnosis criteria.

Kapp and Ne'eman (2020) described this process as;

> ASAN pursued a combination of social, political, and scientific strategies to "lobby" the DSM-5 process. Ultimately, our work was rooted in a simple reality, often obscured given the inscrutable nature of the process of making the DSM: it was written by people, and people can be communicated with, influenced, and convinced, even when they are autism researchers (p.173).

This process indicated that the scientific criteria are not necessarily "purely objective", and how autistic people who are grouped by scientific labels challenge the social categorization by collective power.

Thomas Teo (2018) also criticized the natural kind approach in clinical psychology. He emphasized that human minds are embedded in sociohistorical contexts and that psychological concepts are entrenched in culture and linguistic practice. His stance is supported by two philosophical concepts and examples. First, by adopting Hacking's social kind (1999, 2007) and Baudrillard (1997)'s simulation concept, Teo (2018) explained how clinical psychology drew borders between the normal and the abnormal. He insisted that psychosocial knowledge is not a mirror or even a map of the world but rather a result of interactions occurring in human relationships. This is because psychological statements are not determined by the unique characteristics of mental life but are rather an outcome of politics, culture, and relationships (Gergen, 2009; Teo 2018). Baudrillard (1997) exemplified this socially constructed truth with the contrast of a physical map with a political map of Europe. The border between Austria and Italy is a simulation (social constructed sign) that does not correspond to physical features such as rivers or mountains. But the border, which is real in the sense of political society, has consequences for people living on either side of this simulated line – an outcome of historical and military events because people on either side of the border have different identity, education, economy, food habits or language. This reality has been constructed from the border, which is a simulation but becomes a sign of the real, but not real itself. This example corresponds to Hacking's (1995, 2007, 2009a, 20009b) explanation of the emergence of the autistic people. Once regarded as weird or awkward, people were categorized as autistic by official diagnoses developed in clinical psychology discipline, and this socially constructed categorization came to have real consequences for self-identification and educational trajectories. Hacking (1995, 2007) called this interaction between the categorization and the real consequences of people as a looping effect.

Second, Teo (2018) presented the history of IQ tests as another example of the psychological concept that has become a social reality. When Galton (1962) tried to assess the mental ability, he defined intelligence as a matter of sensory acuity such as reaction times, and Cattell, following Galton, proposed a series of mental tests that included, among others, dynamometer pressure (hand squeeze), least noticeable difference in weight, reaction time for sound, bisection of a line, judgment of time, and number of letters remembered after one hearing. However, with the sociohistorical calls for the oppression of immigrants and minority, intelligence tests were turned from problem sets that average (middle class-schooled, white) children would be able to solve into instruments that were interpreted to assess innate abilities, allowing to identify “feeble-minded” in American society as well as immigrants to justify the oppression of Black Americans via scientific discourses (Gould, 1996). Although psychologists did not keep Galton’s concept of intelligence, they kept his hereditarian, eugenicist, and racist assumptions. This traditional concept of intelligence has become so ingrained in our culture, converted into a seemingly self-evident natural kind, that it is difficult to envision and apply alternative conceptualizations (Gardner, 1983). Also, we still have the linguistic relics from this intelligence test in our daily language. The terms idiots (mental age of 0–2 years), imbeciles (mental age between 3–7 years), or morons (mental age between 7–12 years) (feeble-minded) (e.g., Goddard, 1926, p. 4) were formerly technical terms signifying different mental abilities, and are now used in everyday language, whereby the original definition has been lost.

Therefore, the psychological label cannot be purely objective or scientific because it is deeply embedded in sociohistorical context and it is only expressed by the language which also carries sociocultural values (Danzinger, 1997; Teo, 2018). For a long time, clinical psychology does not pay enough attention to these sociocultural aspects of their diagnosis process (Danzinger, 1997; Teo, 2018). Even though the act of labeling does not only entail the description of a person but the normative judgment with ensuing decisions that affect the person’s life significantly (Hacking, 2007, 2009b; Teo, 2018), the social and ethical consequences of the assessment have hardly been studied (Zumbo & Chan, 2014). Furthermore, although the inclusion of social and ethical consequences of measurement to the validity claim was insisted on by Messick(1995) and Kane (2006), most validity studies do not address the social and ethical consequences of assessing instruments (Zumbo & Chan, 2014). This neglect of social/ethical consequences of the diagnosis process is partly due to the ambiguity of defining the range of those consequences and partly due to the lack of resources to trace the longitudinal consequences on people who were assessed (Zumbo & Chan, 2014).

To contribute to this lack of research on ethical consequences of autism assessment, and to address my final research question on how autism assessment affects the way people conceptualize autism, I reanalyzed the interview conducted for the revalidation of SCQ. For this purpose I conducted a thematic coding of the transcript and analyzed emerging themes to examine how autistic people and their families define autism and how this conceptualization is influenced by autism assessment tools. I also examined what the social and ethical consequences that we should consider when developing autism assessment tools.

Methods

Interview

From the interview for the SCQ revalidation, I chose 16 interviews with ten autistic adults and six parents of autistic children. Teachers' interviews were excluded because they do not go through the diagnosis process even though they fill out assessment instruments for their students' diagnosis. Rather than the opinions on the individual items, I focused on the three general questions; 1) How do you describe autism and why? 2) What are your overall impressions of autism screening tools whenever you fill out the surveys? 3) What were your overall impressions of SCQ? In addition, any comments about their thoughts or experience related to their autistic identity were included in this thematic coding.

