Differentiating ID and ASD: Measurement Matters

Cristan Farmer, Ph.D.

Neurodevelopmental & Behavioral Phenotyping Service
Intramural Research Program

National Institute of Mental Health

Disclosures

I am an employee of the US Federal Government. These views are my own. I have no disclosures.

Collaborators

Audrey Thurm, PhD
Clinical Psychologist
Director, NBPS
National Institute of Mental Health

Somer Bishop, PhD
Clinical Psychologist
Assoc. Prof. of Psychology & Psychiatry
University of California, San Francisco

State of the Field: Differentiating Intellectual Disability From Autism Spectrum Disorder

Audrey Thurm[1], Cristan Farmer[1], Emma Salzman[2], Catherine Lord[3] and Somer Bishop[2]*

frontiers in Psychiatry

REVIEW
published: 30 July 2019
doi: 10.3389/fpsyt.2019.00526

NIH National Institute of Mental Health

Objectives

Measurement Matters

The **scientific and clinical importance** of accurately differentiating ID and ASD.

The **DSM-based guidelines and clinical best practice** for the diagnosis of ID, and of ASD in the context of ID.

Matters of Measurement

The current state of knowledge regarding the use and validity of **tools used to assess ASD symptoms** in the context of ID.

The definition of **measurement invariance** and its relevance to tools used to assess ASD symptoms in the context of ID.

The Diagnosis of Intellectual Disability

DSM-5 Diagnostic Criteria for ID

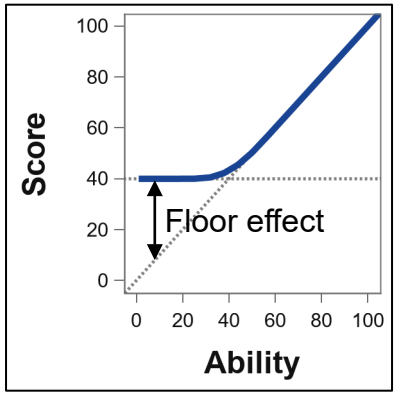
A. **Deficits in intellectual functions**, such as reasoning, problem solving, planning, abstract thinking, judgment, and academic learning and learning from experience, **confirmed by both clinical assessment and individualized, standardized intelligence testing**.

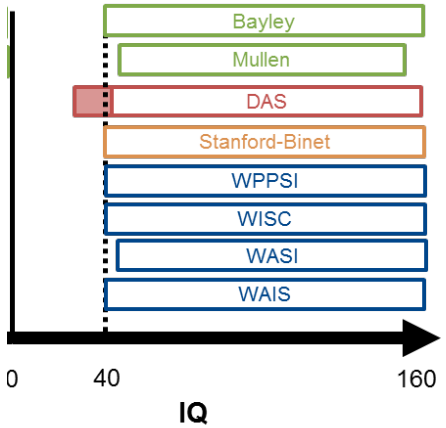
B. Deficits in adaptive functioning that result in failure to meet developmental and sociocultural standards for personal independence and social responsibility. Without ongoing support, the adaptive deficits limit functioning in one or more activities of daily life, such as communication, social participation, and independent living, and across multiple environments, such as home, school, work, and recreation.

C. Onset of intellectual and adaptive deficits during the developmental period.

Limitations of standardized assessment in ID

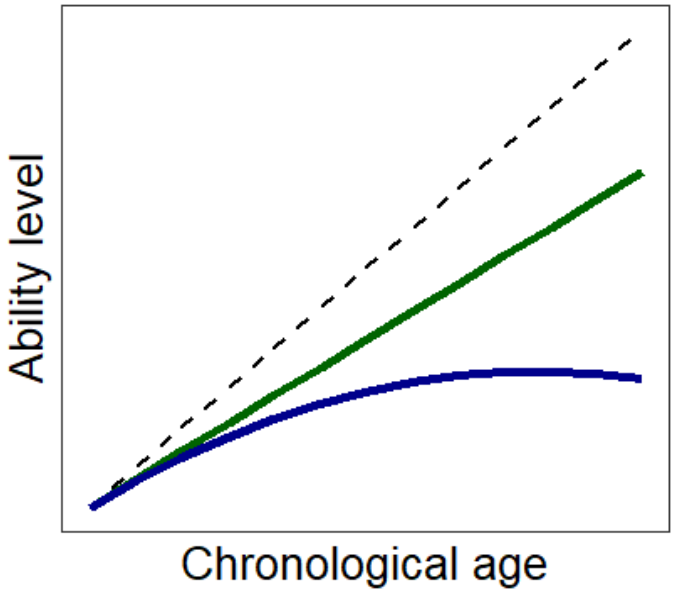
- Because of unreliability, standardized IQ tests do not provide standard score estimates < 40
- Tests appropriate for the developmental level of a person may not be appropriate for their chronological age
- Use of a testing hierarchy and ratio IQ may be indicated (Soorya, Leon, Trelles, & Thurm, 2017)





Top figure illustrates floor effect, wherein individuals with a range of ability get the same score. Bottom figure illustrates the score range covered by popular developmental/IQ assessments.

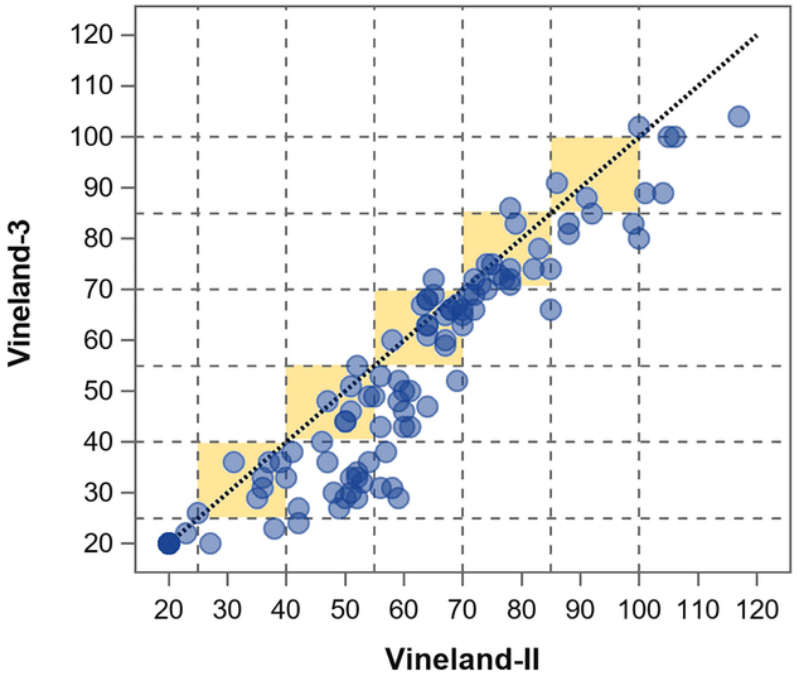
Limitations of standardized assessment in ID



Typical-for-age

Below-average standard scores, but fairly stable over time

Standard scores decrease over time, likely reaching floor



Caution: Norm-referenced scores ("standard scores") often decrease over time, especially among individuals with severe ID, when ability plateaus

Caution: Different measures of the same construct (and even editions of the same measure, see above) can yield different scores for the same person

Left figure illustrates different developmental trajectories. Right figure shows lower scores for the same person on adaptive behavior assessments Vineland-3 versus Vineland-II; Farmer et al. 2020.

DSM-5 Diagnostic Criteria for ID

A. **Deficits in intellectual functions**, such as reasoning, problem solving, planning, abstract thinking, judgment, and academic learning and learning from experience, **confirmed by both clinical assessment and individualized, standardized intelligence testing**.

B. Deficits in adaptive functioning that result in failure to meet developmental and sociocultural standards for personal independence and social responsibility. Without ongoing support, the adaptive deficits limit functioning in one or more activities of daily life, such as communication, social participation, and independent living, and across multiple environments, such as home, school, work, and recreation.

C. Onset of intellectual and adaptive deficits during the developmental period.

DSM-5 ID Severity Classifiers

ID is described as *mild, moderate, severe,* or *profound* based on impairment in adaptive behavior

Severity levels have their own ICD and DSM codes – they are distinct

Domain	Scale	Description
Conceptual Domain	Communication	Speech and listening skills necessary for communicating with others.
	Functional academics	Basic reading, writing, and mathematic skills, along with other abilities needed for daily living.
	Self-direction	Skill necessary for independence, responsible behavior, and self-control.
Social Domain	Leisure	Abilities to plan and participate in leisure activities.
	Social	Skills necessary to have good relations with others.
Practical Domain	Community use	Skills needed to participate in community living.
	Home/school living[a]	Basic abilities necessary to take care of a home or school.
	Health and safety	Skills needed for health care, and to deal with sickness and lesions.
	Self-care	Skills necessary for self-care (e.g., feeding, dressing, hygiene).

