



Improving the Selection, Use and Evaluation of Interim Assessments

National Center for the Improvement of
Educational Assessment

Reidy Interactive Lecture Series (Introduction)

Portsmouth, NH

September 26-27, 2019



The Reidy Interactive Lecture Series

Named for a famous Kentucky educational leader, Ed Reidy, RILS brings together **participants** with a range of expertise to wrestle with difficult challenges in search of practical solutions or promising approaches. Participants are **encouraged to participate!**



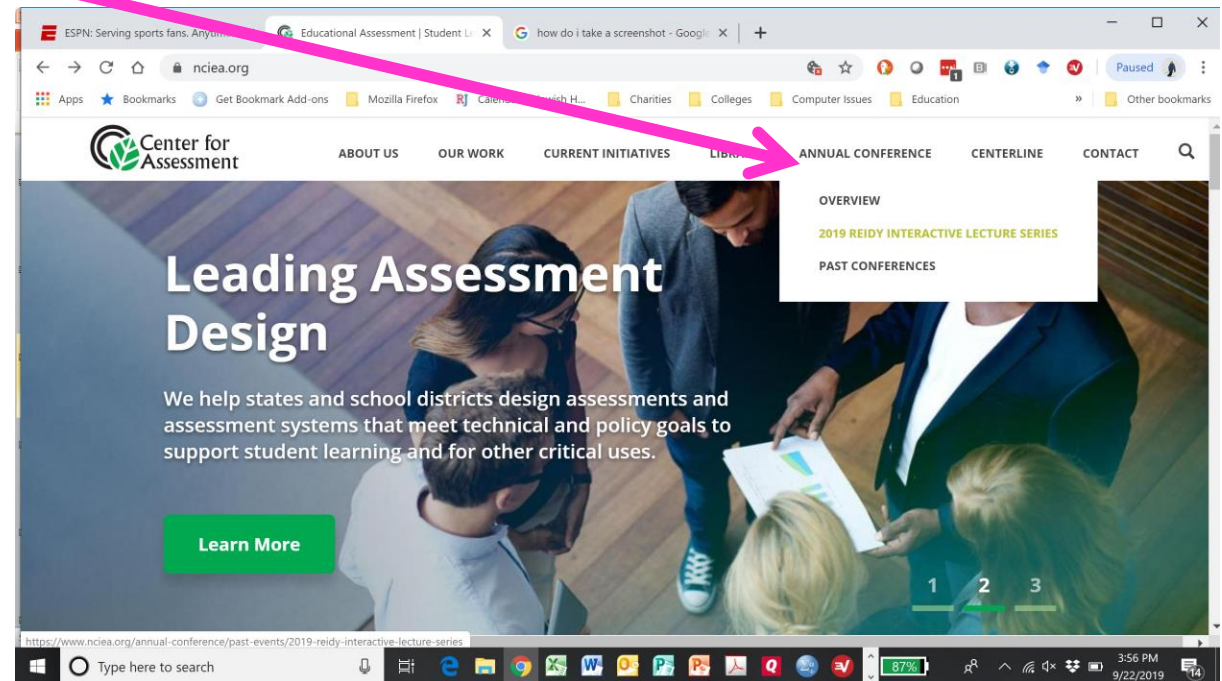
Marion. RILS. September 27, 2018

Overview

- Quick intro to the Reidy Interactive Lecture Series
- Why Interim Assessments?
- Flyover of the Conference

