International Association for Computerized Adaptive Testing
IACAT
Advancing the Science and Practice of Human Assessment

Journal of Computerized Adaptive Testing
Volume 12, Number 1, February 2025
DOI 10.7333/2502-1201088
ISSN 2165-6592

The Impact of Item Bank Transition Rules on Student Ability Estimates and Achievement Level Classifications

Jonghwan Lee[1], Sangdon Lim[2], M. Christina Schneider[1], Garron Gianopulos[1], Luping Niu[2], Sooyong Lee[2], and Seung W. Choi[2]

[1]NWEA

[2]The University of Texas at Austin

This paper describes a hybrid interim-summative computerized adaptive assessment design administered across three academic terms (fall, winter, and spring). Each test event had two phases that either stayed on-grade or moved off-grade dynamically. Using three different transition rule conditions, this design was compared to an on-grade hybrid interim-summative computerized adaptive assessment design using a series of simulations for a Grade 4 and Grade 6 mathematics assessment. A 500- and 800-item bank in each grade was simulated with a normal distribution of items that ranged across achievement levels. Simulees were randomly drawn from a normal distribution. Simulees whose ability estimates met the transition rule requirements were routed to off-grade item banks and test blueprints dynamically. During Phase 1, the algorithm was programmed to administer only on-grade items. At the end of Phase 1, a transition rule was used to determine if the student should be routed to off-grade items and blueprints for Phase 2. If the student ability estimate did not meet the transition rule requirements, the adaptive test continued to administer on-grade items. If the ability estimate met the transition rule requirements, the student was routed to item banks and blueprints at the adjacent lower or higher grade, as appropriate. The results were evaluated based on measurement precision (i.e., RMSE), measurement accuracy (i.e., bias), item exposure rates, and classification accuracy. Results indicated that while more students were routed off-grade when more lenient routing rules were used, most of the resulting evaluation indexes remained similar across all conditions, except for item exposure rates that varied across conditions. More lenient transition rules led to fewer overexposed items. More stringent transition rules maintained continuity with the on-grade achievement level designations. This preliminary evidence indicates that it might be feasible to transition students at the tails of the distributions to an off-grade bank and meet federal requirements. Such a transition integrates a fundamental aspect of interim assessments—going off-grade—with the summative assessment ESSA requirement that proficiency determinations be derived from on-grade items and constraints.

Keywords: *classification accuracy, computerized adaptive tests, multi-phase, off-grade, through-year assessments, transition rules*

A new assessment design that connects interim and summative assessment purposes is beginning to gain momentum in K–12 educational assessment. The United States Department of Education (USED) gave flexibility for different assessment system designs in the final regulations for the Every Student Succeeds Act (ESSA) Volume 81 (USED, 2016) by noting the following:

> States have flexibility to develop new assessment designs, which may include a series of multiple statewide interim assessments during the course of the academic year that result in a single summative assessment score (sometimes described as “modular” assessments (p. 3).

Hybrid interim-summative assessments that result in a single summative score are currently referred to as through-year assessments in the field, and states such as Florida, Nebraska, and Texas began piloting or implementing such designs during the 2022–23 school year. In such an approach, typically scores from the third administration are used for federal accountability.

States intend a through-year assessment to provide a more cohesive depiction of how students develop in a state’s standards using the summative assessment construct of proficiency throughout the school year as a tool for progress monitoring. These assessments are typically characterized as maintaining the same domain-based blueprint throughout the year across grades and across administrations. Proponents of such systems want teachers to have actionable information that indicates a student’s current level of performance and change in performance (growth) across three testing windows to improve classroom teaching and support student learning year-round.

A puzzle that developers of such assessment systems must solve is how to allow off-grade adaptivity, if this is a policy desire, while measuring student performance with on-grade items for accountability purposes. Interim assessment providers have historically allowed students to receive off-grade content because they argued it allows for more precise measurement of student ability, especially at the low end of the scale (Li & Meyer, 2019). Rambo-Hernandez and Warne (2015) argued that a floor and ceiling effect on low- and high-ability students leads to more measurement error when the item bank does not cover a wide range of student ability. They showed that measurement error is minimized when an assessment includes items aligned to students’ ability level, indicating that off-grade testing is a solution for providing higher measurement precision for the low- and high-ability students. Way et al. (2010) noted that using some off-grade items is appropriate when a purpose of the assessment is to measure growth in student knowledge and skills across time. Way et al. argued that such an approach is sensible both from a measurement perspective, and the use of off-grade items aligns to a philosophy of personalized learning. Earlier grade standards are meant to be building blocks to grade-level standards.

Researchers who work in gifted education might have most investigated and advocated for off-grade testing (Achter et al., 1996; Mills & Barnett, 1992; Rambo-Hernandez & Warne, 2015; Stanley & Benbow, 1981; Stanley, 2005; Terman, 1926; VanTassel-Baska et al., 1996). Lohman and Korb (2006) argued that most standardized tests have poor measurement of students who are in the top and bottom decile of ability level and measurement error can lead to inappropriate classroom instruction and decreased student motivation to learn. The research that exists regarding students routing to different paths mainly appears in a multistage testing framework and shows that a great deal of consideration must go into determining the transition rules to achieve maximum efficiency and classification accuracy (Hendrickson, 2007; Svetina et al., 2019; Yan et al., 2016). Recently, Meyer et al. (2023) found that an adaptive algorithm that allows selecting more items at the student's ability estimate from the grade in which the student is enrolled tends to provide a higher final ability estimate than an algorithm that selects items based solely on the student's ability estimate when using the same blueprint but with less focus on selecting items from the grade in which a student is enrolled. Their findings suggest that stakeholders might need to architect careful rules regarding which students receive off-grade items.

Bejar (2016) recommended criterion-referenced routing as a transition rule in the context of summative educational assessments. He suggested that a proficient cutscore could be used as the indicator of how to route a student. Using his suggested model, a student ability estimate could be compared to the proficient level cutscore to determine whether a student should move to a more difficult form or an easier one in the context of a multistage assessment. The Smarter Balanced Assessment Consortium (American Institutes for Research, AIR, 2016) permitted students to be administered off-grade items in a computerized adaptive assessment after the student completed two-thirds of a test event. The transition rule estimated the likelihood a student could be (1) found to be proficient with the below-grade items included in the student's final ability estimate, or (2) denied being recognized as proficient by including above-grade items in the student's final ability estimate. When a criterion-referenced routing approach is applied to an assessment that uses both off-grade and on-grade item banks, such an approach could be used to ensure control regarding which students have access to off-grade items.

Schneider et al. (this volume) described a through-year assessment system prototype that stakeholders conceptualized utilizing criterion-referenced transition rules that serve the following high-level policy goals for a Grade 3–8 assessment system after examining the ESSA regulations (USED, 2016):

1. Summative assessment score interpretations of student ability should underpin the assessment system such that students are able to show comparable ability estimates during different academic terms (fall, winter, and spring). This allows students who meet the cut-score for proficiency to be found in each academic term.
2. Students should have the ability to preserve advanced proficiency so they can move to the next higher adjacent grade when they are individually ready. This means that an advanced student's score for accountability is based on their on-grade responses earlier in the year, whereas the spring administration represents most students' scores for accountability purposes.
3. Students should have multiple opportunities to demonstrate on-grade mastery. Therefore, lower-performing students routed to below-grade items in an academic term should start

each new academic season with on-grade items to quickly determine if those students are now able to access grade-level standards and stay in the item bank of their designated grade. This is done because ESSA regulations and precedent (AIR, 2016) require (1) student proficiency level designations to be based on items administered on-grade, and (2) the proficiency level designation must be made before moving students off-grade to provide supplemental information for teachers. This approach also enhances communication to policymakers regarding which students in a grade are most in need of substantial academic intervention during the year.

The purpose of this paper is to investigate the functionality of the through-year assessment design described by Schneider et al. (this volume).

Research Questions

This study investigated item bank transition rules. The main research questions were:

1. What impact do different transition rule conditions have on the accuracy of student ability estimates for each academic term when used to control if and how students should be administered off-grade items?
2. What impact does allowing off-grade items have on the overall accuracy of student ability estimates for each academic term compared to student ability estimates using only on-grade items?
3. What impact do different off-grade item transition rule conditions have on item exposure and utilization rates compared to when students are only administered on-grade items?

Method

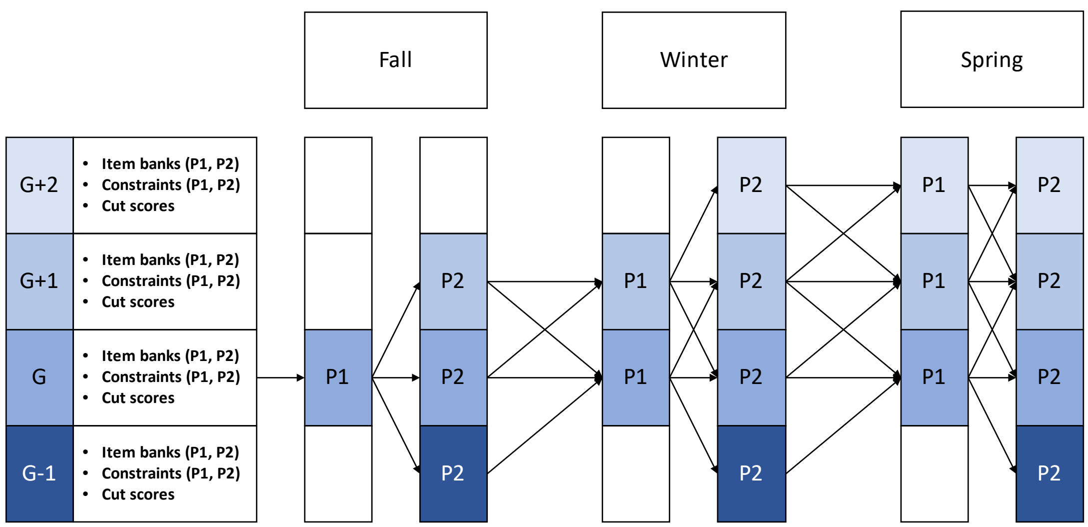
The research questions were examined in a monte-carlo simulation study. Simulated data were used instead of real data for two reasons. First, simulated data separate the effect of model misfit and calibration errors (Bolt, 1999; Davey et al., 1997). Second, stakeholders often desire to see the functioning of new test designs prior to implementation in a pilot.

Test Design

A through-year assessment system has the assumption of being administered three times a year in each academic season (i.e., once in the fall, winter, and spring). The through-year adaptive assessment algorithm we investigated comprised two phases for each test event. Each phase included a grade-specific item bank and set of blueprint constraints. This design differed from the typical multistage adaptive assessment that presents each module within a phase as a fixed form, preassembled prior to testing at different levels of difficulty. In the fall, Phase 1 comprised 25 adaptively selected on-grade items for the student's grade (G) of record that determined a student's path in Phase 2. The student ability (θ) estimate and transition rule at the end of Phase 1 determined whether the student should be routed to an off-grade item bank. Phase 2 comprised sequestered item banks and blueprint constraints associated with a particular grade level that adaptively selected 16 items. At the conclusion of Phase 2, responses to all 41 administered items across phases for the fall test event were used to calculate the student's final θ estimate that was then used to route the student to the appropriate item bank and constraints for Phase 1 of the winter assessment. For this study, the final θ estimate for a student in Phase 2 of each academic season was considered the final ability estimate used to evaluate the performance of the proposed transition rules across time.

In the test design, should a student meet the transition rule criteria, they could be routed to an adjacent grade's item bank and blueprint at the end of each phase. Students who were routed to an adjacent higher grade were constrained to go no farther than two grades above their grade of record ($G + 2$) across all three academic seasons. Students who were routed to a below-grade bank (G – 1) during Phase 2 in an academic season were always returned to the on-grade (G) item bank and blueprint constraints for Phase 1 of the next academic season following ESSA regulations (USED, 2016). Thus, the grade range possible across all academic seasons for a particular grade of record was limited from G – 1 to G + 2, as shown in Figure 1.

Figure 1
Routing of Phases



Source: https://cran.r-project.org/web/packages/maat/vignettes/maat.html

Figure 1 shows three assessment academic seasons: fall, winter, and spring. The assessment within each academic season had two computerized adaptive phases administered as a single test event. The shading in Figure 1 is used to show the grade-level bank and constraints possible for each phase. The arrows show the possible pathways to item banks and blueprint constraints based on the transition rules. For example, the fall has three possible pathways:

1. The on-grade-level Phase 1 paired with an above-grade-level Phase 2.
2. The on-grade-level Phase 1 paired with an on-grade-level Phase 2.
3. The on-grade-level Phase 1 paired with a lower-grade-level Phase 2.

The arrows between the fall and winter administrations show the possible pathways to the bank and constraints that begin Phase 1 of the winter administration, which depends on the student's final θ estimate in the fall and the invoked transition rule.

Simulated Item Banks

Items were simulated to range in difficulty across achievement levels using a normal distribution and included content features found in an existing operational mathematics program used for state accountability purposes. The simulated content included features such as multiple-choice items and technology-enhanced items (e.g., multiple-choice, composite, gap match, graphic gap match, hot text, and text entry), aligned to standards and the corresponding domain, and the use of Webb's (2005) depth of knowledge (DOK). The initial bank size was simulated to be 800 items, but 500 items were also randomly drawn from the original 800 to help gauge the effect on the design functionality in case items must be removed from the bank, which occurs in operational testing programs. Stocking (1988) noted the importance of maintaining the content and statistical characteristics of an item bank as changes are made to ensure that resulting student ability estimates remain comparable. As noted by Schneider et al. (this volume), the simulated item banks were constructed with the assumption that items were aligned to range achievement level descriptors (ALDs; Egan et al., 2012) in sufficient numbers for each achievement level bin. This would allow (1) the blueprints to be met in all achievement levels, and (2) most students in a grade to remain in the grade-level item bank and demonstrate growth by moving into adjacent, higher achievement levels as their ability increased. This approach to creating a sufficiently deep bank was intended to allow most students to remain in the grade-level bank.

