Northwestern EC REACH
EARLY CHILDHOOD RESEARCH ALLIANCE OF CHICAGO

Northwestern | INSTITUTE FOR POLICY RESEARCH

Research Report
February 2026

A CLOSE LOOK AT

# The Kindergarten Individual Developmental Survey (KIDS)

By John Q. Easton with Sophie Magro, Elliot Regenstein, and Terri Sabol
Special Foreword by Diana Rauner and Robin Steans
Special Commentary from James Pellegrino

*Lovett Elementary School, completed renovation*

# TABLE OF CONTENTS



## About the cover photo

This is Lovett Elementary School, located at 6333 West Bloomington Avenue in Chicago. It is one of several elementary schools in the CPS "Dever" architectural design style, named for Mayor William Dever who served from 1923 to 1927. One of the most prominent features of "Dever-style" schools is the large kindergarten room, with its bay window and location in the middle of the first floor, seen in the picture. In addition to being very spacious, the kindergarten room has a partitioned off space for children's hats, coats, and winter boots. There are two bathrooms as part of this comfortable and flexible room. Unfortunately, there is little to no documentation of the "Dever-style" schools in CPS any longer. (While employed by CPS in the late 1980s, Easton visited many of these schools and recalls them vividly.)


Photo reproduced with permission from **AltusWorks, Inc.**, the architectural design firm that planned an extensive structural renovation of Lovett School. https://altusworks.com/about/

Kudos to Dirk Matthews Photography for the excellent photo.

# Foreword

**Diana Mendley Rauner**
President
*Start Early*

**Robin Steans**
President
*Advance Illinois*

Illinois' Kindergarten Individual Development Survey (KIDS) has provided a new dimension to statewide education conversations, offering parents, teachers and policymakers a view of what comprises kindergarten readiness and how our state's children are faring. Its development over the past 15 years has been methodical and cross-disciplinary, but significant challenges remain in implementation. We welcome this paper as a useful contribution to continuous improvement and offer this foreword as context for that larger discussion.

While initially advanced by the advocacy community as a means of focusing attention on the "achievement gap" that exists prior to kindergarten, ISBE leadership and education practitioners embraced a statewide kindergarten assessment as consistent with and supportive of child-led instructional settings in kindergarten. Development of the KIDS instrument followed a long period of study by a diverse stakeholder group of educators, researchers and advocates from the early childhood and K-12 communities. This group identified two key priorities: assessing readiness across multiple domains and using a developmentally appropriate assessment process, which is observational rather than performative. A review of existing instruments found that none met these criteria, so ISBE commissioned WestEd to adapt California's Desired Results Development Profiles (DRDP) to create KIDS. WestEd is a highly regarded research group, and the DRDP has served as the foundation for kindergarten readiness assessments in Colorado, Louisiana, and several other states. KIDS was piloted in Illinois for five years, which included a validity study of 29 items across the three domains that currently comprise the statewide KIDS mandate. However, to reach agreement on statewide adoption, the required items in the survey were reduced to 14.

Successful implementation of any assessment instrument requires structured support to districts and educators and ongoing analysis of whether the tool works in the field as designed. As an observation-based tool, KIDS requires an especially significant amount of support. Unfortunately, KIDS has suffered from a lack of clear ownership and institutional champions since its inception. Funding, training, and administrative control have spanned multiple divisions at ISBE.

Outreach and communication with Illinois' 842 school districts have been limited. ISBE conducted focus groups with teachers, principals, and superintendents that suggest that KIDS is not well understood and rarely used in the classroom to its fullest and intended extent. Commitment to coaching teachers has been intermittent, and problems with data entry persist. In addition, data is not returned to teachers in a timely fashion. And even though Spanish-language survey versions of KIDS are available, English language learners are not always administered the appropriate version.

Yet, while facing multiple challenges, KIDS has transformed our conversations about early childhood, and stakeholders continue to believe that a statewide kindergarten readiness assessment remains essential. Furthermore, our review of the landscape suggests that KIDS remains the best instrument available. KIDS has changed our understanding of the "readiness gap" across our state and communities, leading the Illinois Legislature to join 37 other states in requiring a statewide kindergarten readiness assessment for all students. While many districts use additional assessments in kindergarten, none provide a look at all important domains of development, nor can they be aggregated to provide an understanding of the distribution of readiness across our state. School districts that have embraced KIDS, using full measures

# Foreword

and in some cases multiple administrations across the school year, have found it to be tremendously useful for both teachers and parents. That said, for KIDS to reach its full potential, the challenges in data entry and the timeliness of data reporting will need to be addressed.

In this spirit of continuous learning and improvement, we found this paper helpful as an additional starting point for discussing how to make KIDS more useful to those who administer it. Start Early and Advance Illinois, as organizations, have championed this effort since inception. We write as representatives of our organizations' efforts. We are grateful to CPS for requesting and encouraging this study and are particularly grateful for the authors' advice to release more data to explore questions of validity and reliability. KIDS is a helpful tool with great potential—further analysis might help us understand how to modify or adjust the items to improve validity and provide even greater information for teachers, school and district leaders, and parents.

# Introduction

In this paper, we explain the concept of assessment validity as dependent upon how an assessment is used. This matters to Illinois teachers, principals, school district and state education leaders, because they decide how to use the Kindergarten Individual Development Survey (**KIDS**), such as: to guide classroom instruction and activities in preschools and kindergartens, determine teacher professional development needs within a school, or allocate additional early learning resources across a school district or state.

Illinois State Board of Education (ISBE) provides sound guidance and examples of what uses it considers to be acceptable uses integrating more contemporary thinking about assessment validity. The current views on assessment validity, espoused in the *Standards for Educational and Psychological Testing*[1], expect that assessment providers (private companies contracted by school systems and systems themselves that create their own assessments) provide validity evidence for each recommended use.

A validity argument for a specific use of assessment results also requires robust evidence that we can define the construct we are measuring and using in our decision-making. Given evidence from recent studies of **KIDS**, in the current report, we explore whether three Developmental Areas in the **KIDS 14** assessment are reliably and validly measured. The evidence suggests that **KIDS 14** measures learning and social skills in line with **KIDS** specifications, but **KIDS 14** does not appear to fully differentiate measuring math skills from measuring Language Arts skills. Instead, it appears to measure a more general set of academic learning skills. This has implications for how **KIDS 14** can best be used and suggests how larger-scale quantitative analyses of **KIDS** could strengthen the evidence base for validity across many uses.

The order of the paper is: a discussion of contemporary views of test validity and how they should influence thinking about **KIDS** and its uses; an in-depth description of **KIDS**'s development, content, and scoring; and the findings from early studies of **KIDS**'s validity and reliability. We then recommend new research studies to enhance our understanding of **KIDS** and analyze an example of an ISBE-recommended use of **KIDS** scores and speculate on the evidence base needed to justify such a use. Finally, our paper closes with a commentary by James Pellegrino, Professor Emeritus of Psychology and founding co-director of the Learning Sciences Research Institute at the University of Illinois Chicago.

Given the specialized and sometimes technical nature of the brief, our target readers include district assessment directors; teaching and learning directors; Early Care and Education (ECE) specialists; school and district leaders; fellow early childhood researchers; and staff at ISBE and at the newly formed Illinois Department of Early Childhood (IDEC). Many other educators, including teachers, may also find the information here to be helpful as they consider their own professional development, professional responsibilities, and career aspirations.

## Considering Assessment Validity in Term of Usage

Over the past few decades, testing and assessment experts, test developers, and psychometricians have been developing and espousing increasingly sophisticated views of both reliability and validity. These views take into account the intended interpretive uses of assessment score results as well as the contexts of their use. Contemporary views have shifted from a piecemeal conception of test validity to one that is focused on the uses of assessment and test results. It is inappropriate to ask, **"Is KIDS valid?"** In contrast, a question like, "Do we have sufficient evidence to support the use of **KIDS** for decisions about instructional grouping?" is much more in line with contemporary views of validity.

While ISBE does not refer to **KIDS** as an assessment but rather as a tool to guide teachers, issues of assessment validity are applicable given that **KIDS** is "scored,"

meaning that no matter how it is used, **KIDS** is an assessment of students' capabilities with respect to certain intellectual and socioemotional competencies. Any use of **KIDS** scores requires validity evidence, even when the use is relatively low-stakes and "tool like." To its credit, ISBE clearly spells out appropriate vs. inappropriate uses of **KIDS** data, providing clear examples. We have not, however, seen any explications of the validity evidence needed to support and justify these uses.

The first chapter of the most recent version of the *Standards for Educational and Psychological Testing*[2] focuses on validity and begins with this definition:

> *Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing tests and evaluating tests. The process of validation involves accumulating relevant evidence to provide a sound scientific basis for the proposed score interpretations. It is the interpretations of test scores for proposed uses that are evaluated, not the test itself.*

Although it is not explicit in the above definition, it is vital that we know and thoroughly understand exactly **what** we are intending to measure. Testing experts and researchers in many fields use the word "construct" to describe abstract concepts of skills and abilities that can't be directly observed but can be measured indirectly through tests and assessments. Constructs may range from simple to complex. "Domains" label and organize groups of related constructs.

ISBE uses slightly different language for these terms. According to its developers, **KIDS** measures three constructs, which are the product of 14 separate Measures. What **KIDS** calls a "Measure" appears to correspond to a relatively simple construct. Measures are then grouped into Domains and Developmental Areas, which also represent more complex constructs, like Social and Emotional Development. Constructs are central in thinking about the meaning of test validity, because how we define each construct affects how we evaluate the quality of the available evidence to justify the use of test and assessment results. Traditionally, the term construct validation has been used to describe the process of providing evidence that shows the construct has been accurately defined. (We will address another issue with how **KIDS** uses the term "Measure," relating to scoring, later in the paper.)

It is incumbent upon test developers to define the constructs and domains being assessed.[3] It is equally important that users understand how each construct is defined. This is accomplished through written definitions, detailed explanations, theoretical justifications, and often by comparing and contrasting them to other similar constructs. These detailed definitions guide item developers as they prepare what will become the "operational" definition of the constructs being assessed. As we will see in a few pages in **Table 2 on page 8**, ISBE provides short definitions of the various components of the **KIDS** assessment.

