Shifting the Paradigm of Large-Scale Achievement Assessment or, Help! I'm Lost; Does Anyone Have a Map?

Neal Kingston, Professor, Dept. of Educational Psychology;
Director, Achievement and Assessment Institute, University of Kansas

I had to get to the Lied Lodge in Nebraska City yesterday. But knowing where I was going was not enough.

I did not want to be like Daniel Boone, the famed frontiersman and explorer who when asked if he was ever lost said, "I can't say as ever I was lost, but I was bewildered once for three days" (Faragher, 1993, p. 65).

I actually knew a bit more than where I was going – I knew I was starting in Lawrence, Kansas. I knew even more than that. I knew I have to head west to Topeka then north to Nebraska City. But frankly that was not enough; I really needed a map. Or better yet, a GPS, because there could be detours or traffic jams.

I am an educational researcher – a psychometrician, not a geographical cartographer. Psychometricians apply statistical models to assessment responses and summarize those responses in a way that facilitates valid inferences about psychological and educational constructs underlying those assessments. In the field of psychometrics, the science of testing, at best we use road signs, or lists of standards that a child must master.

Road signs are not sufficient.

Unfortunately, there has been little direct connection between the subject matter we assess and statistical methods we have used to analyze test results and build test forms.

We have one set of methods for identifying the content we put on our tests and a completely unrelated set of methods for analyzing and producing results. Those results are shared on a numerical scale with no inherent meaning, such as 200 –

Figure 1. Section of Google Map showing geographical relationship between Lawrence, KS and Nebraska City, NE.

800, or 1 – 36, or 120 – 170, which are the score scales for the SATs, ACTs, and GREs respectively.

What does a student with a score of 340 on the SAT Mathematics test know? Do two students with the score of 27 on the ACT know the same things? And while the examples I gave are for admission tests that are not intended to provide feedback to improve student learning, the same thing is true about tests intended to provide such information, such as K-12 state tests or tests used to certify examinee knowledge.

Just like maps can help us navigate in the physical world, maps, could help us navigate the cognitive world.

The Dynamic Learning Maps Alternate Assessment was developed at the University of Kansas Achievement and Assessment Institute working closely with The Center for Literacy and Disability Studies at the University of North Carolina Chapel Hill Medical School.

Alternate Assessments are designed for students with significant cognitive disabilities – the approximately 1% of students with the greatest intellectual challenges. These children are often invisible within their greater communities, being educated in separate classrooms and often in separate residential facilities. Most suffer from either intellectual disabilities, usually with IQs below 55, severe autism, or other conditions that interfere with their ability to process information. About ten percent of the alternate assessment population have no formal communication systems. There is high co-morbidity with single and multiple physical disabilities including motor control, deafness, and blindness.

The DLM was administered for the first time in spring 2015 to 15 of 18 states that have collaborated in building the assessment. We expect some additional states will use DLM next year.