Content Constraints

The content constraints used in this study were adapted from an existing Grade 3–8 operational mathematics program used for state accountability purposes. Appendix A presents the specific content constraints for all grades and phases. For each grade, the same proportional blueprint and content constraints were provided for each phase of the test event, resulting in 41 adaptively selected items. For example,

1. Grade *G* + 1 Winter Phase 1 and Grade *G* + 1 Winter Phase 2 differed in the number of items required to satisfy a constraint, but the proportional representation to the state blueprint was the same.
2. The Grade *G* + 1 Winter Phase 1 and Grade *G* + 1 Winter Phase 2 constraints produced the same overall G + 1 blueprint to the existing operational state mathematics program.
3. Similarly, Grade *G* Fall Phase 2, Grade *G* Winter Phase 2, and Grade *G* Spring Phase 2 used the same constraint configuration.
4. The Grade *G* Phase 1 and Grade *G* Phase 2 constraints produced the same overall Grade *G* blueprint to the existing operational state mathematics program for each academic season.

Simulation Procedures

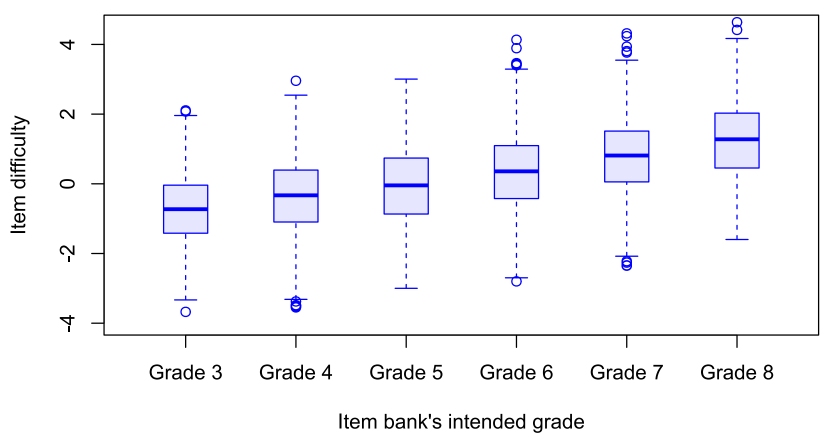
Item parameters. This study used a simulated mathematics item bank for Grades 3–8 with metadata representative of a summative assessment construct (i.e., the inclusion of technology-enhanced items worth more than one point). The item bank size was varied in two levels: 500 and 800. The item banks were generated under a Rasch model (Rasch, 1960) for dichotomous items and a partial credit model (Masters, 1982) for polytomous items.

Table 1 presents the mean and standard deviation (SD) used to generate the item parameters for Grades 3–8. To ensure that the vertical scale was articulated across grades, a lower boundary was set for each grade. For example, the lowest possible *b* parameter for Grade 3 was −4.0 and the lowest possible *b* parameter for Grade 4 was −3.6, as shown in Figure 2. The vertical articulation of item parameters was an assumption of this study's test design, and it contributed to one of the transition rules discussed in the next section.

Table 1
Mean and SD for Normal Distribution

Grade	Mean	SD	Lowest *b* Parameter
3	–0.8	1.1	–4.0
4	–0.4	1.1	–3.6
5	0.0	1.1	–3.2
6	0.4	1.1	–2.8
7	0.8	1.1	–2.4
8	1.2	1.1	–2.0

Figure 2
Vertically Articulated Item Banks Across Grades



Transition rules. Researchers investigated two categories of transition rules to define if and when students were administered off-grade items and blueprints: student-centered and content-centered.

*Student-centered approach***.** The student-centered transition rule was based on using confidence intervals (CIs; Kingsbury & Weiss, 1983; Eggen & Straetmans, 2000) from maximum likelihood estimation (MLE) scoring extended to multiple cutscores (Thompson, 2007). The student-centered approach used CIs obtained from the estimated θ ($\hat{\theta}$) and standard errors from each phase. A CI indicates the range of possible scores where an unknown true score might fall. For example, a 95% CI means that 95 out of 100 times, the true score falls within the range defined by the interval. The comparison between the CIs and pre-specified cutscores of grades determined routing for the next phase. If the upper bound of a CI was below the lowest on-grade cutscore at the end of Phase 1, the routing would be to the below-grade item bank. If the lower bound of a CI was above the highest on-grade cutscore, the routing would be to the above-grade item bank and blueprint. In all other cases, students stayed on-grade. In the current study, three CI ranges were included: 1.00 conditional standard error of measurement (CSEM) (68% CI), 1.64 CSEM (90% CI), and 1.96 CSEM (95% CI).

Content-centered approach. The content-centered transition rule used characteristics of the item bank to determine the student pathway. The content-centered approach identified the student as needing on- or off-grade items by comparing the Phase 1 $\widehat{\theta}$ to a cutscore representing the ceiling and floor of the within-grade item bank. If the student's $\widehat{\theta}$ was either above the 95th percentile of item difficulty or below the 5th percentile of item difficulty in the Phase 1 item bank, the student was routed to the above- or below-grade item bank in the next phase. In subsequent tables and figures, this approach is denoted as a bank-based transition rule.

Based on these two approaches, the following transition rules were evaluated and compared during the simulations in this study:

1. No transition,
2. CI 68,
3. CI 90,
4. CI 95,
5. Bank-based.

True ability distribution. In each replication, simulees' true θ values were randomly drawn from a normal distribution. Because the three academic season administrations typically occur in fall, winter, and spring, real-world student abilities (on average) tend to increase over time on an interim assessment (NWEA, 2019). The means for fall were set to start at –1.0, and –0.2 for Grades 4 and 6, respectively. The means were set to increase by 0.3 in each subsequent administration. The standard deviations were set to 1. Correlations were set to 0.9 for adjacent seasons and 0.8 for sub-adjacent seasons, similar to values reported for a mathematics interim assessment (NWEA, 2019).

Each replication included 1,000 simulees, and each simulee had three true θ values (one for each academic season). Because the data generation mechanism for each true θ was based on a three-variate distribution with between administration correlations, it was not necessarily the case that simulee θs were generated to be always monotonically increasing over academic seasons.

Exposure and Overlap Control

Bank-based exposure control was not implemented in the simulations (i.e., items were not removed as they became more exposed across students). This is consistent with common practice in K–12 educational assessment computerized adaptive assessments. However, overlap control was used across all six phases. This means that a given student should not receive the same item more than once across academic seasons. This is also consistent with common practice in K–12 educational assessment computerized adaptive assessments, when possible. Overlap control was implemented as a soft constraint, penalizing the item information of previously administered items by M = 100. For example, if the original item information to be used for adaptive test assembly was 5.0 for an item at an interim ability estimate, and if that item has been previously administered, the information value was modified to be 5.0 – M = −95.0 for the purpose of adaptive test assembly. This means that previously administered items were technically allowed to be selected, if necessary, to meet blueprint requirements. The anticipation for the item bank was that it was sufficiently deep compared to the number of items that needed to be administered to the student population. The goal was to see each item administered in a relatively small percentage of the test administrations (AIR, 2016).

Adaptive Test Assembly

The adaptive form assembly in each phase was performed using an optimal test design approach with shadow tests (van der Linden & Reese, 1998). The optimal test design approach has an advantage of ensuring that all content requirements are strictly met. The item with the maximum information at the current interim θ estimate was selected to be administered to the simulee.

Ability estimation. Interim and final ability estimates were obtained using MLE, with expected a posteriori (EAP) estimation as a fallback for when MLE was not feasible because of extreme responses (e.g., when all item scores were 0). For the purpose of evaluating classification performance, estimated θs were converted into four achievement level categories using a predefined set of cutscores (Table 2) that approximated those found in an operational testing program.

Based on the test design, Phase 2 $\hat{\theta}$s within each academic season were obtained by combining Phase 1 and Phase 2 responses to ensure that all students received scores similar in measurement precision.

Table 2
Cutscores

Grade	Level 2	Level 3	Level 4
3	-1.47	-0.55	0.48
4	-1.07	-0.15	0.88
5	-0.67	0.25	1.28
6	-0.27	0.65	1.68
7	0.13	-1.05	2.08
8	0.53	1.45	2.48

Performance Evaluation

Estimation of accuracy and precision. For each academic season, final θ estimates were used to evaluate the ability estimation performance of evaluated conditions. Root mean square error (RMSE) and bias were calculated as performance measures. A reliability measure was also calculated. Traditional reliability coefficients from classical test theory consider individual items and depend on all examinees to take common items, whereas students receive different items in an adaptive assessment. Therefore, the marginal reliability was calculated (Samejima, 1994) as

$$\text{reliability} = \frac{\sigma_T^2}{\sigma_X^2} \tag{1}$$

where σ_T^2 is the variance of true θs (i.e., σ_T^2 = 1), and σ_X^2 is the variance of estimated θs.

Classification accuracy. Two measures of classification accuracy were used to evaluate the results: accuracy and Cohen's weighted kappa. Accuracy was computed as the proportion of simulees that had the same achievement level categories between true and estimated θs. Cohen's weighted kappa (1968) was implemented with quadratic weights applied.

Item utilization and exposure rate. The item bank utilization rate was calculated as the number of unique items administered at least once, divided by the number of items in each bank. Item exposure rate was calculated as the number of times an item was administered divided by the number of simulees. This was 1.0 if the item was given to all simulees. A lower rate indicates that the item was not overly exposed to simulees during the test. The exposure rate calculation was based on approaches from AIR (2016) and van der Linden (2003).

Software. The *MAAT* R package (Choi et al., 2022), which implements a multiple administration adaptive testing design, was used to conduct the simulations. *MAAT* is an extension of the R computerized adaptive testing package *TestDesign* (Choi et al., 2021) which performs the adaptive form assembly for each phase in *MAAT*.

Results

Estimation Accuracy and Precision

Table 3 presents the marginal reliability, RMSE, and bias for each condition for fall, winter, and spring. The test-based marginal reliability coefficients ranged from 0.91 to 0.93 across academic seasons across conditions, regardless of bank size. RMSE was 0.27 on average, and the θ bias estimates were near 0 across all conditions. Table 4 presents the reliability of the θ estimates at the end of each phase. The Phase 2 reliabilities represent the estimate based on the 41 administered items. The Phase 1 reliabilities represent the estimate at the time the transition rule decision was made. The Phase 1 marginal reliability coefficients ranged from 0.87 to 0.89 across academic seasons across conditions, regardless of bank size.

Classification Accuracy

Table 5 shows the classification accuracy measures (accuracy and quadratic-weighted kappa) across academic seasons and conditions. The classification measures remained similar across transition rule conditions. This indicates that student routing does not result materially in a loss of accuracy across conditions if using off-grade items and that the transition rules are functioning as intended. Across academic seasons, accuracy decreased slightly (on average 0.84, 0.82, and 0.81 for each season, respectively), whereas weighted kappa increased slightly (0.89–0.90 for each season, respectively). Classifications are most critical to investigate for students who were routed to a below-grade item bank. Appendix B shows the final achievement level designation of students routed to lower-grade items for a simulation. As shown in Appendix B, the more stringent transition rules, CI-95 and bank-based, were preferable to maintain consistency with Phase 1 account-ability classifications.

Table 3
Reliability, RMSE, and Bias by Academic Season

Item Bank Size	Grade	Routing Rule	Fall			Winter			Spring		
			Reliability	RMSE	Bias	Reliability	RMSE	Bias	Reliability	RMSE	Bias
800	4	CI-68	0.928	0.273	-0.001	0.931	0.272	0.000	0.929	0.272	0.000
		CI-90	0.928	0.272	0.000	0.929	0.272	0.000	0.932	0.272	-0.001
		CI-95	0.927	0.271	0.000	0.927	0.272	0.002	0.929	0.272	-0.001
		Bank-based	0.928	0.271	0.002	0.931	0.274	-0.001	0.928	0.274	-0.002
		No routing	0.925	0.273	0.001	0.928	0.274	-0.001	0.922	0.277	0.000
800	6	CI-68	0.927	0.272	0.000	0.931	0.271	-0.001	0.932	0.272	0.001
		CI-90	0.927	0.272	0.000	0.929	0.270	0.000	0.930	0.272	0.000
		CI-95	0.927	0.272	-0.001	0.930	0.272	0.001	0.928	0.272	0.001
		Bank-based	0.926	0.273	-0.001	0.931	0.272	-0.001	0.928	0.273	0.000
		No routing	0.925	0.273	-0.002	0.928	0.273	0.000	0.921	0.275	-0.002
500	4	CI-68	0.925	0.273	0.002	0.930	0.274	0.000	0.929	0.275	-0.001
		CI-90	0.925	0.273	0.002	0.929	0.274	0.000	0.928	0.276	-0.001
		CI-95	0.926	0.273	0.002	0.929	0.275	-0.001	0.926	0.277	0.000
		Bank-based	0.927	0.273	0.002	0.929	0.274	0.000	0.928	0.277	0.001
		No routing	0.922	0.274	0.001	0.919	0.277	0.000	0.914	0.283	-0.001
500	6	CI-68	0.921	0.275	0.000	0.926	0.274	0.000	0.927	0.276	-0.002
		CI-90	0.921	0.275	-0.001	0.925	0.274	-0.001	0.929	0.277	0.000
		CI-95	0.922	0.275	-0.002	0.925	0.276	-0.002	0.924	0.276	0.001
		Bank-based	0.922	0.275	-0.002	0.928	0.277	-0.001	0.923	0.278	-0.001
		No routing	0.920	0.276	-0.002	0.916	0.280	-0.001	0.910	0.285	-0.002

Note. CI-68 = 68% CI; CI-90 = 90% CI; CI-95 = 95% CI; Bank-based = 5th and 95th item bank difficulty; No routing = on-grade item bank only.

