Hypothesis Testing for Adaptive Measurement of Individual Change

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Ji Eun Lee

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Thesis Advisor:
David J. Weiss

June 2015

Acknowledgements

First and foremost I would like to express my deep gratitude to my advisor David J. Weiss. He has guided me how to be a good psychometric researcher. His enthusiasm and keen insight for his research has been very motivating for me. I have learned a lot from him in developing research ideas, designing a study, and good academic writing, thanks to his thoughtful teaching. I also appreciate his support during tough times in graduate school. Without his encouragement and patience, all would not be possible.

I would also like to thank my committee members, Chun Wang, Niels Waller, and Adam Rothman for their time and valuable feedback. I would particulary like to acknowledge Chun Wang, who has been an inspiring researcher and provided constructive suggestions from the planning and throughout the development of the study, as well as encouraged me to finish work. I could accomplish the study more professionally thanks to her help.

My life in graduate school would not been as fruitful without my colleagues on the sixth floor of Elliott Hall: Ben Babcock, Chaitali Phadke, Chris Hulme-Lowe, Dong Gi Seo, Jeff Jones, Leah Feuerstahler, Lian Hortensius, Shannon Von Minden, and Steve Nydick. You have been motivating for many years, helped me to think more deeply and critically. I really enjoyed discussions we have had, academically and non-academically, and more importantly, the coffee-breaks we had were great refreshment to me.

Six years in Minnesota, including severe winters, have been enjoyable and memorable thanks to my friends. Yenna Yoo Bordwell and Lucas Bordwell, thank you for delicious food, coffee, more than enough drink, and heartful cheers. Juri Joeng and Sumi Jung, thank you for sharing time and memory from the first day at Minnesota. Sooyeon Sung, thank you for your warm welcome whenever I needed your big hug. Chulho Kim, thank you for your support especially for helping me to settle down my new life in Chicago and encouraging me to finish thesis writing. Jane Rauenhorst, thank you for helping me to escape from frustration and struggle. People in KTC, I can’t describe my six years in Minnesota without you. It was a great refreshment to play tennis with you and, more importantly, to have happy hour break after the game.

Finally, I wish to thank my parents, sister, and brother, who have always had faith in me, supported me and encouraged me.

Abstract

The significance of individual change has been an important topic in psychology and related fields. This study investigated performance of five hypothesis testing methods—Z, likelihood ratio, score test, and Kullback-Leibler divergence test with uniform and normal prior distributions—and three item selection methods—Fisher information, Kullback-Leibler information and a modified Kullback-Leibler information—as an extension of Finkelman et al.'s (2010) methods to determine the significance of individual change in the context of adaptive measurement of change (AMC). Comparisons between methods were made based on observed Type I error rates and power. Monte Carlo simulation was conducted with the level of item discriminations, bank information shape, bank size, and test length varied. Overall, the Z statistic displayed a better balance of Type I error rates and power than the other four statistics under various conditions. The efficiency of variable-length AMC was evaluated compared to fixed-length AMC based on the number of items saved as well as the precision of decisions.

Table of Contents

List of Tables

List of Figures

Chapter 1:

INTRODUCTION

The measurement of individual change has been an important topic in psychology and related fields. For example, researchers are interested in the existence of treatment effects on attitude change (e.g., Thompson & Haddock, 2012; Walthera et al., 2009). Counseling or clinical psychologists are interested in the effectiveness of intervention (e.g., counseling) by measuring changes in mood, emotion, or level of psychiatric symptoms (e.g., depression) over time (e.g., Ali et al., 2003; Heppner & Heppner, 2003; Perdrixa et al., 2012). In educational research, researchers are interested in the developmental patterns of cognitive abilities (e.g., Cardon et al., 1992; Nation et al., 2010) or professional growth of teachers over time (e.g., Chapman, 2002; Kajander & Mason, 2007). The effect of school variables such as teacher quality on student achievement has also been an interest of educators and researchers (e.g., Brown, Smith & Stein, 1995; Greenwald, Hedges & Laine, 1996; Wright, Horn & Sanders, 1997): since student growth scores are often used for high-stakes decisions such as teacher or school accountability, precise measurement of student growth has been emphasized in the field of education.