Data Analysis

In analyzing this data, I applied Charmaz's (2006) grounded theory approach, where categories are produced inductively through constant comparison between data, emerging categories, and the literature. I used open coding to create first-level codes based on the general topical content. I also identified in vivo codes, selecting keywords or phrases expressed by the participants (Tracy, 2019). I continually compared these first-level code categories with linking, axial codes to identify emergent categories (Charmaz, 2006). When I felt I had reached theoretical saturation, I pulled out key quotations from both the interviews, organizing them according to my developing themes. The two general themes included disorder and difference. The disorder theme has three sub-themes: patient, frustration due to deficits, and lack of respect and understanding. The difference theme has three sub-themes: belief beyond the diagnostic label, celebrating autism, and accommodation.

Results

From the first-level coding, the two general contrasting themes, disorder and difference, were quite evident. Six autistic adults and three parents internalized the deficit views in the medical model on which DSM V and autism assessment instruments are based. These interviewees consistently define autism as a medical condition or neurological disorder that causes dysfunction in their life. On the other hand, four autistic adults and three parents articulated that autism is just a different way of perceiving the world. These interviewees do not fully agree with the way autism assessment instruments describe autistic people and are able to describe how they perceive the world differently and what their strengths are compared with neurotypical people. Under these two major themes, I identified three sub-themes that support the major themes in specific ways.

Disorder

The first major theme was “disorder”. Six autistic adults and three parents fully internalized the medical model represented in SCQ or DSM V and described autism as a medical condition. They defined autism as a disorder that causes challenges and difficulties in their everyday life and social interaction. Also, these interviewees attributed their disorder to neurological dysfunction, and identified themselves and other autistic people as “patients.”

These interviewees believe that autistic screening tools are helpful to understand autistic people’s medical conditions and deficits, and believe that autistic people can identify their challenges through these screening tools. One interviewee said, “I think these surveys are actually good, because I think it is good to get to know the opinions of autistic patients and how they interact and how they are defined, and how they define their own medical condition based on this particular topic.” Another interviewee responded, “It can help any coming autistic patients who are younger than I am. Yeah, better understand yourself, or it's gonna help the parents of a lot of autistic patients to understand and know their children, or each child can change any way in whereby you raise the situation.” These people internalize the criteria of deficits in SCQ or diagnostic tools and understand themselves through those deficits and believe that they need therapy and guidance to navigate this world because their disorder causes so much challenge in every aspect of their life and that their deficits should be corrected by therapies. For example, one interviewee said, “I find it hard to look someone directly in the face while talking with them. So, let's say I find it hard to interact socially”, and another interviewee said that every autistic patient needs lifelong therapies for getting better in their social interaction and communication.

In the following three sub-themes, the patient, frustration due to deficits, and lack of respect and understanding constitute the disorder theme and provide a clearer picture of how the theme of disorder affects autistic people's self-identity and emotional well-being.

The Patient

The patient is the dominant concept that interviewees identify themselves with. As they perceive themselves and other autistic people as patients, they ask for help and treatment for their medical conditions. At the same time, these interviewees fully realize that there is no treatment for the condition of autism, and this understanding leads to the frustration for autistic people because they have a lifelong medical condition without treatment and they cannot escape from this "patient" condition no matter how they make efforts or no matter how much progress they made.

> I was hoping to maybe see a way whereby autistic patients can get through some kind of situations, people that are like myself that don't have parents, or anyone to guide them. And I had to do everything on my own. So I was kind of hoping that there will be a guide, or maybe there will be something like a short list of words autistic patients can do to help them

Therefore, even though autistic people consistently seek help to navigate the world that defines autistic people as patients, they cannot find any help that fully redeems them from this "patient" status.

Frustration due to deficits

This notion of being stuck as a patient leads to the feeling of frustration for autistic people. Due to the list of deficits in assessment instruments, they are well versed in identifying their deficits without an understanding of how clinical criteria distinguishing the ‘normal’ from the ‘abnormal’ are socially constructed. Even for the same mistakes that neurotypical people make or the same situations that neurotypical people are also afraid of, autistic people tend to be more frustrated because they attribute their mistakes or anxiety to their neurological deficits.

As the following example showed, even though the struggles in high school and the anxiety about meeting a woman that an 18-year-old interviewee was interested in are common too for neurotypical people, the interviewee blames himself more as he thinks that his struggles in social life are attributed to his autism and he finds relief from his self-criticism in the diagnosis label.

> At the time I was eighteen, I was miserable. I hated the high school I attended, my self-esteem was almost nonexistent and I felt I had a connection with someone which I eventually found out later was fake. Anyway, I was nervous because of anticipation and also because my therapist is a woman. As I already hinted at, I have struggled trying to socialize with women that I think interest me. When my therapist told me I am on the spectrum, that nervousness transitioned into a sense of relief.

Other examples suggest that autistic people are very aware of their lack of conversation skills because the assessment tools or their language therapists keep emphasizing their mistakes. Even neurotypical people might not fully master some conversation skills such as interrupting properly or talking about political issues in inoffensive ways, but autistic people are more sensitive about their mistakes in conversations and are easily frustrated because through their lifelong assessment process, they have learned that they cannot meet certain expectations established by neurotypical therapists or psychologists.

> Through many years of defending myself and many years of trying to get better, I'm no longer tied down to short phrases or sentences. I do have some difficulties with interrupting properly. I do sometimes break down apart from dots, but I do communicate better than where I was before.

> …um grammar is kind of hard sometimes, you know, should I use this noun? This pronoun? Did I say this? Correct? Oh, did I write this role? What's even worse is when it comes to like subjects, like politics, or like, you know, anything, right? Any subjects? I sometimes don't know.

