HHS Public Access
Author manuscript
Cogn Dev. Author manuscript; available in PMC 2022 April 01.

Published in final edited form as:
Cogn Dev. 2021 ; 58: . doi:10.1016/j.cogdev.2021.101038.

Prerequisite Skills in Cognitive Testing: Innovations in theory and recommendations for practice

Aaron J. Kaat, PhD[1], **Somer Bishop, PhD**[2], **Emma Condy, PhD**[3], **Nancy R. Sullivan, PhD**[4,5], **Latha Soorya, PhD**[6], **Audrey Thurm, PhD**[3]

[1]Department of Medical Social Sciences, Feinberg Medical Center, Northwestern University, Chicago, IL.

[2]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, San Francisco, CA.

[3]Neurodevelopmental and Behavioral Phenotyping Service, National Institute of Mental Health, Bethesda, MD.

[4]Division of Developmental Medicine, Boston Children's Hospital, Boston, MA

[5]Harvard Medical School, Boston, MA.

[6]Deparment of Psychiatry and Behavioral Sciences, Rush University Medical College, Chicago, IL

Abstract

Testing cognitive skill development is important for diagnostic, prognostic, and monitoring purposes, especially for young children and individuals with neurodevelopmental disorders. Developmental tests have been created for infants and toddlers, while traditional IQ tests are often employed beginning in the later preschool period. However, IQ tests rely on developmental skills that are rapidly changing during early childhood. Here, we introduce the idea of *prerequisite skills* in developmental domains, which are discrete skills required for, but not explicitly tested by, traditional IQ tests. Focusing on general cognition, particularly among children with a chronological or mental age under 4 years, may fail to capture important nuances in skill development. New skill-based assessments are needed in general, and in particular for categorization, which is foundational to higher-order cognitive skills. Novel measures quantifying categorization skills would provide a more sensitive measure of development for young children and older individuals with low developmental levels.

Corresponding author: Aaron J. Kaat, Ph.D. Department of Medical Social Sciences, 625 N. Michigan Avenue Suite 2700, Chicago IL, 60611. Phone: 312-503-3477. aaron.kaat@northwestern.edu.

Disclosure Statement
The authors report no conflict of interest.

Compliance with Ethical Standards
This theoretical review article does not include any human or animal experimentation.

Keywords

Categorization; Concept Formation; Cognitive Testing; Neurodevelopmental Disorders; Early Childhood

Introduction

Early childhood is a period of rapid development in cognitive skills. Quantifying cognitive skill development, however, can be a challenge. For individual differences, there has been an emphasis on general cognition (i.e., IQ or developmental quotient), as opposed to discrete skills. However, a limitation of this approach is that the prerequisite developmental skills required for achieving a basal score on a traditional IQ test are still developing during early childhood. Thus, while some IQ tests (e.g. the Stanford-Binet, Roid, 2003; or the Wechsler Preschool and Primary Scale of Intelligence, Wechsler, 2012) are normed down to 2 years of age, the validity of full-scale IQ and other composite scores and their relationship to future intellectual functioning is less meaningful. This is especially the case when performance on individual skills is discrepant within a child or if floor effects occur in one or more domains. Moreover, scores from these tests may be invalid for children with neurodevelopmental disabilities (NDD) or other delays in developmental skills that are explicitly tested, or implicitly required, as part of standard test administration.

Alternatively, there are developmental tests designed for children under 4 years that are less susceptible to floor effects. Developmental tests (e.g. Bayley & Reuner, 2006; Mullen, 1995) also have “cognitive” domains. While these tests cover broad “cognitive” domains, they also fail to comprehensively assess discrete cognitive skills that are rapidly changing during this period. Developmental tests, along with other assessment of specific relevant areas of concern (e.g. motor, language, attention, executive functions) may be used clinically to indicate areas of delay or concern, and assist in formulating a clinical formation and recommendations for intervention, but overall scores are insufficient when quantifying discrete aspects of cognitive development in an individual young child.

There is some support for use of developmental tests in predicting other discrete cognitive abilities in early childhood (Blaga et al., 2009; Lung et al., 2009), and scores remain correlated with later IQ (Girault et al., 2018; Howlin et al., 2014; Månsson et al., 2019). However, absolute scores on an individual level do not show adequate predictive validity (Jenni et al., 2015; O’Shea et al., 2018), especially below the age of 2 years (Girault et al., 2018; Hack et al., 2005; Månsson et al., 2019). This is likely due to the failure to assess the specific cognitive skills experiencing rapid change during this developmental period. It is important to have measures that can both predict later IQ on an individual level, and show potential gains (or declines) in the development of discrete cognitive skills. To address these limitations of developmental and IQ tests, we focus here on prerequisite cognitive skills that are rapidly developing during early childhood, and argue that these need to be assessed within an individual child to accurately quantify cognitive change.