Figure shows ABAS domains/scales which correspond to generally accepted domains of adaptive behavior; from Montero-Zenteno & Fernández-Pinto, 2013

Epidemiology of ID

- Mild ID is familial; more severe forms are largely not
- Moderate – profound ID is *very rare* (<0.1%) in general population
- Genetic etiology associated with more severe forms of ID

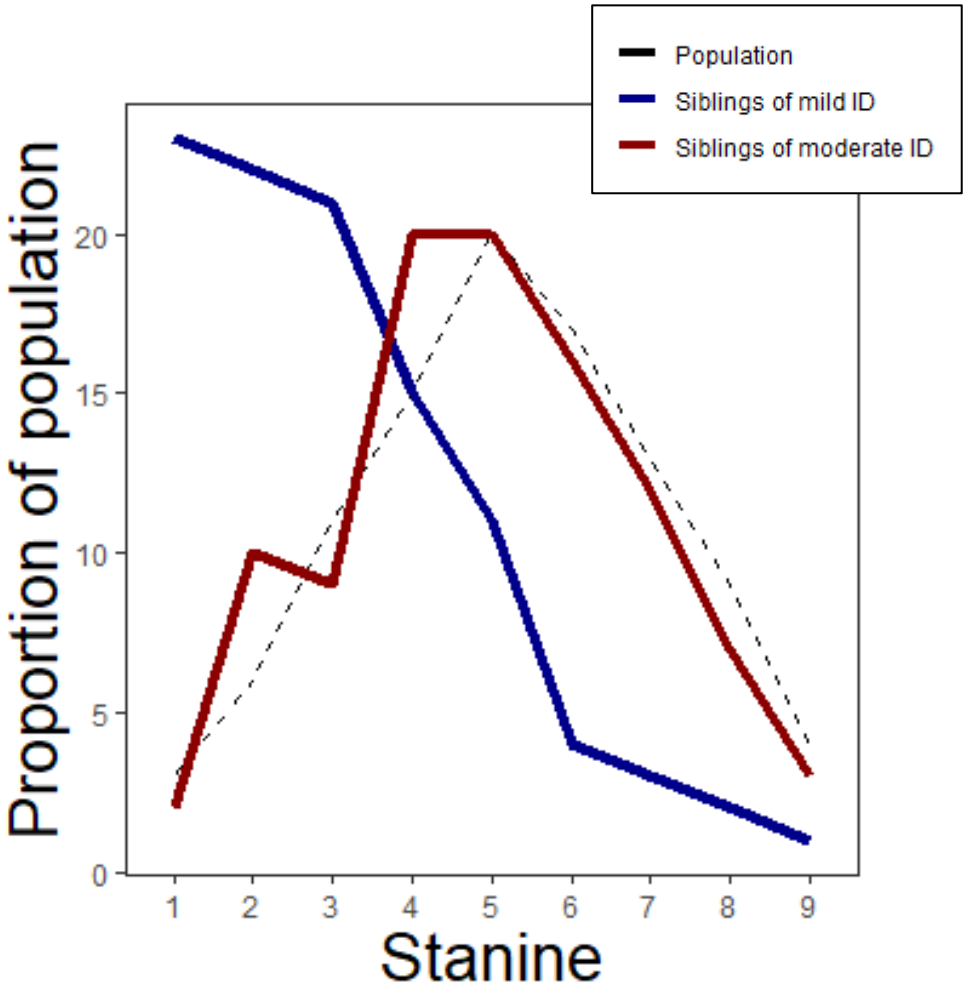


Figure illustrates much higher rate of low IQ among siblings of mild ID, but no elevation among siblings of those with moderate ID. Adapted from Israeli data in Abraham Reichenberg et al. PNAS 2016;113:4:1098-1103

Epidemiology of ID

- Mild ID is familial; more severe forms are largely not
- Moderate – profound ID is *very rare* (<0.1%) in general population
- Genetic etiology associated with more severe forms of ID

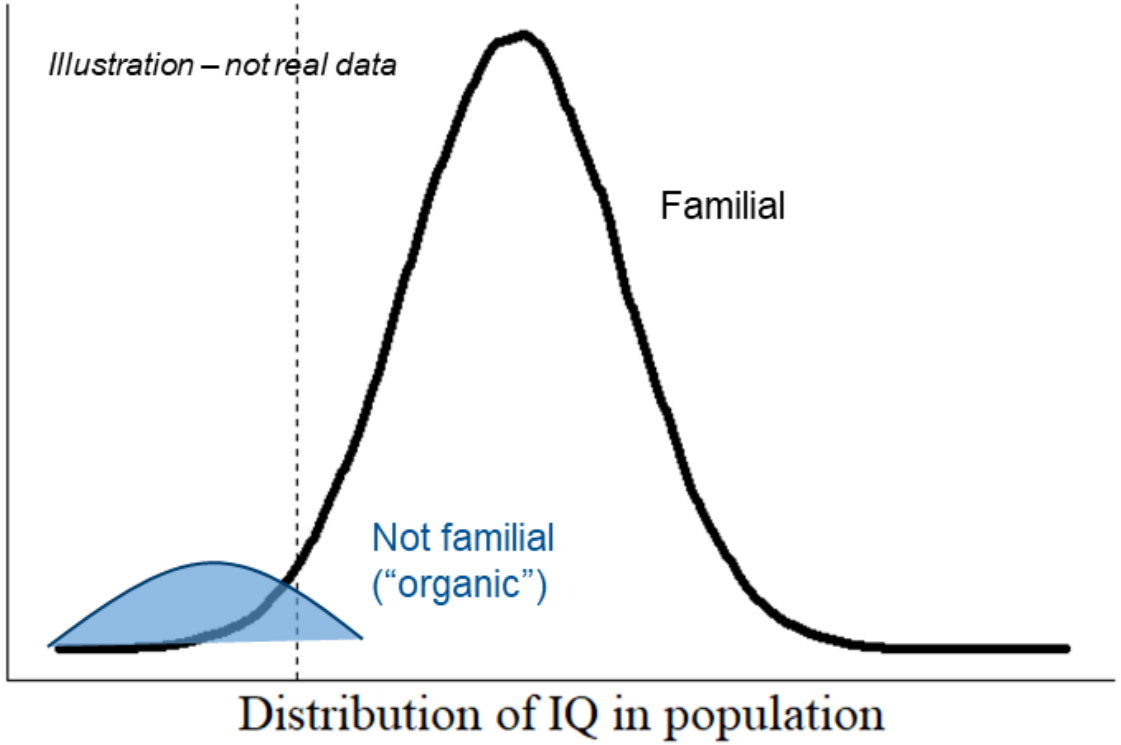


Figure is illustration of the "bump of pathology," which describes individuals whose ID is not familial.

Epidemiology of ID

- Mild ID is familial; more severe forms are not
- Moderate – profound ID is *very rare* (<0.1%) in general population
- Genetic etiology associated with more severe forms of ID

