

Evaluating an Initialization Tool for Student Placement into a Map-Based Assessment

Brooke Nash and W. Jake Thompson

Context of Study

- Adaptive tests can be more efficient than traditional one-test-fits-all approaches.
- Characteristics of students with significant cognitive disabilities (SCD) often requires a shorter test
- The need for additional efficiency with this population drove the development of a new method for placing students with SCD into an assessment system that is based on fine-grained learning maps.

DYNAMIC LEARNING MAPS - OVERVIEW

Essential Elements

- Assessments are based on fine-grained learning map models that represent how academic skills are acquired.
- Within each content area, Essential Elements (EEs) specify the content standards, while the learning map models clarify how students reach those targets.

Linkage Levels

- For each EE, neighborhoods of map nodes, called linkage levels, are identified as assessment targets.
- Assessment items are based on nodes at the five linkage levels:
 - Initial Precursor (IP)
 - Distal Precursor (DP)
 - Proximal Precursor (PP)
 - Target (T)
 - Successor (S)

Assessment Design

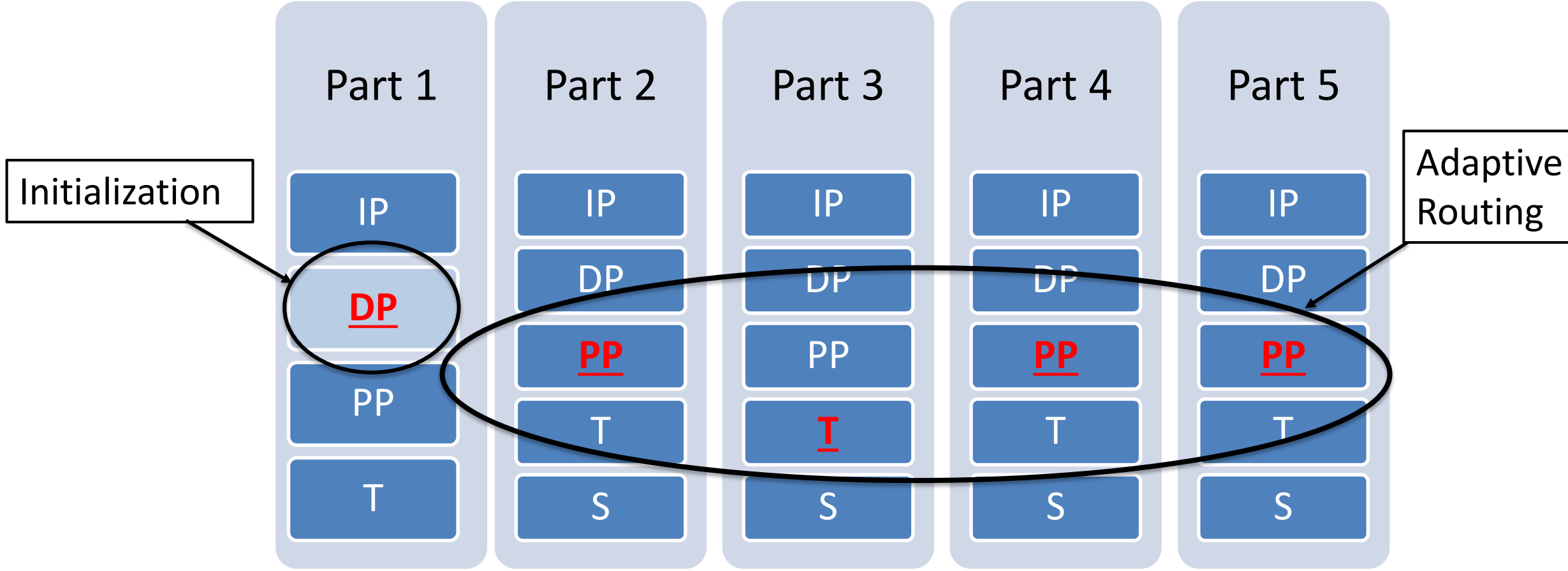
- Two testing models
- Testlets
 - Consist of 3-8 items
 - Assess one linkage level
- Assessments
 - Students typically take between five and seven testlets in each content area.
 - Linkage levels are allowed to vary from testlet to testlet in order to best match students' knowledge and abilities.