Table 4
Reliability Estimate at the End of Each Phase

Item Bank Size	Grade	Routing Rule	Fall		Winter		Spring	
			Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2
800	4	CI-68	0.879	0.928	0.888	0.931	0.883	0.929
		CI-90	0.879	0.928	0.883	0.929	0.887	0.932
		CI-95	0.879	0.927	0.884	0.927	0.882	0.929
		Bank-Based	0.879	0.928	0.887	0.931	0.882	0.928
		No routing	0.879	0.925	0.884	0.928	0.878	0.922
800	6	CI-68	0.878	0.927	0.885	0.931	0.886	0.932
		CI-90	0.878	0.927	0.885	0.929	0.885	0.930
		CI-95	0.878	0.927	0.888	0.930	0.883	0.928
		Bank-based	0.878	0.926	0.885	0.931	0.882	0.928
		No routing	0.878	0.925	0.883	0.928	0.879	0.921
500	4	CI-68	0.876	0.925	0.883	0.930	0.881	0.929
		CI-90	0.876	0.925	0.883	0.929	0.879	0.928
		CI-95	0.876	0.926	0.881	0.929	0.876	0.926
		Bank-based	0.876	0.927	0.883	0.929	0.878	0.928
		No routing	0.876	0.922	0.876	0.919	0.869	0.914
500	6	CI-68	0.870	0.921	0.881	0.926	0.879	0.927
		CI-90	0.870	0.921	0.879	0.925	0.880	0.929
		CI-95	0.870	0.922	0.878	0.925	0.875	0.924
		Bank-based	0.870	0.922	0.878	0.928	0.872	0.923
		No routing	0.870	0.920	0.871	0.916	0.866	0.910

Note. Phase 1 consisted of 25 items, and Phase 2 consisted of 16 items. Phase 2 reliability represents the reliability of ability estimates from both Phases combined (41 items).

Table 5
Classification Accuracy

Item Bank Size	Grade	Routing Rule	Fall		Winter		Spring	
			Accuracy	κ_{weighted}	Accuracy	κ_{weighted}	Accuracy	κ_{weighted}
800	4	CI-68	0.841	0.888	0.823	0.893	0.813	0.897
		CI-90	0.842	0.888	0.821	0.893	0.812	0.896
		CI-95	0.841	0.888	0.823	0.894	0.813	0.897
		Bank-based	0.840	0.888	0.822	0.893	0.812	0.896
		No routing	0.840	0.887	0.825	0.895	0.812	0.896
800	6	CI-68	0.840	0.887	0.826	0.895	0.813	0.897
		CI-90	0.841	0.888	0.826	0.895	0.812	0.896
		CI-95	0.841	0.888	0.824	0.894	0.813	0.896
		Bank-based	0.840	0.887	0.823	0.893	0.812	0.896
		No routing	0.839	0.887	0.823	0.893	0.813	0.897
500	4	CI-68	0.841	0.888	0.823	0.893	0.809	0.894
		CI-90	0.842	0.889	0.823	0.893	0.810	0.895
		CI-95	0.841	0.888	0.821	0.893	0.809	0.894
		Bank-based	0.841	0.888	0.821	0.892	0.809	0.894
		No routing	0.843	0.889	0.822	0.893	0.811	0.896
500	6	CI-68	0.840	0.888	0.822	0.893	0.811	0.896
		CI-90	0.840	0.887	0.824	0.895	0.809	0.894
		CI-95	0.839	0.887	0.823	0.893	0.811	0.896
		Bank-based	0.840	0.888	0.821	0.892	0.809	0.894
		No routing	0.840	0.887	0.821	0.893	0.811	0.895

Note. Grade marked with asterisk (*) is the grade of record. Displayed data is from the 800-item bank condition.

Number of Students Moving On- and Off-Grade

Figure 3 presents student routing diagrams to illustrate the number of students being routed to each pathway when off-grade routing is allowed for grade 6 with transition rule CI-68 as example. Each blue box is a test phase, and every two consecutive boxes represent a test administration. For the Fall test, 645 students remained on-grade, 335 students moved below grade, and 20 students moved to the upper grade from Phase 1 to Phase 2. When the fall test completed, 644 students remained on-grade, 1 student moved from on-grade to the upper grade, and the 335 students who had been moved to the below grade blueprint, constraints, and bank were sent back to the on-grade blueprint, constraints, and bank to begin Phase 1 of the next administration. This process repeated until the spring test concluded. Appendix C shows the routing diagram for Grade 4 and Grade 6 simulees for all transition rules. Comparing the transition rules, the CI-based rules with more conservative criteria had fewer students administered off-grade items. The bank-based transition rule showed the smallest number of students administered off-grade items, effectively making it a stricter rule than CI-based rules.

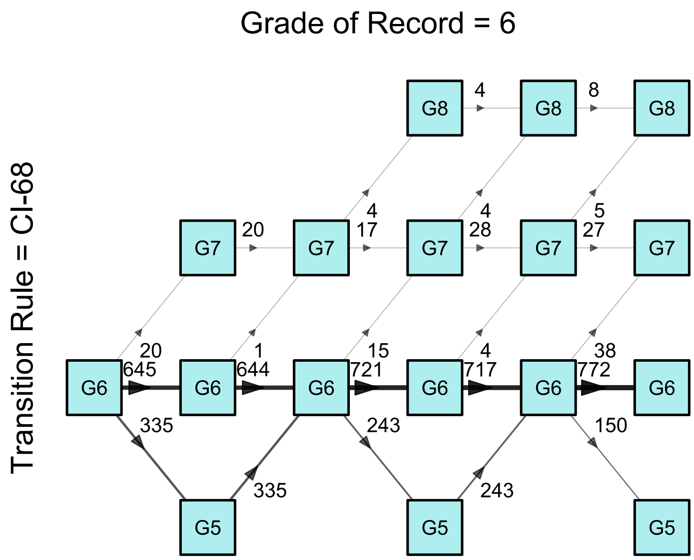
A shared pattern among the routing rules was that as the assessment progressed into later academic seasons, the number of students moving to the below-grade item bank gradually decreased. While this is mainly due to the data generation mechanism in this simulation, where student θs were generated from higher means in later academic seasons, this data generation artifact is based on what is observed in an interim assessment. It is intended in a live student population to capture students' growing abilities in the on-grade standards.

Item Bank Utilization and Exposure Rate

Whether any simulees received an item more than once across phases was also examined. This did not occur. Table 6 presents the item bank utilization rate under different conditions. The value of 1.0 indicates that all items from the bank were utilized. For example, for the Grade 4 bank with the transition rule based on CI-68, the value 0.88 indicates that 88% of the items in the Grade 4 bank were used in the fall simulation. As simulees who met the transition rules were routed to the adjacent off-grade banks in Phase 2, 24% of the Grade 3 items were utilized, whereas only 11% of the Grade 5 items were utilized. The item bank utilization rates differed across transition rules mainly because the number of students who were routed off-grade was dependent on which rule was used. The on-grade item bank utilization rate was not substantively different when transition rules were and were not invoked. The utilization rate for the item bank that stayed on-grade with no transition was 91% versus 88% for the most lenient transition rule. When narrower CI-based transition rules were used (i.e., more lenient routing), higher utilization rates were observed for off-grade banks, as would be expected from more lenient routing rules. The bank-based approach resulted in lower utilization rates for off-grade banks. Combined with results on estimation and classification performance, these results suggest that different routing rules can yield similar levels of reliability while varying item bank utilization.

Figure 4 shows the proportion of items that had specific ranges of exposure rates across conditions. Exposure rates shown in Figure 4 were obtained from items administered throughout all academic seasons combined. In general, conservative transition rules tended to have more

Figure 3
Number of Students Moving On- and Off-Grade



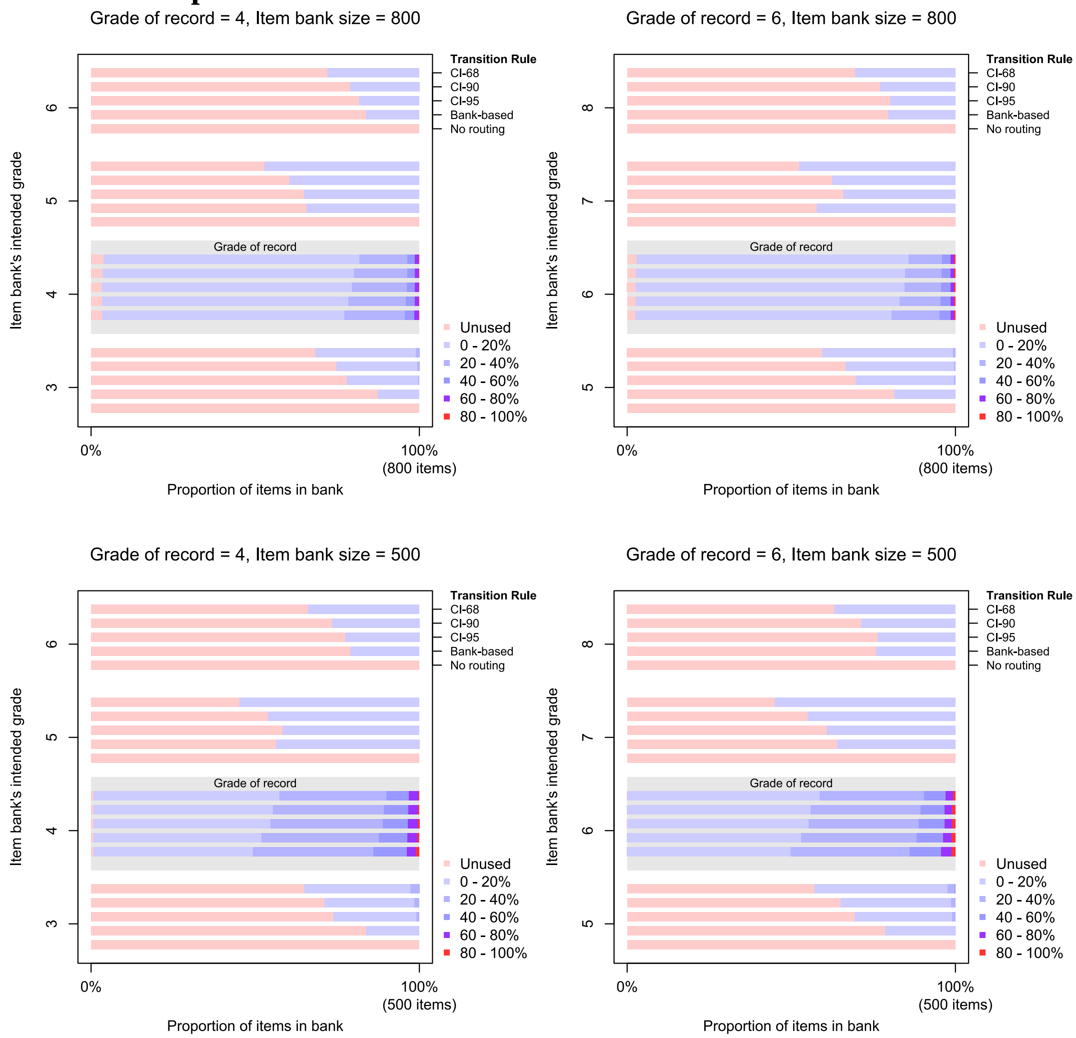
Note. Routing paths that had at least one simulee are displayed.

Table 6
Item Utilization Rates

Routing Rule	Item Bank Grade	Fall	Winter	Spring	Item Bank Grade	Fall	Winter	Spring
CI-68	3	0.242	0.270	0.283	5	0.336	0.358	0.365
	4*	0.884	0.937	0.960	6*	0.894	0.941	0.968
	5	0.109	0.335	0.428	7	0.112	0.326	0.426
	6	0.000	0.071	0.277	8	0.000	0.072	0.300
CI-90	3	0.193	0.210	0.222	5	0.278	0.290	0.296
	4*	0.887	0.941	0.963	6*	0.901	0.946	0.972
	5	0.077	0.263	0.356	7	0.081	0.243	0.335
	6	0.000	0.043	0.208	8	0.000	0.044	0.225
CI-95	3	0.167	0.183	0.193	5	0.256	0.265	0.269
	4*	0.889	0.942	0.964	6*	0.904	0.947	0.973
	5	0.061	0.223	0.315	7	0.067	0.218	0.308
	6	0.000	0.032	0.181	8	0.000	0.037	0.194
Bank-based	3	0.092	0.099	0.110	5	0.154	0.153	0.150
	4*	0.897	0.944	0.965	6*	0.903	0.947	0.972
	5	0.069	0.219	0.306	7	0.099	0.273	0.382
	6	0.000	0.033	0.159	8	0.000	0.049	0.199
No transition	4*	0.905	0.946	0.965	6*	0.911	0.956	0.974

Note. Grades marked asterisk (*) are the grade of record. Displayed data is from the 800-item bank condition.