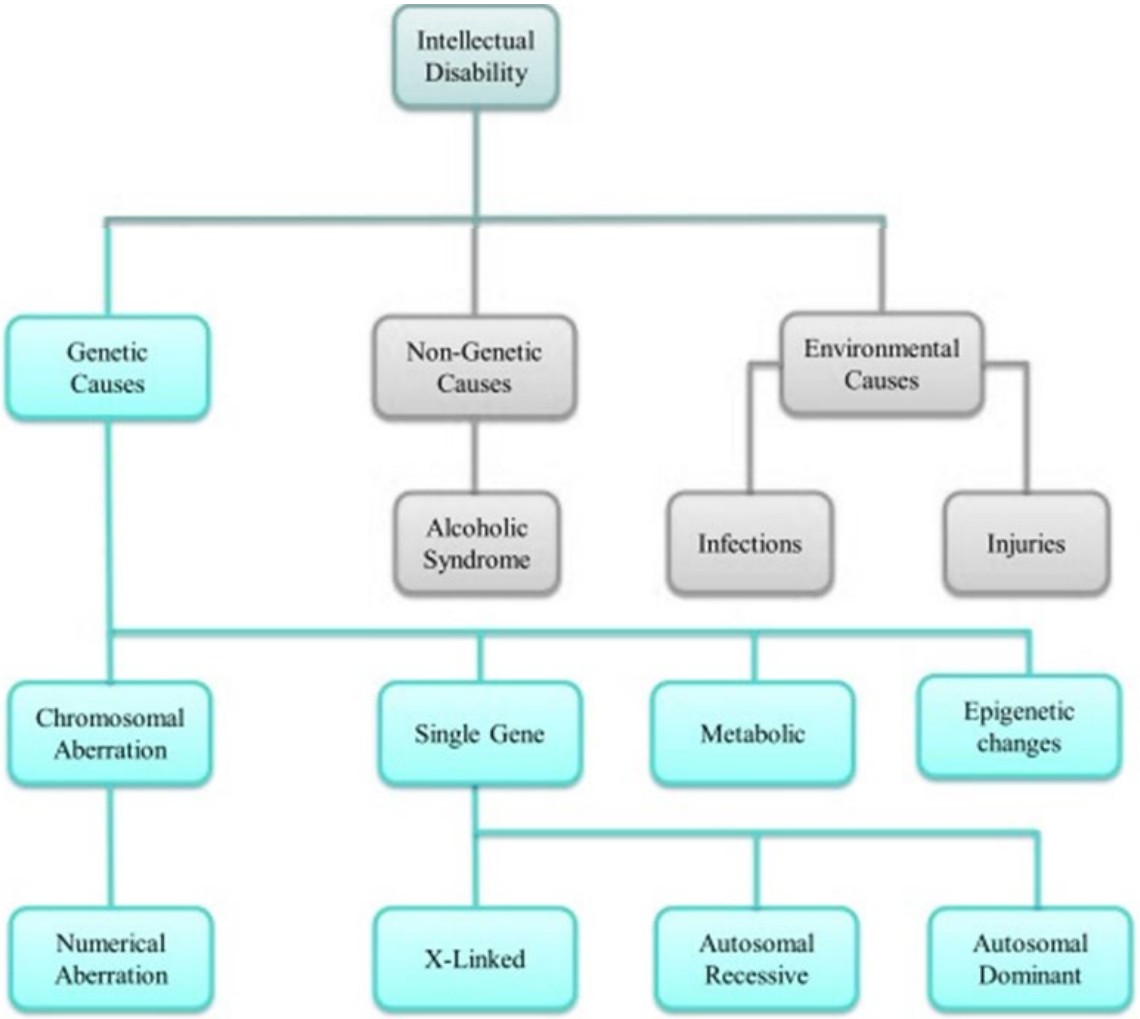


Figure summarizes ID etiologies. Genetic etiology (teal) is more likely to be associated with more severe forms of ID than are non-genetic and environmental (purple; sometimes referred to as "familial/cultural"). Reproduced and adapted from Ilyas et al. (2020) F1000 Research.

Summary points

- Functioning in ID ranges from independent living to 24-hour 1:1 care; these extremes really are distinct disorders
- There are methodological limitations on quantification of IQ below moderate range
- ID below mild range very rare in general population, but in subpopulations with genetic disorders associated with neurodevelopmental problems

Why are ID and ASD difficult to differentiate?

DSM-5 Diagnostic Criteria for ASD

A. Persistent deficits in social communication and social interaction across contexts, not accounted for by general developmental delays …

1. Deficits in social-emotional reciprocity
2. Deficits in nonverbal comm behaviors used for social interaction
3. Deficits in developing and maintaining relationships

B. Restricted, repetitive patterns of behavior, interests, or activities …

1. Stereotyped or repetitive speech, motor movements, or use of objects
2. Insistence on sameness/inflexible adherence to routines, ritualized patterns of verbal or nonverbal behavior
3. Highly restricted, fixated interests / abnormal in intensity or focus
4. Hyper-or hypo-reactivity to sensory input or unusual interest in sensory aspects of environment;

C. Symptoms must be present in early developmental period

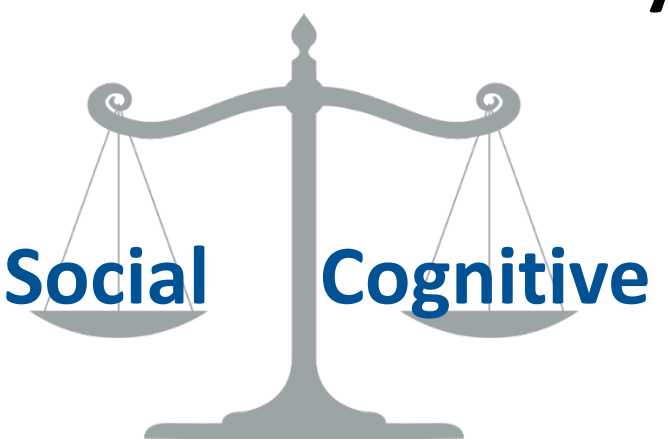
D. Symptoms clinically impair current functioning

E. Disturbances not better explained by ID, DD

DSM-5 Criterion E

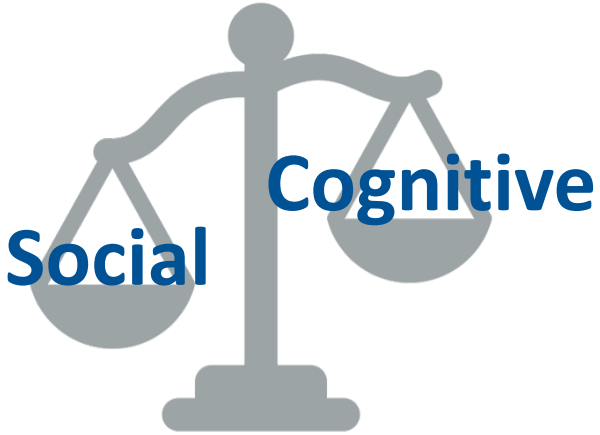
These disturbances are not better explained by intellectual disability or global developmental delay.

Intellectual Disability



Social ability as expected based on developmental level

Autism Spectrum Disorder



Social ability less than predicted by developmental level, plus RRB

Description of social impairments associated with levels of ID

Mild	Communication, conversation, and language are more concrete or immature than the skills of peers. The child may have difficulty accurately understanding the social cues of others. There may be difficulties regulating emotion and behavior compared to peers.
Moderate	Marked differences in social and communicative skills compared to peers. Spoken language is simplistic and concrete. Social judgment and decision making are limited. Friendships with peers are often affected by social or communicative deficits.
Severe	Limited spoken language skills with simplistic vocabulary and grammar. Speech may be single words/phrases. The child understands simple speech and gestures. Relationships are with family members are other familiar people.
Profound	Limited understanding of symbolic communication. The child may understand some simple instructions and gestures. Communication is usually through nonverbal, non-symbolic means. Relationships are usually with family members and other familiar people. Co-occurring physical problems may greatly limit functioning.

Table reproduced from Weis, 2013 Introduction to Abnormal Child and Adolescent Psychology

Cannot expect a child to perform "social" behaviors above their mental age

Table 1. A timeline of gesture development.

10–13 months	12–13 months	15–16 months	18–20 months	2–5 years	School age
Showing Giving Pointing Ritualized request	Representational gestures, play schemes	Gesture or vocal preference	Spoken word preference, gesture-plus-spoken combinations	Speech–gesture integration, beat gestures emerge	Mismatched gesture-plus-spoken combinations
POINT predicts first words	First words emerge		Significant increase in words (types, tokens)	Gesture scaffolds spoken expression and comprehension	Mismatches indexes the transitional knowledge state
Other prelinguistic behaviors include eye contact, joint attention, and turntaking	Gesture serves a complementary function to spoken forms		Increased pointing in combination with spoken words	Transition from BPO to IO gestures	Gesture aids in the transition to concept acquisition
			Transition to empty-handed play schemes	Iconic and beat gestures accompany longer utterances	

Note. BPO = body part as object; IO = imaginary object.

Table shows some expected nonverbal/verbal communication skills by chronological age. Copied from Capone, N. C., & McGregor, K. K. (2004). Journal of Speech, Language, and Hearing Research.

NIH National Institute of Mental Health

Why does it matter?