Agenda, Papers, and Presentations

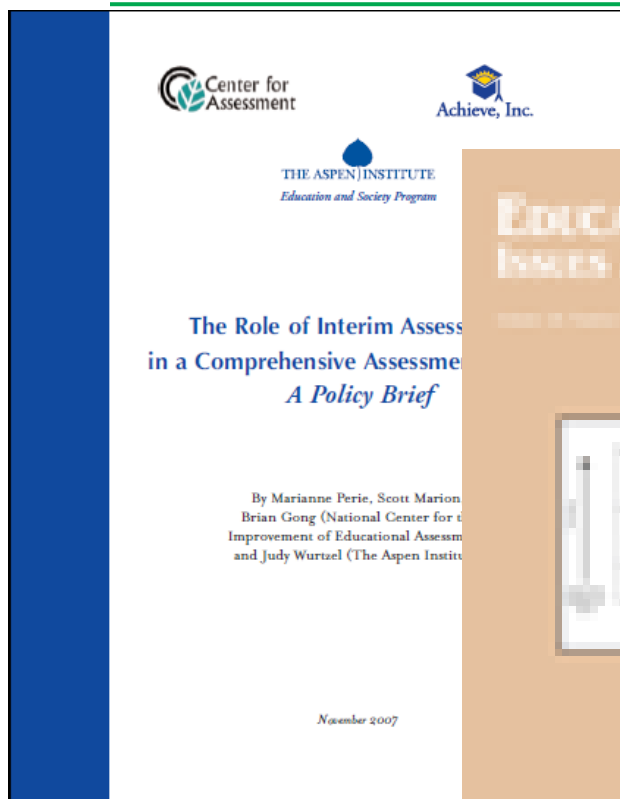
- <https://www.nciea.org/annual-conference/past-events/2019-reidy-interactive-lecture-series>
- Or just go to: www.nciea.org



Who's Here?

- States leaders and assessment/accountability personnel
- Districts leaders and assessment/accountability personnel
- Assessment Companies
- Teachers and Principal
- Consulting Firms/TA Providers/Advocates
- Universities/Research Institutions
- Center for Assessment Board of Directors

The Center's Long History With Interim Assessments

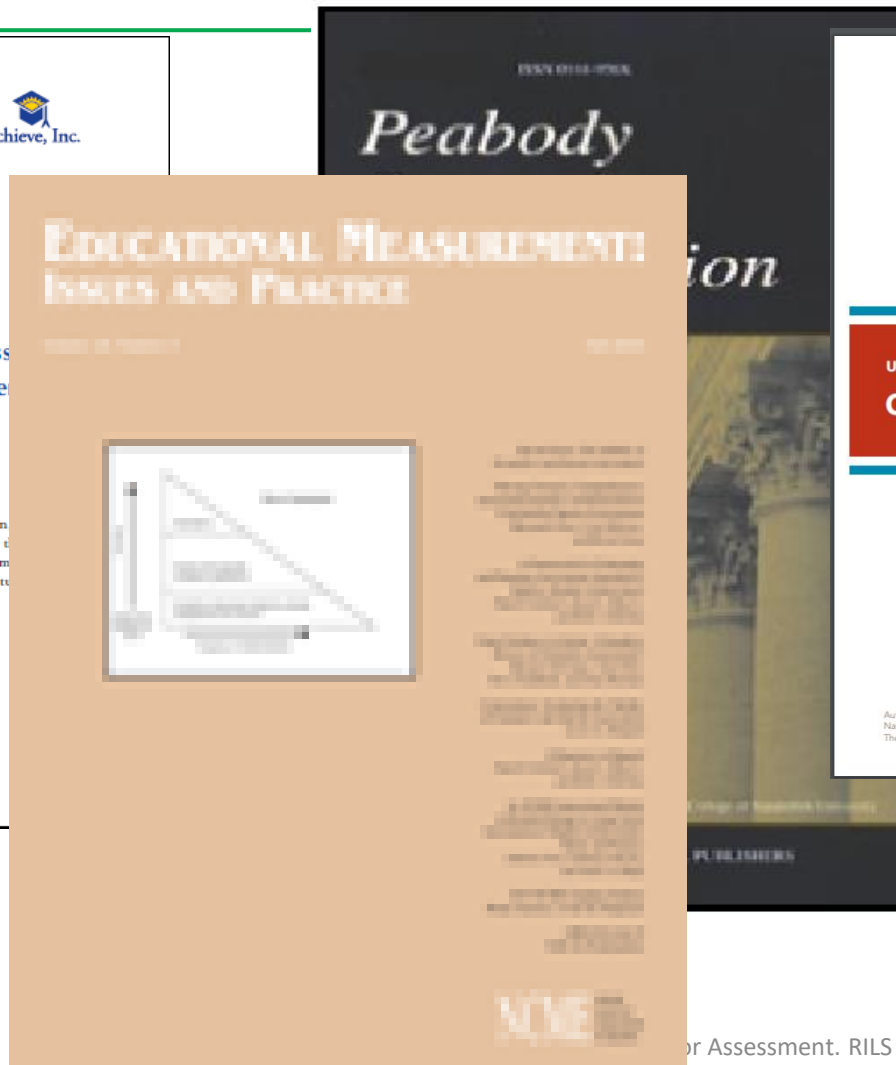


Center for Assessment
Achieve, Inc.
THE ASPEN INSTITUTE
Education and Society Program