We are taking the position that **KIDS** is an assessment, even though it may be most often referred to as a tool for judging student competencies and skills, etc. Because **KIDS** is scored it should be considered an assessment.[4] The scores quantify (that is, assess) important developmental constructs/Measures. Although the Standards explicitly assign test developers the responsibility for ascertaining relevant validity evidence for various uses of an assessment, users should perform due diligence to ensure that even "tool-like" uses meet current standards for validity.

The user's need for and reliance upon validity evidence also relates to whether the intended use is considered "high stakes" or "low stakes." We contend that the higher the stakes the greater the need for strong validity evidence. A high-stakes example would be to use **KIDS** scores to place a student in a lower "track," or in a classroom with lower expectations for student performance. A low-stakes use (such as helping a child interact more easily with his peers) may require less robust validity evidence.

Another important consideration is "scope." Does the use involve a single child, a classroom, a school, or a district? A use involving a curricular decision for an entire school district would need stronger validity evidence than a use that involves creating a new seating plan in one classroom to facilitate communication and use of language. Although the user may not be expected

to create appropriate validity evidence, the user should understand the need for validity evidence to make thoughtful and informed decisions and be able to explain their reasoning for going forward with a use or deciding that the validity evidence base is insufficient for a proposed use.

The concept of test reliability matters, too. Test reliability refers to consistency of the test scores: will the same test produce the same result if administered a second time after the first administration? Will a different teacher or observer give a student essentially the same rating on each of the **KIDS** measures as the original teacher did?

Another conception of test reliability is to ask if the test items are all measuring the same construct. Given **KIDS** attempts to measure three constructs (Domains), the concept of reliability primarily pertains to the three scored Domains which are the product of 14 separate Measures as described below. Reliability of an assessment can be established in a variety of ways, but the most well-known of these are statistical measures of "internal consistency," and Cronbach's alpha is well known for these purposes. It is also important to know the composition of the sample of individuals involved in reliability studies to help users know whether this population included "students like mine" and if reliability estimates are comparable across different groups of students. If estimates of reliability are not similar, group comparisons become unreliable themselves.

Given how important it is to have confidence in the ability of a test or assessment to consistently evaluate well-defined constructs, it is easy to see how reliability is critical in building a validity argument for a particular interpretive use. High levels of internal consistency help convince us that we are measuring coherent constructs, Measures and Domains. Sophisticated statistical techniques such as exploratory and confirmatory factor analysis provide very important information about construct representation and contribute important evidence to test use validity. Such techniques show us the extent to which test items match up to the definitions of the constructs/Measures they are intended to assess. If we think that we are measuring three distinct constructs (here, Domains), but an appropriate statistical technique tells us that the test items cluster together around only two factors, then we become less certain about the identifiability of our constructs. Test developers should then refine the constructs' definitions, and/or go back and develop new items that better measure the constructs. Because **KIDS** measures have only a single item, the concept of internal consistency is not relevant.

In a more expansive view of test validity and test use validity, this type of information about construct representation is crucial for building evidence to support particular uses of assessment results. If we cannot provide statistical evidence that our assessment can differentiate Self-Regulation from Social and Emotional Development, we are not likely to gather sufficient evidence to support using **KIDS** assessment data in ways that maintain that distinction. Nor can we compare one Measure to another without assurance of the reliability of its Domain and Developmental Area. Data analysis that has been traditionally considered strictly as reliability evidence is fundamental to our broader view of validity, as it provides information about the viability of the constructs or measures themselves and the proposed interpretive uses of the instrument's scores.

# What is KIDS?

Now we move to a close look at the **KIDS** assessment itself. **KIDS** is designed as a tool for ongoing documentation of children's development in natural environments. It is intended to support teachers as they observe, document, and reflect on the full continuum of children's learning across the kindergarten year.[5]

To complete **KIDS**, teachers rate students on **14 Measures** in the first 40 days of the school year. Teachers are required to complete **KIDS** implementation training through a variety of methods, including ISBE's professional learning system. (They cannot access the on-line **KIDS** platform until that have completed their training.) They are instructed and advised to use multiple sources of evidence to determine their ratings, with the primary source of information and evidence being their own multiple observations of the students in their classroom. Teachers are also encouraged to consider other sources of information, including family reports and examples of students' work, attitudes, and behaviors. After the teacher believes she has enough evidence to make a rating on a measure, she follows carefully designed rubrics that define each of the 6 possible ratings. The first three ratings (1-3) are considered "Building," and the higher three ratings (4-6) are referred to as "Integrating." **See Appendix Figure 1.B on page 23** for an example of a teacher's rating sheet for one of the 14 Measures. (There is a similar sheet from ISBE for each of the 14 Measures.) The sheet also includes the definition of the measure and examples of student behavior or knowledge in each of the six rating categories. These scoring materials assist the teacher in deciding on the most appropriate category score for each of the 14 Measures.

## Development and History of KIDS

**KIDS** is part of a class of tools called Kindergarten Readiness Assessments (KRAs), which are increasingly used across states as tools to assess children's skills at school entry and throughout the kindergarten year. KRAs typically provide a snapshot of children's development across multiple domains – cognitive, language, social-emotional, and physical – based on teacher reports or observations.[6] According to the National Research Council (2008), KRAs can be an effective tool to help schools and teachers identify individual children's learning needs as well as support instruction and programs. Recent articles have highlighted the importance of ensuring that KRAs are developmentally appropriate, culturally sensitive, and embedded in supports for teaching and guiding children.[7]

**KIDS** was developed to create a new KRA that drew on the latest methodological advances in adult reporting of children's functioning across multiple domains. Commissioned by the Illinois State Board of Education (ISBE), **KIDS** was developed collaboratively among personnel at ISBE; the WestEd Center for Child and Family Studies; the Berkeley Evaluation and Assessment Research Center at the University of California Berkeley; the California Department of Education; and members of the early childhood education community, including kindergarten teachers and school administrators.[8] **KIDS** is based on the Desired Results Developmental Profile – Kindergarten (DRDP-K), a formative child assessment system developed by the California Department of Education with researchers from WestEd and the University of California Berkeley. (See Appendix 1.A (page 22) for a compilation of technical reports regarding the development of DRDP-K and analyses of its technical properties).

Although KRAs are potentially useful in many circumstances, "misuses" are also possible. In 2017, the Chicago-based "champion for early learning," Start Early, prepared a careful and comprehensive review of potentially positive and negative uses of KRAs entitled "Uses and Misuses of Kindergarten Readiness Assessment Results." We recommend this useful guide to KRAs.[9] ISBE hosts the "**KIDS** Advisory Committee" to maintain ongoing conversations about the implementation of **KIDS**, including the content and design of the assessment and the results it produces.[10] We note these as evidence of continuing and ongoing interest in **KIDS** both by ISBE and by the broad early childhood education community in Illinois.

# What does KIDS Measure?

The remainder of this paper focuses solely on **KIDS 14**, the mandated version in Illinois schools. There are two other versions, **KIDS 5** and **KIDS 11**, both of which are longer than **KIDS 14** and cover more Domains.

Although none of the tables in this report include it, **KIDS** has an additional, alternative Language and Literacy Development Domain for English learners. It is called ALT LLD and contains only minor changes from LLD. Two measures are deleted and replaced by two other measures that are appropriate for use in bilingual classrooms. Using content prepared by Chicago Public Schools (CPS), ISBE has prepared a helpful short document called Guidance for Dual Language Learners.[11] ISBE also provides guidance for children with special needs.[12] **Table 1** shows the three Developmental Areas, the four Domains (later reduced to three), and the **14 Measures** in **KIDS 14**. This table is adapted from an ISBE table that appears in many ISBE publications. Our adaptations are intended to make the table more accessible to readers and more compatible with conventional terminology. *The original ISBE table appears as Appendix Table 1.B.*

Teachers score the 14 Measures that are shown in the table.[13] Scoring depends on many sources of evidence and relies heavily on many of the teacher's own observations. Although ISBE provides excellent rubrics

TABLE 1

**KIDS 14: Developmental Areas, Domains, and Measure Names**
(Adapted and Simplified from Illinois State Board of Education, 2023)

| Developmental Areas and Short Names | Domains | Measure Names |
|---|---|---|
| **Developmental Area 1:**<br>• Social and Emotional Development<br>• ATL-REG-SED | Approaches to Learning – Self-Regulation | • Curiosity and Initiative in Learning<br>• Self-Control of Feelings and Behaviors<br>• Engagement and Persistence |
| | Social and Emotional Development | • Relationships and Social Interactions with Familiar Adults<br>• Relationships and Social Interactions with Peers |
| **Developmental Area 2:**<br>• Language and Literacy Development<br>• LLD | Language and Literacy Development | • Communication and Use of Language (Expressive)<br>• Reciprocal Communication and Conversation<br>• Comprehension of Age-Appropriate Text<br>• Phonological Awareness<br>• Letter and Word Knowledge |
| **Developmental Area 3:**<br>• COG: Math<br>• MATH | Cognition, Including Math | • Classification<br>• Number Sense of Quantity<br>• Number Sense of Math Operation<br>• Shapes |

to assist teachers in scoring each Measure, without our (authors) first-hand knowledge or direct reports, ISBE's description of the scoring process suggests that it is cognitively complex and time-consuming. Students who are rated using **KIDS 14** receive three aggregate scores – one for each of the Developmental Areas. There is no total score. (We do not have complete information about the scoring process, but we are investigating the topic and plan to write a short technical brief when we have collected more information.[14])

**Table 2** below extends the information provided in **Table 1** by adding new information, including the definitions of the four domains in **KIDS 14** and one example Measure from each Domain with its definition. Note that Table 1 and 2 display four Domains, whereas only three Domains are scored and reported. The two Domains in the first Development Area (Approaches to Learning – Self-Regulation and Social and Emotional Development) are combined into a single score.