Assessment Delivery

The process for matching students to linkage levels that align with students' skills and abilities involves two steps:

- (1) the assignment of a linkage level for the first administered testlet (i.e., **initialization**), and
- (2) the assignment of a linkage level for all subsequent administered testlets (i.e., **adaptive routing**).

Example Student



Initialization

- Based on teacher responses to “First Contact Survey” which asks about students’:
 - Academic skills in ELA and math
 - Expressive communication skills
- Students are assigned a “complexity band” from the survey responses which translates to a linkage level

Correspondence between Complexity Band and Linkage Level

First Contact Complexity Band	Linkage Level
Foundational	Initial Precursor
Band 1	Distal Precursor
Band 2	Proximal Precursor
Band 3	Target

Note. Successor level testlets are not assigned as a first testlet.

Adaptive Routing

- System automatically calculates the percent correct for the EE of the first testlet and then assigns the next linkage level based on the student's performance.
- Linkage levels for the second and all subsequent testlets are assigned via adaptive routing.

Goals of Testlet Assignment Process

- Initialization:
 - Intended to match students to linkage levels at the beginning of the test.
- Adaptive Routing:
 - Intended to adjust linkage levels based on student performance after the first assigned testlet.

Premise of Study

- If the initialization tool is effective at assigning students to a linkage level that aligns to their knowledge and skills, then adaptive routing is needed less, particularly at the beginning of the assessment.
 - Adaptation should *not* be a function of which complexity band a student is assigned to via initialization.

Purpose of Research

- To further evaluate the DLM initialization process.
 1. What percent of students adapt to a different linkage level between the first and second testlets administered?
 2. Does the linkage level of the first testlet (determined by the initialization process) predict whether or not the second testlet adapts to a different linkage level after controlling for grade level?

Data

- Data from the 2014-15 and 2015-16 end-of-year spring operational testing windows were collected.
- Students were from grades 3 through 8 and high school and included the full range of students eligible for alternate assessment in their state
 - 124,073 students in English language arts
 - 123,848 students in mathematics.

Methods

- RQ 1:
 - The percent of students that adapted between the first and second testlet was calculated for each complexity band by grade and content area.
- RQ 2:
 - A Bayesian logistic regression was conducted where adaptation between the first and second testlet was coded:
 - 0 = did *not* adapt
 - 1 = did adapt
 - Separate analyses were run for ELA and math

Model Specifications

- Linkage levels were dummy coded into 3 IVs
 - Target level was reference group
- Grade level was included as a covariate
- Model Estimation
 - Non-informative normal priors were used
 - Three chains were run, each with 1,000 iterations (500 discarded)
 - This resulted in 1,500 retained iterations
- Replicated data sets were generated at each iteration of the chain in order to facilitate posterior predictive model checks

Descriptive Results - RQ 1

Percent of students in English language arts who adapted between 1st and 2nd testlet by grade, and linkage level of first testlet

Grade	Foundational/ Initial Precursor	Band 1/ Distal Precursor	Band 2/ Proximal Precursor	Band 3/ Target
3	18.5	58.5	85.3	95.1
4	31.5	52.2	55.5	48.6
5	27.5	61.5	76.3	87.8
6	23.8	80.2	75.2	65.3
7	23.0	64.6	64.7	68.2
8	35.6	52.2	56.5	83.8
9	15.3	81.0	78.6	83.1
10	13.5	59.0	37.3	55.6
11	21.0	64.1	60.0	63.9

Descriptive Results - RQ 1 (cont.)