Despite its long history and research in many areas in psychology and related fields, the measurement of change has been a methodologically controversial issue in the psychometric literature (e.g., Bereiter, 1963; Cronbach & Furby, 1970; Lord, 1963, Rogosa, Brand, & Zimowski, 1982; Williams & Zimmerman, 1996; Willet, 1997). A common approach to measuring change is to compute the difference between scores obtained from two occasions, either with repeated measurements or with parallel measurements. The use of difference scores is frequently found in “pre-post” designs in behavioral research (Willet, 1989; Burr & Nesselroade, 1990; Fischer, 2003). This simple approach, however, has been criticized in the psychometric literature for its methodological problems.

First, change scores tend to have lower reliability than the component variables (Bereiter, 1963; Cronbach & Furby, 1970; Hummel-Rossi & Weinberg, 1975; Lord, 1963; Willet, 1988). The reliability of the difference score, $D = Y - X$ is defined in classical test theory (CTT) as

$$\rho_{DD'} = \frac{\sigma_{DD'}}{\sigma_D^2} = \frac{\sigma_X^2 \rho_{XX'} + \sigma_Y^2 \rho_{YY'} - 2\sigma_X \sigma_Y \rho_{XY}}{\sigma_X^2 + \sigma_Y^2 - 2\sigma_X \sigma_Y \rho_{XY}}, \tag{1}$$

where *X* and *Y* are scores at Time 1 and Time 2, respectively. Assuming *X* and *Y* have equal variance and equal reliability, Equation 1 is simplified to

$$\rho_{DD'} = \frac{\rho_{XX'} - \rho_{XY}}{1 - \rho_{XY}} \tag{2}$$

If the correlation between X and Y is positive, which is often the case, the reliability of the change score is smaller than the reliability of *X* (or *Y*).

A negative correlation between the change score and the initial status has also been demonstrated as a problem of difference scores (Bereiter, 1963; Thorndike, 1963). The problem is also referred as the "ceiling effect" because individuals with high Time 1 scores tend to have small change at Time 2, while individuals with low Time 1 scores will likely have larger change at Time 2. This implies that any variable that is positively correlated with Time 1 scores will have a deceptive negative correlation with the difference score (Markus, 1980). Lord (1963) showed that the negative correlation of the difference score with the initial score came from the tendency of regression toward the mean from pretest to posttest measurement. These various methodological concerns have led to distrust in the measurement of change, and Cronbach and Furby (1970) recommended abandoning change measurement.

Nevertheless, a wide range of areas in behavioral science have needed adequate measures of change for their studies and several alternatives to the simple difference scores have been proposed to solve issues in the simple difference score (e.g., Rogosa & Willett, 1985; Willett, 1997; Zimmerman & Williams, 1982). The residual change score (RCS), for example, was designed to be uncorrelated with pretest score (DuBois, 1957). RCS is the residual from the regression of *Y* on *X*, which is defined for person *j* as:

$$R_j = Y_j - \bar{Y} - b_{Y.X}(X_j - \bar{X}), \tag{3}$$

where X_j and Y_j are the scores at Time 1 and Time 2 for person *j*, respectively, $\bar{X}$ and $\bar{Y}$ are the means of scores at Time 1 and Time 2, respectively, and $b_{Y.X}$ is the regression slope of the regression of *Y* on *X*. Manning and Dubois (1962) showed that the RCS had better reliability than the difference score. However, RCS is not a direct measure of true change for an individual between two occasions. Instead, it measures how much an individual changed more or less than the predicted value based on the previous observation, which is difficult to interpret. Also, RCS is group dependent, which makes it less desirable as a measure at the individual level.