## TABLE 2

### KIDS 14: Developmental Areas, Domain Names and Definitions, and Sample Measure Definitions

| Developmental Area | Domains | Domain Definition[15] | Sample Measure Names | Sample Measure Definition |
|---|---|---|---|---|
| **Developmental Area 1:**<br>• Social and Emotional Development<br>• ALT-REG-SED | Approaches to Learning – Self-Regulation | Assesses a child's development of persistence, curiosity, and ability to self-regulate. | Curiosity and Initiative in Learning | Child explores the environment in increasingly focused ways to learn about people, things, materials, and events. |
| | Social & Emotional Development | Assesses a child's development of feelings, behavior, and relationships with nurturing adults and peers. | Relationships and Social Interactions with Familiar Adults | Child develops close relationships with one or more familiar adults (including family members) and interacts in an increasingly competent and cooperative manner with familiar adults. |
| **Developmental Area 2:**<br>• Language and Literacy Development<br>• LLD | Language and Literacy Development | Assesses a child's progress in developing foundational language and literacy skills by observing communication, conversation, awareness of text, and letter and word knowledge. These skills can be demonstrated in any language and in any mode of communication. | Communication and Use of Language (Expressive) | Child's communication develops from nonverbal communication to using language with increasingly complex words and sentences |
| **Developmental Area 3:**<br>• COG: Math<br>• MATH | Cognition, including Math | Assesses a child's number sense, knowledge of shapes, and ability to classify objects through observation, exploration of people and objects, and objects and concepts. | Number Sense of Quantity | Child shows developing understanding of number and quantity. |

## Key Characteristics and Uses of KIDS According to ISBE

**KIDS** has many different uses.[16] These different uses will become more important in our general discussion of the technical concepts of test reliability and test validity as they apply to **KIDS**. We will introduce and describe current expert views on these topics in the following pages. According to ISBE, **KIDS** has the following characteristics:

- **Formative assessment**, which means that its primary purpose is to assess learning *while learning is taking place*. It is intended to provide useful feedback to teachers and families about how to support children and identify individual and common learning gaps.[17]
- **Developmental assessment**, which means that it is focused on whole-child development and can capture change over time, if **KIDS** is completed two or three times per year.
- **Authentic assessment,** which means that it is conducted by teachers who observe children in natural environments and receive supplemental information from parents about their own observations. In this way, it assesses children both in and out of the classroom and uses information that the teacher has not observed directly, but considers behaviors, attitudes, and preferences that occur during the larger part of a child's typical day.
- **Criterion-referenced assessment**, which means that it is anchored in a particular set of standards. ISBE reports state that the "14 metrics (that) are aligned with the Illinois Early Learning and Developmental Standards, Illinois Early Learning Standards – Kindergarten, and the Illinois Learning Standards for English Language Arts, Mathematics, and Science."[18] However, we have not been able to access any material from ISBE on the standards alignment judgment process, nor are there other references to **KIDS** being either a criterion-referenced or standards-based assessment.

Building upon the characteristics noted above, ISBE provided more direct and specific guidance about the use of **KIDS** results in 2023. See **KIDS** FAQs for more information.[19] ISBE lists the following appropriate uses (although, apparently without mention of validity evidence of the types that we previously discussed):

- An informational **tool for teachers** to guide instruction, adapt curriculum, and encourage play-based learning experiences[20]
- An informational **tool for families** as they support their child's unique learning and developmental needs
- A **source of data** to inform policy decisions and leverage funding
- A **tool to foster greater alignment** between early childhood programs, community services, and kindergartens

Since ISBE frequently describes **KIDS** as a tool rather than an assessment, the following uses are discouraged:

- A **diagnostic or achievement measure** for children
- A **tool for enrollment or classroom placement** decisions
- An **indicator of the effectiveness** of individual early childhood providers
- A **tool to evaluate kindergarten teachers**
- An **accountability metric** for schools or classrooms

## Inconsistencies between the Language of KIDS and the Broader Field of Measurement

We now turn to mapping out the terminology used by ISBE to connect to the broader field of measurement. As discussed earlier, the three Developmental Areas, the three Domains (reduced from the original four), **and** the 14 Measures that are included in **KIDS 14** might each be called "Constructs" by psychologists or measurement and assessment specialists. These may constitute sets of knowledge, skills, and abilities that are inferred through a measurement process, as in **KIDS**. Social and Emotional Development is a good example of a construct that is measured by inferences made from observations, self-reports, structured interviews, and other techniques. Good test and assessment practices include the requirement that these domains or constructs be carefully and thoroughly defined ahead of time to guide the assessment development process.

In the general field of testing and assessment, an item refers to a single question or piece of evidence of a student's skill or ability. The word "Measure" is often used synonymously with the term "Scale," a group of achievement or attitudinal questions or statements that teachers respond to which are then scored. For example, we may see a Scale comprising four or five questions (items) about students' sense of "belongingness" in their school. Depending on the source of this scale – a commercially available student survey; a state-administered "healthy schools" survey; a university-developed questionnaire – the assessment administrator judges these responses according to pre-developed algorithms or rubrics and provides a score. The score is usually calculated "automatically" by computer. This score represents how much or how little a student feels that she "belongs" in her school. This quantification process is the heart of measurement.

In **KIDS 14**, each of the three Domains (constructs) includes four or five Measures. Here, the term Measure appears to correspond to what is usually called an "item," since the teacher assigns one rating (score) to each **KIDS** Measure. But teachers score each Measure based on multiple observations of each student and other information. The teacher may keep records of observations in a notebook, checklist, or handout, which form the basis for assigning a score with guidance from detailed rubrics. ISBE provides guidance for collecting and recording evidence in a two-page leaflet called "Strategies for Collecting & Organizing Observations for **KIDS**." ISBE has also created a one-page "handout" for recording a data collection event.[21] These observations and other evidence are akin to what are usually considered items.

Thus, what we might think of as the items that make up the **KIDS** Measures are never preserved or formally recorded. Because this information is not collected, **in effect, a given KIDS measure is a one-item rating scale (1-6). If the teachers' multiple observations were recorded,** they would be what are traditionally called items. The use of Measure and Item as synonyms is different from traditional measurement and assessment practices and thus requires a deep understanding of how **KIDS** operates. The uncertainty in **KIDS** about whether a Measure is an item or a scale compounds the difficulty of ascertaining how many Domains/Constructs the assessment measures. Without seeing the individual "items" teachers use to score the Measures, it becomes impossible to know the extent to which those observations and evidence match up to the constructs (Measures and Domains) they are intended to address.

# Published Studies of KIDS Validity and Reliability

A small body of early published work has examined aspects of the reliability and validity of **KIDS**. We seek to build on this foundation to consider the use of score results for making student grouping decisions, for altering instruction, and/or for reporting student progress to parents, among other uses suggested by ISBE (see section on Uses from ISBE).[22]

The Berkeley Evaluation and Assessment Center and WestEd conducted the first published empirical study of **KIDS 14** data based on a 2014-15 pilot administration in Illinois involving about 26,000 students. This first analysis of preliminary data[23] identified three factors among the 14 Measures with the factors corresponding directly to the three Developmental Areas (Domains) in **KIDS 14**. These results support design and construct representation claims for the three domains. However, a follow-up analysis in 2017-18 from the first large-scale administration of **KIDS** in Illinois documented only two factors (constructs) among the 14 Measures across the three domains[24]. These are **academic knowledge and skills** and **learning and social skills**. The items that tap into content knowledge in either math or language arts fit together in the "knowledge and skills" factor, while the remaining items about communication and relationships are in the second factor, "learning and social skills." In **Table 1** on page 7, all the ATL-REG and SED measures plus the first two measures in LLD (referring to communication) fit in the learning and social skills factor. The remaining LLD and MATH items fit in the academic knowledge and skills factor. Each of the two factors contains seven items.

While we have no direct evidence to support the following speculations, perhaps the seven measures in the learning and social skills factor are derived almost exclusively from teachers' observations, whereas the seven academic knowledge and skills measures are scored from more "objective" data, such as worksheets, factual answers the teachers' questions or quizzes.[25] Despite the findings from this analysis, ISBE continues to score and report three Domains, leaving users with some uncertainty regarding the validity and meanings of the underlying major Domains/constructs that **KIDS** attempts to measure.

Notably, since this initial psychometric work was completed, we know of no other work that has explored the factor structure and construct validity of the **KIDS**. This is because ISBE has not released Measure (item) level data, which means researchers cannot replicate or conduct further analysis of the proposed original construct representation. It is important to recognize that "data reduction"[26] techniques are statistically complex and require highly advanced skills. We, the authors of this paper, do not have sufficient information to judge whether **KIDS** is indeed measuring two or three factors/Domains-constructs. However, if we did have access to individual responses to all 14 measures, we would be able to access the technical expertise to provide convincing results to this challenging problem of validity of construct representation and proper reporting and interpretation of the meaning of the scores obtained from the instrument given various uses as specified by ISBE.

Most importantly for this discussion of **KIDS**, researchers from the Illinois Workforce and Education Collaborative (IWERC have written two excellent reports and a short follow-up memo on **KIDS**.[27,28] The papers primarily focus on important research questions about group differences in test score trends over time (e.g., based on gender), and the extent to which **KIDS** can predict standardized test scores on the Illinois Assessment for Readiness (IAR), the mandated state test, three years later. This has traditionally been called "predictive validity." Using statewide data for two cohorts of students, IWERC researchers matched **KIDS** Domain scores from the fall of 2017 and the fall of 2018 to scores on the mandated state test, the Illinois Assessment of Readiness (IAR), when the kindergarteners were in third grade in 2021 and 2022, respectively. The main result of this work is that scores on the two tests (administered three and a half years apart) are correlated at the level of $r = 0.42$ in English language arts and $r = 0.44$ in math.

This compares to an r = 0.47 in a very similar study conducted in the state of Virginia a few years earlier – the only other published statewide study of which we are aware – that used a KRA focused primarily on language and literacy.[29]

What does the IWERC correlation coefficient tell us? A correlation of 0.42 is usually considered on the low end of "moderate." (After all, r = .0.42 is equivalent to an R square of 0.176, telling us that about 18% of the variability in IAR scores is associated with **KIDS** scores.) But if we consider that the two tests are assessing slightly different constructs, three and a half years apart, among children who are growing rapidly, we can see more value in r = 0.42 and are likely to be more generous in our interpretations.