Percent of students in Mathematics who adapted between 1st and 2nd testlet by grade, and linkage level of first testlet

Grade	Foundational/ Initial Precursor	Band 1/ Distal Precursor	Band 2/ Proximal Precursor	Band 3/ Target
3	10.2	57.4	63.5	47.6
4	12.6	77.8	74.7	64.8
5	25.2	67.8	69.5	78.3
6	19.2	64.7	61.5	61.1
7	13.3	68.2	64.5	73.1
8	26.8	76.9	50.0	65.1
9	20.7	60.0	42.6	44.2
10	16.8	61.5	61.1	54.9
11	16.5	55.6	58.8	40.7

Summary of Results - RQ 1

- Most students who start the test at levels higher than initial precursor adapt to different linkage levels for the second testlet.
- Conversely, the majority of students who start at initial precursor, tend to stay at that level for the second testlet.
 - However, initial precursor level only has one direction to adapt.

Summary of Results - RQ 1 (cont.)

- Across grade levels, there were no consistent patterns.
 - Amount of adaptation did not appear to generally increase or decrease with grade level.
- Across content areas, more students appeared to adapt at the target and initial precursor levels in ELA than in math; however, no patterns were apparent for the distal precursor or proximal precursor groups.

Logistic Regression Results - RQ 2

	English Language Arts				Mathematics		
Coefficient	Mean	SD	95% Credible Interval		Mean	SD	95% Credible Interval
(Intercept)	1.210	0.027	[1.156, 1.259]		0.880	0.025	[0.830, 0.930]
Grade	-0.030	0.003	[-0.035, -0.025]		-0.061	0.003	[-0.066, -0.056]
Initial Precursor vs. Target	-2.126	0.024	[-2.172, -2.080]		-2.058	0.025	[-2.106, -2.006]
Distal Precursor vs. Target	-0.526	0.020	[-0.563, -0.485]		0.200	0.019	[0.164, 0.238]
Proximal Precursor vs. Target	-0.274	0.019	[-0.310, -0.237]		-0.003	0.019	[-0.040, 0.034]

Summary of Results - RQ 2

- Grade level:
 - Students are *less likely* to adapt as grade level increases.
 - For each grade level increase, the change in odds of adapting between the first and second testlet decreases by 0.97.

Summary of Results - RQ 2 (cont.)

- Linkage levels in ELA:
 - In comparison to the target level group, students starting at the:
 - Initial precursor were 0.11 times as likely to adapt
 - Distal precursor were 0.59 times as likely to adapt
 - Proximal precursor were 0.76 times as likely to adapt

Summary of Results - RQ 2 (cont.)