There are a number of key principles underlying the development of the DLM Alternate Assessment, but today I will focus on the use of learning maps.
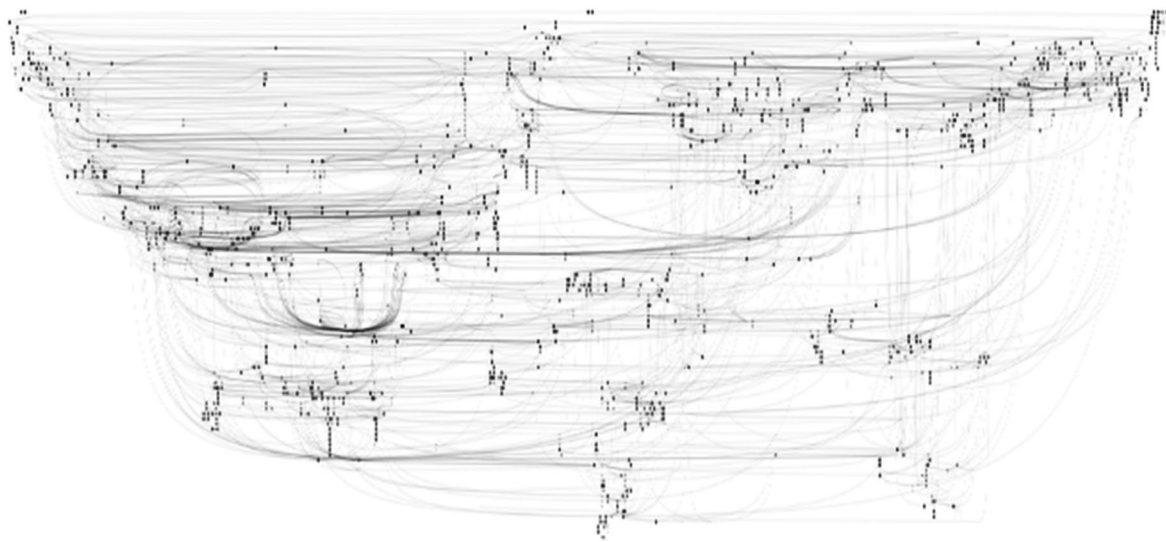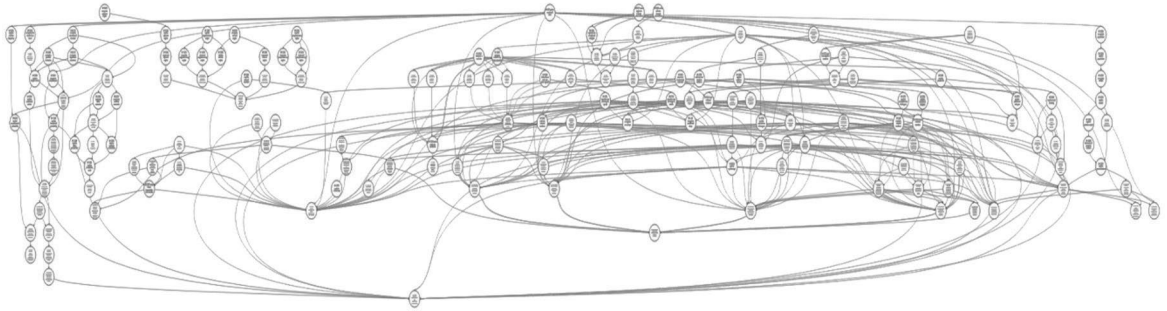

Figure 2. Mathematics learning map

Figure 3. 7th Grade mathematics portion of the learning map

We had a team of about a dozen researchers spend more than two years scouring the research literature for studies showing what we know about how students learn academic content in English Language Arts and Mathematics. We had a handful of experts and hundreds of practicing educators review the maps. We identified a large number of nodes – knowledge, skills, and aspects of cognition: 159 that were foundational to both disciplines, 1,850 in English Language Arts and 2,554 in mathematics. We identified about 10,000 connections or pathways among the nodes. Figure 2 shows the current version of the mathematics learning map.

When looked at in its entirety the information in the maps is hidden. Even broken down it appeared overwhelming, as you can see from them section on constructing understandings of text.

Smaller areas, such as nodes related to the 173 nodes and 243 connections within grade 7 mathematics are much more manageable, but still too complicated to identify detail and facilitate teacher usage.

We were sure it would be off-putting to educators and we would have to hide it inside the code in our software. Luckily the educators who reviewed the map convinced us that we were wrong. In fact,

they wanted to take it home after our review session and start using it in their classrooms.

While we collected the maps and did not allow this at the time, we did create additional structures within the map to help all teachers with their understanding. Within each of English Language Arts and Mathematics, we identified nine conceptual areas, such as constructing understandings of text and calculating accurately and efficiently using simple arithmetic operations. A subset of the map nodes were identified as Essential Elements particularly important learning targets.

For each Essential Element we identified common learning pathways including nodes that supported mastering the targeted essential element. Here is an example related to the Essential Element, "Identify two related points the author makes in an informational text" – ELA EE.RI 3.8. On the left we see where in the map those nodes reside. On the right we have highlighted several important nodes.