The IWERC studies also raise questions about the **KIDS** constructs (domains) because predictions from one domain to another are often similarly strong as predictions within the same domain. For example, **KIDS 14** ATL-REG-SED, LLD, and MATH scores all predict IAR math scores well: Of those children who are "proficient" in Grade 3 math, 43% were kindergarten-ready in ATL-REG-SED, 48% were kindergarten-ready in LLD, and 52% were kindergarten-ready in MATH. Additionally, the domain scores are highly correlated among themselves, suggesting significant overlap. To reiterate the point made previously, a validity argument for a specific use of test or assessment results requires robust evidence that we can define the construct we are measuring and using in our decision making.

Given the evidence provided by the WestEd and IWERC studies, we have many open questions about whether the three Developmental Areas/Major Constructs in **KIDS 14** are reliably and validly measured. It is our position that **KIDS** measures **learning and social skills** in relatively close alignment to the **KIDS** specifications, but **KIDS** does not appear to differentiate math skills from language arts skills. Instead, **KIDS** measures a more general set of **academic and learning skills.**[30] Because of this, we are not confident in uses that differentiate academic skills by content area. On the other hand, we are more confident in uses that rely on what is called Social and Emotional Development in **KIDS** and the set of academic skills that are needed in both math and language and literacy learning.

# Considerations for Continued KIDS Research

First, it is important to recognize that the most recent edition of the *Standards for Educational and Psychological Testing* directs that it is the responsibility of the test developer to provide validity evidence for various recommended uses of a test: in this case, the scores derived from teachers' rating of their students. We will soon look closely at the ISBE recommended uses of **KIDS** scores and speculate on the specific evidence required to ensure the validity of these intended uses. Before examining specific uses, we provide some examples of how some larger scale quantitative analyses could play valuable roles in developing validity profiles across many uses. Here is a starting list of useful studies that could add new information to support validity arguments for a range of specific use cases. It is very likely that ISBE has also asked many similar questions, but we have not seen written materials that confirm this.

## Examples of Large-scale Statistical Analyses

- ISBE should make Measure (item) level data available to qualified researchers. Access to these data would enable researchers to conduct basic, easily interpretable descriptive statistical analyses.[31] These would be to ISBE's own advantage by providing greater insights into the "inner workings" of the **KIDS** assessment. For example, we could create "error bands" around scores that would inform users of how confident they should feel about the accuracy of scores. We could look at typical distributions of scores across the 6 scale score points on individual items and ask whether they were reasonable or not. For example, are some items too "hard" or too "easy"? Are extreme score points almost never used? Do distributions appear to be acceptable for both fall and spring administrations? Access to item-level data would also make it possible to assess the internal consistency reliabilities of Developmental Area and Domain scores. It would also be informative to know what percent of the time scores would match (or be within 1-2 points) on retesting within a short span of time.
- Next, following the descriptive analyses, we need a more complete understanding of the underlying factor structure of **KIDS** and the extent to which these factors correspond to the Domains/constructs/Developmental Areas that **KIDS** intends to measure. If the constructs and statistical factors do not directly correspond, ISBE and its research partners should consider refining or revising the Constructs or re-thinking better ways of measuring and reporting them. The general concept of construct validity is key here. Do the statistical findings line up with the written definitions of the constructs?[33] If not, ISBE could refine the construct definitions or re-think the measures and how they are scored. We must know what we are measuring to build validity arguments for **any** use of **KIDS** scores.
- Is the factor structure stable across different groups of students (by gender, race, language, and income level, for example)? Without this evidence, it is impossible to interpret any comparisons of groups of children to one another.
- If enough teachers use and save the ISBE "handouts" to collect evidence to support their scoring of measures, we could analyze these as if they were items in the 14 Measures. Do these "item facsimiles" scale into a Measure? This would provide very important information about how teachers interpret the written definitions of the Measures and the rubrics used to guide scoring. This would be a complex and technically challenging study, since it would probably require coding the teachers' written remarks on the "handouts." It would also be logistically difficult. See below for a simpler approach to this question.
- How are teacher and student characteristics related to **KIDS** scores? Under this overarching question, there are many important issues. First is "interrater reliability." In other words, to what extent would different teachers observing the same child under the

same conditions score the child similarly or differently? If the scores are very similar, we gain confidence in the assessment. If they differ greatly, we lose confidence in our ability to build a validity argument for certain uses. We should also look for signs of bias in teacher ratings and discover if race match or non-match between student and teacher has a systematic relationship to scores.

- Are there observable "school effects" in **KIDS** ratings? Are average **KIDS** scores similar in schools that serve demographically similar students? What contextual factors within and between schools are related to **KIDS** scores?
- Finally, we must learn more about "growth" in **KIDS** scores from fall to winter to spring. To what extent is growth in **KIDS** scores unrelated to school experiences but primarily attributable to children's age-related maturation? We would expect that much of the changes in **KIDS** scores during the kindergarten year would be the result of physical, social, and emotional maturation. Can we account for this factor in scoring to make fall to spring score comparisons more accurate and separate school influences from other strong influences from family, peers, neighbors, community?

Ultimately, answers to the questions above and others like them will assist test developers and educational researchers build validity cases for many different uses of **KIDS**. They will also assist users in feeling secure about their choices and decisions about how they use **KIDS** scores. Some users may feel that they need more definitive evidence to support a given intended use.

## Examples of Smaller-Scale, Often Qualitative Studies

- How do the **KIDS** Domain or Measure scores relate to other sources of easily accessible information, such as report card grades or attendance?
- How do teachers determine when they have observed students enough to make them feel confident in rating children on each of the 14 measures/items?
- Inquiries into how teachers and school leaders use **KIDS** results could help us learn whether students are ever harmed by these uses.[34] Are scores on **KIDS** being used in ways that might have unintended negative consequences for students in the long run? Are teacher expectations influenced by **KIDS** scores in ways that result in differential attitudes and behaviors towards low-(or high-) scoring students? Or, do teacher expectations and biases influence how they score students on the 14 measures? Despite prohibitions against this use, are students being "tracked" based on **KIDS** findings?
- If Measure-level scores (recall that Measures are equivalent to items in **KIDS**) are interpreted, teachers may make comparisons among Measure scores for an individual child. Such a practice would be questionable if the Measures differ in average difficulty. That is, if some items are inherently more difficult, due to quirks in scoring rubrics or to actual differences in developmental trajectories of different qualities measured, then Measure-to-Measure comparisons may be misleading. Such comparisons, which are a kind of "difference score" are inherently less reliable than separate item scores, sometimes by quite a bit.
- In the previous section, "Examples of Large-Scale Statistical Studies," we suggested collecting all the teachers' "handouts" and other records to understand how a teacher determined a Measure score. Instead of such a big and demanding study, we could instead interview a representative sample of teachers and ask them to describe what data points they collect (and how) in order to determine a score for each of the 14 Measures.
- How much time do teachers spend rating each child on the 14 measures? What is the average amount of total time per child? What share of kindergarten teachers conduct thorough ratings in accordance with ISBE recommendations? What are common "shortcuts" and how often are teachers rating children without any written documentation? How many students are typically excluded from the **KIDS** rating process and for what reasons?
- How does a parent-teacher discussion about **KIDS** scores compare to a discussion about report cards or other parent/teacher discussions?

These are only examples of research studies that would enhance our understanding of **KIDS** and provide additional evidence that could be helpful in building a validity case for a specific use. The major purpose of this memo is to help readers think of these concepts in terms of specific uses for the assessment/tool and not as permanent properties of the instruments themselves.

Returning to the first paragraph of the most recent edition of the *Standards for Educational and Psychological Testing*, the process of validation is about providing evidence to support specific uses of scores and other results from assessments. A critical part of this process is providing evidence that will reassure users that the assessment is measuring the constructs (domains) that it claims to.

To conclude this short memo, we speculate on what evidence we would want to have to decide that **KIDS** is sufficiently valid for a few use examples, including two that ISBE has already commented on.

Here are uses of **KIDS** findings that ISBE considers "appropriate:"

- An informational tool for families as they support their child's unique learning and developmental needs.

What evidence would we need to support the validity of the assessment for this use?

We have determined that because **KIDS** is scored, we consider it an assessment with tool-like properties. Nonetheless, how can we be assured that this is a valid use of **KIDS**?

As we've argued previously, we are more comfortable with the scores from the Social and Emotional Development portion of **KIDS** given that statistical analyses could not differentiate Math Cognition from Language and Literacy Development. Instead, that analysis found a more general Developmental Area of Academic Learning. Therefore, we would focus our discussion on SED and instead of differentiating MATH from LLD consider the commonalities among the measures in the Academic Learning Developmental Area.

Knowing this statistical information, we may turn to qualitative forms of validity data, which we haven't discussed. Is there case study information available to guide the teacher? What is the teacher's own experience over her years of teaching and what have her colleagues shared with her? Has she shared similar information with parents in the past, for example, by discussing examples of student work, test and quiz results, classroom behavior, and report card grades with parents? Have these "information sharing" experiences been valuable? How do parents react to these experiences? When do parents use shared information and to what effect? Does the teacher observe any differences in children's attitudes and behaviors in the classroom? Has the teacher observed negative effects of information sharing with parents, such as parents "blaming" teachers for low scores, suggesting poor preparation? We are suggesting that in this use-case example, documentation of other teachers' experiences plus the teacher's own experience would provide guidance in determining whether this is a valid (and worthwhile) use of **KIDS**. If the teacher has little experience and is "novice," she needs guidance from her colleagues and leaders, and if this is documented in writing, that would be better. In any case, this is a low-stakes use for both child and parents and involves only a single child (narrow scope). However, there is little guidance provided for these "tool-like" uses.

- Using **KIDS** results to inform school improvement.