- **Validity of research:** ASD geneticists are looking to specific genetic conditions with increased rates of ASD for clues about how to derive treatments for ASD symptoms
- **Outcome measures:** If we realize the goal of treatments for the root cause of specific neurodevelopmental problems, such as preventative gene therapy, specific endpoints must be defined
- **Precision medicine**: ID and ASD have different prognosis and interventions

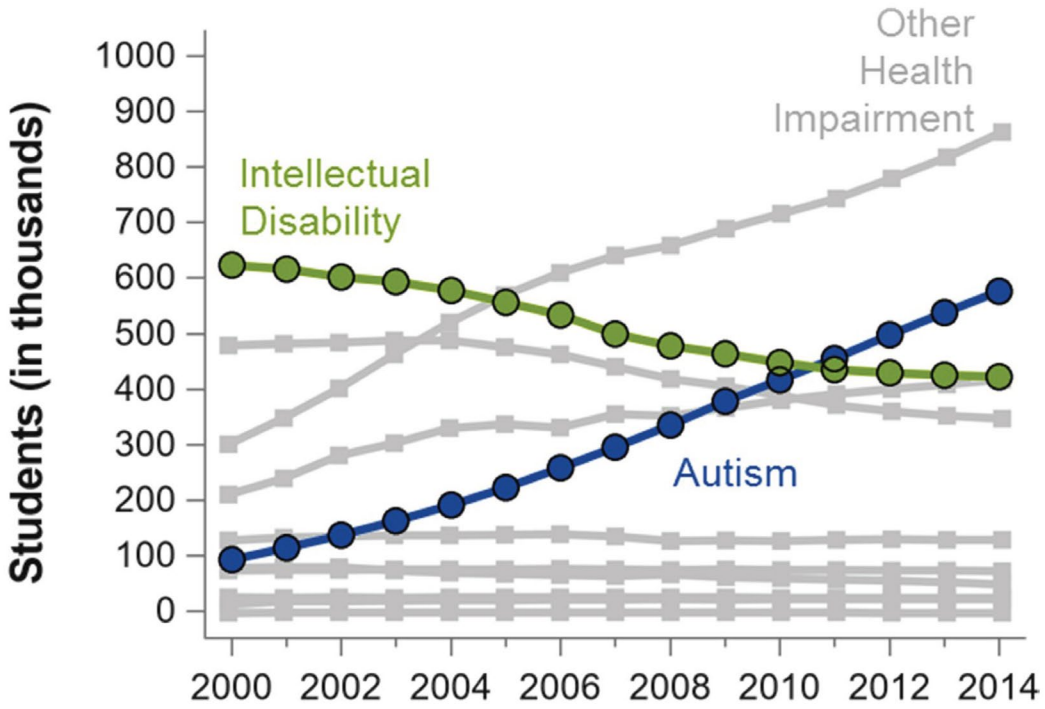
When does it NOT matter?

Some interventions may be indicated regardless of diagnosis

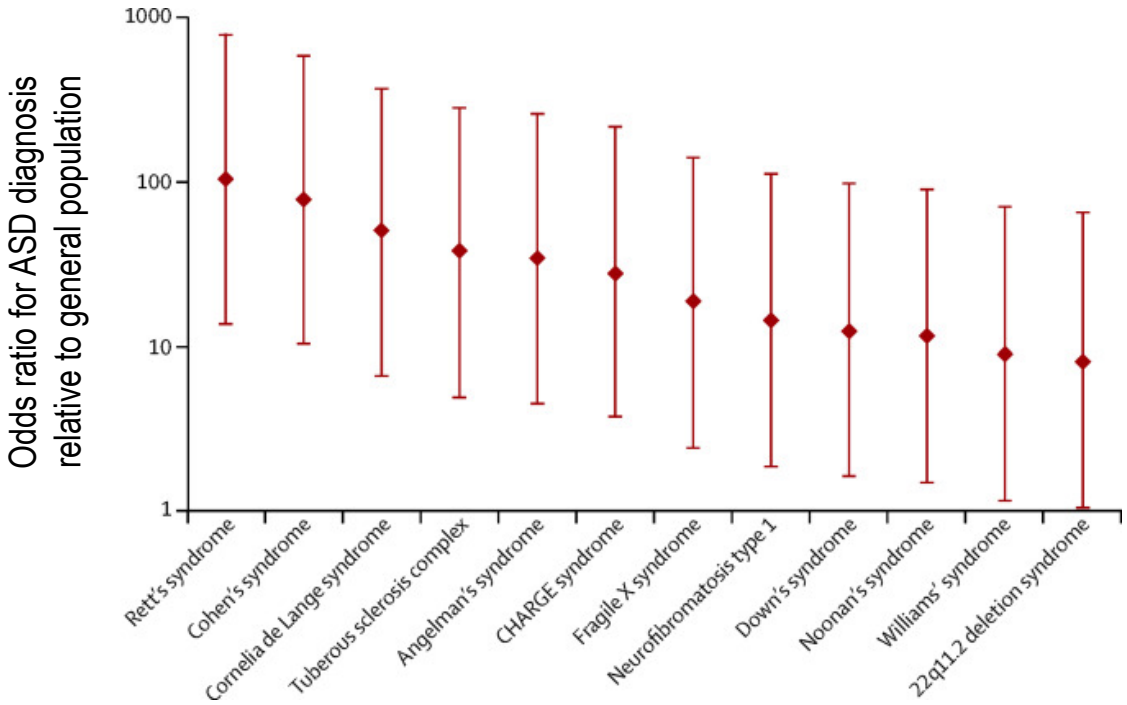
It is my recommendation that [redacted] receive ABA therapy to help advance her development. ABA was developed for children with developmental disabilities, including Autism Spectrum Disorder, but most recently has had the most publicity among children with autism. [redacted] has not been diagnosed with autism, though she has some characteristics of children with autism. During her most recent evaluation, [redacted] developmental delay was too great to determine if she also met criteria for autism. There is research evidence that ABA is also effective at advancing the development of all types of children (not just those with autism).

More information can be found at: https://www.kennedykrieger.org/sites/default/files/library/documents/patient-care/centers-and-programs/down-syndrome-clinic-and-research-center/aba-scientific-support-9-2015.pdf

Recent trends suggest diagnostic substitution



Large increase in ASD /decline in ID prevalence among students receiving special education services. Figure produced from data obtained from the U.S. Department of Education (2016).



Increasing rates of ASD features/diagnosis reported in genetic conditions previously considered to be associated with ID (e.g., Fragile X Syndrome and Williams Syndrome) (Polyak et al., 2015; Richards et al., 2015). Figure reproduced from Richards et al. (2015).

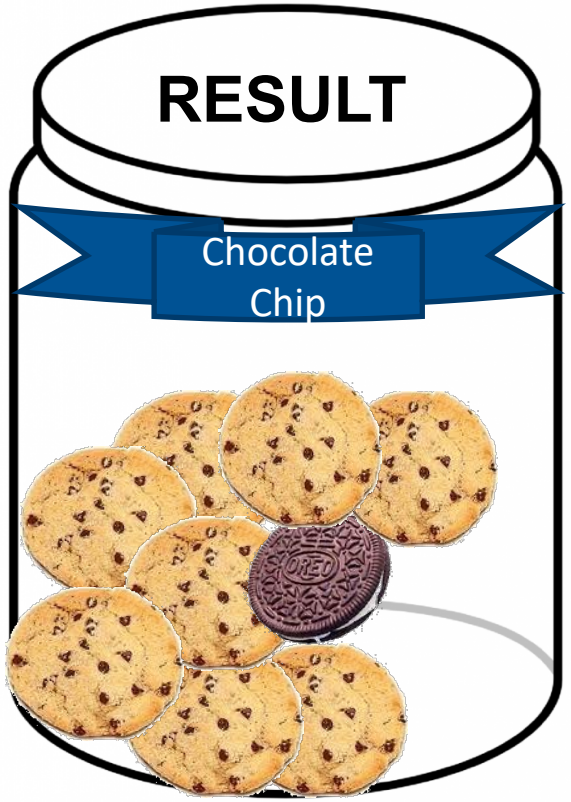
Psychometric principles used to evaluate behavioral measures

Broad concepts: Reliability and validity

Figure illustrates reliability and validity using the target metaphor. Copied from http://media.acc.qcc.cuny.edu/faculty/volchok/Measurement_Volchok/Measurement_Volchok6.html

Broad concepts: Sensitivity & specificity

- **Sensitivity** is proportion of actual positives who test positive
  - *Positive predictive value is probability that positive result is true*
- **Specificity** is proportion of actual negatives who test negative
  - *Negative predictive value is probability that negative result is true*