The ability of a measure to accurately assess cognitive skills may be influenced by a child’s age and, for those with an NDD, the degree and type of delay(s) exhibited. Both traditional

IQ and developmental tests assume that other prerequisite skills are intact at the “expected” level for which the test or item set was developed. Thus, an item set designed to measure visual-spatial skills in 3 year olds would assume that skills in other developmental domains (e.g., language, motor) are at least at a 3 year old level. However, these skills may be unevenly developed (e.g. poor receptive language hindering comprehension of even the most basic test instructions; poor motor abilities limiting basic manual manipulation of materials). For example, children with low receptive language abilities may have sufficient visual-spatial abilities to differentiate distinct matrices, but not understand a direction such as “find the one that is missing.” These are the prerequisite skills that are likely more sensitive to developmental and intervention effects, and the development of which is necessary to be intact in order to successfully complete IQ tests.

As shown in Table 1, developmental tests assess basic and broad skills, while tasks on IQ tests require the integration of multiple skills; what is missing is an assessment of the narrow skills that are prerequisite to the more general cognition captured by an IQ test. To better characterize cognitive development in all children, there is a need for explicit assessments of these prerequisite developmental skills, including assessments that can track incremental changes over time.

Out of Age-Range Testing for Individuals with NDD

The cognitive testing gap described above is particularly evident in assessment of individuals with an NDD, especially intellectual disability. Due to an inability to establish a basal on a traditional IQ test (Munson et al., 2008), clinicians and researchers must “fall back” on tests designed for much younger children (Soorya, Leon, Trelles, & Thurm, 2018). Although use of developmental tests for out-of-age range testing appears to be a valuable approach (and often the only option), at least for obtaining estimates that have been found both to be predictive of future outcomes (Bishop, Farmer, & Thurm, 2015) and correlated with biomarkers of specific conditions (Thurm et al., 2016), there are measurement limitations, including score inflation or deflation, depending on which measure is chosen and what type of score is analyzed (Farmer, Golden, & Thurm, 2016; Farmer et al., 2020). Another problem with relying on developmental tests and/or out of age range testing practices arises when considering prediction of scores or changes of scores over time. The current hierarchy of test choices may result in individuals with NDD (both from “in-range” and “out-of-range” tests) ‘advancing’ from a developmental test to a traditional IQ test—accompanied by significant discontinuity in measurement, especially when the child fails to achieve a basal score on a traditional IQ tests (cf. Farmer et al., 2016; Hessl et al., 2009; Sansone et al., 2014).

In addition, while raw scores on IQ subtests may show sufficient variability within the low range of ability to capture potential change with treatment, converting these raw scores to standard scores poses a challenge. Low performance may be collapsed into only a very few number of standard score categories—or even just one lowest category—making it difficult to detect small (but meaningful) change using standard scores (Thompson et al., 2018). In order to measure true cognitive gains among individuals with low mental ages, tests must include a large enough item set of the discrete prerequisite skills that are actually changing.

Different types of skills will need to be distinguished to the greatest extent possible across tests to increase our ability to interpret changes in test scores.

Concept Formation as a Critical Prerequisite Skill

While there certainly are a host of important prerequisite skills required for successfully completing an IQ test (see Table 1 for examples), we focus on concept formation as a critical prerequisite skill within the categorization domain. Concept formation is critical because a) it is rapidly developing during early childhood; b) it is an explicit or implicit requirement for successful completion of items across multiple domains of traditional IQ tests; and c) young children or children with NDD may have relative strengths in it but specific problems with other discrete skills that may affect their performance on traditional measures of categorization.

Categorization begins to develop based on exemplars, which can be seen in infants as young as 3 months, albeit with looking paradigms that do not involve manipulating objects or finger pointing (Arterberry & Bornstein, 2002; Mareschal, French, & Quinn, 2000). It starts with identification of visual similarities and attentional processes directed toward these similarities (Sloutsky, 2003). Categorization based on perceptual similarities, especially shape or color, emerges early, and as a child develops, extends to categorization based on other features (e.g., size, quantity, motion, location). The ability to match categories, even during the preverbal stage (under 12 months) helps to prime infants for word learning (Pomiechowska & Gliga, 2019), and thus may be critical for advancing language development. Then, although categorization may continue to develop without commensurate language development (Joseph, Tager-Flusberg, & Lord, 2002), verbally-mediated labels become one of the most salient attributes in defining category boundaries (Fairchild, Mathis & Papafragou, 2018). This also highlights the need for assessing prerequisite skills individually, to minimize the negative effect that other skills (e.g., language delay) have on the assessment of categorization.