Finally we pull out one or more nodes at each of five levels. In this case we have two initial precursor nodes "Notice what is new" and "Recognize same;" one distal precursor, "Determine similar or

different based on physical characteristics;" one proximal precursor, "Identify relationships between concrete facts or details;" the targeted essential element, "Identify two related points made by the author of an informational text;" and a successor skill, "Identify one or more reasons supporting a point in the text."

Now we have broken down the map to areas around each 538 English Language Arts and 172 Mathematics essential elements. Each of these mini-maps is relatively easy for a teacher to comprehend, and any individual teacher only needs to be able to use at most one hundred of the mini-maps to guide instruction.

With these maps our teachers are less likely to get lost! Once they decide where they are going and determine where they are starting they have information about the route they should take. And if the special needs of a particular child means that child needs to take a detour from the most common learning pathway, the map shows the teacher how she or he can do that.

For more than a decade K-12 school assessments have been embedded in high-stakes accountability systems. This has led to an epidemic of teaching to tests not designed to be taught to. Moreover schools stop new instruction well in advance of annual testing and focus on test preparation.

The best test preparation should be good instruction. We must return to a paradigm of INSTRUCTION driving assessment, rather than the other way around. To this end, assessments must be developed based on the learning map as opposed to being based on a list of content specifications. To ensure this we created Essential Element Concept Maps to guide our test developers. These concept maps include the five level mini-maps, but also identify key vocabulary, and provide one or more templates for creating assessment tasks for the measured nodes.
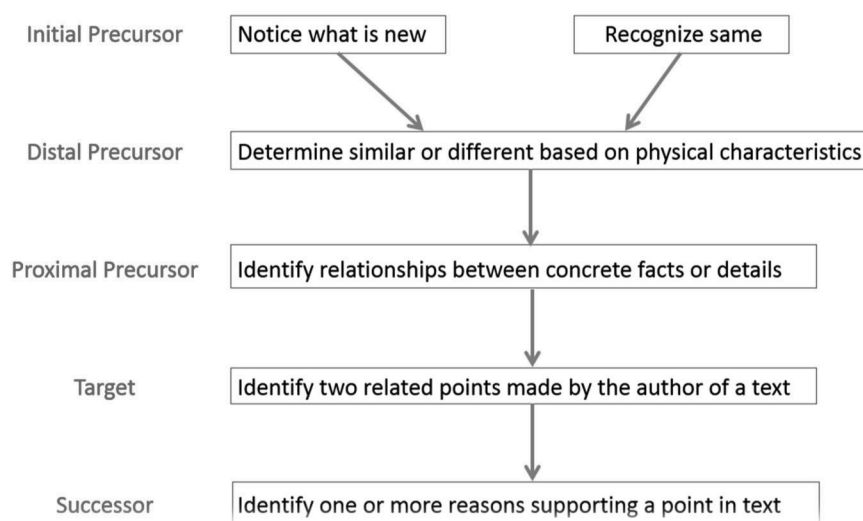

Figure 4. Key nodes related to identifying two related points made by the author of a text.

54

After the tests are developed and administered we need to use statistical models consistent with our learning map. This means we cannot simple add up the number of items a student responded to correctly. We are not interested in a total score. Instead we are interested in which particular nodes a student mastered – a concept that relates directly to the map.

We need a different statistical approach such as Bayesian network analysis. Let me give you an example of how that would work.

Let's say I would like to know if it rained last night. Simple, you might say, just look out the window and see if your lawn is wet. But it is more complicated than that. There is a relationship among three potentially observable events:

- Did it rain last night?
- Was my sprinkler on?
- Is my grass wet?

Based on our prior knowledge we can add probabilities to these three events. In this case perhaps we know the probability of it raining in Lawrence on any random summer night. Let us say that probability is .2.

I must admit, my wife turns the sprinkler system on and off at our house. She also pays much better attention to weather forecasts than I do. If it is not going to rain she turns on the sprinkler 40% of the time – she does not want it on 100% of the time as that would be wasteful or even damage the lawn. However one percent of the time she forgets to turn it off when it is going to rain.