Sensitivity = .80

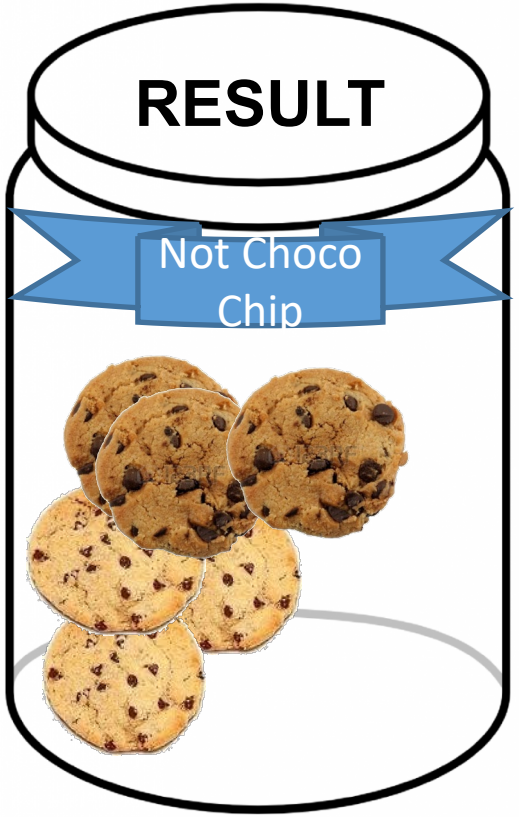
Correct Positives

8 /10 chocolate chip cookies in the chocolate chip jar

Specificity = .90

Correct Negatives

9 / 10 Oreos in the NOT chocolate chip jar



Sensitivity = .70

Correct Positives

7 /10 chocolate chip cookies in the chocolate chip jar

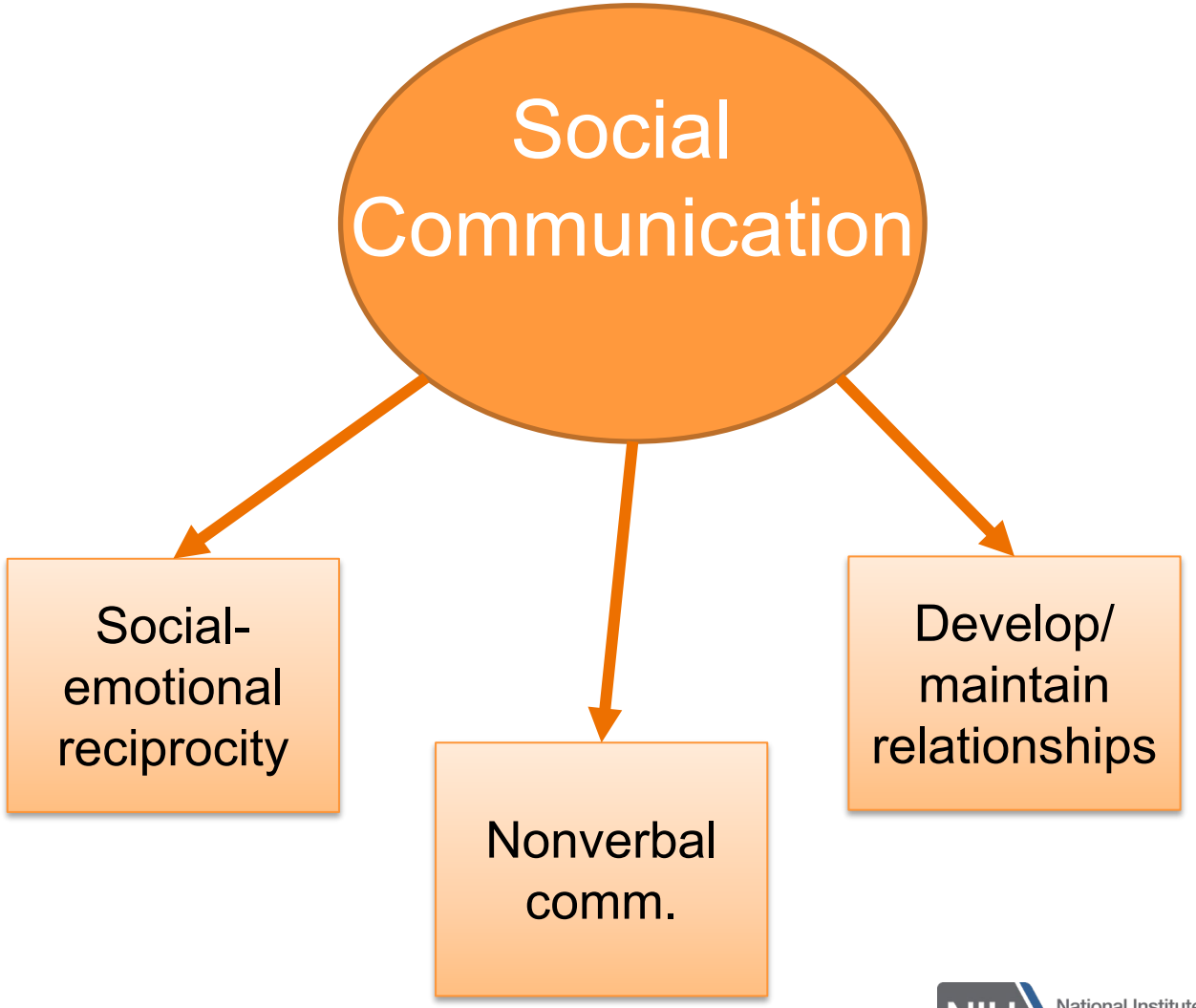
Specificity = .30

Correct Negatives

3 / 10 Carob Chip in the NOT chocolate chip jar

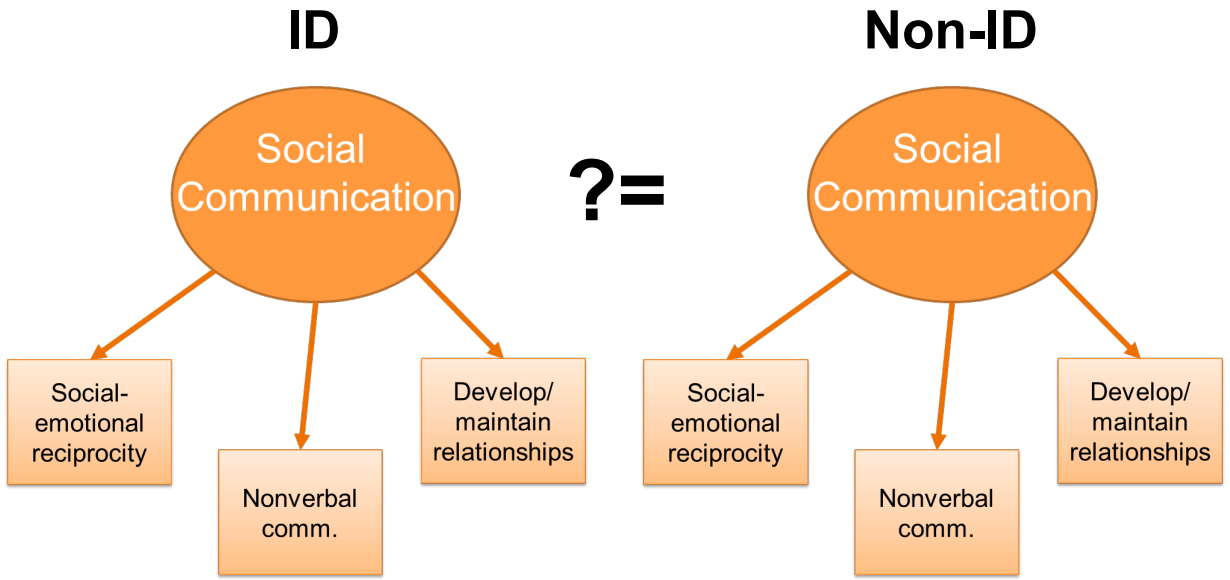
Deep concept: Measurement invariance

Behavioral measures predicated on idea that each item on a scale is caused by some construct which we cannot directly observe



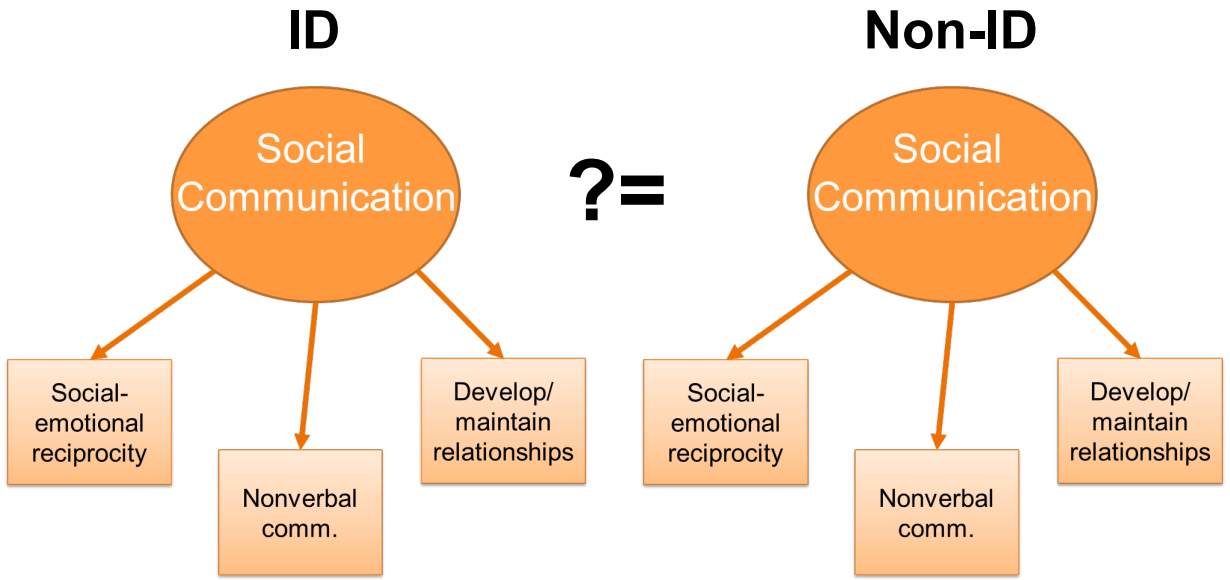
Deep concept: Measurement invariance

Invariance means that scores on the measure mean the same thing for different groups of participants