Measuring Change With Item Response Theory

Some modern approaches conceptualized change measurement with the application of item response theory (IRT) to overcome problems that rose from classical test theory. Fischer (1976, 1983), for example, developed the linear logistic model with relaxed assumptions (LLRA) within the framework of the generalized Rasch model. For Time 1, the probability of person *j* responding to item *i* correctly is defined as

$$P(\theta_{ij}) = \frac{\exp(\theta_{ij} - b_i)}{1 + \exp(\theta_{ij} - b_i)}, \tag{4}$$

where θ_{ij} is ability for person *j* associated with item *i* and b_i is the item difficulty. At Time 2, *m* treatments that are applied to person *j* are taken into accounted in the model, so that the probability of correctly responding to item *i* is defined as

$$P(\theta_{ij}, \eta_m) = \frac{\exp(\theta_{ij} - b_i + \sum_{k=1}^{m} q_{jk}\eta_k + \tau)}{1 + \exp(\theta_{ij} - b_i + \sum_{k=1}^{m} q_{jk}\eta_k + \tau)}, \tag{5}$$

where $q_{j1}, \cdots q_{jm}$ are doses of *m* treatments, η_k is the effect for treatment *m*, and τ is the trend effect, which is independent of the treatments (e.g., natural maturation). In LLRA, the effect parameters η_k and τ can be estimated independent of θ_{ij} using conditional maximum likelihood estimation. Fischer's LLRA is appropriate for group level of change, but not for measuring individual change, because the treatment effects and trend effects are assumed to be constant for everyone across all occasions.

Embretson (1991) has proposed a multidimensional Rasch model for learning and change (MRMLC) for repeated measurements based on the Weiner simplex. In the MRMLC, performance on the *k*th occasion is assumed as a function of *k* abilities (i.e., ability at each of *k* occasions). Then the probability that person *j* responds to item *i* correctly under *k* occasions is defined as

$$P(\boldsymbol{\theta}_j) = \frac{\exp(\sum_{m=1}^{k} \theta_{jm} - b_i)}{1 + \exp(\sum_{m=1}^{k} \theta_{jm} - b_i)}, \tag{6}$$

where $\boldsymbol{\theta}_j$ is the vector of abilities in which θ_{jm} is the ability at $k = m$, that is the response for item *i* administered at occasion *k* is a function of all θs up to *k* occasions. The MRMLC measures individual change; however, it is restricted to the unrealistic assumption of equal item discriminations. Also the model complexity increases as more measurements are taken: when repeated measurement is used, more parameters are estimated at later time points with the same set of items. Thus the plausibility of the model decreases as more replications of measurement are made.

Mellenbergh and van den Brink (1998) proposed an item response change model also based on the Wiener simplex. In their item response change model, the probability of person *j* giving a correct answer to the *i*th item at the *k*th occasion is defined as

$$P_{ijk} = \frac{\exp(\gamma'_{ij1} + \gamma'_{j2} + \ldots + \gamma'_{jk})}{1 + \exp(\gamma'_{ij1} + \gamma'_{j2} + \ldots + \gamma'_{jk})}, \tag{7}$$

where γ'_{ij1} is the *j*th person's expected response to the *i*th item at the initial occasion, which is defined as the log odds of a correct and incorrect response of item *i* at the first occasion. $\gamma'_{j2}, \cdots, \gamma'_{jk}$ are item change parameters for occasions $2, \cdots, k$, respectively, and are constant for all items per occasion. The maximum likelihood estimator of the change parameter is

$$\hat{\gamma}'_{j2} = \ln(n_p / n_n), \tag{8}$$

and the estimated variance is

$$Var(\hat{\gamma}'_{j2}) = \frac{(n_p + n_n)}{n_p n_n}, \tag{9}$$

where n_p and n_n are frequencies of test items changed in the positive (i.e., incorrect in the pretest and correct in the posttest) and negative (i.e., correct in the pretest and incorrect in the posttest) directions. However, their item change model only has an item-specific parameter γ'_{ij1}.