The Role of Interim Assessment in a Comprehensive Assessment: A Policy Brief

By Marianne Perie, Scott Marion, Brian Gong (National Center for the Improvement of Educational Assessment), and Judy Wurtzel (The Aspen Institute)

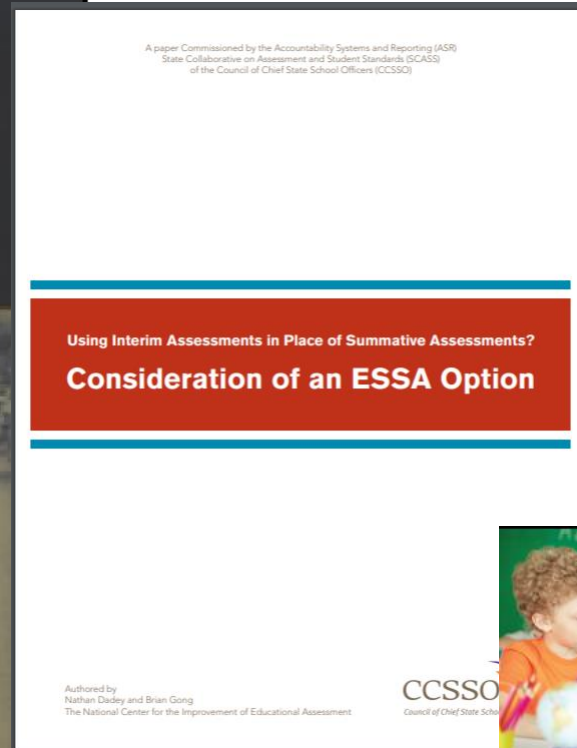
November 2007



Peabody

EDUCATIONAL MEASUREMENTS: ISSUES AND PRACTICES

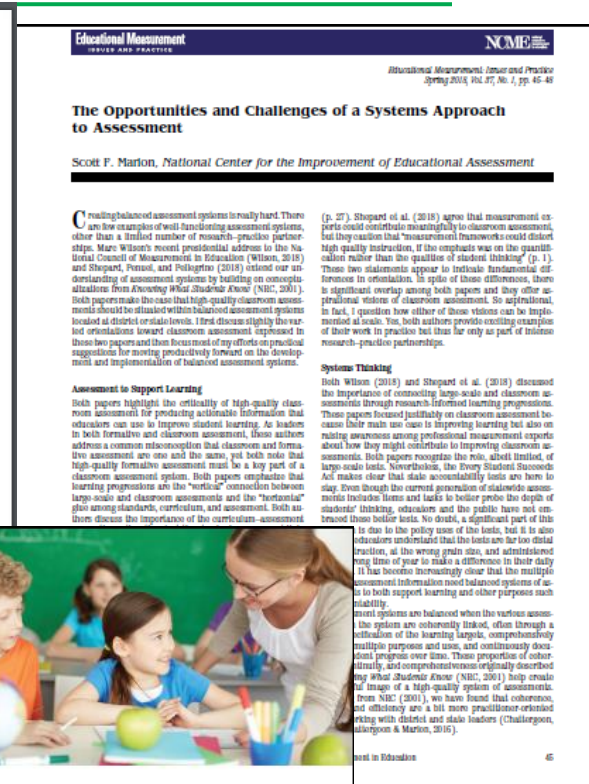
NCME



A paper Commissioned by the Accountability Systems and Reporting (ASR) State Collaborative on Assessment and Student Standards (SCASS) of the Council of Chief State School Officers (CCSSO)

Using Interim Assessments in Place of Summative Assessments? Consideration of an ESSA Option

Authored by Nathan Daxley and Brian Gong
The National Center for the Improvement of Educational Assessment