Typically by 18 months, categorization moves from stimuli with which the child has had experiential knowledge to include novel stimuli and hypothetical representations (Meltzoff, 1990). At that point, toddlers start to be able to make predictions about and classify stimuli they have not previously seen. Simultaneously, toddlers also improve in their ability to classify objects based on lessening inclusiveness in a category (Bornstein & Arterberry, 2010). By 3 years of age, children can categorize based on more abstract (semantic category) concepts (Bovet, Vauclair, & Blaye, 2005), which further contributes to language development (Jones & Smith, 2002; McClelland & Rogers, 2003). These semantic categories may be based on any number of similarities in both objects and actions (e.g., “animals” as a broad category, “dogs” as a more specific one; or “things that move” or “sleeping” as potential actions). Insofar as categorization skills are central to nonverbal cognitive processes, deficits in categorization are commonly observed in NDDs, including specific genetic syndromes (e.g., Down syndrome; Klinger & Dawson, 2001; Phillips et al., 2014), intellectual disability (Gligorović & Buha, 2013), and autism spectrum disorder (Klinger & Dawson, 2001).

Despite the importance of categorization to cognitive functioning, we lack a unifying measurement approach that quantifies the transition from the very basic perceptual category matching assessed in developmental tests to more advanced semantic category matching and sorting, without becoming either a test of language acquisition or cognitive flexibility. The key cognitive skill is *not* the number or content of categories known, but rather the ability to use knowledge about categories and attributes to identify which stimuli are most alike. We propose that direct assessment of concept formation may provide more valid and sensitive means of measuring change in this prerequisite cognitive skill.

Concept formation has been defined as "the search for and listing of attributes that can be used to distinguish exemplars from non-exemplars of various categories" (Bruner, Goodnow, & Austin, 1967). As mentioned above, identical match-to-sample skills emerge early in infancy. However, the ability to attend to the most relevant stimuli attributes with lessening salience and/or greater competing attributes—i.e., concept formation—is rarely directly assessed. Thus, concept formation bridges categorization abilities on developmental tests and subtests on traditional IQ tests (e.g., picture concepts, nonverbal reasoning). It is most common for IQ tests to synthesize concept formation with additional tasks in order to assess matrix reasoning (Curie et al., 2016), memory, set shifting (Zelazo, Carlson, & Kesek, 2008) and other executive function (EF) paradigms, or other "higher-order" cognitive abilities, including verbal questions about how items are similar, as opposed to concept formation on its own.

There is a need to develop and norm measures that would show variability in skills such as concept formation among older children with NDDs and in younger children who are developing these skills as expected for their age. Such a test could not, and should not, replace the use of comprehensive IQ (or developmental) tests, but could be used for research purposes when evaluating specific cognitive skills, and as a separate test to augment more comprehensive neurodevelopmental assessments. This is similar to how vocabulary and motor tests are often added to neurodevelopmental assessments when there is a clinical question of whether these skills may be relative (or absolute) strengths or weaknesses. In addition, tests of specific prerequisite skills such as concept formation could be used to determine which more comprehensive test may be most appropriate—especially for young children (and older children with NDDs) where cognitive delays or deficits are strongly suspected. A certain level of skill on such a test may indicate if the child is ready to progress (i.e., will likely receive a score above the floor) to an IQ test where the prerequisite skill is required (as shown in Table 1).

New concept formation measures must also provide solutions to the limitations of existing measures. When prerequisite motor and language skills are not achieved, existing measures fail. For instance, in populations with minimal verbal abilities (Kasari et al., 2013), there is often poor receptive language, hindering the comprehension of even the most basic test instructions. In genetic conditions associated with NDD such as Rett Syndrome, wherein both verbal abilities and fine motor skills for basic manual manipulation of materials are limited, the tests need to be modified further. For these reasons, even tests that are explicitly developed to be "nonverbal" (e.g., the Leiter International Performance Scale, Roid et al., 2013) may not be suitable if they rely on understanding of complex gestures or require motor precision involving manipulating small cards or objects. New paradigms for assessing concept formation may need to simplify or allow modification of typical response modalities, such as tablet based responses (versus picture card) or eye tracking, which appears successful for administration of basic categorization tasks such as simple shape matching (Clarkson et al., 2017).

Strategies for Testing Concept Formation and other Prerequisite Skills

Requirements for a Prerequisite Skills Test.