Lack of respect and understanding

It is notable that these autistic people and parents who internalize deficit views in the medical model also expressed their anger and fear about a society that does not respect autistic people. Some autistic people accept the notion that they are patients with inherent deficits, but they do acknowledge that they have strengths and they express anger about the lack of respect from society. They indicated that neurotypical people do not understand their individual differences and do not show respect for autistic people's opinions.

> we autistic patients, we are very brilliant, our brain, our brain moves at a faster or slower pace. So if it depends on, it actually depends on the individual. …autism also affects each person differently. Like I said, it can either make the brain move faster, or it can make the brain slower… I have a friend of mine that is very intelligent, and as he is an autistic patient, but due to its his autism, people do not take us seriously.

> …people are not taking me for the kind of person I am. Because I have autism. I'm not supposed to speak up. I'm not deserve to be listened to.

The next quote shows the offensiveness and stigmatized judgment the autistic people and their families went through in their diagnosis process. Considering that the interviewee successfully acquired a bachelor's degree and was looking for a paid internship when I conducted the interview, it is obvious that the neurotypical professionals including doctors and therapists did not understand the potential and strength of autistic people. These layers of lack of understanding and mistreatment throughout their life might cause more frustration, anger, or anxiety in autistic people's social life.

> The first was when I was three. I have no recollection of it, but my dad remembers taking me to an older doctor who told him that I should be put in a mental hospital. For the doctor who told my dad that I should be in a mental hospital, let's just say (mildy) that I tell him he is a naïve fool.

Furthermore, the last quote indicated how autistic people feel regarding their lack of agency in the decision-making process. Many autistic assessment tools are designed to be filled out by their caregivers, and not much research has been done about how autistic people feel this lack of agency for their decision-making process in every life stage. As the following quote suggests, thinking that autistic people might be forced to sign consent forms requiring the autistic person's signature, more attention should be paid to the anxiety that autistic people feel from the pressure of social norms and rules without considering their feelings.

> I once saw a(n) (autistic) kid who was being pressured by his parents to give into agreeing something he wasn't clear cut for, just to affirm the suggestion that his doing

> just fine. I think the world around us will be better and safer with lesser pressure from people. (parentheses added)

Difference

The other major theme was “difference”, which suggests some interviewees perceive autism as difference or diversity rather than a disorder or deficit. Rather than internalizing the medical model embedded in the assessment tools, these people feel proud of their ways of processing the world. These people do acknowledge that their brains process the world differently but they do not think this difference is a disorder. They can articulate how they perceive the world differently from the way neurotypical people do. One interviewee said, “To me, it's more of an ability than an inability.”, and explained how his love and persistence for organic chemistry and computers which might be boring for other people has yielded his thriving life. Instead, they are proud of how their brains process the world and are not afraid of speaking up about the differences between the neurotypical and neurodivergent worlds.

> Autistic people tend to view things as very compartmentalized and very logical. Which can be a problem in a world that is, which as you may imagine, is a problem in a world that very often, it does not conform to neat categories and is most definitely not logical…routine is very important to most autistic people I believe because it is a for one thing, it is something that can be counted on in an otherwise really confusing world. And it and for another thing it is, it is something that conforms to the way autistic people think the world is supposed to be that is, that is supposed to make sense. Supposed to be able to be categorized..

This “difference” theme is specified in the following three sub-themes; belief beyond the diagnostic label, celebrating autism, and accommodation.

Belief beyond the diagnostic label

As the autistic adults I interviewed are over 21, some people were diagnosed more than 20 years ago. At that time, the knowledge and assessment process of autism was not as advanced, and parents had to encounter several labels that did not represent their children correctly or doctors who could not understand how special their children are. However, some parents had a firm belief in their children and helped their children to understand what strengths they have. One mother shared her diagnosis experiences;

> When Josh was first diagnosed the diagnosis was paranoid schizophrenia. I was so disturbed, I had taken some psychology in college, and this didn't seem correct to me. So, I went to the library and began to read more. In my reading I discovered Autism and he ticked 9 of the 10 boxes. So, I got in touch with the Regional Center and he was diagnosed correctly there. Not much was known in those long ago days of 1978, but we did get good occupational services for him.

Josh's mother added that her son had several labels; paranoid schizophrenia, central nervous system disorder, and autism. However, she did not accept the first doctor's diagnosis and tried to find the diagnosis that could explain Josh's challenges and therapists who could help him. At the same time, she avoided using any labels on Josh and kept reminding Josh of his strengths. Her belief in Josh's ability and her dedication to providing childhood experiences that are not different from neurotypical peers contributed to consolidating Josh's pride in himself, which is well presented in Josh's comment.

> Honestly, as a child, I was kind of proud of it. Yeah. I felt it made me special. I didn't even really understand the meaning of the word dysfunction. I knew I was different from others. Although I did not really understand how I knew, I knew, because Mom

> and Dad told me so that I was smarter than most of my peers are better at memorizing things.

Another mother shared her experience at the Regional Center when his son was diagnosed. The following quote implies that sometimes autistic children's diagnosis process and service allocation can be subject to their funding situation, but the mother was determined to take full responsibility for raising her child and to have faith in his potential whatever institutions and professionals label her son.

> . ..So this was 30 years ago. And I felt as if I was on trial, and it was absolutely, or at least at Regional Center. And I understand they were dealing with funding issues, but I really got the sense that whatever was happening, I was responsible for Luke ..

These mothers did not confine their children to the label of autism or other neurological disabilities. They respect their children's differences and find ways to accommodate their needs for their thriving life in the community. These beliefs and respect contributed to forming positive identity and pride in their autistic children and led their children to celebrate autism rather than mask or deny their identity.