Another use of **KIDS** that should be appropriate is using its results to inform school improvement processes. When schools are subject to improvement requirements pursuant to federal and state accountability laws, they are expected to use data to diagnose the issues where improvement is needed. Historically this work has focused primarily on the assessment results in grades 3-8 that form the core of state accountability under the federal Every Student Succeeds Act (and prior to that No Child Left Behind). But even though **KIDS** results should not be used to determine a school's accountability status, they can be used to guide that school's improvement efforts.

It is worth emphasizing that **KIDS** results should not be used to judge the quality of schools – or preschools. **KIDS** has not been validated for accountability purposes, and the nature of the data collected is fundamentally different from the data collected in the

Illinois Assessment of Readiness that is used in state accountability. In order for teachers to be honest and accurate in their completion of the tool, it is essential that **KIDS** not have any stakes in grading the teacher's school.[35]

But once a school has been designated for improvement, that school's **KIDS** data can be incredibly helpful to developing an improvement plan. For example, if the school's **KIDS** scores show that children are coming into kindergarten at a strong developmental level but then falling off track between kindergarten and third grade, that information should be useful in focusing improvement resources on the K-3 years.

In many schools designated for improvement, however, the more likely scenario is that children will be coming into kindergarten struggling in multiple domains. If that is the case, it indicates that for the school to achieve sustainable long-term improvement, it will likely need to have a strategy for improving kindergarten readiness. That strategy might include expanding the school's own early childhood offerings, partnering with other community providers, or some combination thereof; extensive discussion at the local level will undoubtedly be needed to determine the correct path forward. But using **KIDS** to highlight the need to strengthen the community-level support for early childhood as part of the school improvement process is a promising and underutilized strategy.

# Conclusion and Priority Follow-up Actions

We did not want this short paper to repeat many easily available descriptions of **KIDS**. There are several excellent descriptions from ISBE and the meticulous technical work of the teams at WestEd and IWERC also provide good descriptive information about this set of tools/assessments.

Instead, we focused on the concepts of reliability and validity as they are related to **KIDS**. Our most basic and important argument is that validity is not an inherent property of any test or assessment. Test validity must be judged by how the test is used and the claims that are made based on that use. It is incumbent on test developers to provide evidence for a range of specific uses. It is not meaningful to ask, "is the state math test valid?" It *is* meaningful to ask, "do we have sufficient evidence to claim validity for deciding whether a child should be required to repeat third grade because of his scores on the state test?"

This shift in thinking lifts the concept of "construct validity" to greater prominence. First and foremost, we need certainty that we are measuring what we claim to measure. This is the central component of validity and being able to back our claims with credible evidence is essential when we argue that a specific use of an instrument is acceptable and valid in this instance. Various components of "reliability" become central to validity arguments. An evidence case for the validity argument for a specific use can range from simple to complex, as we tried to illustrate in our examples.

The *Standards for Educational and Psychological Testing*[36] remains the most authoritative source of information for arguing for or against what constitutes "valid uses" of **KIDS** scores.

**KIDS** is a thoughtfully constructed, potentially valuable tool to help guide teachers' evaluations of kindergarteners' proficiencies. Measurements of young children are inherently unstable (unreliable), and so one strength of **KIDS** is its reliance on multiple observations over a period of time. Nonetheless, **KIDS** can still offer no more than "snapshots" of children over a limited period based on teachers' perceptions of their skills. Children differ in their patterns of maturation, and those behind at one point may surge ahead later. With these caveats, as a low-stakes resource to help guide teachers and parents/guardians toward a fuller understanding of children's growing competencies, **KIDS** may be of significant value. Yet the difference between low-stakes and high-stakes uses can be hard to discern. We hope that this paper can help users think through the validity demands across the multiple potential uses of **KIDS**.

Finally, we would like to repeat and emphasize what we see as the most important and highest priorities for ensuring the integrity and appropriate uses of **KIDS**.

- First, we ask again that ISBE release more technical information about **KIDS** including the student-level teacher ratings on all 14 measures. This will help researchers and their practice partners gain a greater understanding on the Domains that **KIDS** measures. As we've noted, there are two conflicting results from two studies that investigated the underlying "factors" in **KIDS**. At EC*REACH we collaborate with our partners in the Office of Early Childhood Education at CPS. They are eager to know more about the assessments administered to young children so that they may know more about what each of them measure to be more efficient and more attentive to some instruments than to others. Part of such a study would include attention to whether the factors (domains) are stable across different groups of children. In technical terms, this is called measurement invariance. Related to this need for access to "raw" data, we have not been able to obtain technical details about the scoring process used to create the three Domain scores in **KIDS**. We wish that ISBE was able to be more open and transparent about this type of technical information.
- Neither we nor our colleagues have a strong grasp on how **KIDS** is used in classrooms, by school leaders, and the administrators who supervise and guide them. We strongly recommend a robust "implementation study" that would provide evidence on how to improve usage of **KIDS** results. Anyone doing research on **KIDS** implementation and usage should consider both "practice" uses and "policy"

uses. Practice uses include how teachers apply information from **KIDS** results to their classroom practices, how groups of teachers work together to understand their students' strengths and weaknesses, and how school leaders prioritize professional development, curriculum, and instructional strategies.

- Finally, like the proposed implementation study described above, we would like to see an exploration of whether and how the validity requirements differ between "practice uses" and "policy uses." This is a complex question, given how practice and policy are so often intertwined. Considering "high-stakes" vs. "low-stakes" uses is useful here, moving beyond the **KIDS** assessment to the many other assessments used in CPS and Illinois, an exploration of the type of use. Here, we propose comparing practice and policy uses for **KIDS**, but there may also be other relevant classifications of "use types." Ultimately, research on **KIDS** uses and validity evidence could increase our understanding of the validity evidence needed to support how a broad range of assessments are used by districts across Illinois, including CPS.

# Commentary

**James W. Pellegrino**
*Emeritus Professor of Psychology and Learning Sciences, University of Illinois Chicago*

*Member of ISBE Assessment and Accountability Technical Advisory Committee*

Assessment of what students know and can do is a significant aspect of educational practice. It occurs in many forms – tests, surveys, observations, questionnaires, etc. – and for individuals from the very young to adults, and in a variety of educational and other settings. Regardless of who is being assessed, what is being assessed, or how the assessment is transacted and reported, it is essential that the assessment yield information that is valid for the intended interpretive use by whomever is the designated user of the assessment results. This research brief illustrates critical issues regarding what validity is by using the case study of the **KIDS** assessment that the Illinois State Board of Education (ISBE) has adopted for use by Illinois' early childhood educators. Overall, I applaud the thoughtfulness of this paper's discussion of validity in general and as applied to **KIDS** given **KIDS**' potentially significant role in promoting the cognitive and socio-emotional development of our youngest students.

As described in the brief, considerable thought, effort, and resources were invested in the development of this instrument. Furthermore, ISBE provides resources to help early childhood educators use the assessment tool to make inferences about critical aspects of young children's cognitive and socio-emotional development. The current brief asks the question whether evidence exists to support interpretation of the metrics derived from **KIDS** for individual children at specific points in their early education journey, and for various interpretive uses. Do the metrics have the intended meaning regarding a child's status for the specific cognitive and socio-emotional competencies that **KIDS** is purportedly assessing? Are the metrics meaningful and usable? In essence, are the metrics valid for what is a primary use by educators – to provide important insights that an educator and/or a child's parents can use to promote further development of a child's particular competencies.

But the brief goes well beyond such formative uses of **KIDS** results for individual students to aggregations of results across students within and across classrooms or other possible interpretive uses. This includes going so far as to use the scores in a summative way to compare classrooms and/or predict subsequent educational outcomes. Are such comparative and predictive uses also valid? What is the evidence to support validity claims for such uses? A major contribution of this brief is to bring to the foreground questions about what evidence exists or is needed to support a variety of potential uses that extend well beyond a teacher using **KIDS** to help her understand and act upon a child's status for an important developmental competency. The brief discusses several sources of data that could and should be obtained to support the validity argument for various aggregated uses of the scores obtained from the **KIDS** instrument. All their suggestions are worthy of further consideration by ISBE.

Personally, the most important of the possible inquiries articulated by the authors is how teachers use the instrument and its metrics to guide their educational practice in the classroom. It would be a shame if the considerable investment by ISBE in development of **KIDS** and promotion of its use was for naught if evidence revealed that **KIDS** was not being used in ways that are appropriate and valid by some substantial portion of early childhood educators. This would be important and actionable information and is one example of the validity related critical inquiries that ISBE should consider pursuing for the **KIDS** assessment program.

# References


**American Psychological Association., National Council on Measurement in Education., & Joint Committee on Standards for Educational and Psychological Testing. (2014).** *Standards for educational and psychological testing.* American Educational Research Association.

**BEAR Center & WestEd. (2015).** *Preliminary **KIDS** Analysis: Three subscales Using a Subset of Measures Across Domains.* University of California, BEAR Center.