Educational Measurement: Issues and Practices
NCME

The Opportunities and Challenges of a Systems Approach to Assessment

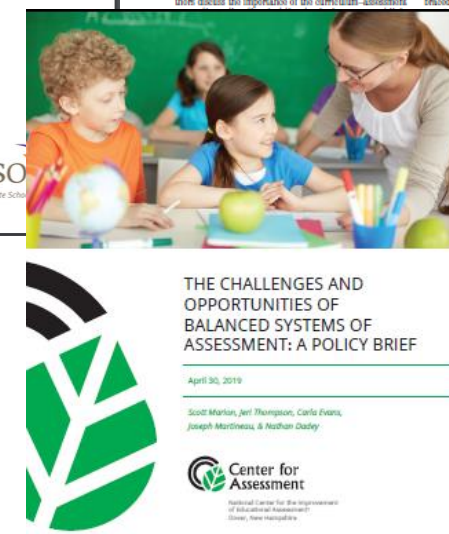
Scott F. Marion, National Center for the Improvement of Educational Assessment

Creating balanced assessment systems is really hard. There are few examples of well-functioning assessment systems, other than a limited number of research-practice partnerships. Marc Wilson's recent presidential address to the National Council of Measurement in Education (Wilson, 2018) and Shepard, Porock, and Philpotts (2018) extend our understanding of assessment systems by building on conceptual advances from *Knowing What Students Know* (NIC, 2011). Both papers make the case that high-quality classroom assessments should be situated within balanced assessment systems located at district or state levels. I first discuss slightly the varied orientations toward classroom assessment expressed in these two papers and then focus most of my efforts on practical suggestions for moving productively forward on the development and implementation of balanced assessment systems.

Shepard et al. (2018) agree that measurement experts could contribute meaningfully to classroom assessment, but they caution that "measurement frameworks could distort high-quality instruction, if the emphasis was on the quantification rather than the quality of student thinking" (p. 1). These two statements appear to indicate fundamental differences in orientation. In spite of these differences, there is significant overlap among both papers and they offer aspirational visions of classroom assessment. So aspirational, in fact, I question how either of these visions can be implemented at scale. Yes, both authors provide exciting examples of their work in practice but thus far only as part of isolated research-practice partnerships.

Systems Thinking
Both Wilson (2018) and Shepard et al. (2018) discussed the importance of connecting large-scale and classroom assessments through research-informed learning progressions. These papers focused justifiably on classroom assessment because their main use case is improving learning but also on raising awareness among professional measurement experts about how they might contribute to improving classroom assessments. Both papers recognize the role, albeit limited, of large-scale tests. Nevertheless, the *Every Student Succeeds Act* makes clear that state accountability tests are here to stay. Even though the current generation of statewide assessments include items and tasks to better probe the depth of students' thinking, educators and the public have not embraced those better tests. No doubt, a significant part of this is due to the policy uses of the tests, but it is also educators' understanding that the tests are too distal, fraction, at the wrong grain size, and administered long into the year to make a difference in their duty. It has become increasingly clear that the multiple assessment information used balanced systems of use to both support learning and other purposes such as accountability.

Interim systems are balanced when the various assessment systems are coherently linked, often through a delineation of the learning targets, comprehensively multiple purposes and uses, and continuously documented progress over time. These properties of coherence, and comprehensiveness originally described by *Knowing What Students Know* (NIC, 2011) help create a map of a high-quality system of assessments. From NIC (2011), we have found that coherence and efficiency are a bit more practitioner-oriented (along with district and state leaders (Challenger, Alving, & Marion, 2016)).



THE CHALLENGES AND OPPORTUNITIES OF BALANCED SYSTEMS OF ASSESSMENT: A POLICY BRIEF

April 30, 2019

Scott Marion, Jeff Thompson, Carlo Evans, Joseph Martineau, & Nathan Daxley

Center for Assessment
National Center for the Improvement of Educational Assessment
www.ncei.org

Our Board's Long History As Well



December 2009
NB-51

Graduate School
of Education
University
of Pennsylvania

Can Interim Assessments be Used for Instructional Change?

Margaret E. Goertz, Leslie Nabors Oláh, and Matthew Riggan

The past ten years have witnessed an explosion in the use of interim assessments by school districts across the country. A primary reason for this rapid growth is the assumption that interim assessments can inform and improve instructional practice and thereby contribute to increased student achievement. Testing companies, states, and districts have become invested in selling or creating interim assessments and data management systems designed to help teachers, principals, and district leaders make sense of student data, identify areas of strengths and weaknesses, identify instructional strategies for targeted students, and much more. Districts are keeping their interim tests even under pressure to cut budgets (Sawchuk, 2009). The U.S. Department of Education is using its Race to the Top program to encourage school districts to develop formative or interim assessments as part of comprehensive state assessment systems.

