



One Lay of the Land: Preliminary Findings from Two High Level Scans

Nathan Dadey & Calvary Diggs

The National Center for the Improvement of Educational Assessment

September 26th, 2019

Reidy Interactive Lecture Series

Portsmouth, NH



This work is licensed under a
Creative Commons Attribution 4.0
International License.

The landscape of interim assessments has changed substantially over the last 20 years

due, in part to,

▶ Policy

▶ Technology

▶ Scale

The goal of this work is to share **two high level scans** that hopefully help us get a better understand that landscape.

1 State Supported Interim Assessment Options
Summary of interim assessments supported by state department of educations, as reported on department websites.

2 Review of Available Literature on Interim Assessments
A review of published literature dealing with the use of interim assessments.

Summary of State Supported Interims

State Supported Interim Assessments

- Increasingly, states are supporting schools and districts by providing access to or guidance on interim assessments.
 - We think this is reflective of a shift towards systems of assessments. That is, many states are now considering how assessments they can influence can work together.
- Given this, goal of this work is to understand what types of assessments are being supported by examining all **50 state department of education websites**.

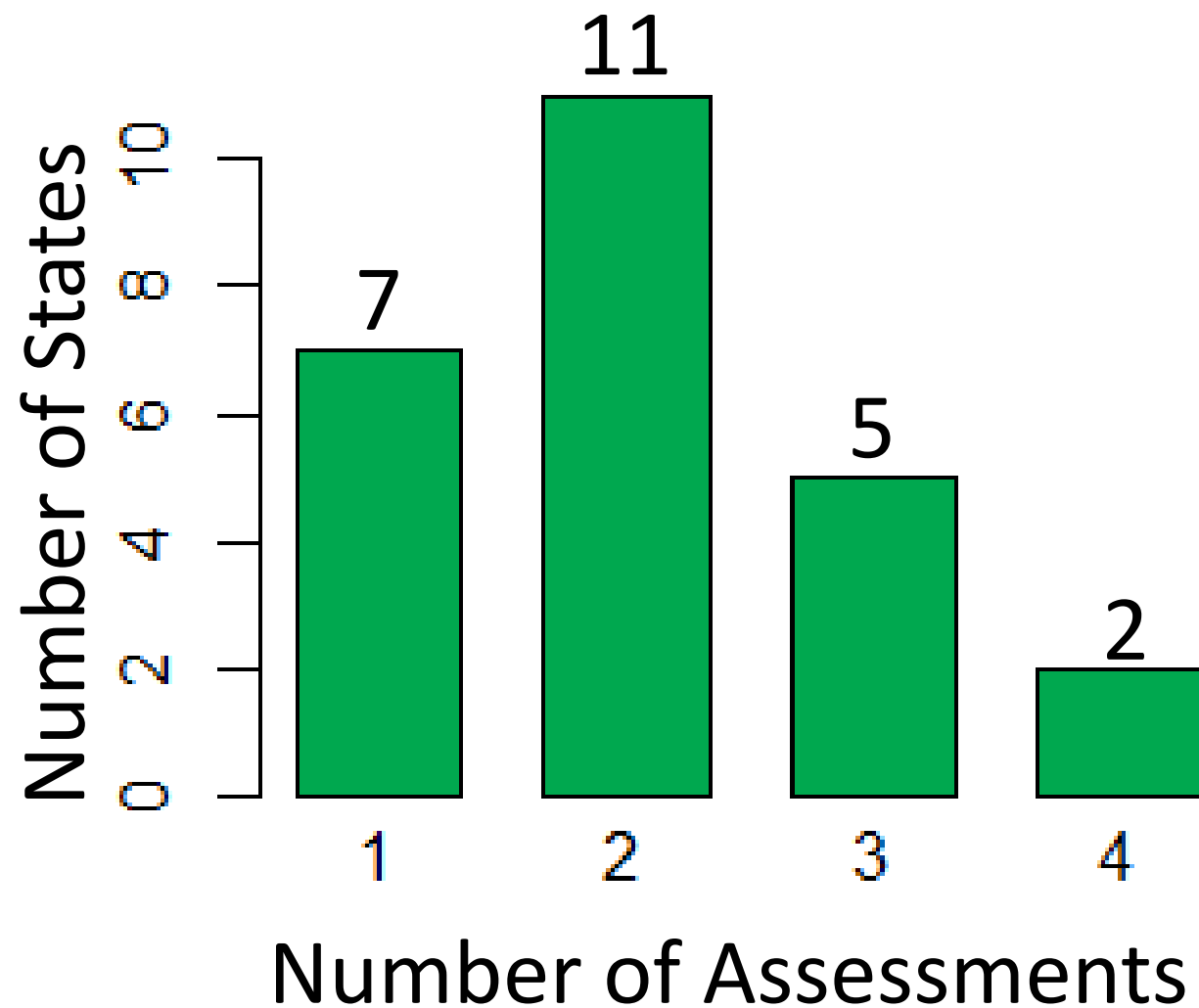
Note: a previous version of this work was shared at the 2018 NCME Classroom Assessment Conference and the 2018 Reidy Interactive Lecture Series.