Mellenbergh and van den Brink (1998) further restriced the item-specific parameter, γ'_{ij1}, in Equation 7 as a linear combination of q-dimensional latent traits for measuring individual change such as

$$\gamma'_{ij1} = b_i + a_{i1}\theta_{j1} + \ldots + a_{iq}\theta_{jq}, \tag{10}$$

where $a_{i1}, \cdots, a_{iq}$ are item discriminations and $\theta_{j1}, \cdots, \theta_{jq}$ are latent traits for q dimensions. Note that the change parameters of later occasions, $\gamma'_{j2}, \cdots, \gamma'_{jk}$ in Equation 7 depend on frequencies of test item change in positive and negative directions (Equation 8) rather than θs at those occasions. This model can be considered as a generalization of Embretson's (1991) MRMLC to two parameters, but the model is still limited to two parameters. Also, the model includes more parameters to estimate given the same items as more measurements are made, so that parameter estimation accuracy in later time points would decrease.

Other Recent Methods for Measuring Change

Other modern approaches to the measurement of change include a model-based estimation approach in analyzing longitudinal data with multiple observations rather than traditional pre-post testing data (e.g., Bryk & Raudenbush, 1987; Rogosa & Willet, 1985). The most commonly used approach to modeling change is latent growth curve models (Collins, 2006). Growth modeling approaches attempt to explain the structure of individual variability as well as change at the group level. In modeling latent growth trajectories, advanced statistical techniques such as hierarchical linear modeling, structural equation modeling, or time series have been applied (e.g., Bryk & Raudenbush, 1987; McArdle & Epstein, 1987; Von Eye, 1990). However, most of the approaches in the measurement of individual change have been limited to the use of measuring instruments that are developed based on CTT.

According to CTT, tests are usually constructed with high discriminating items with difficulties (i.e., proportion correct) around 0.5 to maximize the test reliability (Crocker & Algina, 2006). In addition, many applications of psychological testing either repeat the same measuring instrument or use parallel forms when measuring individual change over multiple occasions, which unavoidably leads to less precise measure of individual change over time (Von Minden, 2011; Weiss & Von Minden, 2011). Tests based on CTT have peaked test information functions with a narrow range of item difficulty. When viewed from an IRT perspective, under repeated measurement with parallel forms of the test (or the same test), the peak of the test is fixed, while examinees' latent traits become off-target as they change, and thus larger measurement error will be obtained at later time points. This issue of the quality of measurement has rarely been addressed in the psychometric literature (Embretson, 1996; Kang & Waller, 2005; Von Minden, 2011; Weiss & Von Minden, 2011).

Measuring Change With CAT

Recent studies have investigated advantages of computerized adaptive testing (CAT) in estimating individual change (Kim-Kang & Weiss, 2007, 2008; Finkelman et al., 2010; Weiss, 2011; Weiss & Von Minden, 2011). CAT is a type of computer-based assessment in which items are successively administered based on the performance of an examinee. Application of CAT has increased in various measurement domains (Fliege et al., 2005; Simms & Clark, 2005). CAT is composed of the following five components for a test administration (Weiss & Kingsbury, 1984; Thompson & Weiss, 2011):

Basic Elements of CAT

Item bank. CAT requires a pre-calibrated item bank, which is typically developed and calibrated based on IRT (Weiss & Kingbury, 1984). A main advantage of using IRT in CAT is to have items and examinees on the same scale (Birnbaum, 1968), so that it becomes possible to select optimal items given an examinee's ability/trait level. One important aspect of a CAT bank, in addition to content and the number of items, to be considered in the bank development is the test/bank information function (TIF). For a given IRT model, the item information function for item *i* can be defined as

$$I_i(\theta) = \frac{[P_i(\theta)']^2}{P_i(\theta)Q_i(\theta)}, \tag{11}$$

where $P_i(\theta)'$ is the first derivative of $P_i(\theta)$ with respect to θ. $P_i(\theta)$ is the probability of a keyed response for item *i* given θ and $Q_i(\theta) = 1 - P_i(\theta)$. $P_i(\theta)$ in IRT is called an item response function (IRF) and is defined for the three-parameter logistic (3PL) IRT model as