Celebrating autism

Autistic people who do not identify themselves as patients are proud of their identity and empowered by their different ways of processing the world. They understand their strengths, they can articulate their differences, and they also ask questions about the ways neurotypical people behave.

Their positive self-identity as an autistic person is based on the support and respect of their parents. Several autistic interviewees stated that they feel special rather than dysfunctional because their parents kept reminding them of how special they are. An autistic person said how he felt empowered when her mother said that he is special, which shows that the respect and expectation from their families might insulate autistic people from internalizing deficit views and help them to thrive in the neurotypical world with pride and confidence.

> I just remembered my mom smiling at me. And she said, You're a special kid. I didn't understand it, then. But she told me that I'm special. I kind of came across my medical report, but I got no doubt. I am really empowered..

Another interviewee, Luke's mother emphasized that the diagnostic label does not limit their potential to thrive in their whole life. She firmly believed that autistic people can blossom as they mature with due respect and appropriate accommodations.

> I just wanted to add one small thing that what I've recognized with Luke is that he continues to mature to grow, continues to learn to, continues to, yes, diversify his abilities and his communication, his communication, it's like he's continuing to blossom. And so as the mom of a child diagnosed with autism, what you may experience now is not what you will experience at 20 or 25, or 30, or 33.

Furthermore, the interview indicated that the attitude to celebrate autism is more noticeable when autistic people have autistic friends who understand them. It is ironic that his psychologist does not understand his attitude to celebrate autism.

> To me, autism is more of an ability than an inability. I feel like many people don't understand what it really means. And even my psychologist, for instance, when I tell her, she doesn't quite get it. And I, I feel like unless I connect with my friends, who have Aspergers, Actually, my friend Nick, who also has Asperger's. He really

> understands me and I feel like it's those that are on the spectrum that understand each other usually.

Finally, autistic people also observe and question neurotypical people’s behaviors, which are also socially inappropriate. When talking about item 38 in SCQ, (If you come into a room and start talking to her/him without calling her/his name, does she/he usually look up and pay attention to you?), Luke pointed out that there are some neurotypical people who do not pay attention to his presence or intentionally ignore him.

> All my siblings do interact with me. Um, however, my sister Allison is married to Bill and he's not from our family. And I feel like sometimes he does not respond to me like when I say hi. I mean, he'll show me stuff but like, yeah, sometimes it takes him a moment or not even like, sometimes he just ignores me, which is weird. He's always on his phone whenever he doesn't see me.

In this quote, Luke might indicate that neurotypical people do not always have desirable social skills and challenge the notion that “lack of social skill” is the hallmark of autistic people, which should be corrected through remedial education.

Accommodation

The interview did not address the issues of accommodation as a major topic, but Josh’s mom and Josh did mention how an appropriate accommodation can yield a dramatic change for Josh’s growth. Replacing handwriting with typing is now a typical example of an accommodation, but it should be noted that Josh and his mom indicated that using a computer really liberated Josh from the nightmare of the homework and helped him to keep pursuing his academic path.

> As a child, and even now as an adult, he writes, well, very legibly, but very laboriously. So Josh draws almost each letter of his name, or whatever he's writing. So as he said, homework just took forever because he was so slow. And when he was maybe 11, or 12, or so we got a computer. And he sat down and he could he learned to type right away. Yeah. And he was able to do his homework and minutes before it had taken hours.

Josh remembered that his teachers often told him to concentrate on the school work when he needed so much time doing his work. However, as this example suggests, the efforts to understand autistic people's challenges and provide timely and appropriate accommodation can make a difference instead of identifying individual deficits and blaming them.

Discussion and Limitation

The interview revealed that some autistic people understand themselves through the assessment tools and attribute their life challenges to individual deficits. These autistic people who regard themselves as patients with medical conditions show the frustration of remaining as incurable patients as well as anxiety of being left alone without any guidance. However, other autistic people and parents, who did not confine themselves to this autistic assessment criteria, celebrate their difference, have friends or families who understand them without judgment and are more satisfied with their life.

This result supports the notion of social kind that classification of people and diagnosis process do influence the way people identify themselves and cannot be regarded as purely objective activity (Hacking, 1995, 2007). Instead, the distinction made by the autism assessments has reified deficits and abnormality, and imposed the sociocultural notion that autistic people should be fixed but cannot be fixed completely. Therefore, this finding supports the rationale that it is worth paying attention to the social and ethical consequences of assessment on people being assessed if the assessment process yields any consequences that were not intended when designing the assessment instruments (Kane, 2006; Messick,1995).

Among contemporary validity discussions, following Messick's tradition, Kane (2006) emphasized the articulated purpose of assessment and the consideration of social and ethical consequences of the assessment instrument in the validation process when the instrument and its results are used beyond this initial purpose. Most validity studies of SCQ have presented that the purpose of the instrument is the precise diagnosis, making a clear distinction between autism and other developmental disabilities. However, the interview I conducted indicated that SCQ and other assessment instruments do more than make the distinction between autistic people and non-autistic people. Assessment instruments and their use not only give autistic people the diagnostic label but reify their differences as deficits that need to be fixed. This reification might lead autistic people to suffer from anxiety, anger, or autistic burnout, a syndrome conceptualized as resulting from chronic life stress and a mismatch of expectations and abilities without adequate suppor*t* (Raymaker et al., 2020) because autistic people identify themselves as incurable patients and they feel that these deficits cannot be completely fixed to meet the criteria of the diagnostic instruments. The notion of incurable patients supports the justification of giving up their agency and relying on the guidance of neurotypical people, which reinforces the hierarchy between clinical professionals and autistic people.