25

state websites had
information on
34 unique interim
assessments.

Notes: The definition of interim is used quite broadly here (e.g., included early literacy screeners). Each lists of approved assessments are counted as a single assessment in the above counts. At least five states had materials that addressed the idea of systems of assessments: [LA](#), [MI](#), [RI](#), [VT](#) and [WY](#).

The number of assessments per state ranged from 1 to 4, with a median of 2.



Most states (22)
provided assessments
directly to districts,
often through online
platforms.

The remaining 3
states provided lists of
approved assessments:

- Early Literacy: [CO](#), [MI](#) & [SC](#)
- K-12 Math & ELA: [SC](#)

Distinguishing further among the assessments
involves considering the
interaction of **content** and **time**.

I.e., How, if at all, is the content domain divided up
and assessed over time?

Note: the paper *Matching Instructional Uses with Interim Assessment Designs* (Gong, 2019) dives deeply into this issue and focuses on the claims. Gong ([2010](#)) also explores this issue, as does Dadey & Gong ([2017](#)).

Multiple terms are floating around in the literature to address these types of design choices.

Mini-Interim

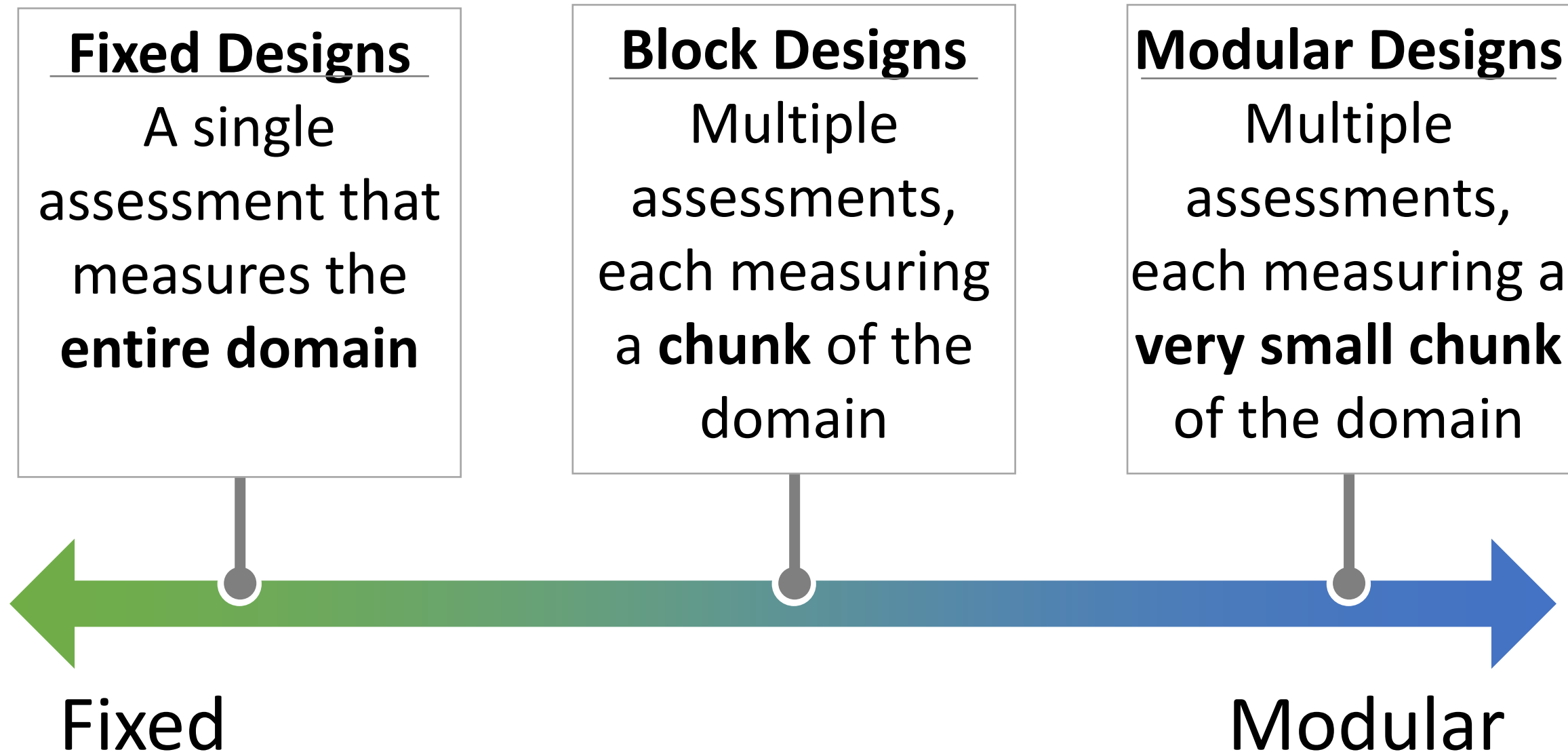
Modular

Different Blueprint

Mini-Summative

Micro-Modular

Same Blueprint



Design	Number of		
	States	Assessments	Unique Assessments
Fixed	19	23	14
Block	13	15	6
Modular	8	11	11

Notes: List of approved assessments are excluded in the above counts.

Literature Review of Interim Assessment Use