Much of the rhetoric around interim assessments paints a rosy picture. Supporters argue that these tests will provide data on student understanding; teachers' analysis of this data will in turn lead to greater differentiation of instruction and better teaching of content, leading to improved student learning. Much of the belief in the potential of interim assessments comes from the body of research on formative assessment, particularly those studies showing that "short-cycle" formative assessments—largely those that are based on information collected by teachers in their classrooms and that provide feedback to teachers within a single class period—are a powerful means to improve the quality of teaching and raise student performance (Black & William, 1998; Crooks, 1988; Hattie & Timperley, 2007; Natriello, 1987).

However, this research has not focused on interim assessments, but rather on practices that are embedded within classroom instruction.

Very little research exists on how interim assessments are actually used by individual teachers in classrooms, by principals, and by districts. Some recent studies surveyed teachers about their use of test data in instruction. Many of these teachers reported that interim test results helped them monitor student progress and identify skill gaps for their students, and led them to modify curriculum and instruction (cf. Christman, et al., 2009; Chase & White, 2008; Stecher, et al., 2008). These studies, however, did not examine how individual teachers actually analyzed and used these data to inform their classroom practice, the policy conditions that supported teachers' ability to use interim assessment data to improve instruction, or the interaction of interim assessments with other classroom assessment practices. Our study begins to fill that vacuum.

The purpose of this exploratory study was to examine the use of interim assessments and the policy supports that promote their use to change instruction, focusing on elementary school mathematics. We use the term "interim assessments" to refer to assessments that: a) evaluate student knowledge and skills, typically within a limited time frame; and b) the results of which can be easily aggregated and analyzed across classrooms, schools, or even districts (Perie, Marion, & Gong, 2009). Our study looked at how 45 elementary school teachers in a purposive sample of 9 schools in 2 districts used interim assessments in mathematics in 2006-07. The study focused on teachers' use of data in a cycle of instruction-

This Policy Brief was derived from an extensive research report titled, *From Testing to Teaching: The Use of Interim Assessments in Classroom Instruction*. Visit CPRE's website (www.cpre.org) to download a free copy.

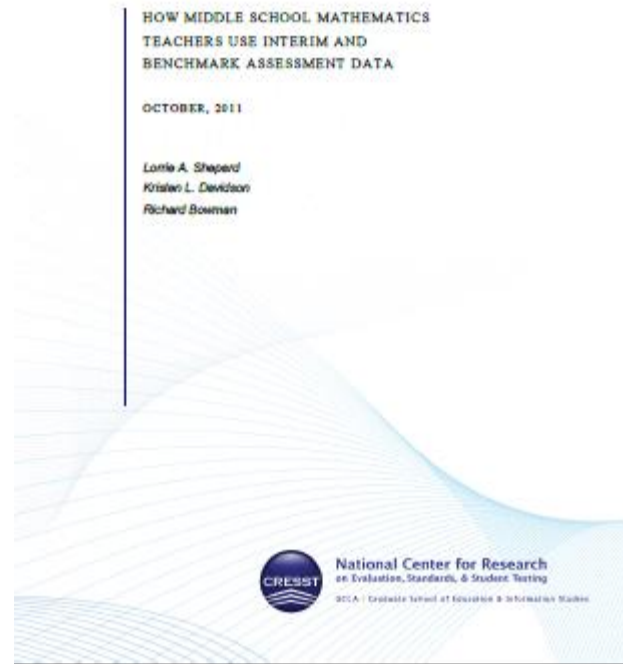
Center for Policy Research in Education
University of Pennsylvania
Teachers College
Columbia University
Harvard University
Stanford University
University of Michigan
University of Wisconsin-Madison
Northwestern University