Figure 4
Exposure Rate Distributions from All Seasons Combined



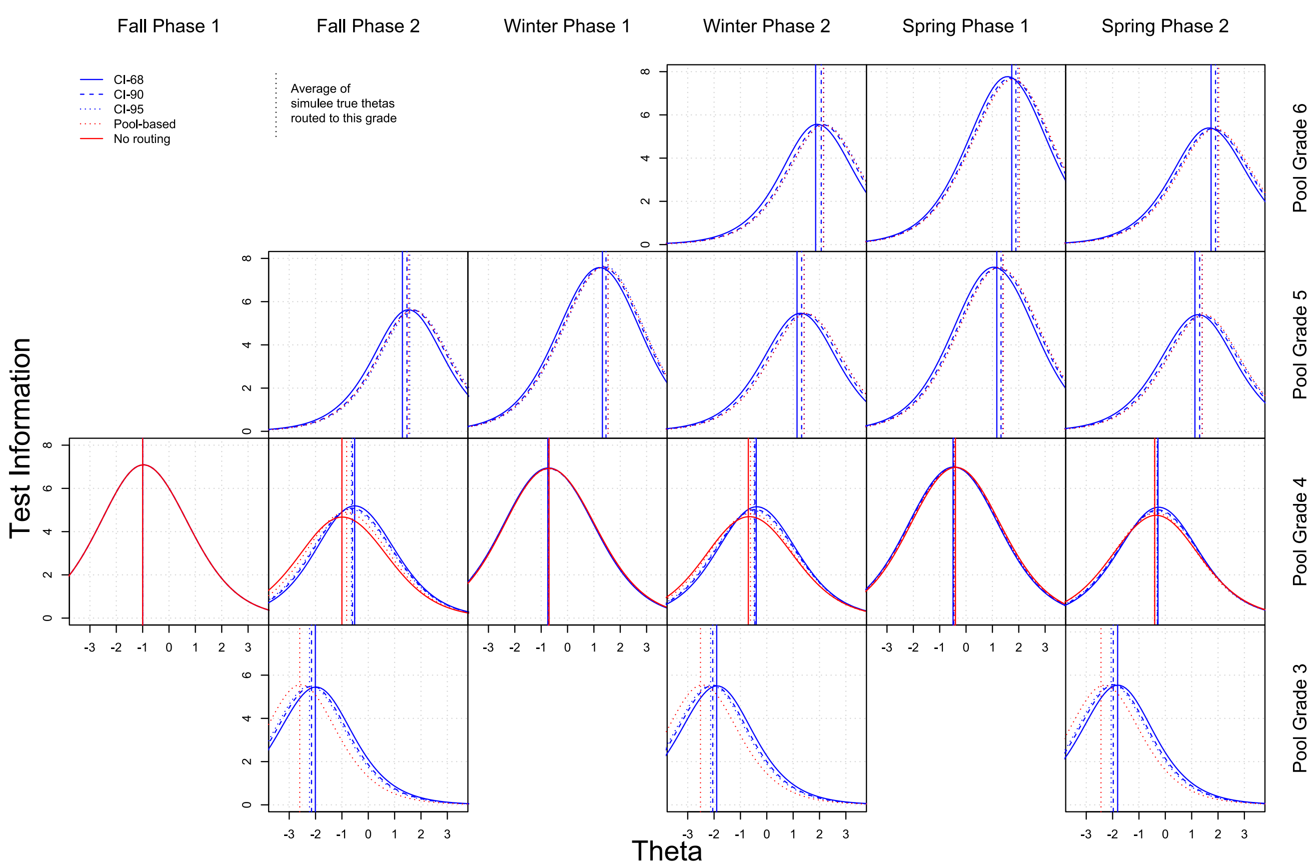
overexposed items in on-grade banks, as would be expected from conservative routing rules. The bank-based routing rule, being the strictest rule (excluding the no routing condition), had the most overexposed items in on-grade banks. These results suggest that when maintaining adequate item exposure rates is a concern, implementing an off-grade routing rule can provide an increase in items with adequate exposure rates (0%–20%) in the on-grade bank and a decrease in overexposed items (> 20%) in the on-grade bank while maintaining similar levels of θ estimation performance, if this is the primary focus of stakeholders.

Test Information by Phase

Figure 5 shows the average test information for each phase across academic seasons for simulees who did and did not meet the transition rule criteria across conditions for the 800-item bank. Patterns were similar across the examined conditions, so only the data from the Grade 4 800-item bank condition is displayed. Throughout all six phases and item bank grades, adaptive tests had close-to-optimal test information at simulees' true θs. This is consistent with performance expected from shadow-test-based optimal test assembly.

There was some variation in test information by routing rule. First, in the on-grade item banks (the row labeled "Bank Grade 4"), using lenient routing rules tended to yield higher test information and using conservative routing rules tended to yield lower test information. One possible cause of this pattern is that when a student stays on-grade in Phase 1 and Phase 2 in a single academic season, the form assembled in Phase 2 is subject to overlap control (i.e., items already administered in Phase 1 should not be included in Phase 2), which would lead to a less informative form in Phase 2. In contrast, when a student is routed off-grade in Phase 2, the form in Phase 2 is allowed to be assembled from the off-grade item bank, which has a more informative set of items for the simulee. Second, more conservative transition rules (i.e., CI-95, bank-based) tended to route students with more extreme θs to off-grade banks, which is consistent with how such routing rules would be expected to behave. This led to adaptive forms in off-grade banks having test information curves that were shifted toward more extreme θs.

Figure 5
Average Test Information Across Seasons by Routing Rule



Note. Test information functions shown in each cell is an average obtained from all adaptive tests administered in that season/phase/grade combination. Displayed data is from the [Grade of Record = 4, Item Bank Size = 800] condition.

Discussion

We investigated (1) the impact different transition rule conditions had on the accuracy of student θ estimates for each academic term when used to control if and how students should be administered off-grade items, and (2) how these results compared to estimates that were outcomes of simulees who were only administered on-grade items. Student-centered transition rules and the content-centered (bank-based) transition rule were compared to a no transition rule condition using 500-item and 800-item banks. Student θ estimates and classification accuracy estimates for achievement levels remained similar across all conditions. This preliminary evidence indicates that it might be feasible to transition students at the tails of the distributions to an off-grade bank, which supports more precise ability estimates. The design approach integrates a fundamental aspect of interim assessments—going off-grade—while meeting the ESSA requirement for proficiency determinations being derived from on-grade items and constraints.

The reliability of θ estimates and bias were sufficiently robust during Phase 1 as to meet criteria for making proficiency determinations. Adequate reliability, when the criterion-referenced transition decision is made (Bejar, 2016) and an assurance that students who are routed to off grade items are neither prohibited from being designated from proficient based on above-grade items nor included as proficient based on below-grade items (AIR, 2016) were among the considerations that needed to be investigated with the design prototype. The results from this simulation study indicated that the design met these criteria. Simulees routed to below-grade items using the more stringent transition rules maintained the same achievement level status when final θ estimates were aggregated across Phase 1 and Phase 2 (see Appendix B). Transitioning select students to off-grade items and communicating that information to teachers in carefully designed score reports should support teachers in better understanding the set of standards in which the student is currently functioning. Lohman and Korb (2006) argued that such information would better serve instructional decision making and student motivation to learn. The impact different conditions had on item exposure and utilization rates were also investigated. Across transition rules (including all students staying in the on-grade item bank), on-grade item utilization rates were similar. As expected, more items were exposed when the bank comprised 500 items than when the bank comprised 800 items. The off-grade item utilization and exposure rates were most similar for CI-90 and CI-95.

The content-centered approach resulted in the lowest off-grade item utilization rates because this transition rule routed the fewest students to off-grade items. The content-centered rule is optimal for states who would want to route students to off-grade items only when student abilities are above or below the on-grade items in the bank. Using such a rule could serve to support the reliability of the fall administration when the on-grade item bank might be too difficult for students early in the year. The simulation results showed that different transition rules lead to variations in the number of students routed to off-grade banks. However, if student ability trajectories increase across academic seasons in the same way as simulated in this study (and similar to what is observed in an interim assessment), always starting Phase 1 with on-grade items for students who were previously routed off-grade should also support reducing the number of students who move off-grade, given the findings of Meyer et al. (2023).

Limitations

As with any simulation study, these findings should be interpreted in context of the item banks used in the simulation. The item banks in the current study were generated to be vertically articulated and to function with state assessment content constraints. These data allowed a wide range of items to be available for the adaptive test assembly across simulees routed to on- and off-grade item banks. It is common for item banks for adaptive assessments to have areas where more items are needed for optimal adaptivity. The way items were simulated likely allowed the current adaptive test assembly simulation to maintain a high degree of reliability across the examined conditions. It would be useful to investigate the design using the item banks and constraints available from an on-grade through-year testing program to examine if the findings differ significantly from the results of this study when transition rules are applied. It would also be wise to administer the design to students in a small-scale pilot to verify the functionality. Which and if transition rules should be implemented is ultimately a state policy decision, but this study does lead to the conclusion that the quality of the item bank is likely a driving factor for ensuring that students are not moved to off-grade items due to areas of sparseness in the item bank.

This simulation modeled student change in ability over time (i.e., student growth), which was implemented as academic season changes in true θs. The design prototype investigated was able to accurately estimate each academic season θs, regardless of whether transition rules were used. Through-year assessments need to investigate if item banks are sufficient to accurately estimate student θs for each academic season. The shadow forms adaptively assembled for each test event had close-to-optimal test information for each academic season true θs (see Figure 5). These results show what might be feasible, but the results are also predicated on the notion that the field can engineer test score interpretations to help teachers understand how students are growing in the complexity and difficulty of the state standards (Schneider et al., this volume). Also of importance was that the data generation mechanism for the academic season true θs did not constrain θs to be monotonically increasing from fall, winter, and spring. Not all students increase in ability across the year. Still, different methods of generating parameters for academic season true θs might lead to different results.

References

Achter, J. A., Lubinski, D., & Benbow, C. P. (1996). Multipotentiality among the intellectually gifted: “It was never there and already it's vanishing.” *Journal of Counseling Psychology, 43*(1), 65–76. CrossRef

American Institutes for Research (AIR). (2016, October). *Smarter Balanced summative assessments simulation results*. WebLink

Bejar, I. (2016). Past and future of multistage testing in educational reform. In D. Yan, A. A. Von Davier, & C. Lewis, (Eds.), *Computerized multistage testing: Theory and applications*. CRC Press.

Bolt, D. M. (1999). Evaluating the effects of multidimensionality on IRT true-score equating. *Applied Measurement in Education, 12*(4), 383–407. CrossRef

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213–220. CrossRef

Choi, S. W., Lim, S., Niu, L., & Lee, S. (2022). *MAAT: Multiple administrations adaptive testing*. R package version 1.0.2.9000. WebLink

Choi, S. W., Lim, S., & van der Linden, W. J. (2021). TestDesign: An optimal test design approach to constructing fixed and adaptive tests in R. *Behaviormetrika*. CrossRef

Davey, T., Nering, M. L., & Thompson, T. (1997). *Realistic simulation of item response data*. ACT Research Report Series 97-4. WebLink

Egan, K. L., Schneider, M. C., & Ferrara, S. (2012). Performance level descriptors: History, practice and a proposed framework. In G. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 79–106). Routledge.

Eggen, T. J. H. M, & Straetmans, G. J. J. M. (2000). Computerized adaptive testing for classifying examinees into three categories. *Educational and Psychological Measurement, 60*, 713–734. CrossRef

Hendrickson, A. (2007). An NCME instructional module on multistage testing. *Educational Measurement: Issues and Practice , 26*(2), 44–52. CrossRef

Kingsbury, G. G., & Weiss, D. J. (1983). A comparison of IRT-based adaptive mastery testing and a sequential mastery testing procedure. In D. J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing* (pp. 237–254). Academic Press.

Li, S., & Meyer, J. P. (2019). *Simulation study for evaluating MAP® Growth™ grade-level item banks*. NWEA. Weblink

Lohman, D. F., & Korb, K. A. (2006). Gifted today but not tomorrow? Longitudinal changes in ability and achievement during elementary school. *Journal for the Education of the Gifted, 29*(4), 451–484. CrossRef

Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149–174.

Meyer, J. P., Hu, A., & Li, S. (2023). *Content proximity spring 2022 pilot study*. NWEA Research Brief. WebLink

Mills, C. J., & Barnett, L. B. (1992). The use of the Secondary School Admission Test (SSAT) to identify academically talented elementary school students. *Gifted Child Quarterly, 36*(3), 155–159. CrossRef

NWEA. (2019). *MAP® Growth™ technical report*. WebLink

Rambo-Hernandez, K. E., & Warne, R. T. (2015). Measuring the outliers: An introduction to out-of-level testing with high-achieving students. *Teaching Exceptional Children, 47*(4), 199–207. CrossRef

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. MESA Press.

Samejima, F. (1994). Estimation of reliability coefficients using the test information function and its modifications. *Applied Psychological Measurement, 18*(3), 229–244. CrossRef

Schneider, M. C., Choi., S. W., & Lewis, D. (this volume, 2025). Design considerations and reporting solutions for a multiple administrations adaptive testing system. *Journal of Computerized Adaptive Testing, 12(1),* 88-122. CrossRef

Stanley, J. C. (2005). A quiet revolution: Finding boys and girls who reason exceptionally well and/or verbally and helping them get the supplemental educational opportunities they need. *High Ability Studies 16*(1), 5–14. CrossRef

Stanley, J. C., & Benbow, C. P. (1981). *Using the SAT to find intellectually talented seventh graders*. College Board Review.

Stocking, M. L. (1988, May). *Some considerations in maintaining adaptive test item pools*. Educational Testing Service. *WebLink*

Svetina, D., Liaw, Y. L., Rutkowski, L., & Rutkowski, D. (2019). Routing strategies and optimizing design for multistage testing in international large-scale assessments. *Journal of Educational Measurement 56*(1), 192–213. *CrossRef*

Terman, L. M. (1926). *Genetic studies of genius...: Vol. I. Mental and Physical Traits of a Thousand Gifted Children*. Stanford University Press.