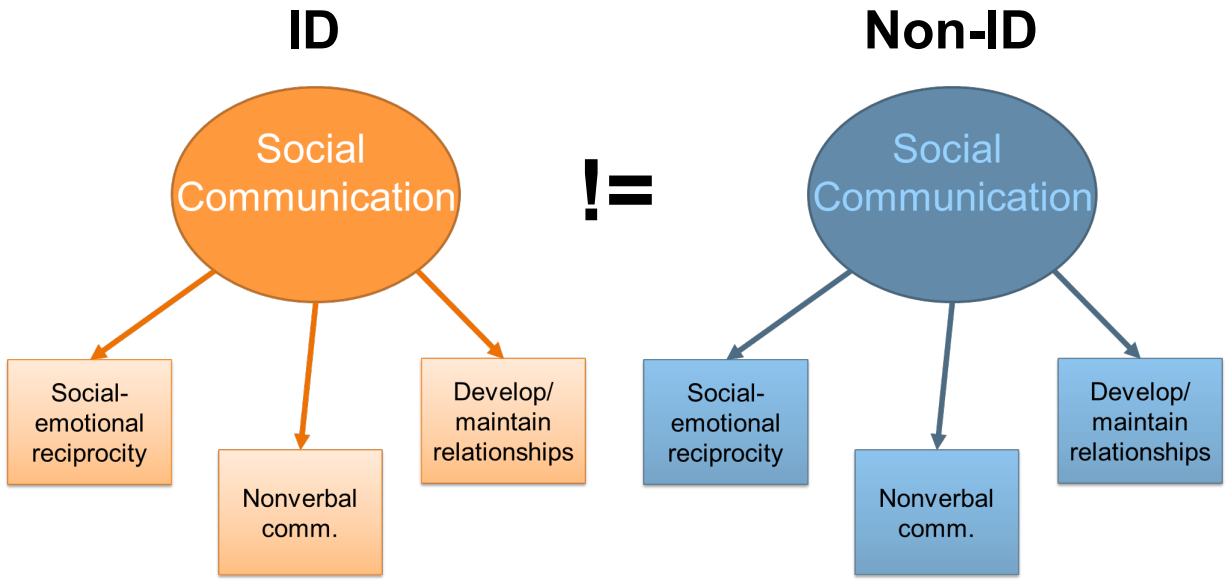
Deep concept: Measurement invariance

- **Configural:** Pattern of loadings on factors is similar between groups
- **Metric**: Strength of loadings is similar between groups
- **Scalar**: For same underlying level of construct, both groups receive same score



Deep concept: Measurement invariance

If invariance is not supported, scores cannot be compared between groups – this issue cannot be corrected by “controlling for” group



Points to remember

"Reliability" and "validity" are not monolithic; they are composed of many types of evidence. Evidence for some aspects of validity may be strong and for other aspects, weak.

Psychometric properties are not discrete; validity is not a "yes or no" question. Properties are supported by evidence which is characterized by its degree of strength.

Psychometric properties are not inherent to a scale. They vary based on the context of usage (e.g., population, setting), and must be generalized with caution. What is favorable (e.g., sensitivity more important than specificity) depends on the usage.

Psychometric limitations may not be obvious to the casual user. Threats to validity may be difficult to detect without careful study, and seemingly esoteric aspects of the psychometric profile may have significant consequences for the scientific value of data.

Tools used to measure ASD symptoms and their use in ID

Who is missing from test development?

People with ID historically excluded from ASD research generally (explicitly with IQ requirements or by proxy with genetic/ motoric/ comorbidity exclusions)

Sensory/motor impairment more common w/ID

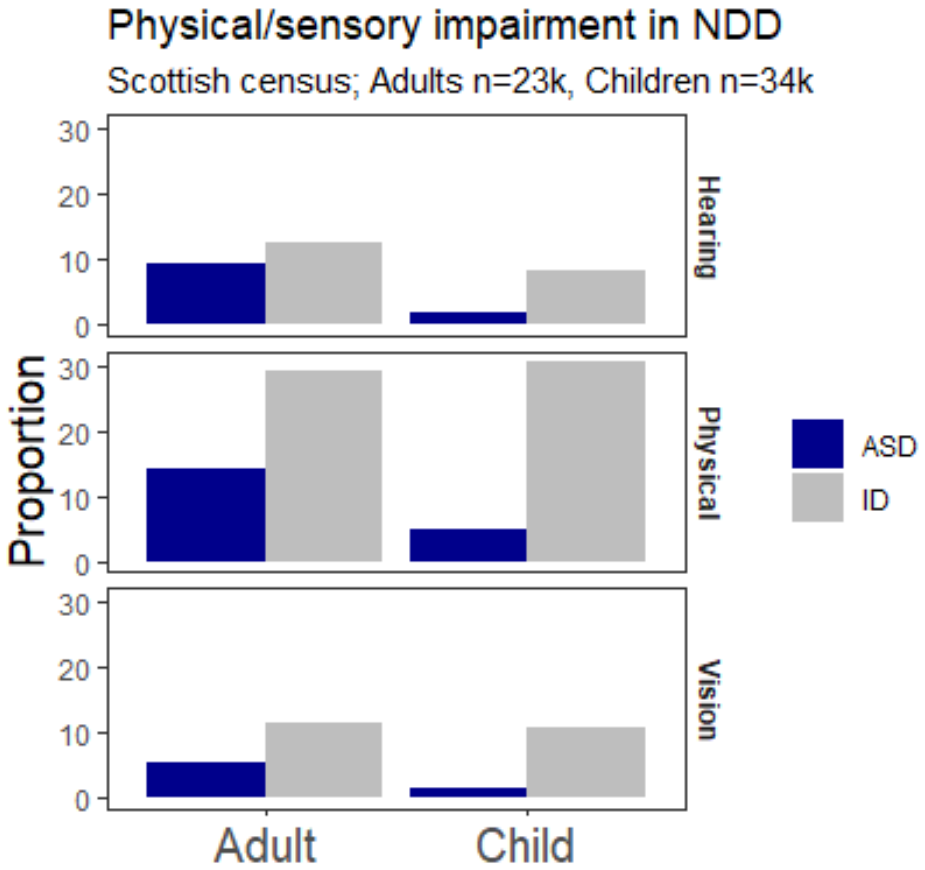


Figure shows relative low prevalence of sensory and physical impairment among people with ASD in a population-based study. Created from Kinnear et al. (2020). *Journal of Applied Research in Intellectual Disabilities.*