“School districts across the country have [rapidly] adopted interim and benchmark assessments... Thus, it is especially important that well-conceived, empirical studies of the effects of such programs be carried out.”

(Shepard, [2010](#))

Our suspicion is that research on the uses of interim assessments has been emerging.

The goal of this project is to provide a **review** of this body of literature, with an emphasis on how the assessments are **used**.

Literature Review

4,000



141



20

articles eligible for inclusion after cleaning, based on searches of key terms & assessment names within ASP and ERIC (which limits the review).

articles eligible for coding, after applying [inclusion/exclusion rules](#), within additional articles included with certainty.

articles coded thus far.

Studies by Year

(N = 141)

Number of Studies

30

20

10

0

2001

2002

2003

2004

2005

2006

2007

2008

2009

2010

2011

2012

2013

2014

2015

2016

2017

2018

Year



Focusing on Use for the Coded Studies

We coded each discrete use within each of the 20 studies as:

1. **Instructional, Evaluative or Predictive¹** following Perie, Marion, & Gong ([2009](#)), and then
2. Created **additional codes with each category**, to better describe the specific ways in which the assessments are use following a grounded approach (following Hook, [2015](#)).

¹Notes: In cases where predictions are used in a clearly instructional or evaluate manner, we have retained them within the predictive category.

Uses by Study (n=20)

20

Studies contained at least one
instructional use.

16

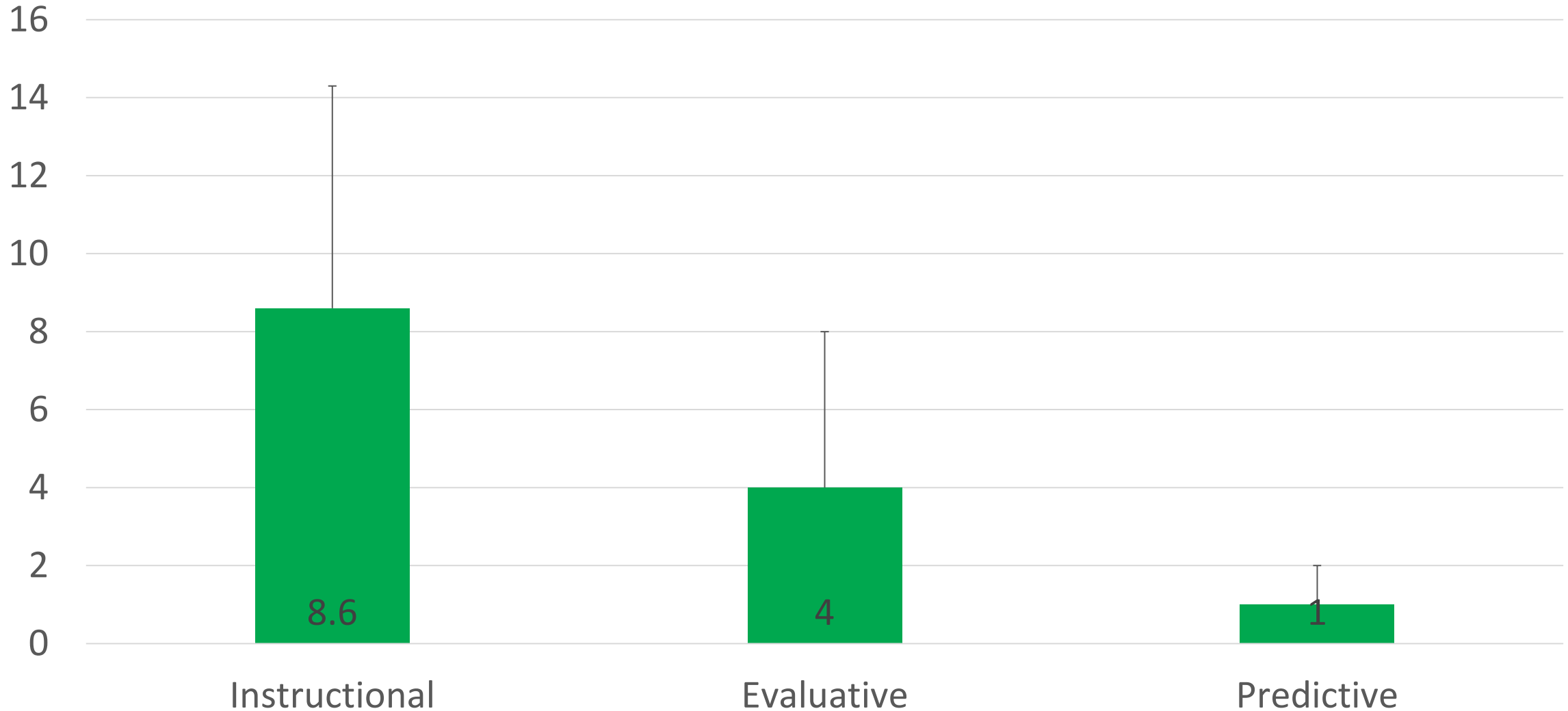
Studies contained at least one
evaluative use.