Based on the rationale that autistic people are incurable patients who are not capable enough to fill out the assessment questionnaire even after they become adults, clinical professionals and their parents hardly question the practice of filling out assessment questionnaires by the caregivers. In terms of social justice, this practice should be reconsidered because it might deprive an autistic individual of their agency in making their own life decisions (Carey et al., 2019; Nicolaidis et al., 2011). Moreover, in the measurement perspective, the practice should be questioned due to the measurement error from caregivers' misunderstanding of autistic people's communicative and behavioral intentions (Morsbach & Prinz, 2006). Considering that even self-report practice is questioned due to response bias, incorrect memory, or item misunderstanding, it is worth paying attention to the current practice that autistic people's direct input is limited and the questionnaire is filled out by the caregivers who might not fully understand autistic people (Nicolaidis et al., 2020).

In addition to depriving autistic people's agency and compromising measurement accuracy, the practice of acquiring responses from caregivers reduces the opportunity to examine response processes directly from autistic people. Messick (1995) emphasized the systematic response process analysis to minimize under-construct representation and construct irrelevant variation. In the interview, autistic people stated that they are not fully understood by neurotypical people and are better understood by autistic peers. This remark corroborates the position that autistic attributes might not be fully represented in the items that are developed by non-autistic people. Currently, the studies of response process are rare not just for the autism assessment validation studies but for the validation studies in general (Zumbo & Chan, 2014), but if researchers make more efforts to refine questionnaire items to be more accessible to the autistic community and to incorporate autistic people's perspectives into the item development, the autism assessment tools may be constructed on the more realistic autism attributes, not on the attributes assumed by the people who have not experienced what autism is. The autism assessment tools with autistic people's perspectives not only capture the autism attributes in scientifically precise ways but also be more useful in identifying autistic people's daily challenges and needs.

However, these implications should be considered with caution due to the following limitations. First, these findings are based on the interview with only 16 people, who do not represent the full diversity of identities and geographical region, race, age, or gender. If I had conducted the interview with more people, the results might have been different. Second, this result is based on one-time interviews. If multiple interviews had been conducted with the interviewees, it might have been possible to acquire more in-depth narratives from each individual. Lastly, with the exception of two, most interviewees were able to read, write, and verbalize their ideas without difficulties, which inevitably excluded the voice of autistic people who have limited expressive language. Future research should pay more attention to the autistic population who have limited expressive language or use alternative communication devices to examine how these people think about the diagnosis process and how they define themselves through this process.

Final Conclusion

As the definition of autism has been changing, there is a need to revalidate autism assessment instruments. The conceptual shift from the medical model to the neurodiversity model, more accumulated autistic narratives, and increasing awareness of ethical consequences of autism assessment diagnosis call for rethinking current autism assessment instruments. With these needs, this study reveals that autism assessment instruments are not founded on a solid theory to define autism, do entail subjective and value-laden language, and exclude autistic people’s perspectives and participation. Also, this study suggests that the ethical consequences of autism assessment should be paid more attention to because autism assessment does affect the way autistic people identify themselves. With these findings, in the future, more studies should investigate how to develop a new autism assessment based on the neurodiversity model, how to refine items to incorporate autistic people’s first-person accounts, and how to make the assessment instrument more accessible to autistic people to collect their response directly. By paying more attention to these validity issues, we can assess autism more ethically and understand autism more correctly, not just categorizing and differentiating them from other developmental disabilities.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: American Psychological Association.

Angoff, W. H. (1988). Validity: An evolving concept. In H. Wainer & H. Braun (Eds.), Test validity (pp. 19–32). Hillsdale: Lawrence Erlbaum Associates.

Barnard-Brak, L., Richman, D. M., Chesnut, S. R., & Little, T. D. (2016). Social Communication Questionnaire scoring procedures for autism spectrum disorder and the prevalence of potential social communication disorder in ASD. *School Psychology Quarterly*, *31*(4), 522.

Baron-Cohen, S., Ashwin, E., Ashwin, C., Tavassoli, T., & Chakrabarti, B. (2009). Talent in autism: hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1522), 1377-1383.

Bascom, J. (2012). Loud hands: Autistic people, speaking. Washington, DC: Autistic Self Advocacy Network.

Baudrillard, J. (1997). Simulacra and simulation (S. Glaser, Trans.). Ann Arbor, MI: University of Michigan Press. (Original work published 1981).

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.

Bingham, W. V. (1937). Aptitudes and aptitude testing. New York: Harper.

Bleuler, E. (1950[1911]) Dementia Praecox or the Group of Schizophrenias. New York: International Universities.

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2004). The concept of validity. Psychological review, 111(4), 1061.

Booth, J. (2016). Autism equality in the workplace: Removing barriers and challenging discrimination. Jessica Kingsley Publishers.

Brooks, W. T., & Benson, B. A. (2013). The validity of the social communication questionnaire

in adults with intellectual disability. *Research in Autism Spectrum Disorders*, *7*(2), 247-255.

Carey, A. C., Block, P., & Scotch, R. K. (2019). Sometimes allies: Parent-led disability organizations and social movements. *Disability Studies Quarterly*, *39*(1).

Chapman, R. (2019). Neurodiversity theory and its discontents: Autism, schizophrenia, and the

social model of disability. The Bloomsbury companion to philosophy of psychiatry, 371.

Chapman, R. (2020). Neurodiversity, disability, wellbeing. *Neurodiversity studies: A new critical*

paradigm, 57-72.