**Bowdon, J., Dahlke, K., Yang, R., Pan, J., Marcus, J., & Lemieux, C. (2019).** *Children's knowledge and skills at kindergarten entry in Illinois: Results from the first statewide administration of the Kindergarten Individual Development Survey.* https://files.eric.ed.gov/fulltext/ED599357.pdf

**Chicago Public Schools. (2023).** *Kindergarten individual development survey (**KIDS**) fact sheet 2022-2023. Chicago Public Schools.* https://docs.google.com/document/d/1MZsjFnhuDNaAtp2pv61VfSuXyVsBgqMxFnCeCNFqI7I/view?tab=t.0

**Herring, W. A., Bassok, D., McGinty, A. S., Miller, L. C., & Wyckoff, J. H. (2022).** Racial and socioeconomic disparities in the relationship between children's Early Literacy Skills and third-grade outcomes: Lessons from a kindergarten readiness assessment. *Educational Researcher, 51*(7). https://doi.org/10.3102/0013189x221091535

**Illinois State Board of Education. (n.d.).** ***KIDS** frequently asked questions.* https://www.isbe.net/Documents_**KIDS**WebsiteResources/**KIDS**_FAQ.pdf

**Illinois State Board of Education. (n.d.).** *Strategies for collecting observations.* https://www.isbe.net/Documents_**KIDS**WebsiteResources/Strategies_for_Collecting_Observations.pdf

**Illinois State Board of Education. (n.d.).** *Summary of **KIDS** Measures and Reports.* Illinois State Board of Education. https://www.isbe.net/Documents_KIDSWebsiteResources/Summary_of_**KIDS**_Measures_and_ Reports.pdf

**Illinois State Board of Education. (2016).** *Illinois School Readiness Initiative.* https://www.isbe.net/Documents/**KIDS**-IL-School-Readiness-Initiative.pdf

**Illinois State Board of Education. (2017a).** ***KIDS**: Guidance for Dual Language Learners.* https://www.isbe.net/Documents_**KIDS**WebsiteResources/LLD_Guidance.pdf

**Illinois State Board of Education. (2017b).** ***KIDS**: Participation or Exemption Guidance for Children with Special Needs.* https://www.isbe.net/Documents_**KIDS**WebsiteResources/**KIDS**_Exempt_Special_Needs_Guidance.pdf

**Illinois State Board of Education. (2023a).** *2022–23 School Year Illinois Kindergarten Individual Development Survey (**KIDS**) report.* Illinois State Board of Education. https://www.isbe.net/Documents/IL-**KIDS**-Report-2022-2023.pdf

**Illinois State Board of Education. (2023b).** *Kindergarten individual development survey: User's guide and instrument.* https://www.isbe.net/Documents/**KIDS**-User-Guide-Instrument.pdf

**Illinois Workforce and Education Research Collaborative (n.d.).** ***KIDS** socioemotional domain also predicts Grade 3 test scores.* https://dpi.uillinois.edu/wp-content/uploads/2024/12/**KIDS**-SEL-1-pager.pdf

**Kiguel, S., Cashdollar, S., & Bates, M. (2024a).** Inequity in the early years: Student development trajectories from Kindergarten to Grade 3. *Kindergarten Readiness in Illinois Series.* Illinois Workforce and Education Research Collaborative (IWERC), Discovery Partners Institute, University of Illinois. https://dpi.uillinois.edu/wp-content/uploads/2024/10/**KIDS**-Report-2.pdf

**Kiguel, S., Cashdollar, S., & Bates, M. (2024b).** Trends and disparities in readiness using the Kindergarten Individual Development Survey (**KIDS**). *Kindergarten Readiness in Illinois Series.* Illinois Workforce and Education Research Collaborative (IWERC), Discovery Partners Institute, University of Illinois. https://files.eric.ed.gov/fulltext/ED662128.pdf

**Lesaux, N. K., & Jones, S. M. (2016).** *The leading edge of early childhood education: Linking science to policy for a new generation* (Vol. 113, No. 12, pp. 2705-2738). Harvard University Press, Cambridge

**Regenstein, E., Connors, M., Romero-Jurado, R., & Weiner, J. (2017).** *Uses and Misuses of Kindergarten Readiness Assessment Results. Start Early Policy Conversations.* https://www.startearly.org/resource/uses-and-misuses-of-kindergarten-readiness-assessment-results/

# APPENDIX A

## Compilation of Peer Reviewed Research Studies of the Desired Results Developmental Profile - Kindergarten

Professor Mark Wilson kindly prepared this compilation of seven peer reviewed studies of the Desired Results Developmental Profile – Kindergarten (DRDP–K). Wilson is Professor at the School of Education at the University of California, Berkeley. He also directs the Berkeley Evaluation and Assessment Research Center (BEAR). BEAR, WestEd, and the California Department of Education collaborated to create DRDP-K. These same collaborators were also influential in the development of KIDS. As an "offspring" of DRDP-K, KIDS was adapted and revised to meet the needs of ISBE and many concerned individuals and organizations who foresaw the need for a Kindergarten Readiness Assessment for Illinois. The seven papers below cover a range of topics related to the development, use, and technical properties of DRDP-K.

Each of the papers listed below is accompanied by a "url" that connects to a specific web address where the paper is located. They may be read on-line or downloaded. These were checked for accuracy on January 20, 2026.

**California Department of Education. (2018).** *DRDP (2015) 2017-2018 Differential Item Functioning (DIF) Analyses Report.* Report prepared for the California Department of Education. [DIF]. https://www.draccess.org/DIFanalysesReport.html

**Chen-Gaddini, M.; Sussman J., Newton, E., Ruiz Jimenez, G. S., Kriener-Althen, K., Gochyyev, P., Draney, K., & Mangione, P. (2022).** *DRDP Technical Report: Validity in Relation to External Assessments of Child Development.* WestEd. https://www.desiredresults.us/sites/default/files/docs/resources/research/DRDP%20EV%20report%20with%20designed%20template_0628_ADA%20%281%29.pdf

**Chen-Gaddini, M.; Sussman J., Newton, E., Ruiz Jimenez, G. S., Kriener-Althen, K., Gochyyev, P., Draney, K., & Mangione, P. (2022).** *DRDP Technical Report for Early Infancy Through Kindergarten: Interrater Reliability.* WestEd. [Interrater Reliability]. https://www.desiredresults.us/sites/default/files/docs/resources/research/DRDP%20IRR%20report%20with%20designed%20template_0628_ADA%20%281%29.pdf

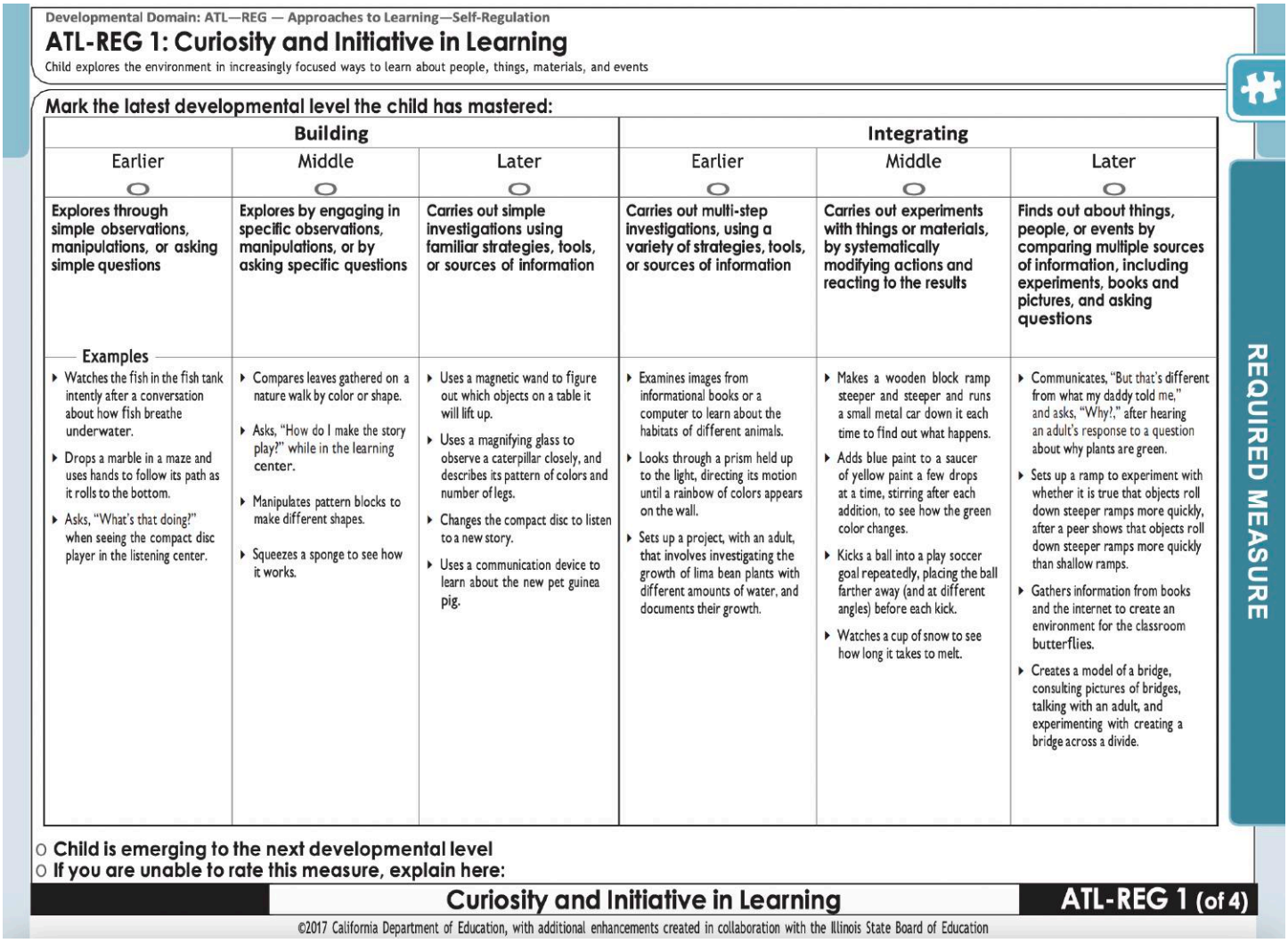
**Draney, K., Sussman, J., Kriener-Althen, K., Newton, E. K., Gochyyev, P., & Mangione, P. (2022).** *DRDP technical report for early infancy through kindergarten: structural validity and reliability information for the desired results.* Report prepared for the California Department of Education. https://www.desiredresults.us/sites/default/files/docs/resources/research/DRDP%20early%20infancy%20thru%20K%20tech%20report%2020211216_FINAL%20ADA_V3.pdf

**DRDP Collaborative Research Group. 2018. Technical Report for the Desired Results Developmental Profile (2015).** *Report prepared for the California Department of Education.* https://www.desiredresults.us/sites/default/files/docs/resources/research/DRDP2015_Technical%20Report_20180920_clean508_0.pdf

# APPENDIX B

## FIGURE 1.B

### One of the 14 sheets that guides the teacher's rating [37]



Developmental Domain: ATL—REG — Approaches to Learning—Self-Regulation

**ATL-REG 1: Curiosity and Initiative in Learning**

Child explores the environment in increasingly focused ways to learn about people, things, materials, and events