In order to develop a useful direct assessment which would fill the gaps described above, several requirements must be met. Such measures must: 1) be standardized in both chronologically young children (less than 2 years) and older individuals with an NDD (based on the existing gaps this would include a chronological age above 5 years but mental age below 4 years), 2) be responsive to change—that is, to allow growth to be shown even when children are significantly cognitively delayed, 3) minimize the requirement of integration of multiple domains of prerequisite skills (i.e., minimizing motor precision on non-motor tasks, and 4) be motivating based on use of state-of-the-art technology.

Standardization.—A common strategy for norming cognitive tests is to use regression-based norms (potentially with spline-based or other polynomial terms; Zachary & Gorsuch, 1985). In this way, lower performance is observed within a domain, but likely at a younger age. Most norming studies exclude or only minimally-represent individuals with an NDD, and thus standard scores within the intellectual disability range (IQ < 70) are based off of comparisons to performance of much younger individuals. As opposed to this status quo, standardization of tests focused on prerequisite cognitive skills must explicitly sample groups of children with NDD and cognitive deficits who may be appropriate for this test, as well as typically developing children who exhibit age-appropriate skills in the domains included in the test. This approach would allow observed scores across the ability range for all individuals for whom the test is appropriate. While oversampling individuals with NDD would not be appropriate for age-based norms, a second goal (described next) would be measuring responsiveness to change. In this way, oversampling allows better definition of change-sensitive scores (Farmer et al., 2020) for the intended purpose of testing.

Responsiveness to Change.—The ability of measures such as those targeting prerequisite skills to be responsive to change is largely based on the psychometric properties used in their development. Interpretation of change on standardized tests is obfuscated by the norms used to interpret them. On the one hand, while increases or decreases in raw scores suggest more or less ability, raw scores are not on an interval-scale of measurement and thus pose analytic problems. On the other hand, standard scores on both developmental and IQ tests *should not* change over time in the absence of some intervention or injury. Increases or decreases in standard scores suggest a different developmental trajectory than what would be expected for same-age peers. For this reason, change-sensitive scores have been developed for many tests (Farmer et al., 2020).

Insofar as development of prerequisite skills for IQ tests occurs rapidly within a relatively short time frame for typically-developing children, the ability to detect this learning is of great importance. Unlike traditional developmental tests or IQ tests, tests of prerequisite skills should be able to be administered at a higher frequency without being burdened by practice effects. Alternative forms or computer adaptive testing (CAT) administration may help mitigate these concerns. CAT in particular holds promise, as it can target the appropriate skill difficulty for an individual, regardless of chronological age. The item selection algorithm would need to ensure representation of a wide variety of specific categories at each level, though, given that the focus would be on whether the child has learned the skill of attending to the relevant attributes of an item in the presence of decreasing salience. The acquisition of a specific category is less relevant—humans are very good at categorization, as discussed above—the important skill to assess is whether the child can perform a task such as matching based on identifying the attributes of objects in common. In repeated assessment, then, CAT scoring procedures must be able to indicate both 1) when real change—that is, acquisition more than would be expected by chance or by any lingering practice effect—has occurred, and 2) when the rate of change is faster, slower, or commensurate with age-based expectations.

Minimal input of unrelated transitional skills.—Modifications to traditional standardized IQ test procedures are often required for individuals with discrepant/uneven skill profiles in order to indicate responses, such as when traditional means (e.g., pointing, verbal responses) are not possible (Thompson et al., 2018; Warschausky et al., 2012). Therefore, in a test specifically designed to assess concept formation, it must minimize interference from other prerequisite skills (e.g., complex verbal language or motoric coordination).

For individuals with significant motor impairments, reduction in the required input of unrelated transitional motor skills is necessary. Designing assessments to allow multiple response modalities, such as including a touch/scan response (Thompson et al., 2018) would be beneficial in minimizing the interference of unrelated motor deficits. Additionally, standardized tests that have been designed specifically for computer-based assessment use have also been modified to accommodate individuals with disabilities (Magasi et al., 2017). While many still require manual manipulation of some sort, the use of eye tracking as the output to determine the target answer is also now being explored (Chard, Roulin, & Bouvard, 2014; Tager-Flusberg et al., 2017).

In addition to motor demands, when testing both young children with delayed and/or uneven language skills, and certainly for older individuals with an NDD, it is also imperative that complex receptive language demands be minimized. Tests that are truly nonverbal in nature exist, and include use of pantomime or imitation in teaching trials for understanding of task directions (Roid et al., 2013). Modifications such as these will need to be embraced in the development of an assessment of concept formation.