10

Studies contained at least one
predictive use.

Number of Coded Discrete Uses by Study

Average Number of discrete uses within study



Top Five Discrete Instructional Uses

Five Most Frequent Instructional Uses Evident in the Research on Interim Assessments

	Use Description	<i>k</i>	Example
1.	Broad claim to modify or improve instruction	14	A total of 86% of teacher reported modifying instruction based on interim assessment results (Clune & White, 2008).
2.	Identify students to provide additional support	13	Results were used to identify students for supplemental instruction (e.g., software, working with volunteers, afterschool tutoring; Shepard et al [2011]).
3.	Identify what content to reteach	10	The administration hoped teachers would reteach with new strategies (Bulkley et al., 2010).
4.	Improve score on the summative assessment	10	Guide schoolwide improvement efforts to meet AYP (especially in low-performing schools; Bulkley et al., 2010).
5.	Differentiate instruction	9	Identifying students with similar patterns of performance on the assessment and using that to constructs groups to differentiate instruction (Blanc et al., 2010)

Note. Table created based on coding a total of 20 studies featuring interim assessments.

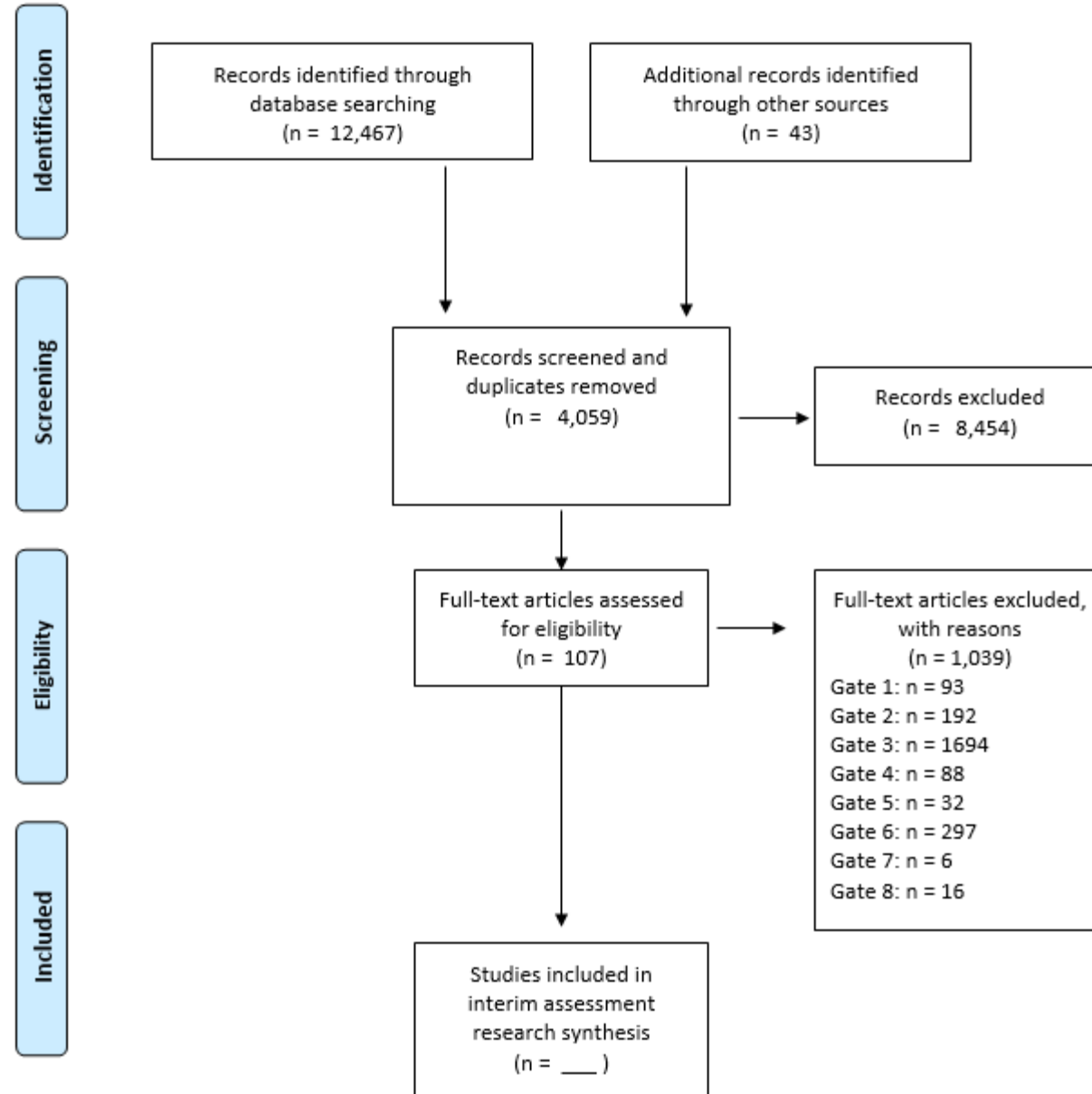
Preliminary Conclusions

- Interim assessments are used in a multitude of ways.
- Description of assessment use is widespread, examination of efficacy is not.
- Much of the way in which use is describe is unspecific, meaning evaluation is difficult even when it is conducted as we have little idea about *how* the interim assessments are used.

Supplemental Slides



PRISMA 2009 Flow Diagram



Inclusion Criteria

1. English language, conducted in the US, published since 2000
2. Conducted with a population from K-12 settings
3. An assessment was a key component of the study and functioned as an IV
4. The assessment was academic in nature
5. The assessment was administered by school or affiliated staff for school use (i.e., not administered by parents, clinics, outside consultants)