$$P_i(\theta) = c_i + (1 - c_i)\frac{1}{1 + \exp\left[-Da_i(\theta - b_i)\right]}, \tag{12}$$

where D = 1.7 and c_i is the guessing parameter. The TIF is the sum of item information available in a bank,

$$I(\theta) = \sum_{i=1}^{n} I_i(\theta), \tag{13}$$

where *n* is the number of items in an item bank. A CAT item bank is developed to obtain a TIF similar to a target TIF that is appropriate for the purpose of the test. For example, a flat TIF is targeted if the purpose of the test is to measure examinees at different ability level equally precisely. If a test classifies candidates based on a cut-score, then a higher TIF around the cut-score is preferred.

Starting point. In CAT, items are selected based on the previous responses to items administered. However, how to select the first item needs to be specified in advance since no response is observed at the initial point of item selection. Several options are available to assign the starting θ estimate. One option is to assign a fixed value, for example $\theta = 0$. This approach is straightforward to implement, however it could result in items at the fixed difficulty becoming overly exposed. When it is necessary to enhance test security, the first item can instead be randomly selected from a pre-specified θ range, for example between −0.5 to 0.5. If some prior information related to examinee ability is available, such as scores from a previous year or scores from a highly correlated test, then those previous scores can be used to assign an individualized starting point to each examinee.

Item selection. CAT requires sequential item selection to be made following each observed item response and interim θ estimate. A classical item selection method is to select the next item that provides the maximum Fisher information in Equation 11 at the current θ estimate among all items that have not yet been administered (Weiss, 1982). The maximum information criterion has been demonstrated to have two main disadvantages: inaccuracy in θ estimates early the test and overexposure for highly discriminating items. Errors in the first few θ estimates are generally large in the early stages of CAT. Since the maximum information criterion selects items based on the current θ estimate, the performance of the selected item may not be optimal at the true ability level. Also, selection based on maximum information often leads to items with high discrimination at a given θ level administered more often than the other items, which raise concerns of test security and/or low bank utilization. Several item selection criteria have been introduced to deal with problems in the maximum information item selection. See van der Linden & Pashley (2010), for example, for details of new approaches.

Scoring. CAT estimates θ each time before the next item is selected based on the administered items and observed responses at that point in the test. The maximum likelihood estimator (MLE) is a commonly used θ estimation method in IRT-based CAT. A likelihood function, $L(\theta | \mathbf{u})$, is obtained by multiplying IRFs of correct and incorrect responses for each item given θ,

$$L(\theta | \mathbf{u}) = \prod_{i=1}^{k} P_i(\theta)^{u_i} Q_i(\theta)^{1-u_i} , \tag{14}$$

where k is the number of items administered so far, $\mathbf{u}$ is a response vector and u_i is the ith element of $\mathbf{u}$. The MLE is defined as the θ value that maximizes the likelihood function in Equation 14. The MLE has been reported to have a smaller bias function but larger standard error than those of Bayesian estimation methods—maximum a posteriori (MAP) or expected a posteriori (EAP) (e.g., Wang, Hanson & Lau, 1999; Wang & Vispoel, 1998; Weiss, 1982). One disadvantage of MLE in relation to CAT is that no finite MLE estimate exists when responses are either all correct or all incorrect, in which case an arbitrary value, e.g., −4 or 4 is usually assigned. This is particularly an issue in CAT since it results in inaccurate estimates for the first item, and possibly for the first few items until responses are mixed. Therefore, modifications are implemented when MLE is used. For example, a stepwise increase (e.g., +3) or a stepwise decrease (e.g., −3) can be used for the current θ estimates until responses are mixed. Alternative θ estimation methods — MAP, EAP, or weighted likelihood estimate (WLE) — can be temporarily implemented until non-mixed responses are observed. See van der Linden & Pashley (2010), for example, for detailed description and comparisons of alternative θ estimation methods.