REQUIRED MEASURE

**Mark the latest developmental level the child has mastered:**

| Building | | | Integrating | | |
|---|---|---|---|---|---|
| Earlier ○ | Middle ○ | Later ○ | Earlier ○ | Middle ○ | Later ○ |
| **Explores through simple observations, manipulations, or asking simple questions** | **Explores by engaging in specific observations, manipulations, or by asking specific questions** | **Carries out simple investigations using familiar strategies, tools, or sources of information** | **Carries out multi-step investigations, using a variety of strategies, tools, or sources of information** | **Carries out experiments with things or materials, by systematically modifying actions and reacting to the results** | **Finds out about things, people, or events by comparing multiple sources of information, including experiments, books and pictures, and asking questions** |
| **Examples** | | | | | |
| ▸ Watches the fish in the fish tank intently after a conversation about how fish breathe underwater.<br>▸ Drops a marble in a maze and uses hands to follow its path as it rolls to the bottom.<br>▸ Asks, "What's that doing?" when seeing the compact disc player in the listening center. | ▸ Compares leaves gathered on a nature walk by color or shape.<br>▸ Asks, "How do I make the story play?" while in the learning center.<br>▸ Manipulates pattern blocks to make different shapes.<br>▸ Squeezes a sponge to see how it works. | ▸ Uses a magnetic wand to figure out which objects on a table it will lift up.<br>▸ Uses a magnifying glass to observe a caterpillar closely, and describes its pattern of colors and number of legs.<br>▸ Changes the compact disc to listen to a new story.<br>▸ Uses a communication device to learn about the new pet guinea pig. | ▸ Examines images from informational books or a computer to learn about the habitats of different animals.<br>▸ Looks through a prism held up to the light, directing its motion until a rainbow of colors appears on the wall.<br>▸ Sets up a project, with an adult, that involves investigating the growth of lima bean plants with different amounts of water, and documents their growth. | ▸ Makes a wooden block ramp steeper and steeper and runs a small metal car down it each time to find out what happens.<br>▸ Adds blue paint to a saucer of yellow paint a few drops at a time, stirring after each addition, to see how the green color changes.<br>▸ Kicks a ball into a play soccer goal repeatedly, placing the ball farther away (and at different angles) before each kick.<br>▸ Watches a cup of snow to see how long it takes to melt. | ▸ Communicates, "But that's different from what my daddy told me," and asks, "Why?," after hearing an adult's response to a question about why plants are green.<br>▸ Sets up a ramp to experiment with whether it is true that objects roll down steeper ramps more quickly, after a peer shows that objects roll down steeper ramps more quickly than shallow ramps.<br>▸ Gathers information from books and the internet to create an environment for the classroom butterflies.<br>▸ Creates a model of a bridge, consulting pictures of bridges, talking with an adult, and experimenting with creating a bridge across a divide. |

○ **Child is emerging to the next developmental level**
○ **If you are unable to rate this measure, explain here:**

**Curiosity and Initiative in Learning** **ATL-REG 1 (of 4)**

Like California's DRDP-K, **KIDS** is available to schools from ISBE in three different versions, as shown in **Appendix Table 1.B.**[38] The versions differ primarily in the number of domains that they contain; the number of domains assessed directly affects the time required for teachers to make observations, collect data, and judge each child's level of development on each.

The State of Illinois requires all schools to administer the shortest version of the assessment, **KIDS 14**, shown in Column 2 of Table 1.B. Column 3 shows **KIDS 5**, which is longer than **KIDS 14**, but shorter than **KIDS 11**, the most complete and comprehensive version. It represents five of the 11 Domains and 29 Measures. **KIDS 11** is the longest and most comprehensive version of the **KIDS** assessment. It contains 11 Domains of School Readiness that are fully aligned with Illinois Learning Standards. There are 55 Measures embedded in the complete set of 11 Domains.

TABLE 1.B

## Three Versions of KIDS with their corresponding Learning Domains [39]

| Learning Domains | KIDS 14 – ISBE REQUIRED VERSION | KIDS 5 | KIDS 11 |
|---|---|---|---|
| **Approaches to Learning – Self-Regulation** (ATL-REG) | ATL-REG-SED Subset (the first two domains are combined here [40]) | ATL-REG Domain | ATL-REG Domain |
| **Social and Emotional Development** (SED) | | SED Domain | SED Domain |
| **Language and Literacy Development** (LLD) | LLD Subset | LLD Domain | LLD Domain |
| **Cognition, Including Math and Science** (COG:MATH, COG:SCI)[41] | COG:MATH Subset | COG:MATH Domain | COG:MATH and COG:SCI Domains |
| **Physical Development** (PD) | | PD Domain | PD Domain |
| **Health** (HLTH) | | | HLTH Domain |
| **History – Social Science** (HSS) | | | HSS Domain |
| **Visual and Performing Arts** (VPA) | | | VPA Domain |
| **English Language Development** (ELD) | | | ELD Domain |
| **Language and Literacy Development in Spanish** (SPAN) | | | SPAN Domain |

# Endnotes


1. **American Psychological Association., National Council on Measurement in Education., & Joint Committee on Standards for Educational and Psychological Testing. (2014).** *Standards for educational and psychological testing.* American Educational Research Association. Note that the standards are now being revised.
2. **Ibid.**
3. Please note that the Validity Chapter in *Standards for Educational and Psychological Testing* includes several specific standards arranged by cluster. For example, Cluster 1," Establishing Intended Uses and Consequences," describes one major standard with seven subsections. We do not refer to any of these by number, but our descriptions that follow include many clusters and standards and their components.
4. The precise psychometric scoring techniques have not been publicly disclosed because of the "proprietary" nature of **KIDS** scoring. Scoring and reporting of **KIDS** is conducted by an external contractor. There is some public documentation describing in general terms the use of Item Response Theory in scoring. **Illinois State Board of Education. (n.d.).** *Summary of **KIDS** Measures and Reports.* Illinois State Board of Education. https://www.isbe.net/Documents_**KIDS**WebsiteResources/Summary_of_**KIDS**_Measures_and_Reports.pdf
5. **Illinois State Board of Education. (2023a).** *2022–23 School Year Illinois Kindergarten Individual Development Survey (**KIDS**) report.* Illinois State Board of Education. https://www.isbe.net/Documents/IL-**KIDS**-Report-2022-2023.pdf
6. **Lesaux, N. K., & Jones, S. M. (2016).** *The leading edge of early childhood education: Linking science to policy for a new generation* (Vol. 113, No. 12, pp. 2705-2738). Harvard University Press, Cambridge.
7. **Ibid.**
8. **Illinois State Board of Education. (2016).** *Illinois School Readiness Initiative.* https://www.isbe.net/Documents/**KIDS**-IL-School-Readiness-Initiative.pdf
9. **Regenstein, E., Connors, M., Romero-Jurado, R., & Weiner, J. (2017).** Uses and Misuses of Kindergarten Readiness Assessment Results. *Start Early Policy Conversations.* https://www.startearly.org/resource/uses-and-misuses-of-kindergarten-readiness-assessment-results/
10. **See** https://www.isbe.net/kidsadvisory
11. **Illinois State Board of Education (2017a).** *Guidance for Dual Language Learners.* https://www.isbe.net/Documents_**KIDS**WebsiteResources/LLD_Guidance.pdf
12. **Illinois State Board of Education. (2017b).** ***KIDS**: Participation or Exemption Guidance for Children with Special Needs.* https://www.isbe.net/Documents_**KIDS**Website Resources/**KIDS**_Exempt_Special_Needs_Guidance.pdf
13. **Illinois State Board of Education, 2023.** *Kindergarten individual development survey: User's guide and instrument.* https://www.isbe.net/Documents/**KIDS**-User-Guide-Instrument.pdf
14. ISBE published a short, undated, document called An Overview on Measures and Data Reporting that includes a very brief description of the IRT modelling used to score the three Development Areas/Domains. https://www.isbe.net/Documents_**KIDS**WebsiteResources/Summary_of_**KIDS**_Measures_and_Reports.pdf
15. Domain definitions are copied from page 3, https://www.isbe.net/Documents/IL-**KIDS**-Report-2022-2023.pdf
16. These bullet points reflect the uses that ISBE says are intended in the **User's guide: Illinois State Board of Education. (2023b).** *Kindergarten individual development survey: User's guide and instrument.* https://www.isbe.net/Documents/**KIDS**-User-Guide-Instrument.pdf
17. ISBE notes that **KIDS 11**, the longest version, may be used for summative assessment. We limit our discussion here to **KIDS 14**, the mandated version.
18. See page 3, **Illinois State Board of Education, 2023.** https://www.isbe.net/Documents/IL-**KIDS**-Report-2022-2023.pdf
19. **Illinois State Board of Education. (n.d.).** ***KIDS** frequently asked questions.* https://www.isbe.net/Documents_**KIDS**WebsiteResources/**KIDS**_FAQ.pdf
20. Play-based and playful learning is occasionally discussed or mentioned in ISBE reports and in external writings related to **KIDS**. We believe that this is an important idea worthy of deeper exploration. EC*REACH will soon turn greater attention to this topic in future writings, but we are not going to consider it here.
21. **Illinois State Board of Education. (n.d.).** *Strategies for collecting observations.* https://www.isbe.net/Documents_**KIDS**WebsiteResources/Strategies_for_Collecting_Observations.pdf
22. See pages 5-6, **Illinois State Board of Education, 2023.** https://www.isbe.net/Documents/IL-**KIDS**-Report-2022-2023.pdf?utm_source=chatgpt.com
23. **BEAR Center & WestEd. (2015).** *Preliminary **KIDS** Analysis: Three subscales Using a Subset of Measures Across Domains.* University of California, BEAR Center.

24. **Bowdon, J., Dahlke, K., Yang, R., Pan, J., Marcus, J., & Lemieux, C. (2019).** *Children's knowledge and skills at kindergarten entry in Illinois: Results from the first statewide administration of the Kindergarten Individual Development Survey.* https://files.eric.ed.gov/fulltext/ED599357.pdf

    The statistical methods used by these researchers are complex and we are unable to evaluate them. We refer readers to the original paper for more technical details.