Motivating Technology for Testing Transitional Skills.—As implied in the other requirements, the optimal test modality for concept formation will likely need to utilize technology. Tablets or other screen-based systems for administration and scoring may be beneficial. Not only will this allow for CAT-administration but also minimize motor and receptive language demands if the appropriate user interface is utilized. Tablet-based or other computerized assessment increases the ability to test special populations, such as infants and toddlers, and individuals with physical or neurodevelopmental disabilities, among others (Hessl et al., 2016; Tulsky & Heinemann, 2017; Twomey et al., 2018). Tablet-based testing is growing in use in a variety of populations (Raiford et al., 2014; Rentz et al., 2016; Twomey et al., 2018), along with its use in preschool educational setting, including preschool children with NDD (Chmiliar, 2017). Additionally, children and individuals with NDD may be motivated to use electronic devices for other reasons. Technology is expected to increase motivation and compliance with testing (Piaw, 2012), since use of tablets appears to be intrinsically rewarding for many young children and use of computer assistance in teaching children with NDD to perform visual matching tasks has already been shown to be effective (Hu et al., 2019).

Already tablet-based assessment is becoming common. The Wechsler tests have been adapted for tablet-based administration (both for examiner- and examinee-facing stimuli; Noland, 2017). Other tests were developed specifically for electronic administration, such as the NIH Toolbox (Gershon et al., 2013) and CANTAB (Fray & Robbins, 1996). Electronics and smart devices are becoming ubiquitous in modern society. Even among infants and toddlers, it is possible to use tablets to assess cognitive skills (Twomey et al., 2018). Electronic administration also necessitates a higher consistency and uniformity in administration, minimizing administrator effects when paradigms are implemented in different research or clinical settings.

Conclusions

Developmental tests and IQ tests use different conceptual frameworks and test different abilities. Part of this is due to the active development of individual skills during early childhood. Scores on developmental tests are poor predictors of later intelligence (Hack et al., 2005; Månsson et al., 2019). IQ tests assume that the skills measured by a developmental test have been mastered. However, prerequisite developmental skills such as concept formation bridge developmental and IQ tests. New tests are necessary to adequately evaluate these skills. This is especially true for concept formation, given that complex categorization rules are necessary for most domains on a traditional IQ test, but only the most basic level of matching is assessed on developmental tests. Creating tests which assess skills required to even obtain a basal on a traditional IQ test will benefit both younger children and older individuals with disabilities. This has a clear practical benefit of increasing interpretability of scores for young children and those with intellectual disability or other NDDs. New measures should be more sensitive to change and thus capture meaningful differences in a time of rapid transition.

Acknowledgements:

Research reported in this manuscript was partially supported by the National Institutes of Health under grant 1U24OD023319 and through the Intramural Research Program ZICMH002961. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Arterberry ME, & Bornstein MH (2002). Variability and its sources in infant categorization. Infant Behavior and Development, 25(4), 515–528.

Bayley N, & Reuner G (2006). Bayley Scales of Infant and Toddler Development: Bayley-III (Vol. 7): Harcourt Assessment, Psych. Corporation San Antonio, Tex, USA.

Bishop SL, Farmer C, & Thurm A (2015). Measurement of nonverbal IQ in autism spectrum disorder: scores in young adulthood compared to early childhood. Journal of Autism and Developmental Disorders, 45(4), 966–974. doi:10.1007/s10803-014-2250-3 [PubMed: 25239176]

Blaga OM, Shaddy DJ, Anderson CJ, Kannass KN, Little TD, & Colombo J (2009). Structure and continuity of intellectual development in early childhood. Intelligence, 37(1), 106–113. doi:10.1016/j.intell.2008.09.003 [PubMed: 20046219]

Bornstein MH, & Arterberry ME (2010). The development of object categorization in young children: hierarchical inclusiveness, age, perceptual attribute, and group versus individual analyses. Developmental Psychology, 46(2), 350–365. doi:10.1037/a0018411 [PubMed: 20210495]

Bovet D, Vauclair J, & Blaye A (2005). Categorization and abstraction abilities in 3-year-old children: a comparison with monkey data. Animal Cognition, 8(1), 53–59. doi:10.1007/s10071-004-0226-y [PubMed: 15300466]

Bruner J, Goodnow J, & Austin G (1967). A study of thinking New York: Science Editions.

Chard M, Roulin JL, & Bouvard M (2014). Visual habituation paradigm with adults with profound intellectual and multiple disabilities: a new way for cognitive assessment? Journal of Applied Research in Intellectual Disability, 27(5), 481–488.