CRESST REPORT 807

HOW MIDDLE SCHOOL MATHEMATICS TEACHERS USE INTERIM AND BENCHMARK ASSESSMENT DATA

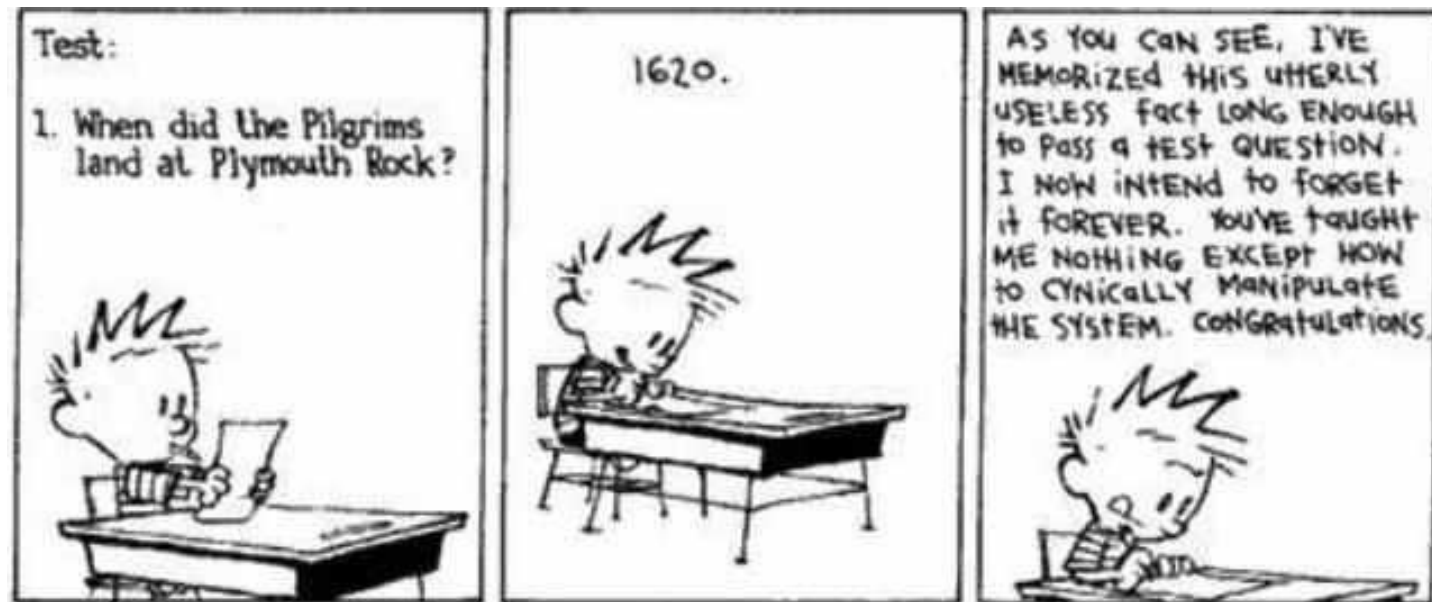
OCTOBER, 2011

Lorrie A. Shepard
Kristen L. Davidson
Richard Bowman



Why Interim Assessments for RILS?

- We've seen a massive increase of interim assessments in the past 20 years.
- We and others have raised questions about many aspects of interim assessment technical quality, use and utility.
- Are all these district and school leaders wrong?
- Clearly not, but even our colleagues in the interim assessment business would acknowledge the considerable range in implementation quality and use.



What's Coming?

- This is truly a hands-on and minds-on conference
- State and district **case studies** to help ground us in reality
- Deep examination and experience with a **set of tools** designed to help district/school leaders improve the selection, interpretation, and use of interim assessments
- Collegial and challenging conversations among participants and audience members (this is an **interactive** conference!)

What's Coming: Day 1 Agenda

8:45	Balanced Assessments Systems, Interim Scan, and Goals for RILS
9:45	Table Discussion – Purpose and Uses of Interim Assessments
10:45	Break
11:00	Case Study: Wyoming
12:00	Lunch
1:00	Introduction to the Toolkit
1:15	Case Study: School Administrative Unit #9, New Hampshire
1:45	Phase I of the Toolkit: Identifying Gaps and Needs
3:00	Break
3:15	Phase II of the Toolkit: Determining Assessment Features and Characteristics
4:45	Adjourn

Thanks!

- To all of you for joining us!
- To Juan and Erika for dedicated leadership and endless planning
- To LauraLee, Erin, and Sandi for being organizational ninjas
- To our Board of Trustees for support and wisdom
- To my amazingly brilliant colleagues for being the “Center”



Tonight

- Please join us this evening for cocktails and hors d'oeuvres at 5:15 right in this room