Coleman, M. (Ed.). (2009). The neurology of autism. Oxford University Press.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative* *analysis*. sage.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52(4), 281.

Danziger, K. (1997). *Naming the mind: How psychology found its language*. Sage.

Dinishak, J. (2021). Autistic autobiography and hermeneutical injustice. *Metaphilosophy*, *52*(5), 556-569.

Dinishak, J. (2016). The deficit view and its critics. Disability Studies Quarterly, 36(4).

Evans, B. (2013). How autism became autism: The radical transformation of a central concept of child development in Britain. *History of the human sciences*, *26*(3), 3-31.

Galton, F. (1962). Hereditary genius: An inquiry into its laws and consequences. Cleveland: World. (Original work published 1869).

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York, NY: Basic Books.

Gergen, K. J. (2009). Relational being: Beyond self and community. New York, NY: Oxford University Press.

Goddard, H. H. (1926). Feeble-mindedness: Its causes and consequences. New York, NY: Macmillan. (Original work published 1914).

Gould, S. J. (1996). The mismeasure of man (Revised and expanded). New York, NY: Norton.

Guilford, J. P. (1946). New standards for test evaluation. Educational and Psychological Measurement, 6(4), 427–438.

Hacking, I. (1995). The looping effects of human kinds.

Hacking, I. (2007, April). Kinds of people: Moving targets. In *Proceedings-British Academy* (Vol. 151, p. 285). Oxford University Press Inc..

Hacking, I. (2009 a). Autistic autobiography. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1522), 1467-1473.

Hacking, I. (2009 b). Humans, aliens & autism. Daedalus, 138(3), 44-59.

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. Nature neuroscience, 9(10), 1218-1220.

Haslam, N. (2014). Natural kinds in psychiatry: Conceptually implausible, empirically questionable, and stigmatizing. Classifying psychopathology: Mental kinds and natural kinds, 11-28.

Humphreys, K., Hasson, U., Avidan, G., Minshew, N., & Behrmann, M. (2008). Cortical patterns of category-selective activation for faces, places and objects in adults with autism. Autism Research, 1(1), 52-63.

Jaarsma, P., & Welin, S. (2012). Autism as a natural human variation: Reflections on the claims of the neurodiversity movement. Health care analysis, 20(1), 20-30.

Kane, M. T. (2001). Current concerns in validity theory. Journal of Educational Measurement,

38, 319–342.

Kane, M. T. (2006). Validation. In R.L. Brennan (Ed.), Educational measurement, 4th edition (pp. 17-64). Santa Barbara: Greenwood Publishing Group.

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

Kapp, S. (2019). How social deficit models exacerbate the medical model: autism as case in point. Autism Policy & Practice, 2(1), 3-28.

Kapp, S. K., Gillespie-Lynch, K., Sherman, L. E., & Hutman, T. (2013). Deficit, difference, or

both? Autism and neurodiversity. Developmental psychology, 49(1), 59.

Kapp, S. K., & Ne'eman, A. (2020). Lobbying autism's diagnostic revision in the DSM-5. In *Autistic Community and the Neurodiversity Movement* (pp. 167-194). Palgrave Macmillan, Singapore.

Kurth, J. A. (2015). Educational placement of students with autism: The impact of state of residence. Focus on Autism and Other Developmental Disabilities, 30(4), 249–256. doi:10.1177/1088357614547891

Latif, S. (2016). THE ETHICS AND CONSEQUENCES OF MAKING AUTISM SPECTRUM

DISORDER DIAGNOSES. *Re-thinking Autism: Diagnosis, Identity and Equality*,

288.

Lord, C. (2011) Epidemiology: How common is autism? Nature 474(7350): 166–168.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). Autism diagnostic observation schedule (ADOS) Los Angeles. *CA: Western Psychological Services*.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of autism and developmental disorders*, *24*(5), 659-685.

Lotter, V. (1966) ‘Epidemiology of Autistic Conditions in Young Children’, Social Psychiatry 1: 124–37.

MacFarlane, J. R., & Kanaya, T. (2009). What does it mean to be autistic? Inter-state variation in special education criteria for autism services. Journal of Child and Family Studies, 18(6), 662–669. doi:10.100710826-009-9268-8

Markus, K. A., & Borsboom, D. (2013). *Frontiers of test validity theory: Measurement, causation, and meaning*. Routledge.

Marvin, Alison R., Daniel J. Marvin, Paul H. Lipkin, and J. Kiely Law. "Analysis of Social Communication Questionnaire (SCQ) screening for children less than age 4." *Current developmental disorders reports* 4, no. 4 (2017): 137-144.

Maul, A. (2018). Validity. In Bruce Frey (Ed.) The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation (1771-1775). Thousand

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13–103).

New York: American Council on Education and Macmillan.

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons'

responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741.

Milton, D. E. (2012). On the ontological status of autism: the 'double empathy problem'. 36 Disability & Society, 27(6), 883-887.

Morsbach, S. K., & Prinz, R. J. (2006). Understanding and improving the validity of self-report

of parenting. *Clinical Child and Family Psychology Review*, *9*(1), 1-21.

Nacewicz, B. M., Dalton, K. M., Johnstone, T., Long, M. T., McAuliff, E. M., Oakes, T. R., ... &

Davidson, R. J. (2006). Amygdala volume and nonverbal social impairment in adolescent

and adult males with autism. Archives of general psychiatry, 63(12), 1417-1428.

Nadesan, M. H. (2013). Constructing autism: Unraveling the truth and understanding the social. Routledge.