Thompson, N. A. (2007). A practitioner's guide for variable-length computerized classification testing. *Practical Assessment, Research, and Evaluation, 12*(1), 1. *CrossRef*

U.S. Department of Education (USED). (2016). *Final regulations* (Docket ID ED-2016-OESE-0032). *WebLink*

Van der Linden, W. J. & Reese, L. M (1998) A model for optimal constrained adaptive testing. Applied Psychological Measurement 22(3), 259-270. *WebLink*

van der Linden, W. J. (2003). Some alternatives to Sympson-Hetter item-exposure control in computerized adaptive testing. *Journal of Educational and Behavioral Statistics, 28*(3), 249–265. *CrossRef*

VanTassel-Baska, J., Benbow, C. P., & Lubinski, D. (1996). *Contributions of the talent-search concept to gifted education* (pp. 236–245). Johns Hopkins University Press.

Way, W. D., Twing, J. S., Camara, W., Sweeney, K., Lazer, S., & Mazzeo, J. (2010). *Some considerations related to the use of adaptive testing for the common core assessments*. Unpublished manuscript.

Webb, N. L. (2005). *Web Alignment Tool (WAT): Training manual (Draft Version 1.1)*. Wisconsin Center for Education Research, Council of Chief State School Officers.

Yan, D., Von Davier, A. A., & Lewis, C. (Eds.). (2014) *Computerized multistage testing: Theory and applications*. CRC Press.

Acknowledgements and Assistance

The MAAT simulator and design prototype was funded by NWEA.
MAAT is open-source and maintained by Seung Choi at the University of Texas.
We also would like to acknowledge Kelly Rivard for her editorial support.

Author Address

Jonghwan Lee Email: jay.lee1@pearson.com

Citation

Lee. J., Lim, S., Schneider, M.C., Gianopulos, G., Niu, L., Lee., S., & Choi, S.W. (2025).
The impact of item bank transition rules on student ability estimates
and achievement level classifications,
Journal of Computerized Adaptive Testing, 12(1), 88-122.

Appendix A: Constraints

Table A-1
Grade 3 Math Constraints

Constraint ID	Type	Condition	Number of Items Range	
			Phase 1	Phase 2
1	Number		25	16
2	Number	ITEM_TYPE == "Polytomous"	4–7	4–5
3	Number	STANDARD == "MA 3.1.1.a" & DOK %in% c(1, 2)	2–3	1–2
4	Number	STANDARD == "MA 3.1.1.b" & DOK %in% c(1, 1)	0–1	0–1
5	Number	STANDARD == "MA 3.1.1.c" & DOK %in% c(1, 1)	0–1	0–1
6	Number	STANDARD == "MA 3.1.1.d" & DOK %in% c(1, 2)	1–2	1–2
7	Number	STANDARD == "MA 3.1.1.e" & DOK %in% c(1, 2)	0–1	0–1
8	Number	STANDARD == "MA 3.1.1.f" & DOK %in% c(1, 2)	1	1
9	Number	STANDARD == "MA 3.1.1.g" & DOK %in% c(1, 2)	1	1
10	Number	STANDARD == "MA 3.1.1.h" & DOK %in% c(1, 2)	0–1	0–1
11	Number	STANDARD == "MA 3.1.1.i" & DOK %in% c(1, 2)	0–1	0–1
12	Number	DOMAIN == "NR"	5–6	5
13	Number	STANDARD == "MA 3.1.2.a" & DOK %in% c(1, 1)	0–1	0–1
14	Number	STANDARD == "MA 3.1.2.c" & DOK %in% c(1, 2)	0–1	0–1
15	Number	STANDARD == "MA 3.1.2.e" & DOK %in% c(1, 1)	0–1	0–1
16	Number	STANDARD == "MA 3.1.2.f" & DOK %in% c(1, 2)	0–1	0–1
17	Number	DOMAIN == "NO"	3	1–2
18	Number	UDOMAIN == "NUM"	9	7
19	Number	UDOMAIN == "NUM" & ITEM_TYPE == "Polytomous"	1–9	1–7
20	Number	STANDARD == "MA 3.2.1.a" & DOK %in% c(1, 2)	0–1	0–1
21	Number	STANDARD == "MA 3.2.1.b" & DOK %in% c(1, 2)	0–1	0–1
22	Number	DOMAIN == "AR"	1	0–1
23	Number	STANDARD == "MA 3.2.2.b" & DOK %in% c(1, 2)	1–2	1
24	Number	DOMAIN == "AP"	1–2	1
25	Number	STANDARD == "MA 3.2.3.a" & DOK %in% c(2, 2)	0–1	0–1
26	Number	STANDARD == "MA 3.2.3.b" & DOK %in% c(2, 2)	0–1	0–1
27	Number	DOMAIN == "AA"	1–2	0–1
28	Number	UDOMAIN == "ALG"	4	2
29	Number	UDOMAIN == "ALG" & ITEM_TYPE == "Polytomous"	1–4	1–2
30	Number	STANDARD == "MA 3.3.1.a" & DOK %in% c(1, 1)	2	1–2
31	Number	STANDARD == "MA 3.3.1.b" & DOK %in% c(1, 2)	0–1	0–1
32	Number	STANDARD == "MA 3.3.1.c" & DOK %in% c(1, 2)	1–2	1
33	Number	DOMAIN == "GC"	3	2
34	Number	STANDARD == "MA 3.3.3.a" & DOK %in% c(1, 2)	0–1	0–1
35	Number	STANDARD == "MA 3.3.3.b" & DOK %in% c(1, 1)	0–1	0–1
36	Number	STANDARD == "MA 3.3.3.c" & DOK %in% c(2, 2)	1	0–1

37	Number	STANDARD == "MA 3.3.3.e" & DOK %in% c(1, 1)	1	1
38	Number	STANDARD == "MA 3.3.3.g" & DOK %in% c(1, 2)	0–1	0–1
39	Number	STANDARD == "MA 3.3.3.h" & DOK %in% c(1, 3)	0–1	0–1
40	Number	DOMAIN == "GM"	4	2
41	Number	UDOMAIN == "GEO"	7	4
42	Number	UDOMAIN == "GEO" & ITEM_TYPE == "Polytomous"	1–7	1–4
43	Number	STANDARD == "MA 3.4.1.a" & DOK %in% c(2, 3)	2	1–2
44	Number	STANDARD == "MA 3.4.1.b" & DOK %in% c(1, 2)	1	0–1
45	Number	DOMAIN == "DR"	3	2
46	Number	STANDARD == "MA 3.4.2.a" & DOK %in% c(2, 2)	2	1
47	Number	DOMAIN == "DA"	2	1
48	Number	UDOMAIN == "DTA"	5	3
49	Number	UDOMAIN == "DTA" & ITEM_TYPE == "Polytomous"	1–5	1–3
50	SUM	POINTS	29–32	20–21

Table A-2
Grade 4 Math Constraints

			Number of Items Range	
Constraint ID	Type	Condition	Phase 1	Phase 2
1	Number		25	16
2	Number	ITEM_TYPE == "Polytomous"	4–7	4–5
3	Number	STANDARD == "MA 4.1.1.a" & DOK %in% c(1, 2)	0–1	0–1
4	Number	STANDARD == "MA 4.1.1.c" & DOK %in% c(1, 1)	0–1	0–1
5	Number	STANDARD == "MA 4.1.1.d" & DOK %in% c(1, 1)	0–1	0–1
6	Number	STANDARD == "MA 4.1.1.e" & DOK %in% c(1, 1)	0–1	0–1
7	Number	STANDARD == "MA 4.1.1.f" & DOK %in% c(1, 1)	0–2	0–1
8	Number	STANDARD == "MA 4.1.1.g" & DOK %in% c(1, 1)	0–1	0–1
9	Number	STANDARD == "MA 4.1.1.h" & DOK %in% c(1, 1)	1–2	1
10	Number	STANDARD == "MA 4.1.1.k" & DOK %in% c(1, 2)	0–1	0–1
11	Number	DOMAIN == "NR"	4–6	3–4
12	Number	STANDARD == "MA 4.1.2.b" & DOK %in% c(1, 1)	0–1	0–1
13	Number	STANDARD == "MA 4.1.2.c" & DOK %in% c(1, 1)	0–2	0–1
14	Number	STANDARD == "MA 4.1.2.d" & DOK %in% c(1, 1)	0–1	0–1
15	Number	STANDARD == "MA 4.1.2.f" & DOK %in% c(1, 1)	0–1	0–1
16	Number	STANDARD == "MA 4.1.2.g" & DOK %in% c(1, 1)	0–1	0–1
17	Number	DOMAIN == "NO"	5–6	3–4
18	Number	UDOMAIN == "NUM"	10	7
19	Number	UDOMAIN == "NUM" & ITEM_TYPE == "Polytomous"	1–10	1–7
20	Number	STANDARD == "MA 4.2.1.a" & DOK %in% c(1, 2)	1–2	0–1
21	Number	DOMAIN == "AR"	1	0–1

22	Number	STANDARD == "MA 4.2.2.a" & DOK %in% c(1, 2)	2–3	1–2
23	Number	DOMAIN == "AP"	2	1–2
24	Number	STANDARD == "MA 4.2.3.a" & DOK %in% c(2, 2)	1–2	0–1
25	Number	STANDARD == "MA 4.2.3.b" & DOK %in% c(2, 2)	1–2	0–2
26	Number	DOMAIN == "AA"	3	1–2
27	Number	UDOMAIN == "ALG"	6	4
28	Number	UDOMAIN == "ALG" & ITEM_TYPE == "Polytomous"	1–6	1–4
29	Number	STANDARD == "MA 4.3.1.b" & DOK %in% c(1, 2)	0–1	0–1
30	Number	STANDARD == "MA 4.3.1.c" & DOK %in% c(1, 2)	0–1	0–1
31	Number	STANDARD == "MA 4.3.1.d" & DOK %in% c(2, 3)	0–1	0–1
32	Number	STANDARD == "MA 4.3.1.e" & DOK %in% c(1, 2)	0–1	0–1
33	Number	STANDARD == "MA 4.3.1.f" & DOK %in% c(1, 2)	0–1	0–1
34	Number	STANDARD == "MA 4.3.1.g" & DOK %in% c(1, 2)	0–1	0–1
35	Number	STANDARD == "MA 4.3.1.h" & DOK %in% c(1, 2)	0–1	0–1
36	Number	DOMAIN == "GC"	4	2
37	Number	STANDARD == "MA 4.3.3.a" & DOK %in% c(1, 2)	0–1	0–1
38	Number	STANDARD == "MA 4.3.3.c" & DOK %in% c(1, 1)	0–1	0–1
39	Number	DOMAIN == "GM"	1	1
40	Number	UDOMAIN == "GEO"	5	3
41	Number	UDOMAIN == "GEO" & ITEM_TYPE == "Polytomous"	1–5	1–3
42	Number	STANDARD == "MA 4.4.1.a" & DOK %in% c(2, 2)	0–2	0–2
43	Number	DOMAIN == "DR"	1–2	1–2
44	Number	STANDARD == "MA 4.4.2.a" & DOK %in% c(2, 2)	2–3	0–2
45	Number	DOMAIN == "DA"	2	1–2
46	Number	UDOMAIN == "DTA"	4	2
47	Number	UDOMAIN == "DTA" & ITEM_TYPE == "Polytomous"	1–4	1–2
48	SUM	POINTS	29–32	20–21

Table A-3
Grade 5 Math Constraints

			Number of Items Range	
Constraint ID	Type	Condition	Phase 1	Phase 2
1	Number		25	16
2	Number	ITEM_TYPE == "Polytomous"	4–7	4–5
3	Number	STANDARD == "MA 5.1.1.a" & DOK %in% c(1, 2)	1–2	0–1
4	Number	STANDARD == "MA 5.1.1.b" & DOK %in% c(1, 2)	0–1	0–1
5	Number	STANDARD == "MA 5.1.1.c" & DOK %in% c(1, 1)	0–1	0–1
6	Number	STANDARD == "MA 5.1.1.d" & DOK %in% c(1, 2)	0–2	0–1
7	Number	STANDARD == "MA 5.1.1.e" & DOK %in% c(1, 1)	0–1	0–1
8	Number	DOMAIN == "NR"	4–6	2–3