Chmiliar L (2017). Improving learning outcomes: The iPad and preschool children with disabilities. Frontiers in Psychology, 8, 660. doi: 10.3389/fpsyg.2017.00660 [PubMed: 28529493]

Clarkson T, LeBlanc J, DeGregorio G, Vogel-Farley V, Barnes K, Kaufmann WE, & Nelson CA (2017). Adapting the Mullen Scales of Early Learning for a Standardized Measure of Development in Children With Rett Syndrome. Intellectual and Developmental Disabilities, 55(6), 419–431. [PubMed: 29194024]

Curie A, Brun A, Cheylus A, Reboul A, Nazir T, Bussy G, … des Portes V (2016). A Novel Analog Reasoning Paradigm: New Insights in Intellectually Disabled Patients. PLoS ONE, 11(2), e0149717. doi:10.1371/journal.pone.0149717 [PubMed: 26918704]

Fairchild S, Mathis A, & Papafragou A (2018). Linguistic cues are privileged over non-linguistic cues in young children’s categorization. Cognitive Development, 48, 167–175. doi: 10.1016/j.cogdev.2018.08.007

Farmer C, Golden C, & Thurm A (2016). Concurrent validity of the differential ability scales, second edition with the Mullen Scales of Early Learning in young children with and without neurodevelopmental disorders. Child Neuropsychology, 22(5), 556–569. doi:10.1080/09297049.2015.1020775 [PubMed: 25833070]

Farmer C, Kaat A, Thurm A, Anselm I, Akshoomoff N, Bennett A, … Miller J (2020). Person ability scores as an alternative to norm-referenced scores as outcome measures in studies of neurodevelopmental disorders. American Journal of Intellectual and Developmental Disabilities, 125(6), 475–480.

Fray PJ, & Robbins TW (1996). CANTAB battery: proposed utility in neurotoxicology. Neurotoxicology and Teratology, 18(4), 499–504. [PubMed: 8866544]

Gershon RC, Wagster MV, Hendrie HC, Fox N, Cook KF, & Nowinski CJ (2013). NIH Toolbox for Assessment of Neurological and Behavioral Function. Neurology, 80(11 Supplement 3), S2–S6. doi:10.1212/WNL.0b013e3182872e5f [PubMed: 23479538]

Gligorović M & Buha N (2013) Conceptual abilities of children with mild intellectual disability: Analysis of Wisconsin Card Sorting Test performance, Journal of Intellectual & Developmental Disability, 38(2), 134–140, DOI: 10.3109/13668250.2013.772956 [PubMed: 23510029]

Girault JB, Langworthy BW, Goldman BD, Stephens RL, Cornea E, Reznick JS, … Gilmore JH (2018). The Predictive Value of Developmental Assessments at 1 and 2 for Intelligence Quotients at 6. Intelligence, 68, 58–65. doi:10.1016/j.intell.2018.03.003 [PubMed: 30270948]

Hack M, Taylor HG, Drotar D, Schluchter M, Cartar L, Wilson-Costello D, … Morrow M (2005). Poor Predictive Validity of the Bayley Scales of Infant Development for Cognitive Function of Extremely Low Birth Weight Children at School Age. Pediatrics, 116(2), 333–341. doi:10.1542/peds.2005-0173 [PubMed: 16061586]

Hessl D, Nguyen DV, Green C, Chavez A, Tassone F, Hagerman RJ, … Hall S (2009). A solution to limitations of cognitive testing in children with intellectual disabilities: the case of fragile X syndrome. Journal of Neurodevelopmental Disorders, 1(1), 33–45. doi:10.1007/s11689-008-9001-8 [PubMed: 19865612]

Hessl D, Sansone SM, Berry-Kravis E, Riley K, Widaman KF, Abbeduto L, … Gershon RC (2016). The NIH Toolbox Cognitive Battery for intellectual disabilities: three preliminary studies and future directions. Journal of Neurodevelopmental Disorders, 8(1), 35. doi:10.1186/s11689-016-9167-4 [PubMed: 27602170]

Howlin P, Savage S, Moss P, Tempier A, & Rutter M (2014). Cognitive and language skills in adults with autism: a 40-year follow-up. Journal of Child Psychology and Psychiatry, 55(1), 49–58. doi:10.1111/jcpp.12115 [PubMed: 23848399]

Hu X, Lee GT, Tsai Y-T, Yang Y, & Cai S (2019). Comparing Computer-Assisted and Teacher-Implemented Visual Matching Instruction for Children with ASD and/or Other DD. Journal of Autism and Developmental Disorders. doi:10.1007/s10803-019-03978-2

Jenni OG, Fintelmann S, Caflisch J, Latal B, Rousson V, & Chaouch A (2015). Stability of cognitive performance in children with mild intellectual disability. Developmental Medicine and Child Neurology, 57(5), 463–469. doi:10.1111/dmcn.12620 [PubMed: 25363202]

Jones SS, & Smith LB (2002). How children know the relevant properties for generalizing object names. Developmental Science, 5(2), 219–232.