Ne'eman, A. (2010). The future (and the past) of Autism advocacy, or why the ASA's magazine,

The Advocate wouldn't publish this piece. Disability Studies Quarterly, 30(1).

Nicolaidis, C. (2012). What can physicians learn from the neurodiversity movement?. AMA

Journal of Ethics, 14(6), 503-510.

Nicolaidis, C., Raymaker, D., McDonald, K., Dern, S., Ashkenazy, E., Boisclair, C., ... & Baggs, A. (2011). Collaboration strategies in nontraditional community-based participatory research partnerships: lessons from an academic–community partnership with autistic self-advocates. Progress in Community Health Partnerships, 5(2), 143.

Nicolaidis, C., Raymaker, D. M., McDonald, K. E., Lund, E. M., Leotti, S., Kapp, S. K., ... & Hunter, M. (2020). Creating accessible survey instruments for use with autistic adults and people with intellectual disability: Lessons learned and recommendations. Autism in Adulthood, 2(1), 61-76.

Novikova, S. I., Richman, D. M., Supekar, K., Barnard-Brak, L., & Hall, D. (2013). NDAR: A model federal system for secondary analysis in developmental disabilities research. In *International review of research in developmental disabilities*(Vol. 45, pp. 123-153). Academic Press.

Office of Special Education Programs. (2008). State level data files child count 1991–2006. Washington, D.C: U.S. Department of Education. Retrieved October 13, 2008 from https://www.ideadata.org/StateLevelFiles.asp.

Oliver, M. (1990). The politics of disablement. Basingstoke: MacMillan.

Orsini, M., & Smith, M. (2010). Social movements, knowledge and public policy: the case of autism activism in Canada and the US. Critical Policy Studies, 4(1), 38-57.

Ortega, F. (2009). The cerebral subject and the challenge of neurodiversity. BioSocieties, 4(4), 425-445.

Piaget, J. (1923) ‘La pense´e symbolique et la pense´e de l’enfant’ [Symbolic Thought and Infantile Thought], Archives de psychologie 18: 273–304.

Piaget, Jean (1929) The Child’s Conception of the World. London: Kegan Paul.

Raymaker, D. M., Teo, A. R., Steckler, N. A., Lentz, B., Scharer, M., Delos Santos, A., ... & Nicolaidis, C. (2020). “Having all of your internal resources exhausted beyond measure and being left with no clean-up crew”: Defining autistic burnout. *Autism in adulthood*, *2*(2), 132-143.

Safer-Lichtenstein, J., Hamilton, J., & McIntyre, L. L. (2021). School-Based Autism Rates by State: An Analysis of Demographics, Political Leanings, and Differential Identification. Journal of Autism and Developmental Disorders, 51(7), 2271–2283. doi:10.100710803-020-04700-3 PMID:32926306

Safer-Lichtenstein, J., & McIntyre, L. L. (2020). Comparing autism symptom severity between children with a medical autism diagnosis and an autism special education eligibility. Focus on Autism and Other Developmental Disabilities, 35(3), 186–192. doi:10.1177/1088357620922162

Shakespeare, T. (2006). The social model of disability. *The disability studies reader*, *2*, 197-204.

Silberman, S. (2015). *Neurotribes: The legacy of autism and the future of neurodiversity*.

Penguin.

Sireci, S. G. (2009). Packing and unpacking sources of validity evidence. In R. W. Lissitz (Ed.),

The concept of validity: Revisions, new directions and applications (pp. 19–37). Charlotte: Information Age Publishing Inc.

Slaney, K. (2017). *Validating psychological constructs: Historical, philosophical, and practical dimensions*. Springer.

Teo, T. (2018). *Outline of theoretical psychology*. Palgrave Macmillan UK.

Timimi, S., McCabe, B., Runswick-Cole, K., & Mallett, R. (2016 a). What have we learned from the science of autism. Re-thinking autism: Diagnosis, identity and equality, 30-48.

Timimi, S., McCabe, B., Runswick-Cole, K., Mallett, R., & Timimi, S. (2016 b). Autism screening and diagnostic tools. *Re-thinking autism: Diagnosis, identity and equality*, 159-183.

Timimi, S., Milton, D., Bovell, V., Kapp, S., & Russell, G. (2019). Deconstructing diagnosis: Four commentaries on a diagnostic tool to assess individuals for autism spectrum disorders. *Autonomy (Birmingham, England)*, *1*(6).

Tracy, S. J. (2019). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. John Wiley & Sons.

Tsai, L. Y. (2004). Autistic disorder: Textbook of child and adolescent psychiatry (pp. 261–263).

Arlington, VA: American Psychiatric Publishing.

Waterhouse, L. (2013). Rethinking autism: Variation and complexity. Academic Press.

Wei, T., Chesnut, S. R., Barnard-Brak, L., & Richman, D. (2015). Psychometric analysis of the

Social Communication Questionnaire using an item-response theory framework: implications for the use of the lifetime and current forms. *Journal of Psychopathology and Behavioral Assessment*, *37*(3), 469-480.

Wilson, M. (2004). *Constructing Measures: An Item Response Modeling Approach: An Item Response Modeling Approach*. Routledge.

Wu, M., & Adams, R. J. (2013). Properties of Rasch residual fit statistics. *Journal of Applied Measurement*.

Zumbo, B. D., & Chan, E. K. (2014). Validity and validation in social, behavioral, and health sciences (Social Indicators Research Series, Vol. 54. *Cham: Springer International Publishing*.