9	Number	STANDARD == "MA 5.1.2.a" & DOK %in% c(1, 1)	0–2	0–1
10	Number	STANDARD == "MA 5.1.2.b" & DOK %in% c(1, 1)	0–2	0–1
11	Number	STANDARD == "MA 5.1.2.c" & DOK %in% c(1, 2)	0–1	0–1
12	Number	STANDARD == "MA 5.1.2.d" & DOK %in% c(1, 1)	0–2	0–1
13	Number	STANDARD == "MA 5.1.2.g" & DOK %in% c(1, 1)	1–2	1–2
14	Number	STANDARD == "MA 5.1.2.h" & DOK %in% c(1, 1)	0–2	0–1
15	Number	STANDARD == "MA 5.1.2.j" & DOK %in% c(1, 1)	0–2	0–1
16	Number	DOMAIN == "NO"	5–6	4–5
17	Number	UDOMAIN == "NUM"	10	7
18	Number	UDOMAIN == "NUM" & ITEM_TYPE == "Polytomous"	1–10	1–7
19	Number	STANDARD == "MA 5.2.1.a" & DOK %in% c(1, 1)	2–3	1
20	Number	DOMAIN == "AR"	2	1
21	Number	STANDARD == "MA 5.2.2.a" & DOK %in% c(1, 2)	3	2
22	Number	DOMAIN == "AP"	3	2
23	Number	STANDARD == "MA 5.2.3.a" & DOK %in% c(2, 3)	1	1
24	Number	DOMAIN == "AA"	1	1
25	Number	UDOMAIN == "ALG"	6	4
26	Number	UDOMAIN == "ALG" & ITEM_TYPE == "Polytomous"	1–6	1–4
27	Number	STANDARD == "MA 5.3.1.a" & DOK %in% c(1, 1)	0–2	0–2
28	Number	STANDARD == "MA 5.3.1.b" & DOK %in% c(1, 1)	0–2	0–2
29	Number	STANDARD == "MA 5.3.1.c" & DOK %in% c(2, 3)	0–2	0–2
30	Number	DOMAIN == "GC"	3	1
31	Number	STANDARD == "MA 5.3.2.b" & DOK %in% c(1, 1)	1	1
32	Number	DOMAIN == "GO"	1	1
33	Number	STANDARD == "MA 5.3.3.b" & DOK %in% c(1, 2)	0–1	0–1
34	Number	STANDARD == "MA 5.3.3.c" & DOK %in% c(1, 2)	0–1	0–1
35	Number	DOMAIN == "GM"	1	1
36	Number	UDOMAIN == "GEO"	5	3
37	Number	UDOMAIN == "GEO" & ITEM_TYPE == "Polytomous"	1–5	1–3
38	Number	STANDARD == "MA 5.4.2.a" & DOK %in% c(1, 3)	1–3	1–2
39	Number	STANDARD == "MA 5.4.2.b" & DOK %in% c(2, 3)	1–3	1–2
40	Number	DOMAIN == "DA"	4	2
41	Number	UDOMAIN == "DTA"	4	2
42	Number	UDOMAIN == "DTA" & ITEM_TYPE == "Polytomous"	1–4	1–2
43	SUM	POINTS	29–32	20–21

Table A-4
Grade 6 Math Constraints

			Number of Items Range	
Constraint ID	Type	Condition	Phase 1	Phase 2
1	Number		25	16
2	Number	ITEM_TYPE == "Polytomous"	4–7	4–5
3	Number	STANDARD == "MA 6.1.1.a" & DOK %in% c(1, 2)	0–2	0–1
4	Number	STANDARD == "MA 6.1.1.b" & DOK %in% c(1, 1)	0–1	0–1
5	Number	STANDARD == "MA 6.1.1.c" & DOK %in% c(1, 2)	0–2	0–1
6	Number	STANDARD == "MA 6.1.1.d" & DOK %in% c(1, 2)	0–1	0–1
7	Number	STANDARD == "MA 6.1.1.g" & DOK %in% c(2, 2)	0–2	0–1
8	Number	STANDARD == "MA 6.1.1.h" & DOK %in% c(1, 2)	0–1	0–1
9	Number	STANDARD == "MA 6.1.1.i" & DOK %in% c(1, 1)	0–1	0–1
10	Number	DOMAIN == "NR"	4–5	3–4
11	Number	STANDARD == "MA 6.1.2.a" & DOK %in% c(1, 1)	0–1	0–1
12	Number	STANDARD == "MA 6.1.2.c" & DOK %in% c(1, 1)	0–1	0–1
13	Number	STANDARD == "MA 6.1.2.d" & DOK %in% c(1, 1)	0–1	0–1
14	Number	STANDARD == "MA 6.1.2.e" & DOK %in% c(2, 2)	0–1	0–1
15	Number	DOMAIN == "NO"	3	1–2
16	Number	UDOMAIN == "NUM"	7	5
17	Number	UDOMAIN == "NUM" & ITEM_TYPE == "Polytomous"	1–7	1–5
18	Number	STANDARD == "MA 6.2.1.a" & DOK %in% c(1, 2)	1	1
19	Number	DOMAIN == "AR"	1	1
20	Number	STANDARD == "MA 6.2.2.a" & DOK %in% c(1, 1)	0–2	0–1
21	Number	STANDARD == "MA 6.2.2.b" & DOK %in% c(1, 2)	0–1	0–1
22	Number	STANDARD == "MA 6.2.2.c" & DOK %in% c(1, 1)	0–2	0–1
23	Number	STANDARD == "MA 6.2.2.d" & DOK %in% c(1, 2)	0–2	0–1
24	Number	STANDARD == "MA 6.2.2.e" & DOK %in% c(1, 1)	0–2	0–1
25	Number	STANDARD == "MA 6.2.2.f" & DOK %in% c(2, 2)	0–1	0–1
26	Number	STANDARD == "MA 6.2.2.g" & DOK %in% c(1, 2)	0–2	0–1
27	Number	DOMAIN == "AP"	5	3
28	Number	STANDARD == "MA 6.2.3.b" & DOK %in% c(2, 2)	0–2	0–1
29	Number	STANDARD == "MA 6.2.3.c" & DOK %in% c(2, 2)	1–2	0–1
30	Number	STANDARD == "MA 6.2.3.d" & DOK %in% c(2, 2)	0–2	0–1
31	Number	DOMAIN == "AA"	3	1
32	Number	UDOMAIN == "ALG"	9	5
33	Number	UDOMAIN == "ALG" & ITEM_TYPE == "Polytomous"	1–9	1–5
34	Number	STANDARD == "MA 6.3.1.a" & DOK %in% c(1, 2)	1	0–1
35	Number	DOMAIN == "GC"	1	0–1
36	Number	STANDARD == "MA 6.3.2.a" & DOK %in% c(1, 1)	0–1	0–1
37	Number	STANDARD == "MA 6.3.2.c" & DOK %in% c(1, 2)	0–1	0–1

38	Number	STANDARD == "MA 6.3.2.d" & DOK %in% c(2, 2)	0–1	0–1
39	Number	DOMAIN == "GO"	2–3	1–2
40	Number	STANDARD == "MA 6.3.3.a" & DOK %in% c(2, 2)	1	0–1
41	Number	STANDARD == "MA 6.3.3.b" & DOK %in% c(2, 2)	0–1	0–1
42	Number	STANDARD == "MA 6.3.3.c" & DOK %in% c(2, 2)	0–1	0–1
43	Number	DOMAIN == "GM"	2	1
44	Number	UDOMAIN == "GEO"	5	3
45	Number	UDOMAIN == "GEO" & ITEM_TYPE == "Polytomous"	1–5	1–3
46	Number	STANDARD == "MA 6.4.2.a" & DOK %in% c(2, 2)	1–3	0–1
47	Number	STANDARD == "MA 6.4.2.b" & DOK %in% c(2, 3)	1–3	0–1
48	Number	STANDARD == "MA 6.4.2.c" & DOK %in% c(1, 2)	1–3	0–1
49	Number	STANDARD == "MA 6.4.2.d" & DOK %in% c(2, 3)	1–2	0–1
50	Number	DOMAIN == "DA"	4	3
51	Number	UDOMAIN == "DTA"	4	3
52	Number	UDOMAIN == "DTA" & ITEM_TYPE == "Polytomous"	1–4	1–3
53	SUM	POINTS	29–32	20–21

Table A-5
Grade 7 Math Constraints

			Number of Items Range	
Constraint ID	Type	Condition	Phase 1	Phase 2
1	Number		25	16
2	Number	ITEM_TYPE == "Polytomous"	4–7	4–5
3	Number	STANDARD == "MA 7.1.2.a" & DOK %in% c(1, 2)	1–3	0–2
4	Number	STANDARD == "MA 7.1.2.b" & DOK %in% c(1, 2)	0–2	0–1
5	Number	STANDARD == "MA 7.1.2.d" & DOK %in% c(1, 2)	1	0–1
6	Number	STANDARD == "MA 7.1.2.e" & DOK %in% c(2, 2)	1	0–1
7	Number	DOMAIN == "NO"	5	3
8	Number	UDOMAIN == "NUM"	5	3
9	Number	UDOMAIN == "NUM" & ITEM_TYPE == "Polytomous"	1–5	1–3
10	Number	STANDARD == "MA 7.2.1.a" & DOK %in% c(1, 2)	1–2	0–1
11	Number	STANDARD == "MA 7.2.1.b" & DOK %in% c(2, 2)	0–1	0–1
12	Number	DOMAIN == "AR"	2	1
13	Number	STANDARD == "MA 7.2.2.a" & DOK %in% c(1, 1)	0–1	0–1
14	Number	STANDARD == "MA 7.2.2.b" & DOK %in% c(1, 1)	0–1	0–1
15	Number	STANDARD == "MA 7.2.2.c" & DOK %in% c(1, 2)	0–1	0–1
16	Number	STANDARD == "MA 7.2.2.d" & DOK %in% c(1, 1)	0–1	0–1
17	Number	STANDARD == "MA 7.2.2.e" & DOK %in% c(1, 2)	0–1	0–1
18	Number	DOMAIN == "AP"	3–5	2–3
19	Number	STANDARD == "MA 7.2.3.a" & DOK %in% c(1, 1)	0–1	0–1

20	Number	STANDARD == "MA 7.2.3.b" & DOK %in% c(2, 2)	0–1	0–1
21	Number	STANDARD == "MA 7.2.3.c" & DOK %in% c(2, 2)	0–1	0–1
22	Number	STANDARD == "MA 7.2.3.d" & DOK %in% c(2, 2)	0–1	0–1
23	Number	STANDARD == "MA 7.2.3.e" & DOK %in% c(2, 2)	0–1	0–1
24	Number	STANDARD == "MA 7.2.3.f" & DOK %in% c(2, 2)	0–1	0–1
25	Number	DOMAIN == "AA"	3–5	2–3
26	Number	UDOMAIN == "ALG"	9	6
27	Number	UDOMAIN == "ALG" & ITEM_TYPE == "Polytomous"	1–9	1–6
28	Number	STANDARD == "MA 7.3.1.a" & DOK %in% c(2, 2)	1	1
29	Number	DOMAIN == "GC"	1	1
30	Number	STANDARD == "MA 7.3.3.a" & DOK %in% c(2, 2)	0–2	0–1
31	Number	STANDARD == "MA 7.3.3.b" & DOK %in% c(2, 2)	0–2	0–1
32	Number	STANDARD == "MA 7.3.3.c" & DOK %in% c(1, 2)	0–2	0–1
33	Number	DOMAIN == "GO"	4	2
34	Number	UDOMAIN == "GEO"	5	3
35	Number	UDOMAIN == "GEO" & ITEM_TYPE == "Polytomous"	1–5	1–3
36	Number	STANDARD == "MA 7.4.2.a" & DOK %in% c(2, 2)	2	1
37	Number	DOMAIN == "DR"	2	1
38	Number	STANDARD == "MA 7.4.3.b" & DOK %in% c(2, 3)	0–1	0–1
39	Number	STANDARD == "MA 7.4.3.c" & DOK %in% c(2, 3)	0–1	0–1
40	Number	STANDARD == "MA 7.4.3.e" & DOK %in% c(1, 2)	0–1	0–1
41	Number	STANDARD == "MA 7.4.3.f" & DOK %in% c(1, 2, 3)	0–1	0–1
42	Number	STANDARD == "MA 7.4.3.g" & DOK %in% c(2, 3)	0–1	0–1
43	Number	STANDARD == "MA 7.4.3.h" & DOK %in% c(1, 2)	0–1	0–1
44	Number	DOMAIN == "DA"	4	3
45	Number	UDOMAIN == "DTA"	6	4
46	Number	UDOMAIN == "DTA" & ITEM_TYPE == "Polytomous"	1–6	1–4
47	SUM	POINTS	29–32	20–21

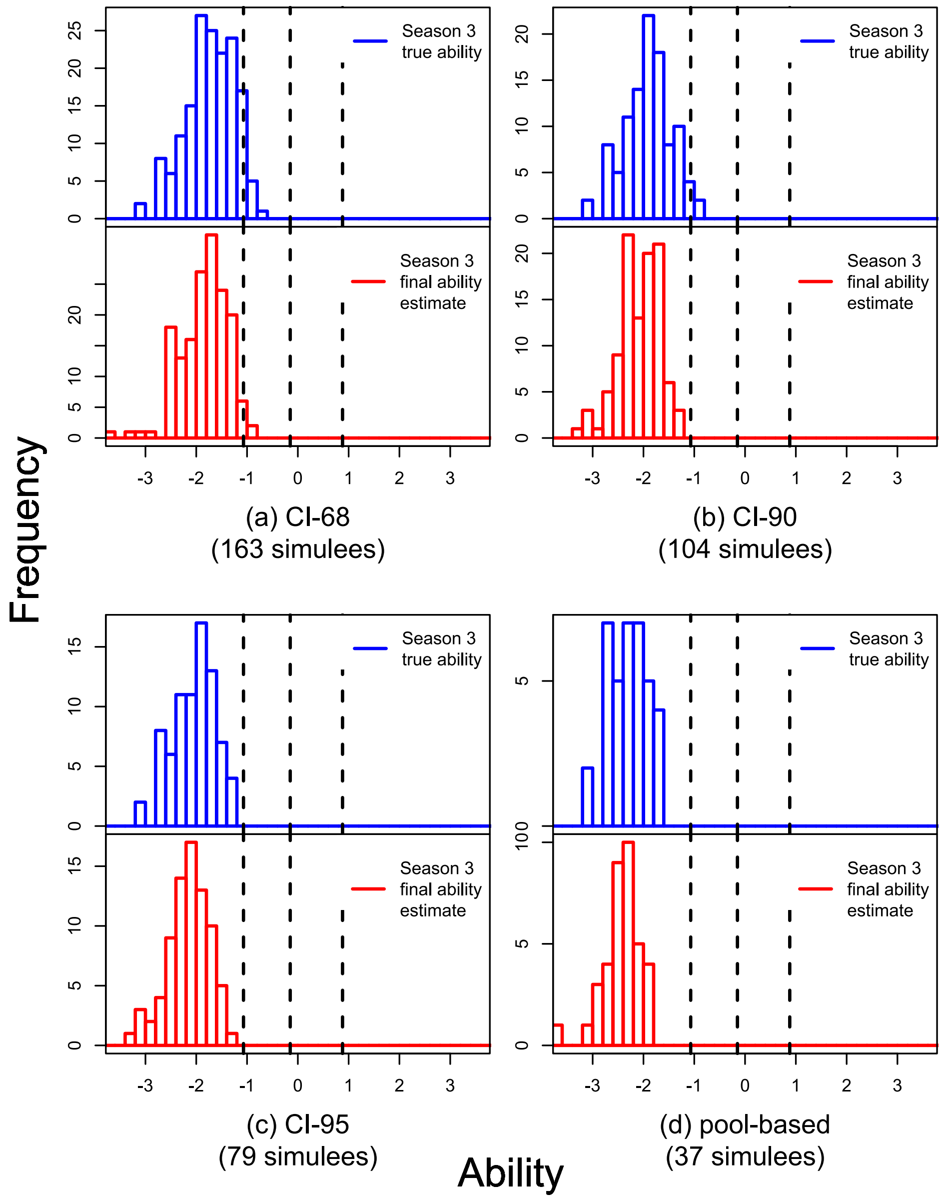
Table A-6
Grade 8 Math Constraints

			Number of Items Range	
Constraint ID	Type	Condition	Phase 1	Phase 2
1	Number		25	16
2	Number	ITEM_TYPE == "Polytomous"	4–7	4–5
3	Number	STANDARD == "MA 8.1.1.a" & DOK %in% c(1, 1)	0–1	0–1
4	Number	STANDARD == "MA 8.1.1.b" & DOK %in% c(1, 1)	0–1	0–1
5	Number	STANDARD == "MA 8.1.1.d" & DOK %in% c(1, 2)	1–2	0–1
6	Number	DOMAIN == "NR"	2–3	1–2
7	Number	STANDARD == "MA 8.1.2.a" & DOK %in% c(1, 1)	0–1	0–1