Termination. CATs are terminated based on a pre-specified termination criterion. One popular criterion terminates CAT after a pre-determined number of items are taken by an examinee (i.e., fixed-length termination), which is similar to conventional paper-and-pencil tests. CAT can also administer different numbers of items adaptively to each examinee. The most common variable-length CAT administers items until a pre-specified level of precision (i.e., standard error) is obtained (Weiss & Kingsbury, 1984). Alternative termination methods terminate a test when, for example, a minimal level of change in the θ estimate is observed (Gialluca & Weiss, 1979; Maurelli & Weiss,1981); or when no items remaining in a bank have more than a minimal level of information at the current θ estimate (Weiss & Kingsbury, 1984; Hart et al. 2006). Variable-length CATs often impose constraints on a minimum and/or maximum test length to have practical justification for complaints from low performers with a test that is too short or to prevent all items in a bank to be administered (Thompson & Weiss, 2011).

Adaptive Measurement of Change

CAT has potential advantages to better capture the actual amount of change than conventional tests. Conventional tests based on CTT typically have a peaked test information function. If repeated measures (or parallel forms) are used to measure change, test information functions at later time points are fixed at Time 1 and the examinee's ability will become off-target as it changes. As a result, measurement precision for conventional tests will be decreased as an examinee's latent trait changes. However, this is not the case for CAT. Since CAT administers items that adapt to each examinee's θ level, CAT can measure change equally precisely across θ levels at each time point.

May and Nicewander (1998) evaluated the scale distortion problem (i.e., ceiling effect) in the difference score with conventional tests and adaptive testing. Difference scores measured in three metrics were compared: (1) difference in number-correct score from conventional tests, (2) difference in θ scored using IRT with conventional tests, and (3) difference in θ measured with adaptive testing. The results showed that the θ metric had smaller scale distortion than number-correct score, and θ measured with adaptive testing had smaller scale distortion than θ measured with conventional tests. However, this study had limited implications for measuring individual change with adaptive testing. First, average difference scores in the three metrics were compared at the group level — high, medium and low performance groups, instead of comparisons being made at the individual θ level. Second, the three metrics were not directly comparable: the difference in θ was multiplied by 10 to be similar to the difference in number-correct scores. Therefore, this scale adjustment only revealed the pattern of change, whether scales in each metric were parallel to the true change or not; the recovery of true change was not quantified. Third, the sample size of each exam was only 50, and the difference scores were averaged only based on the one-third of the sample size, which made it difficult to generalize the findings.

Kim-Kang and Weiss (2007, 2008) referred to CAT applied in measuring change as the adaptive measurement of change (AMC). In the AMC procedure, measurements are taken at each time point using CAT. Similar to regular CAT, CAT banks are constructed at each time point according to the purpose of the tests, or the same bank with a high and wide range of item information is used. The θ estimates from the previous time point are used, if available, as initial θ to begin the CAT at the later time point. θs are estimated after each item is administered, and items are often selected to maximize the Fisher information at the current θ estimate. The CATs are based on pre-specified termination criteria. For fixed-length termination, AMC is terminated after a pre-determined number of items are administered. AMC used in later time points can be terminated when significant change is observed (i.e., variable-length termination). The difference in CAT θ estimates between the two occasions is defined as a measure of change.

Previous studies compared the precision of conventional tests and AMC in measuring individual change. In Kim-Kang and Weiss (2008), individuals were measured at two time points with conventional tests and AMCs. Three change estimates were used for conventional tests: number-correct score converted to the θ metric using the test information function, a residual change score, and IRT θ estimates. The average discrimination conditions were varied as low ($\bar{a} = 0.5$), medium ($\bar{a} = 1.0$) and high ($\bar{a} = 1.5$). Weiss and Von Minden (2011) further extended the work of Kim-Kang and Weiss (2008) to five time points with five growth curves including linear and curvilinear growth patterns. The recovery of individual true θ was evaluated based on bias and root mean square error of θ estimates at each time point for each growth curve. The results of both studies showed that AMC consistently better estimated true growth with small error, while errors in conventional tests increased when tests were off-target of the examinee ability range. As was expected, the precision of measuring individual change decreased in conventional tests as examinees changed at later time points.