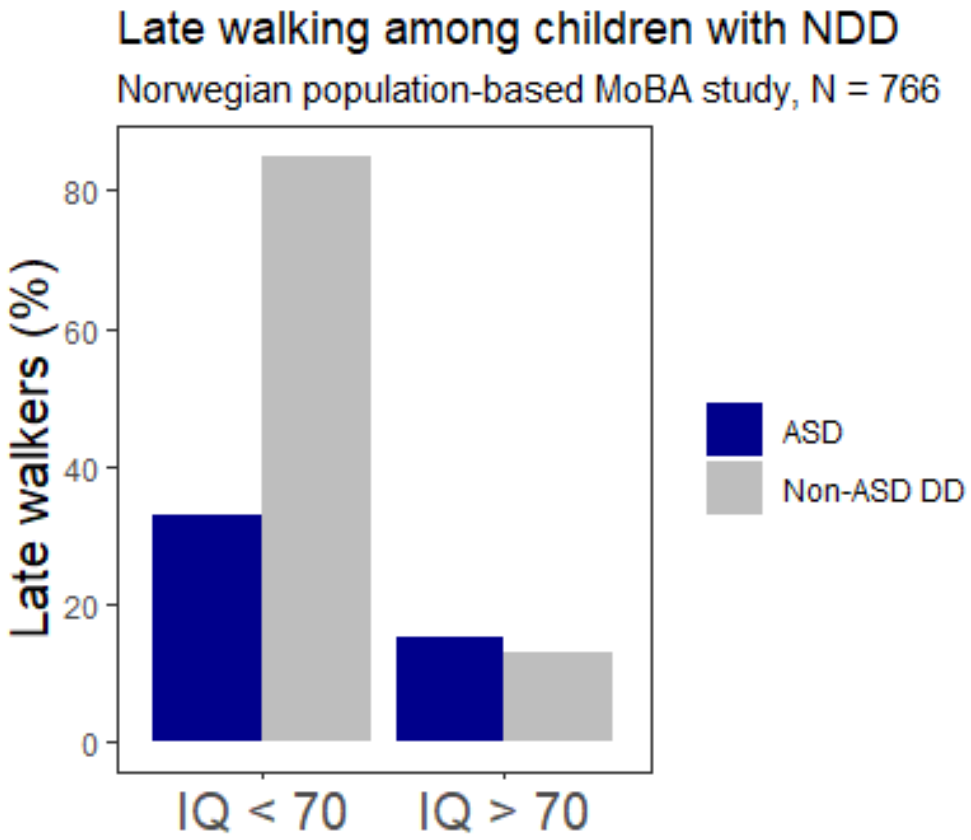


Figure shows preserved walking milestone among children with ASD and ID. Havdahl, Farmer…Bishop. (In press). *JCPP.* See also Bishop, Thurm, Farmer, & Lord. (2016). *Pediatrics* and Bishop, Farmer …Thurm. (2017). *Am J Psychiatry*.

Language impairment more common w/ID

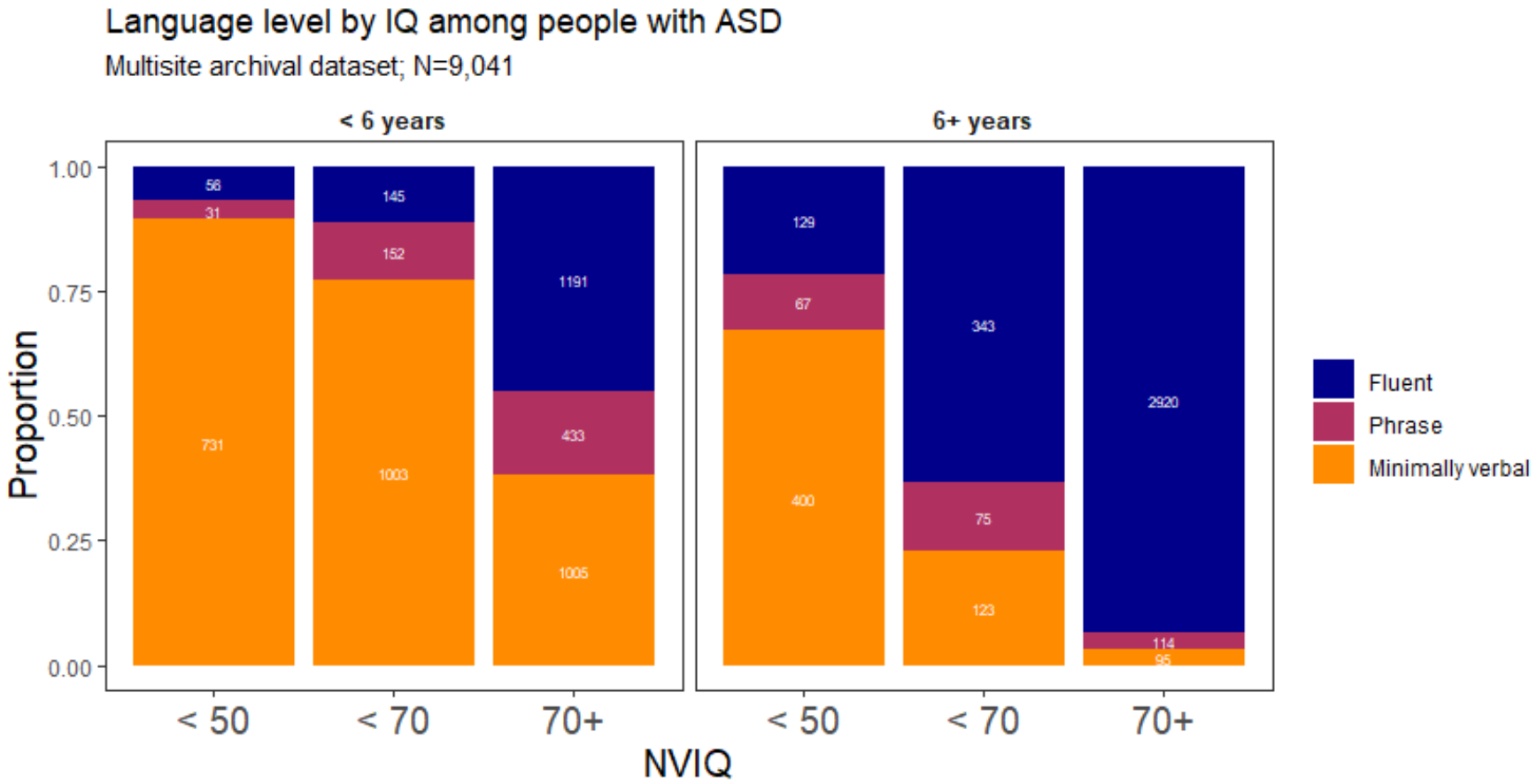


Figure illustrates lower rate of phrase speech, indexed by ADOS overall language level, among children with ASD and low IQ compared to children with ASD and average IQ (unpublished data from multisite integrative data analysis study shown with permission from Somer Bishop, PhD).

Common ASD diagnostic tools

- Autism Diagnostic Interview-Revised (ADI-R)
- Autism Diagnostic Observation Schedule (ADOS)
- Social Responsiveness Scale (SRS)

Even experts using the combined "gold standard measures" overestimate ASD in ID

Diagnostic Criteria		Sensitivity (95% CI)[c]	Specificity (95% CI)[c]
ADI-R and ADOS (AUT)	U.S.: 3+	82.0% (78–85)	86.0% (83–89)
	CAN: 3+	77.2% (70–83)	75.0% (60–86)
	U.S.: <36	80.9% (74–86)	87.0% (81–93)
	U.S.: MR	91.1% (83–98)	50.0% (31–75)

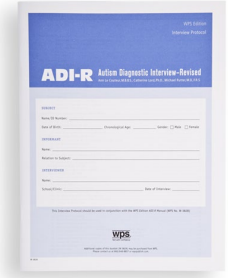
MR = child over 3y with NV mental age <= 18m

"The ADI-R and ADOS provide unique and critical overlapping information that **informs clinical judgments** in making an ASD diagnosis…Additional testing and alternative hypotheses must be considered when … the instrument results deviate from clinical impression."

Table illustrates low specificity of gold standard ASD diagnostic battery among people with ID. Excerpted from Table 2 of Risi, S., Lord, C., et al. (2006). JAACAP.