Finally I have a table of conditional probabilities of the grass being wet. If the sprinkler was off and it did not rain the probability of a wet lawn is 0 – there is no morning dew in my neighborhood. If the sprinkler was not on but it rained there is an 80% chance the grass will be wet – the other 20% of the time it will have dried by morning. If the sprinkler came on but it did not rain, the probability of a wet lawn is .9 – the sprinkler comes on late enough so it is not likely it would be dry by the time I leave the house, even on a sunny warm morning. Finally, if the sprinkler was on AND it did rain the probability of a wet lawn is .99.

What I want to know is if my lawn is wet in the morning, what is the probability that it rained?

Sir Thomas Bayes, an English Presbyterian minister and statistician who lived in the early 1700s comes to our rescue. Though he never published his notes, a remarkable theorem he invented was discovered after his death. That theorem is as follows.

We want to know the probability it rained given the lawn is wet, but what we

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

have is the probability the lawn is wet given it rained. Bayes theorem lets us calculate the desired probability, which is .36.

$$P(R = T|G = T) = \frac{.016038}{.44838} = .3577$$

The same logic applies when we are trying to estimate the probability a student has mastered a node in the learning map given we know which of the three test items the student answered correctly. This is significantly more difficult than in the rain example, both because we cannot

directly observe node mastery and because we have such a large number of nodes for which probabilities must be estimated.

Luckily, there is a family of statistical estimation methods we can use to estimate mastery probabilities for large numbers of observable and non-observable nodes. It is called Markov Chain Monte Carlo estimation, often abbreviated MCMC. We can do that with multiple

In addition, reports based on learning maps could be dynamic and show us student progress over time.

Finally, although we started our learning map with the goal of improving instruction and assessment for students with significant cognitive disabilities, there is no reason this approach would not work equally as well for all students.

We have had one follow up grant to expand the map in mathematics to make

Area	Essential Element	Level Mastery				
		1	2	3	4 (Target)	5
Determining Critical Elements of Text	RL.3.1	Attend to object characteristics	Identify familiar people, objects, places, and events	Answer who and what questions and identify details in a familiar story	Answer who and what questions about story details	Answer who, what, when, and where questions about story details
	RL.3.2	Seek absent objects	Identify familiar people, objects, places, and events	Associate details with events in a familiar story	Associate details with events in diverse stories	Recount diverse stories with key details
	RL.3.3	Identify feeling states in self	Identify feeling words	Identify the feelings of characters in familiar stories	Identify the feelings of characters in a story	Identify character feelings and relate to actions
	RL.3.5	Express interest in book sharing	Differentiate between text and pictures	Identify details and beginning and end of a familiar story	Determine the beginning, middle, and end of a familiar story with a logical order	Identify beginning and end of a story

Figure 5. Section of prototype score report where shading indicates level of mastery attained

nodes each with multiple connected items.

Once we have estimated mastery of nodes we can provide reports that, based on focus groups we have had with parents and educators, provide information that is concrete, understandable, and actionable. No more trying to figure out what it means for your child to have a score of 27.

These reports can even suggest what next area of study would best serve as a next instructional focus – either for the class as a whole or for any individual student. That is, the use of learning maps makes it much easier for teachers to personalize instruction.

it cover the K-12 content necessary for all students and to provide better map visualization tools. We have just applied for another grant that would complete the build out of the map in English Language Arts and expand the mapping software capabilities to allow the tagging of instructional resources, professional development materials, and formative assessments to the nodes (or groups of nodes) in the map. This could allow commercial entities to license the map and software to deliver their own content. A third recent NSF grant application would allow us to map out the components of the data sciences at KU. Also, a medical certification

organization has expressed interest in using this approach in the determination of clinical judgment skills.

Maps organize information in ways familiar to most people. The map metaphor is powerful. The possibilities of this approach cross disciplines.

Of important note in these times of diminishing federal research support in the social sciences, this line of research has been supported by a federal agency (U.S. Department of Education), a large foundation (Bill and Melinda Gates Foundation) and a consortium consisting of 18 state departments of education. This latter source of funding appears to be underrecognized and underutilized by research universities.

References:
1. Faragher, J.M. (1993). Daniel Boone: The life and legend of an American pioneer. New York: Holt Paperbacks.