25. An astute reviewer pointed this out to us.

26. Data reduction refers to many potential methods of simplifying multiple test items or survey responses into a smaller number of variables, such that each new variable correctly groups items or questions that appear to be measuring the same variable (construct). Ideally the test or survey designer would have these constructs in mind ahead of time and items or questions would be designed to capture these constructs. But that isn't always the case.

27. The Illinois Workforce and Education Research Collaborative (IWERC) was created in 2020 to study education research questions that speak to the entire state. IWERC is housed in the University of Illinois System. It has completed several major research studies addressing statewide concerns, including teacher shortages; access to and enrollment in computer education courses; and achievement test losses and recovery due to COVID in Illinois. https://dpi.uillinois.edu/applied-research/iwerc/

28. We highly recommend these papers for both their empirical findings and for their descriptions of **KIDS** and it uses.

    **Kiguel, S., Cashdollar, S., & Bates, M. (2024a).** Inequity in the early years: Student development trajectories from Kindergarten to Grade 3. *Kindergarten Readiness in Illinois Series.* Illinois Workforce and Education Research Collaborative (IWERC), Discovery Partners Institute, University of Illinois.

    **Kiguel, S., Cashdollar, S., & Bates, M. (2024b).** Trends and disparities in readiness using the Kindergarten Individual Development Survey (**KIDS**). *Kindergarten Readiness in Illinois Series.* Illinois Workforce and Education Research Collaborative (IWERC), Discovery Partners Institute, University of Illinois.

    **Illinois Workforce and Education Research Collaborative (n.d.).** ***KIDS*** *socioemotional domain also predicts Grade 3 test scores.* https://dpi.uillinois.edu/wp-content/uploads/2024/12/**KIDS**-SEL-1-pager.pdf

29. **Herring, W. A., Bassok, D., McGinty, A. S., Miller, L. C., & Wyckoff, J. H. (2022).** Racial and socioeconomic disparities in the relationship between children's Early Literacy Skills and third-grade outcomes: Lessons from a kindergarten readiness assessment. *Educational Researcher, 51*(7). https://doi.org/10.3102/0013189x221091535

30. It is possible that the 14 measures are being" clumped" together for other reasons that the actual content in them. They might be differentiated because of the type of evidence used in scoring the measures, or some non-obvious set of similarities or differences.

31. To the best of our knowledge, ISBE only releases **KIDS** scores for the three major Domains: ALT-REG-SED; LLD; and MATH. If scores were available for all 14 Measures inside those Domains, researchers would be able to learn considerably more about the technical properties of **KIDS**. And new insights may guide usage of **KIDS** scores.

32. We are not sure whether the Developmental Areas/Learning Domains/Major Constructs have detailed written definitions. The 14 Measures/Items each have a one sentence definition. The four Domains also have one sentence definitions. See Tables 1 and 2.

33. **Kiguel et al. (2024b).** observed a high correlation between age in months and **KIDS** scores in the fall. "Maturation" is called a "threat to validity" in non-experimental research studies.

34. This concept is called "consequential validity." Are there long or short-run harmful or unfair consequences of the use of **KIDS** scores?

35. **Regenstein, E. et al. (2017).**

36. **American Psychological Association (2014).**

37. **Illinois State Board of Education, 2023b.**

38. The shortened forms of **KIDS** (i.e., **KIDS 14** and **KIDS 5**) do not correspond to the shortened versions of DRDP-K (i.e., DRDP Essential, which contains 33 measures, and DRDP Fundamental, which includes 37 measures). However, **KIDS 11** and the comprehensive version of DRDP-K contain the same 55 measures.

39. This table is slightly revised from the original ISBE table that appears in multiple publications. We made these revisions to make our descriptions below easier to write and read. There is one more revision to be made. The row with MATH and COG:SCI should be turned into two rows, so that MATH and SCI each have their own row.

40. In **KIDS 14**, ATL-REG and SED Domains are combined into a Developmental Area, ALT-REG-SED. LLD and MATH are also called Developmental areas, so that KID14 captures these three Developmental Areas.

# Acknowledgements

First and foremost, we thank the CPS Office of Early Childhood Education, and its Chief Officer, **Leslie McKinily** for encouraging this work. McKinily and her staff are keenly interested in the assessments administered in grades P-2 and ask excellent questions about the strengths and weaknesses of the many different assessments that young students take. This paper is a small contribution to the OECE's insistence on making data- and evidence-based decisions.

We were extraordinarily fortunate for advice, recommendations, and direction from several nationally recognized testing and assessment experts. These professors are among the most influential thinkers about assessment and psychometrics of the late 20th century and the early 21st century.

**Edward Haertel** (Professor Emeritus, Stanford University) was brave and kind enough to read the first complete draft of the paper and steer us away from misleading or confusing language and keep us up to date on contemporary validity theory.

**Mark Wilson** (Professor, Berkeley School of Education, UC Berkeley) provided very helpful information about the early days of **KIDS** and its relationship to the Desired Results Developmental Profile – Kindergarten (DRDP-K). Wilson was also generous to share unpublished technical memos that we found very useful.

**James Pellegrino** (Professor Emeritus and Founding Co-Director, Learning Sciences Research Institute, University of Illinois Chicago) contributed greatly to this paper. His detailed trenchant feedback, specific thoughts on adding or subtracting content, and detailed editorial suggestions were amazingly helpful.

**Peter Leonard** is Director of Student Assessment and Multi-Tiered Support Systems at CPS and is an active member of the National Council on Measurement in Education. His contributions reflect his technical expertise paired with the practical experience of developing and overseeing a balanced assessment system for CPS.

We were very fortunate to have a wonderful group of readers and reviewers. We received important critical responses to successive drafts of the paper. Since all these reviewers provided important feedback, they are listed in alphabetic order here:

**Meg Bates** (Illinois Workforce and Education Collaborative); **Angela Farwig** (Illinois Action For Children) **Rachel Mayo** (Office of Early Childhood Education, CPS. In addition to their unique contributions, co-authors **Sophie Magro, Elliot Regenstein** (Foresight Law and Policy), and **Terri Sabo**l provided extensive feedback on all parts of this paper.

Many of our colleagues at EC*REACH provided invaluable and essential assistance in completing this paper. **Diane Schanzenbach,** EC*REACH Faculty Co-Director, proofread this paper with exceptional attention to detail regarding consistencies for grammar, punctuation, and content.

**Isabel Farrar**, Senior Program Administrator, helped at every step in the writing and production of the paper. **Jordan Herron**, Partner Engagement Manager, helped guide all our design decisions. Undergraduate assistants, **Hannah Pease** and **Sarah Serota** helped with the text, footnotes, and references. Easton thanks his friend and colleague **Barton Dassinger**, Principal, Chavez Multi-cultural Elementary School, for his ongoing advice, feedback, guidance, encouragement, and insider knowledge of "all things CPS."

In the summer of 2025, while the major content of this paper was written, the entire Assessment staff at ISBE was engaged in a high priority task and therefore unable to review our work to date and offer suggestions and request corrections. In the fall of 2025, when the important task was completed, two influential ISBE staff members – Rae Clementz, Director of Assessment, and Jason Heifer, Special Advisor to the State Superintendent – were able to engage with us.

We were fortunate to engage **Maureen Kelleher,** veteran newspaper reporter and education writer and editor. She did a superb job in making this paper more readable and coherent.

The authors are proud of this paper's stylish appearance. We thank **Jeff Hall** for designing it so beautifully. Jeff has helped EC*REACH create its visual "identity". We plan to continue to use to Jeff in the future jh.design@rcn.com

**John Q. Easton**, Corresponding Author john.easton@northwestern.edu

**Cite as: Easton, J.Q., Magro, S., Pellegrino, J., Rauner, D., Regenstein, E., Sabol. T., & Steans, R. (2026).** *A Close Look at the Kindergarten Individual Development Survey (KIDS).* Chicago, IL: Early Childhood Research Alliance of Chicago

## Author Information

**John Q. Easton**
*Interim Executive Director*
EC*REACH
Institute for Policy Research
Northwestern University
john.easton@northwestern.edu

**Sophia Magro**
*NIMH T32 Postdoctoral Fellow*
Institute for Innovations in Developmental Sciences
Feinberg School of Medicine
Northwestern University
sophia.magro@northwestern.edu

**James Pellegrino**
*Professor Emeritus*
*Founding Co-Director*
Learning Sciences Research Institute
University of Illinois Chicago
pellegjw@uic.edu

**Diana Mendley Rauner**
*President*
Start Early
drauner@startearly.org

**Elliot Regenstein**
*Partner*
Foresight Law + Policy
Elliot.Regenstein@FLPAdvisors.com

**Terri J. Sabol**
*Associate Professor*
Human Development and Social Policy
School of Education and Social Policy
*Faculty Fellow*
Institute for Policy Research
*Faculty Director*
EC*REACH
Northwestern University
terri.sabol@northwestern.edu

**Robin Steans**
*President*
Advance Illinois
rsteans@advanceillinois.org




## EC*REACH Funders in Alphabetical Order

**Annie E. Casey Foundation**
**Anonymous**
**Anonymous**
**The Brinson Foundation**
**Calico Foundation**
**CME Group Foundation**
**Crown Family Philanthropies**
**Foundation for Child Development**
**Lewis-Sebring Family Foundation**
**Schreiber Philanthropy**
**The Spencer Foundation**
**Start Early**
**Steans Family Foundation**
**ViVO Foundation**

Northwestern
EC★REACH
EARLY CHILDHOOD RESEARCH
ALLIANCE OF CHICAGO

626 Library Place
Evanston IL 60208
**Phone** (847) 467-0548
**Web** ecreach.org

## OUR MISSION

The mission of Early Child Research Alliance of Chicago (EC*REACH) is to unite diverse partners across sectors, agencies, and institutions to co-construct and conduct action-oriented research that promotes equitable solutions for early childhood policy and practice in Chicago. EC*REACH is hosted by the Institute for Policy Research at Northwestern University.

To learn more, visit www.ecreach.org.