Joseph RM, Tager-Flusberg H, & Lord C (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. Journal of Child Psychology and Psychiatry, 43(6), 807–821. doi:10.1111/1469-7610.00092 [PubMed: 12236615]

Kasari C, Brady N, Lord C, & Tager-Flusberg H (2013). Assessing the Minimally Verbal School-Aged Child With Autism Spectrum Disorder. Autism Research, 6(6), 479–493. doi:10.1002/aur.1334 [PubMed: 24353165]

Klinger L, & Dawson G (2001). Prototype formation in autism. Development and Psychopathology, 13(1), 111–124. doi:10.1017/S0954579401001080 [PubMed: 11346046]

Lung F-W, Shu B-C, Chiang T-L, Chen P-F, & Lin L-L (2009). Predictive validity of Bayley scale in language development of children at 6–36 months. Pediatrics International, 51(5), 666–669. doi:10.1111/j.1442-200X.2009.02844.x [PubMed: 19419503]

Magasi S, Harniss M, Tulsky DS, Cohen ML, Heaton RK, & Heinemann AW (2017). Test accommodations for individuals with neurological conditions completing the NIH Toolbox—Cognition Battery: An evaluation of frequency and appropriateness. Rehabilitation Psychology, 62(4), 455. [PubMed: 29265866]

Månsson J, Stjernqvist K, Serenius F, Ådén U, & Källén K (2019). Agreement between Bayley-III measurements and WISC-IV measurements in typically developing children. Journal of Psychoeducational Assessment, 37(5), 603–616.

Mareschal D, French RM, & Quinn PC (2000). A connectionist account of asymmetric category learning in early infancy. Developmental Psychology, 36(5), 635–645. [PubMed: 10976603]

McClelland JL, & Rogers TT (2003). The parallel distributed processing approach to semantic cognition. Nature Reviews Neurosciences, 4(4), 310–322. doi:10.1038/nrn1076

Meltzoff AN (1990). Towards a developmental cognitive science. The implications of cross-modal matching and imitation for the development of representation and memory in infancy. Annals of the New York Academy of Sciences, 608, 1–31; discussion 31–37.

Mullen EM (Ed.) (1995). Mullen Scales of Early Learning. Circle Pines, MN American Guidance Service.

Munson J, Dawson G, Sterling L, Beauchaine T, Zhou A, Elizabeth K, … Abbott R (2008). Evidence for latent classes of IQ in young children with autism spectrum disorder. American Journal of Mental Retardation, 113(6), 439–452. doi:10.1352/2008.113:439-452 [PubMed: 19127655]

Noland RM (2017). Intelligence testing using a tablet computer: Experiences with using Q-interactive. Training and Education in Professional Psychology, 11(3), 156–163. doi:10.1037/tep0000149

O'Shea TM, Joseph RM, Allred EN, Taylor HG, Leviton A, Heeren T, … Kuban KCK (2018). Accuracy of the Bayley-II mental development index at 2 years as a predictor of cognitive impairment at school age among children born extremely preterm. Journal of Perinatology, 38(7), 908–916. doi:10.1038/s41372-017-0020-8 [PubMed: 29808002]

Phillips BA, Conners FA, Merrill E, & Klinger MR (2014). Rule-based category learning in Down syndrome. American Journal of Intellectual and Developmental Disabilities, 119(3), 220–234. doi:10.1352/1944-7558-119.3.220

Piaw CY (2012). Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong? Procedia-Social and Behavioral Sciences, 64, 655–664.

Pomiechowska B, & Gliga T (2019). Lexical Acquisition Through Category Matching: 12-Month-Old Infants Associate Words to Visual Categories. Psychological Science, 30(2), 288–299. doi:10.1177/0956797618817506 [PubMed: 30575444]

Raiford SE, Holdnack J, Drozdick L, & Zhang O (2014). Q-interactive® Special Group Studies: The WISC®–V and Children with Intellectual Giftedness and Intellectual Disability.

Rentz DM, Dekhtyar M, Sherman J, Burnham S, Blacker D, Aghjayan SL, … Chenhall T (2016). The feasibility of at-home iPad cognitive testing for use in clinical trials. The Journal of Prevention of Alzheimer's disease, 3(1), 8.