Testing Significance of Individual Change

In addition to estimating the amount of change, researchers and practitioners are also interested in whether the observed change is significantly different from the previous status. Clinical psychologists are often interested in whether a patient's psychiatric symptom has been significantly reduced or increased; counseling psychologists are interested in whether depression has been significantly reduced after intervention (e.g., Falloon et al., 1985; Gagne & Toye, 1994; Smits et al., 2008). Since the effect of instruction on a student's achievement or knowledge level is one of the primary concerns in education, teachers are interested in whether students have significantly improved (i.e., learned) over a semester (e.g., Fennema et al., 1996; Lei & Zhao, 2007). Thus, determining the significance of individual change has a wide range of practical implications.

Using IRT, Fischer (2001, 2003) has applied Clopper-Pearson confidence intervals in testing the hypothesis of no change $H_0 : \delta = 1$ against $H_a : \delta \neq 1$. When change in Time 2 is expressed as $\theta_2 = \theta_1 + \eta$, where θ_1 and θ_2 indicate θ at Time 1 and Time 2, respectively, the parameter δ is defined as $\delta = \exp(\eta)$. Thus, it becomes equivalent to testing $H_0 : \eta = 0$ against $H_a : \eta \neq 0$. In Fischer's method, $\hat{\delta}$ is a conditional maximum likelihood estimator, and confidence intervals are constructed using the conditional probability of posttest raw score given the sum of pretest and posttest total raw scores. This approach has been applied to conventional tests and is restricted to a family of Rasch models (i.e, dichotomous Rasch, Rasch rating scale, and Rasch partial credit models). Also, conditional maximum likelihood estimation and conditional probabilities in Fischer's application require item responses and item parameters from both pretest and posttest, which makes it difficult to apply in live adaptive testing.

The first attempt at hypothesis testing in AMC used confidence intervals (Kingsbury & Weiss, 1983) for the hypotheses of no change; $H_0 : \theta_2 = \theta_1$ and $H_a : \theta_2 \neq \theta_1$, indicating that θ at Time 2 is no different from θ at Time 1. Confidence intervals around two estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ are constructed, and the significance of change is determined when the two confidence intervals do not overlap. The confidence intervals are estimated as

$$CI = \hat{\theta}_t \pm 2 \cdot SE(\hat{\theta}_t), \tag{15}$$

where $\hat{\theta}_t$ indicates θ estimate at Time *t*, and $SE(\hat{\theta}_t)$ indicates the standard error of $\hat{\theta}_t$ which is calculated as the inverse of the square root of the second derivative of the negative log likelihood function with respect to θ evaluated at $\hat{\theta}_t$ (i.e., observed test information at $\hat{\theta}_t$):

$$SE(\hat{\theta}_t) = \left. \frac{1}{\sqrt{-\partial^2 \ln L(\theta) / \partial \theta^2}} \right|_{\theta = \hat{\theta}_t} . \tag{16}$$

However, this approach has been demonstrated to be too conservative: although the desired Type I error rate was 0.05, the observed Type I error rate based on the confidence interval method never reached 0.01 (Kim-Kang & Weiss, 2008; Finkelman et al., 2010).

Guo and Drasgow (2010) proposed two hypothesis testing methods, a *Z*-test and a likelihood ratio test (LR) identifying cheating on unproctored internet tests (UIT). The change in ability estimates between the UIT and the proctored verification test, both of which were adaptive tests, was tested to detect suspicious cheating with a hypotheses of $H_0 : \theta_u = \theta_v$ and $H_a : \theta_u > \theta_v$.