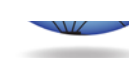
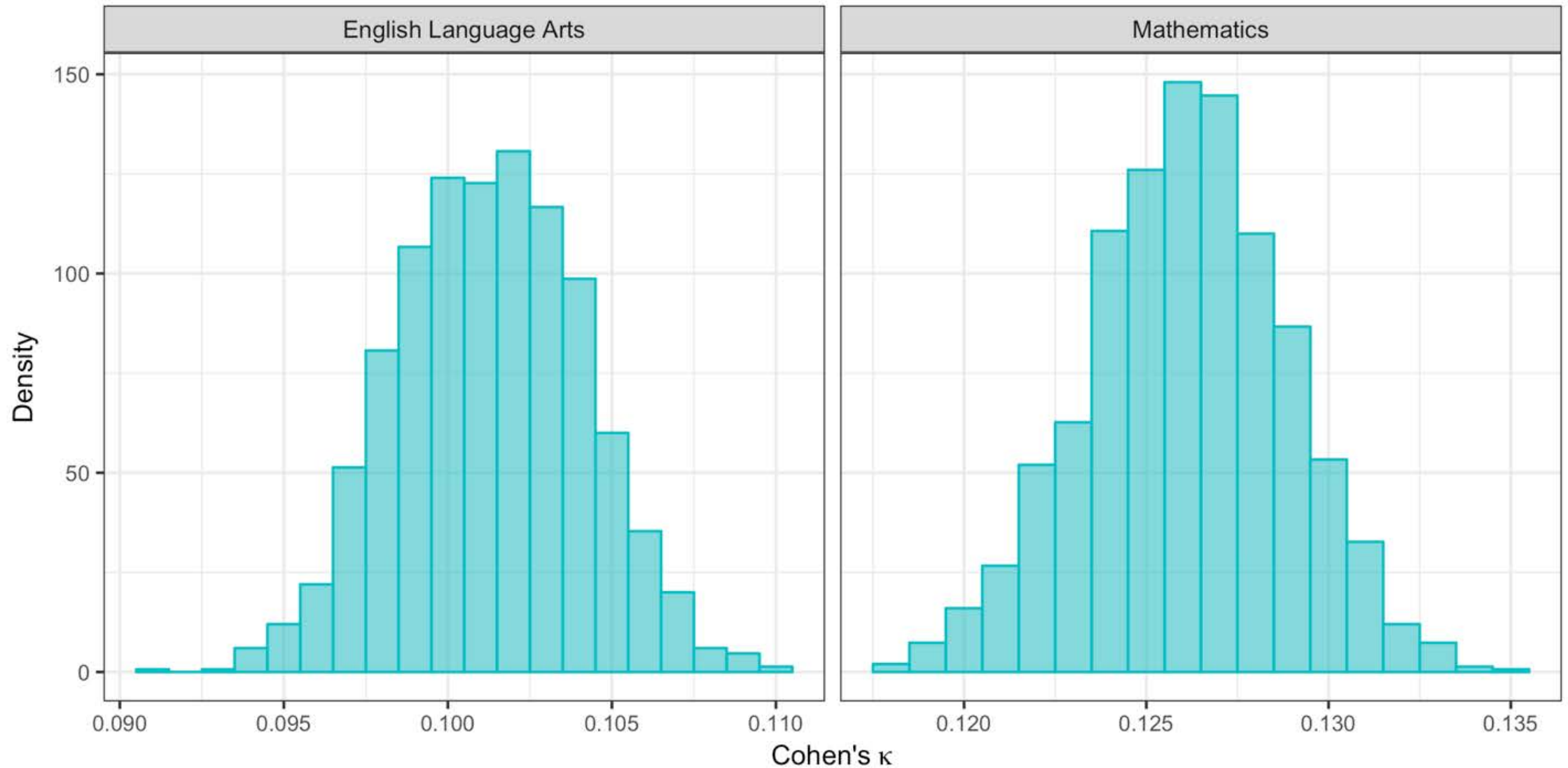
- Linkage levels in Math:
 - In comparison to the target level group, students starting at the:
 - Initial precursor were 0.13 times as likely to adapt
 - Distal precursor were 1.22 times as likely to adapt
 - Proximal precursor were slightly less likely to adapt

Model Fit Results

- Cohen's Kappa was calculated between the observed data and each of the simulated data sets.

Content Area	Mean Kappa	Range of Kappa
ELA	0.10	0.09 – 0.11
Math	0.13	0.11 – 0.13

Distributions of Cohen's Kappa



Discussion

- On average, roughly 65% of students who start at the upper three linkage levels adapt between the first and second testlet.
- Possible explanations:
 - Initial precursor only has one direction to adapt
 - Target level is comprised of students who test at the successor level
 - Proximal and distal precursor levels may represent a more diverse population of students
 - Items are easy

Discussion (cont.)

- Grade and most linkage levels were statistically significant predictors of adaptation status.
 - However, large sample sizes detect small differences
- Poor overall model fit suggests that the initialization process is not able to adequately explain the variance in adaptation.

Next Steps

- Additional analyses are needed to better understand the factors that contribute to adaptation.
 - Other student characteristics
 - Routing thresholds
- Future research is planned to evaluate the direction in which students are adapting between the first and second testlet.

THANK YOU!

For more information, please contact:

bnash@ku.edu

or

Go to: www.dynamiclearningmaps.org