8	Number	STANDARD == "MA 8.1.2.b" & DOK %in% c(1, 1)	0–1	0–1
9	Number	STANDARD == "MA 8.1.2.c" & DOK %in% c(1, 1)	0–1	0–1
10	Number	STANDARD == "MA 8.1.2.e" & DOK %in% c(1, 2)	0–1	0–1
11	Number	DOMAIN == "NO"	3–4	2–3
12	Number	UDOMAIN == "NUM"	6	4
13	Number	UDOMAIN == "NUM" & ITEM_TYPE == "Polytomous"	1–6	1–4
14	Number	STANDARD == "MA 8.2.1.a" & DOK %in% c(1, 2)	1–2	0–1
15	Number	STANDARD == "MA 8.2.1.b" & DOK %in% c(1, 2)	0–2	0–1
16	Number	STANDARD == "MA 8.2.1.c" & DOK %in% c(1, 1)	0–1	0–1
17	Number	STANDARD == "MA 8.2.1.d" & DOK %in% c(1, 2)	0–1	0–1
18	Number	DOMAIN == "AR"	3–4	2–3
19	Number	STANDARD == "MA 8.2.2.a" & DOK %in% c(1, 1)	1–2	0–1
20	Number	STANDARD == "MA 8.2.2.b" & DOK %in% c(1, 2)	0–1	0–1
21	Number	DOMAIN == "AP"	1–2	1–2
22	Number	STANDARD == "MA 8.2.3.a" & DOK %in% c(1, 1)	0–1	0–1
23	Number	STANDARD == "MA 8.2.3.b" & DOK %in% c(2, 2)	0–1	0–1
24	Number	STANDARD == "MA 8.2.3.c" & DOK %in% c(2, 3)	0–2	0–1
25	Number	DOMAIN == "AA"	3	2–3
26	Number	UDOMAIN == "ALG"	8	5
27	Number	UDOMAIN == "ALG" & ITEM_TYPE == "Polytomous"	1–8	1–5
28	Number	STANDARD == "MA 8.3.1.a" & DOK %in% c(2, 2)	0–2	0–1
29	Number	STANDARD == "MA 8.3.1.b" & DOK %in% c(1, 2)	1–2	0–1
30	Number	DOMAIN == "GC"	1–2	1–2
31	Number	STANDARD == "MA 8.3.2.a" & DOK %in% c(2, 2)	0–1	0–1
32	Number	STANDARD == "MA 8.3.2.b" & DOK %in% c(1, 2)	0–1	0–1
33	Number	STANDARD == "MA 8.3.2.c" & DOK %in% c(1, 2)	0–1	0–1
34	Number	DOMAIN == "GO"	1–2	1–2
35	Number	STANDARD == "MA 8.3.3.b" & DOK %in% c(2, 3)	0–2	0–1
36	Number	STANDARD == "MA 8.3.3.c" & DOK %in% c(1, 2)	0–1	0–1
37	Number	STANDARD == "MA 8.3.3.d" & DOK %in% c(2, 2)	0–2	0–1
38	Number	DOMAIN == "GM"	3–4	2–3
39	Number	UDOMAIN == "GEO"	7	5
40	Number	UDOMAIN == "GEO" & ITEM_TYPE == "Polytomous"	1–7	1–5
41	Number	STANDARD == "MA 8.4.1.a" & DOK %in% c(1, 1)	1–3	1
42	Number	DOMAIN == "DR"	1–3	1
43	Number	STANDARD == "MA 8.4.2.a" & DOK %in% c(2, 2)	1–3	1
44	Number	DOMAIN == "DA"	1–3	1
45	Number	UDOMAIN == "DTA"	4	2
46	Number	UDOMAIN == "DTA" & ITEM_TYPE == "Polytomous"	1–4	1–2
47	SUM	POINTS	29–32	20–21

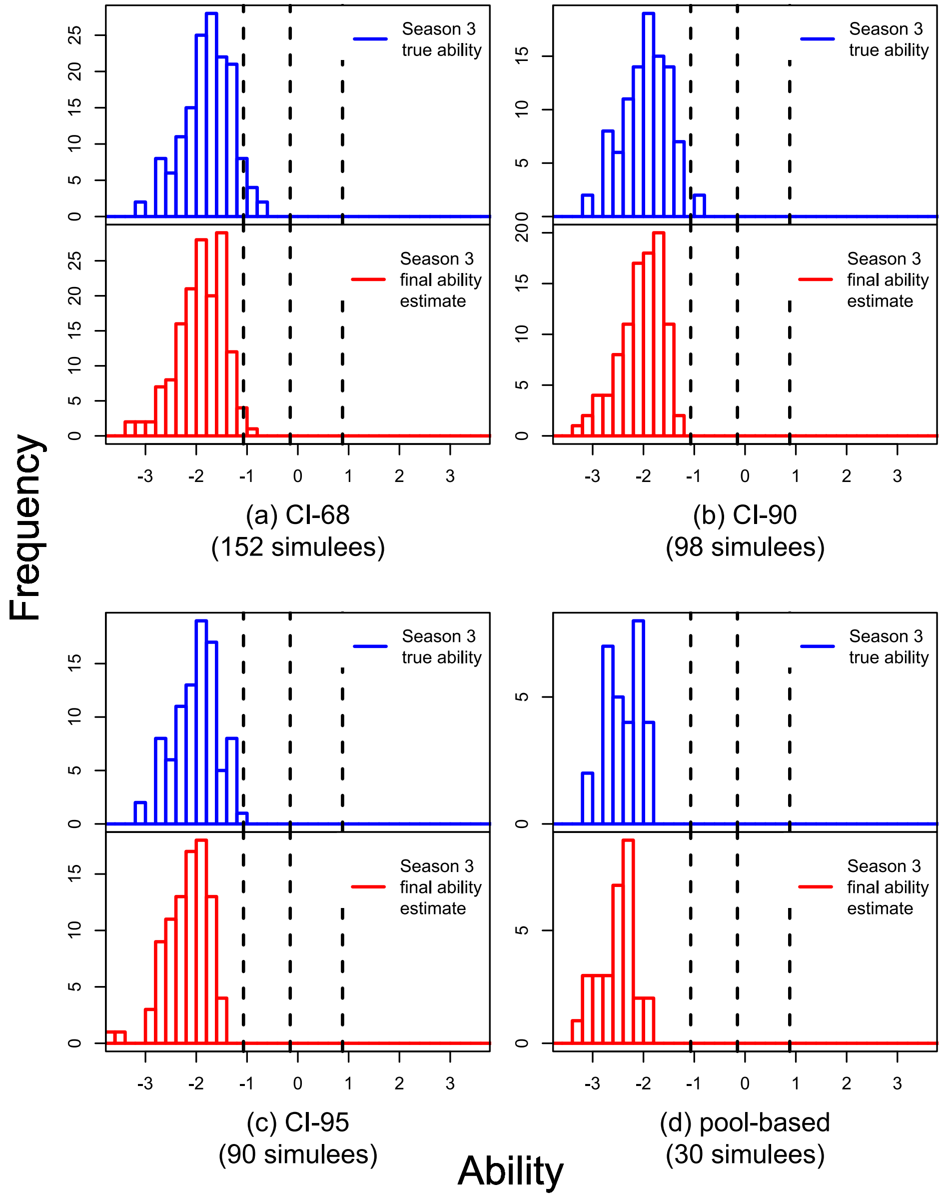
Appendix B: Ability Distributions

Figure B-1
Ability Distributions of Students Routed to Below-Grade in Season 3 (800-Item Banks)



Note. Displayed data are from one replication (four simulations, 1,000 students each) where the grade of record was Grade 4 and item banks had sizes of 800 in each grade. The number of simulees shown under panels (a) to (d) are the number of students who were routed to below-grade (Grade 3) in Season 3 Phase 2. The vertical dotted lines show the cutscores: the first was used for routing below, and the third was used for routing above (the middle cut was not used for routing purposes).

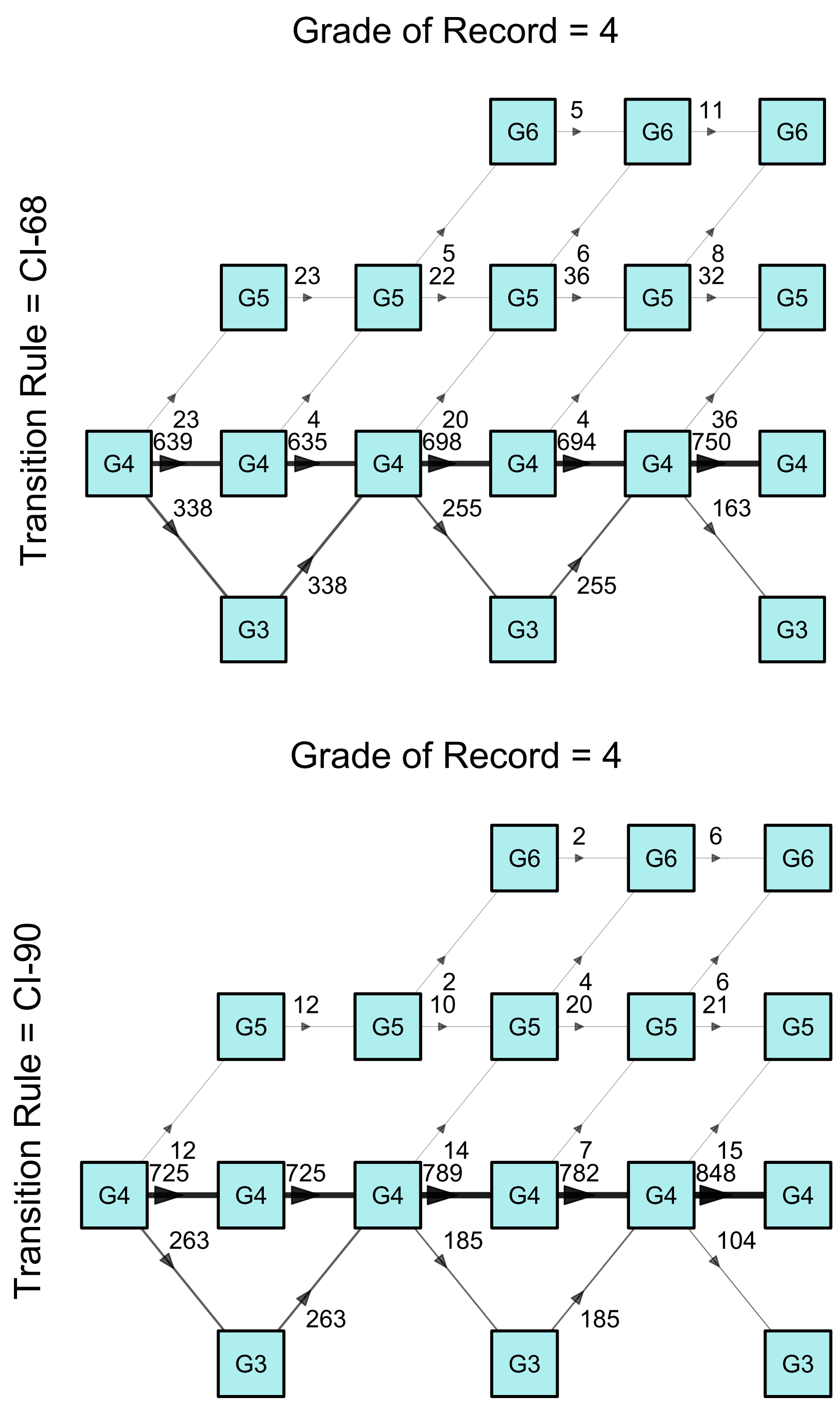
Figure B-2
Ability Distributions of Students Routed to Below-Grade in Season 3 (500-Item Banks)