Appendix A

Interview Protocol

A. General topic-related questions
    1. How do you describe autism and why?
    2. What are your overall impressions of autism screening tools whenever you fill out the surveys?
        a. How do you feel when you fill out the survey?
        b. Is there any experience that is memorable for you either in a positive or a negative way?
        c. What do you think is the purpose of these autism screening tools?

B. Questions specific to SCQ (Current form)?
    1. What were your overall impressions of this questionnaire?
    2. Were there any items that might be unclear to someone filling it out?
    3. When you encounter items like item 9 , how do you determine "inappropriate"?. (probing: how do you distinguish "normal" from "abnormal" or "socially appropriate" from "socially inappropriate"?)
    4. Do you think if you understood your child/student's communicative intention enough to answer these items (ex. 21, 22 spontaneously) when your child was about 4-5? (follow up question: how do you know that is what your child meant?)

5. In your opinion, were there any phrases or words that might offend a caregiver filling this out? (follow up : in your opinion, were there any phrases or words that may be culturally biased?)
6. What is this survey missing? What aspects of autism are missing?
7. What is this survey overemphasizing? In your opinion, what aspects of autism are overemphasizing?
8. Please read this item; "Sometimes your child covers his/her ears to block out painful noises like vacuum cleaners or people talking too much or too loud." Do you think there are any differences in the way this item is written from the items that you read in SCQ?

Appendix B

Social Communication Current Form

SCQ_Current_Answer Sheet
Name of the subject
DOB
Interview Date
Age
Gender
Name of Respondent
Relation to Subject

Thank you for taking the time to complete this questionnaire. Please answer each question by selecting *yes* or *no*. A few questions ask about several related types of behavior; please select *yes* if any of these behaviors were present during the past 3 months. Although you may be uncertain about whether some behaviors were present or not, please answer *yes* or *no* to every question on the basis of what you think.

Item

1. Is she/he now able to talk using short phrases or sentences? If no, skip to question8.
2. Do you have a to and fro "conversation" with her/him that involves taking turns or building on what you have said?
3. Does she/he ever use odd phrases or say the same thing over and over in almost exactly the same way (either phrases that she/he hears other people use or ones that she/he makes up)?
4. Does she/he ever use socially inappropriate questions or statements? For example, does she/he ever regularly ask personal questions or make personal comments at awkward times?
5. Does she/he ever get her/his pronouns mixed up (e.g., saying *you* or *she/he* for *I*)?

6. Does she/he ever use words that she/he seems to have invented or made up her/himself; put things in odd, indirect ways; or use metaphorical ways of saying things (eg., saying *hot rain* for *steam*)?
7. Does she/he ever say the same thing over and over in exactly the same way or insist that you say the same thing over and over again?
8. Does she/he have things that she/he seems to do in a very particular way or order or rituals that she/he insists that you go through?
9. Does her/his facial expressions usually seem appropriate to the particular situation as far as you can tell?
10. Does she/he ever use your hand like a tool or as if it were part of his/her own body (e.g., pointing with your finger or putting your hand on a doorknob to get you to open the door?)
11. Does she/he ever have any interests that preoccupy her/him and might seem odd to other people (e.g., traffic lights, drainpipes, or timetables)?
12. Does she/he ever seem to be more interested in parts of a toy or object (e.g., spinning the wheels of a car), rather than in using the objects as it was intended?
13. Does she/he ever have any special interests that are unusual in their intensity but otherwise appropriate for his/her age and peer group (e.g., trains or dinosaurs)?
14. Does she/he ever seem to be unusually interested in the sight, feel, sound, taste, or smell of things or people?
15. Does she/he ever have any mannerisms or off ways of moving her/his hands or fingers, such as flapping or moving her/his fingers in front of her/his eyes?

16. Does she/he ever have any complicated movements of her/his whole body, such as spinning or replete bouncing up and down?
17. Does she/he ever injure her/himself deliberately, such as by biting her/his arms or banging her/his head?
18. Does she/he ever have any objects (other than soft toys or comfort blanket) that she/he has to carry around?
19. Does she/he have any particular friends or a best friend?
20. Does she/he ever talk with you just to be friendly (rather than to get something)?
21. Does she/e ever spontaneously copy you(or other people) or what you are doing (such as vacuuming, gardening, or mending things)?
22. Does she/he ever spontaneously point things around her/him just to show you things (not because she/he wants them)?
23. Does she/he ever use gestures, other than pointing or pulling your hand, to let you know what she/he wants?
24. Does she/he nod her/his head to indicate *yes*?
25. Does she/he shake her/his head to indicate *no*?
26. Does she/he usually look at you directly in the face when doing things with you or talking with you?
27. Does she/he smile back if someone smiles at her/him?
28. Does she/he ever show you things that interest her/him to engage your attention?
29. Does she/he ever offer to share things other than food with you?
30. Does she/he ever seem to want you to join in her/his enjoyment of something?
31. Does she/he ever try to comfort you if you are sad or hurt?

32. If she/he wants something or wants help, does she/he look at you and use gestures with sounds or words to get your attention?
33. Does she/he show a normal range of facial expressions?
34. Does she/he ever spontaneously join in and try to copy the actions in social games, such as *The Mulberry Bush* or *London Bridges Is Falling Down*?
35. Does she/he play any pretend or make-believe games?
36. Does she/he seem interested in other children of approximately the same age whom she/he does not know?
37. Does she/he respond positively when another child approaches her/him?
38. If you come into a room and start talking to her/him without calling her/his name, does she/he usually look up and pay attention to you?
39. Does she/he ever play imaginative games with another child in such a way that you can tell that each child understands what the other is pretending?
40. Does she/he play cooperatively in games that need some form of joining in with a group of other children, such as hide-and seek or ball games?