6. The assessment can be described as an “interim assessment”
 - i. multiple measurements points or used mid-way through the term
 - ii. assessment data were used for a purpose that may be broadly described as evaluative, instructional, or predictive
 - iii. assessment data can be (or were) aggregated to represent the entire grade, school, district, or state
 - iv. if not aggregated, data would still be relevant to a purpose (evaluative, instructional, predictive) at a higher level (e.g., school, district, state)
7. The interim assessment is/was commercially available or was developed by a school or system(s)
8. The study evaluates an interim assessment using a traditional experimental or quasi-experimental design, observational methods, or recollections.

Exclusion (The inclusion criteria were designed to identify studies and assessments that fall under our current typology of an, “interim assessment.” In the current conceptualization, interim is used to denote assessments that are not clearly used for formative or summative purposes. However, there are various ways that each of these concepts could be defined, which, in turn, would provide a different approach to the inclusion and exclusion process. This study aims to provide a wide, scoping search of the literature. Exclusion criteria for both summative and formative assessments were defined as follows,

Focal assessment is clearly used for formative purposes (i.e., teacher administered, informal, conducted frequently, and data can not be aggregated and meaningfully generalized beyond the individual classroom)

Focal assessment is clearly used for summative purposes (i.e., an end of the year state test)

The following criteria were set to allow studies who used interim assessments for traditionally summative or formative purposes, thereby allowing the study to explore the range of their utility for a variety of applications.

Method

- Literature Review
 - Rapid review (Grant et al., 2009)
 - Four means of identification
- Inclusion Criteria
 - “Interim assessment”
 - Additional Gate
- Coding
 - General characteristics
 - Discrete uses