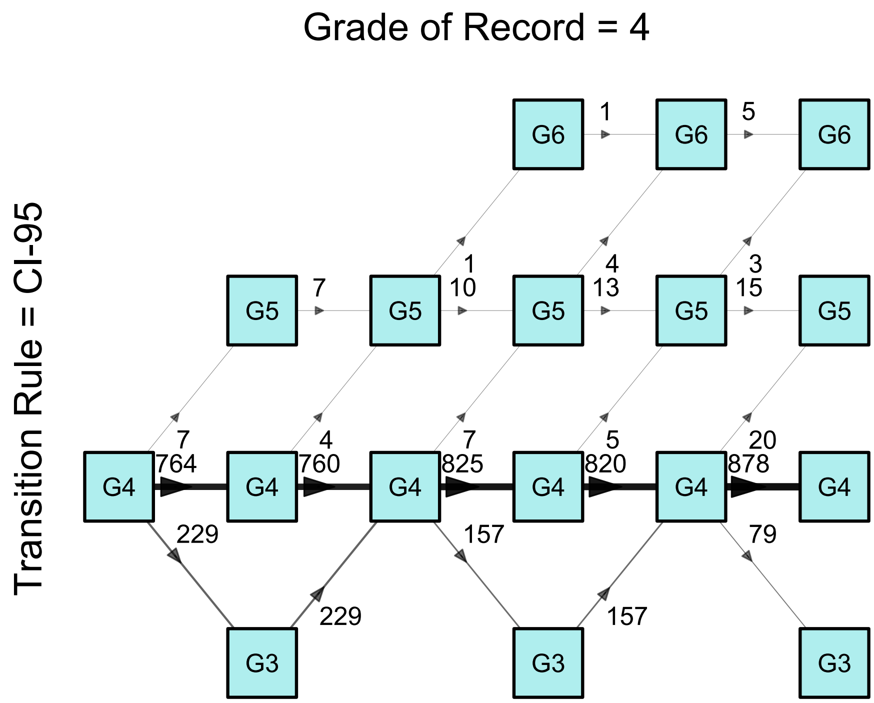


Note. Displayed data are from one replication (four simulations, 1,000 students each) where the grade of record was Grade 4 and item banks had sizes of 500 in each grade. The number of simulees shown under panels (a) to (d) are the number of students who were routed to below-grade (Grade 3) banks and constraints in Season 3 Phase 2. The vertical dotted lines show the cutscores: the first was used for routing below, and the third was used for routing above (the middle cut was not used for routing purposes).

Appendix C

Figure C-1
Number of Students Moving On- and Off-Grade (Grade 4 Simulation)




Grade of Record = 4
Transition Rule = CI-95
G6
1
G6
5
G6
G5
7
G5
1
10
G5
4
13
G5
3
15
G5
G4
7
764
G4
4
760
G4
7
825
G4
5
820
G4
20
878
G4
229
157
79
229
157
G3
G3
G3

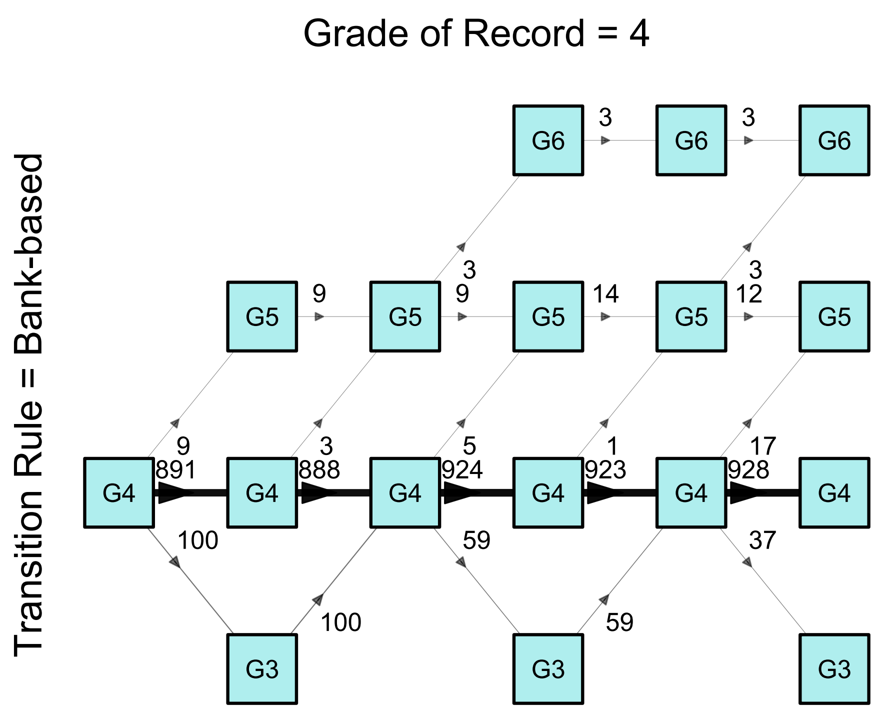
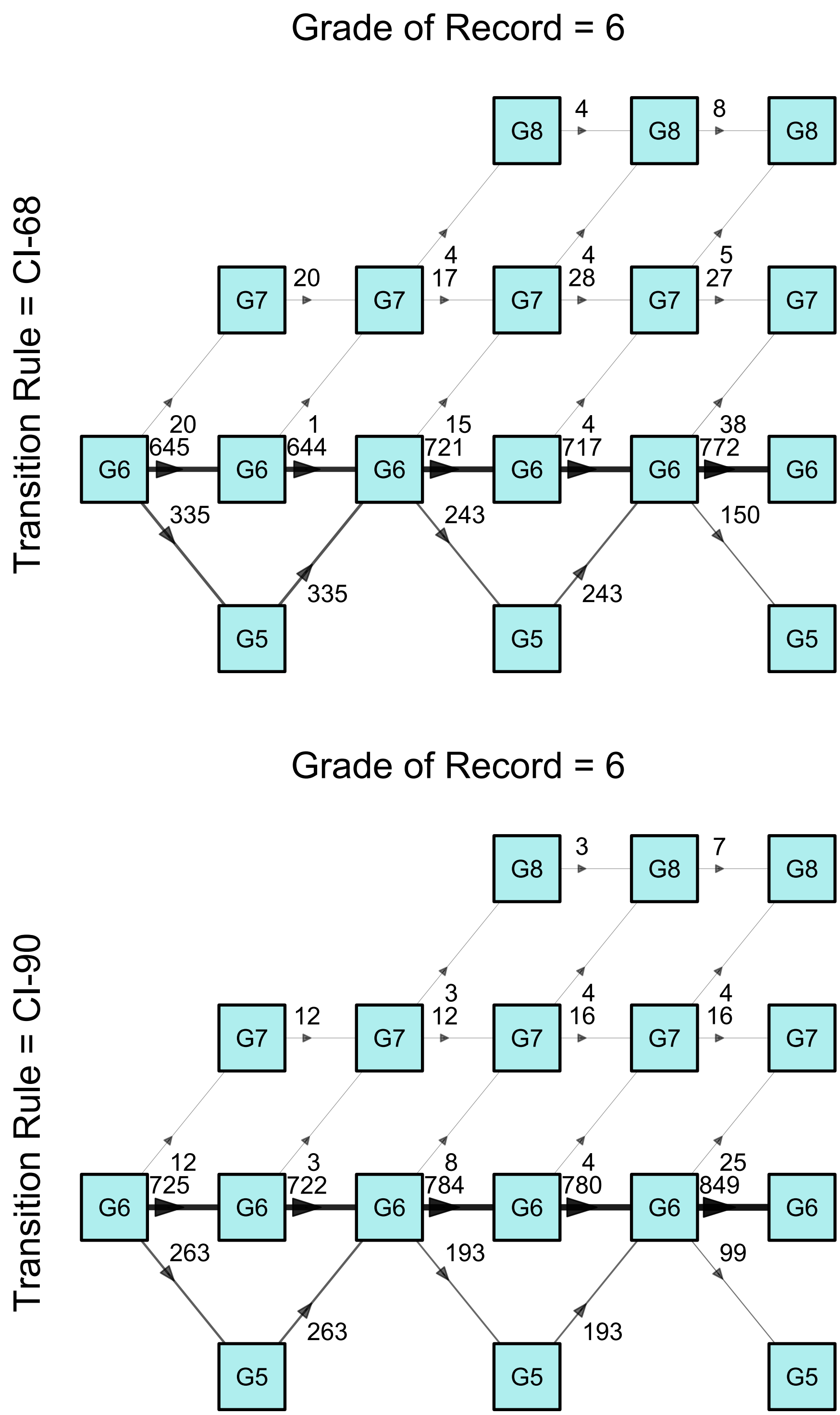
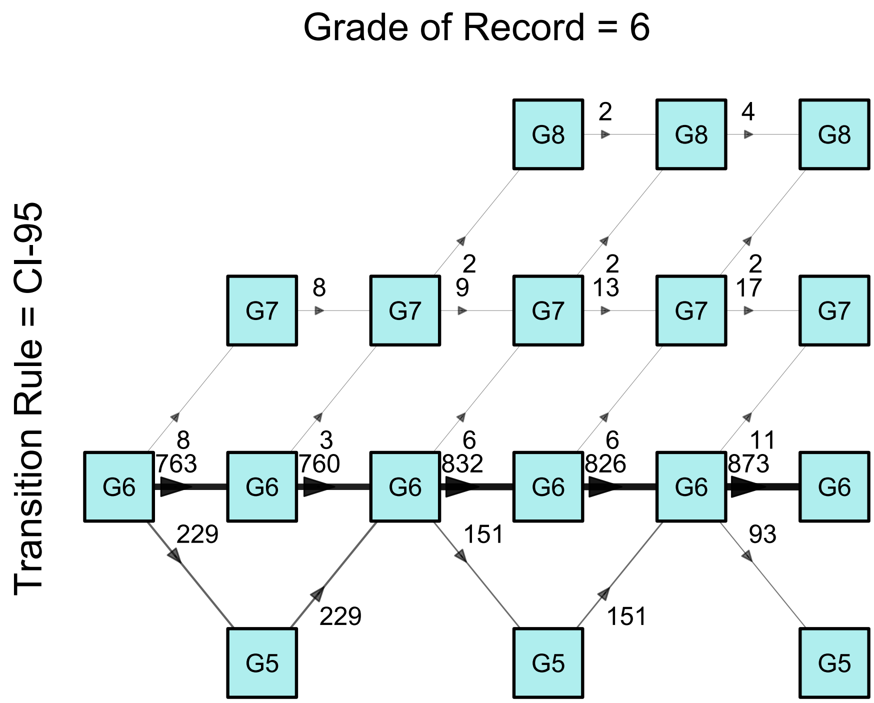
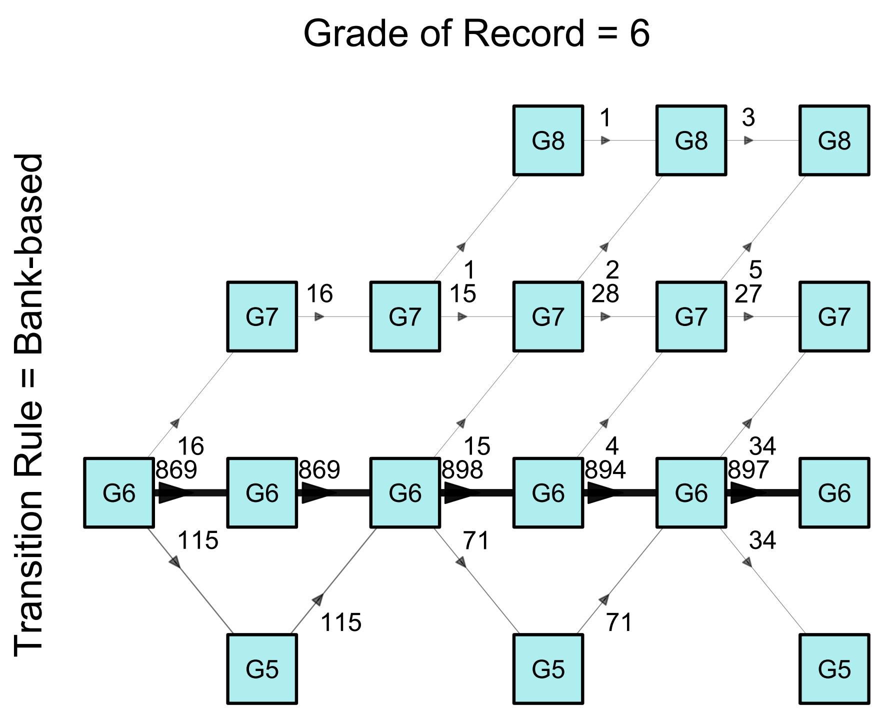

Grade of Record = 4
Transition Rule = Bank-based
G6
3
G6
3
G6
G5
9
G5
3
9
G5
14
G5
3
12
G5
G4
9
891
G4
3
888
G4
5
924
G4
1
923
G4
17
928
G4
100
59
37
100
59
G3
G3
G3

Figure C-2
Number of Students Moving On- and Off-Grade (Grade 6 Simulation)




Grade of Record = 6
Transition Rule = CI-95
G8 2 G8 4 G8
G7 8 G7 9 G7 13 G7 17 G7
2 2 2
8 3 6 6 11
G6 763 G6 760 G6 832 G6 826 G6 873 G6
229 151 93
229 151
G5 G5 G5


Grade of Record = 6
Transition Rule = Bank-based
G8 1 G8 3 G8
G7 16 G7 15 G7 28 G7 27 G7
1 2 5
16 15 4 34
G6 869 G6 869 G6 898 G6 894 G6 897 G6
115 71 34
115 71
G5 G5 G5