NIH National Institute of Mental Health

Common ASD diagnostic tools over-identify people with ID



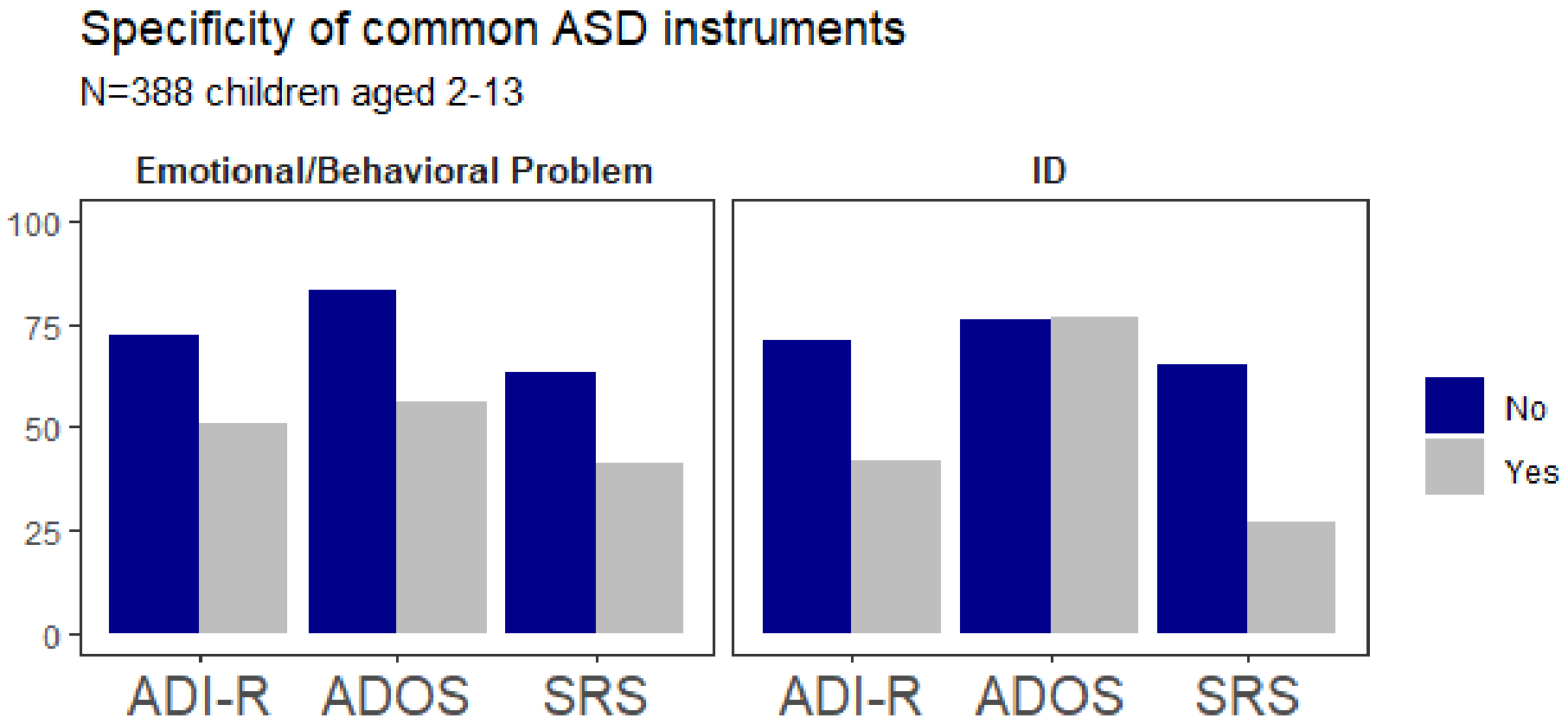


Figure illustrates reduced specificity of the Autism Diagnostic Interview-Revised, Autism Diagnostic Observation Schedule, and the Social Responsiveness Scale among children with emotional/behavioral problems (on CBCL) and those with intellectual disability. Produced from Havdahl et al. (2016) JAACAP.

Common ASD tools have different meaning among people with and without ID

SRS-2 is not invariant across people with and without ID (Sturm et al., 2017)

- Integrative archival analysis of 21k youth with ASD
- 49 of 65 SRS-2 items exhibited substantial Differential Item Functioning (DIF), based on age, sex, language, and/or IQ

Distribution of SRS-2 short form latent scores in Simons Simplex Collection

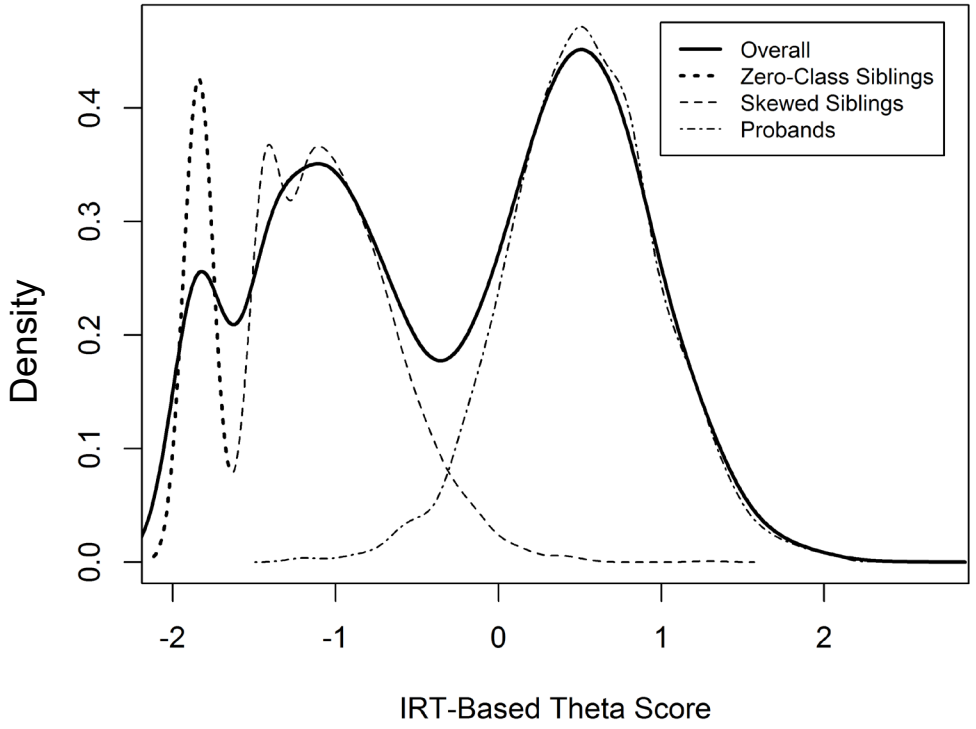


Figure illustrates non-normal distribution of latent construct measured by SRS total score, which suggests multiple populations and a more complex psychometric profile that has not been explored. Reproduced from Kaat & Farmer (2017) JCPP.

Putting it all together: Guidelines and best practice for diagnosing ASD in ID

Clinical guidelines for the assessment of ASD in ID

Separate basic social communication skills from more advanced skills that may be impaired across a range of neurodevelopmental disorders (incl. ID).

Social Communication Skills Most Specific to ASD

Early-emerging nonverbal behaviors (e.g., eye contact, facial expressions, gestures)

Shared enjoyment and participation in simple, back and forth games (e.g., peek-a-boo)

Clinical guidelines for the assessment of ASD in ID

What is the individual's cognitive ability and general developmental level?

Chronological age is also relevant, in that older individuals have had more opportunity to learn behaviors; may modify appearance of mental age.

Clinical guidelines for the assessment of ASD in ID

What are the individual's motor and sensory impairments (e.g., vision, hearing)?

Do these impairments better explain social-communication abilities?

Clinical guidelines for the assessment of ASD in ID

Is the clinician able to judge whether social impairments are commensurate with general developmental level?

Not Applicable ≠ Impairment

Clinical guidelines for the assessment of ASD in ID

Is the developmental trajectory consistent with ASD?

The sudden manifestation of social communication deficits and repetitive behaviors beyond the early childhood period is unlikely.

Clinical guidelines for the assessment of ASD in ID

Is it wise to administer a standardized assessment (e.g., ADI, ADOS, SRS)?

Interpretations of the data may be significantly limited by characteristics commonly observed in individuals with ID.

Manuals discourage use for <18 months mental age, or without clinical suspicion.

Clinical guidelines for the assessment of ASD in ID

Does the multidisciplinary team agree?

Clinicians with primary training in ASD may be biased towards ASD diagnosis. Complexity of ID presentation may require substantial consultation (medical problems, neurological issues, motor impairments, etc.).

Suggested guidelines for diagnosis of ASD in severe-to-profound ID

Avoid

- Sensory deficits or physical impairments disallow adequate assessment
- Temporary factors (seizure, recent general decline, onset of psychiatric changes) disallow adequate assessment
- Recent and continued development of necessary motor skills

Consider Waiting

- Child is <5 years and severity of ID is not established

Proceed

- ASD criteria clearly met, including Criterion E
- ASD symptoms seem to drive (lack of compliance) estimates of ID symptoms (IQ and adaptive behavior)

Summary

Take-home messages

- The accurate diagnosis of ASD in ID is important for treatment planning and prognosis.
- The valid measurement of ASD symptoms in the context of ID is a fundamental issue in parsing whether genetic conditions are truly associated with ASD.
- Psychometric properties are not inherent to the instrument, but instead correspond to a given population and setting. No ASD screener, diagnostic, or severity measure has favorable psychometric properties when used in ID.
- ASD diagnostic measures cannot be used without clinical judgment; clinical judgment requires understanding developmental status (e.g., severity of ID) and associated behavioral expectations. People make diagnoses, not scales.

Acknowledgements

Cristan.Farmer@NIH.gov

- Audrey Thurm, PhD
- Somer Bishop, PhD
- Cathy Lord, PhD
- Aaron Kaat, PhD