Roid GH (2003). Standford-Binet Intelligence Scales, fifth Edition. Itasca, IL: Riverside Publishing.

Roid GH, Miller LJ, Pomplun M, & Koch C (2013). Leiter-3. International Performance Scale-hird Edition. Wood Dale: Stoelting.

Sansone SM, Schneider A, Bickel E, Berry-Kravis E, Prescott C, & Hessl D (2014). Improving IQ measurement in intellectual disabilities using true deviation from population norms. Journal of Neurodevelopmental Disorders, 6(1), 16. doi:10.1186/1866-1955-6-16 [PubMed: 26491488]

Sloutsky VM (2003). The role of similarity in the development of categorization. Trends in Cognitive Sciences, 7(6), 246–251. doi:10.1016/S1364-6613(03)00109-8 [PubMed: 12804690]

Soorya L, Leon J, Trelles MP, & Thurm A (2018). Framework for assessing individuals with rare genetic disorders associated with profound intellectual and multiple disabilities (PIMD): the example of Phelan McDermid Syndrome. Clinical Neuropsychology, 32(7), 1226–1255. doi:10.1080/13854046.2017.1413211

Tager-Flusberg H, Plesa Skwerer D, Joseph RM, Brukilacchio B, Decker J, Eggleston B, … Yoder A (2017). Conducting research with minimally verbal participants with autism spectrum disorder. Autism, 21(7), 852–861. doi:10.1177/1362361316654605 [PubMed: 27354431]

Thompson T, Coleman JM, Riley K, Snider LA, Howard LJ, Sansone SM, & Hessl D (2018). Standardized Assessment Accommodations for Individuals with Intellectual Disability. Contemporary School Psychology, 22(4), 443–457. doi:10.1007/s40688-018-0171-4 [PubMed: 30420939]

Thurm A, Tierney E, Farmer C, Albert P, Joseph L, Swedo S, … Porter FD (2016). Development, behavior, and biomarker characterization of Smith-Lemli-Opitz syndrome: an update. Journal of Neurodevelopmental Disorders, 8, 12. doi:10.1186/s11689-016-9145-x [PubMed: 27053961]

Tulsky DS, & Heinemann AW (2017). The clinical utility and construct validity of the NIH Toolbox Cognition Battery (NIHTB-CB) in individuals with disabilities. Rehabilitation Psychology, 62(4), 409–412. doi:10.1037/rep0000201 [PubMed: 29265861]

Twomey DM, Wrigley C, Ahearne C, Murphy R, De Haan M, Marlow N, & Murray DM (2018). Feasibility of using touch screen technology for early cognitive assessment in children. Archives of Disease in Childhood. doi:10.1136/archdischild-2017-314010

Warschausky S, Van Tubbergen M, Asbell S, Kaufman J, Ayyangar R, & Donders J (2012). Modified Test Administration Using Assistive Technology:Preliminary Psychometric Findings. Assessment, 19(4), 472–479. doi:10.1177/1073191111402458 [PubMed: 21467093]

Wechsler D (2012). Wechsler preschool and primary scale of intelligence—fourth edition: The Psychological Corporation San Antonio, TX.

Zachary RA, & Gorsuch RL (1985). Continuous norming: implications for the WAIS-R. Journal of Clinical Psychology, 41(1), 86–94. doi:10.1002/1097-4679(198501)41:1<86::aid-jclp2270410115>3.0.co;2-w [PubMed: 3973045]

Zelazo PD, Carlson SM, & Kesek A (2008). Development of executive function in childhood. In Nelson CA & Luciana M (Eds.), Handbook of Developmental Cognitive Neuroscience (2nd ed., pp. 553–574). Cambridge, Mass.: MIT Press.

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- Traditional IQ tests require prerequisite skills, which they do not assess.
- There should be assessments to capture the rapid development of these skills.
- Categorization by perceptual or semantic attributes is a key prerequisite skill.
- Categorization is also foundational for numerous higher-order cognitive skills.

Table 1:

Example Skills at Each Level

	Tasks common to developmental tests	**Discrete skills prerequisite to IQ tests**	**Tasks common to IQ tests**
Language	One-word receptive vocabulary	Following 3-step unrelated instructions	Verbal analogies; Comprehension
Motor	Gross and fine motor milestones	Hand-eye and bilateral coordination for object manipulation	Visuospatial processing (e.g., block design)
Categorization	Identical match-to-sample	Sorting object by where they are most likely to be found (i.e., concept formation)	Nonverbal Reasoning Semantic classification; Working memory; Matrices; Analogies (verbal or nonverbal)