The *Z* test in Guo and Drasgow (2010) was defined as

$$Z = \frac{\hat{\theta}_u - \hat{\theta}_v}{\sqrt{I_u(\hat{\theta}_u)^{-1} + I_v(\hat{\theta}_v)^{-1}}}, \tag{17}$$

where $\hat{\theta}_u$ and $\hat{\theta}_v$ are θ estimates from the UIT and the proctored verification test, respectively. $I_u(\hat{\theta}_u)$ and $I_v(\hat{\theta}_v)$ are test information of the UIT and the proctored verification tests, respectively, each evaluated at its MLE. The test information can be defined in general as the expected curvature of the likelihood function,

$$I(\theta) = -E\left[\frac{d^2 \ln L(\theta \mid \mathbf{u})}{d\theta^2}\right]. \tag{18}$$

The Z statistic was compared to a standard normal distribution to determine cheating.

The LR statistic in Guo and Drasgow (2010) was based on the likelihoods of the response vectors from the UIT and the proctored verification test,

$$\Lambda = \frac{L(\mathbf{u})L(\mathbf{v})}{L(\mathbf{u},\mathbf{v})}, \tag{19}$$

where $L(\mathbf{u})$, $L(\mathbf{v})$, $L(\mathbf{u},\mathbf{v})$ are the likelihood of observing the responses in the UIT, the proctored verification test, and the two response vectors together, each of which was defined as

$$L(\mathbf{u}) = \int \prod_{i=1}^{n_u} P_i(\theta)^{u_i} \left[1 - P_i(\theta)\right]^{1-u_i} f(\theta) d\theta, \tag{20}$$

$$L(\mathbf{v}) = \int \prod_{j=1}^{n_v} P_j(\theta)^{v_j} \left[1 - P_j(\theta)\right]^{1-v_j} f(\theta) d\theta, \tag{21}$$

$$L(\mathbf{u},\mathbf{v}) = \int \left\{ \prod_{i=1}^{n_u} P_i(\theta)^{u_i} \left[1 - P_i(\theta)\right]^{1-u_i} \right\} \left\{ \prod_{j=1}^{n_v} P_j(\theta)^{v_j} \left[1 - P_j(\theta)\right]^{1-v_j} \right\} f(\theta) d\theta. \tag{22}$$

The integrations were numerically approximated using θs from −4 to 4 with a 0.1 increment. The critical value was empirically determined using a simulation to have Type I error approximately

0.01 since the distribution of the likelihood ratio in Equation 19 is unknown. The Z and LR statistics were evaluated and compared based on observed Type I error and power in detecting suspicious cheating (i.e., decrease in θ in the verification test from UIT).

The methods in Guo and Drasgow (2010) were limited by the fact that the two statistics were not parameterized based on the hypotheses being tested. A *Z* statistic is typically defined under the null hypothesis to have an approximately standard normal distribution. The *Z* statistic in Guo and Drasgow (2010), however, was defined using MLE in each test instead of under the null hypothesis, and thus it is questionable whether the *Z* statistic approximates a standard normal distribution as was used in their study. Similarly, a LR statistic is typically defined based on the null and the alternative hypothesis. However, the LR statistic in Guo and Drasgow (2010) is the ratio between the likelihoods of the observed responses, without taking into account the hypotheses. Although an empirically defined critical value was used, the performance of the LR statistic is not mathematically sound either. The results in the study support the skepticism on the functionality of the two statistics: only less than half of θs examined in the study reached 0.01, while more than a third of θs were close to 0 for both of the statistics.

Finkelman et al. (2010) introduced two new hypothesis testing methods that are similar in name to those of Guo and Drasgow (2010) but are more firmly rooted in appropriate statistical theory. The *Z*-test used the standardized difference in MLE θ estimates and the significance of the test statistic is determined from the standardized normal distribution, and their LR statistic evaluates two likelihood functions under the null and alternative hypotheses of change.

The Z test. In testing the null hypothesis of no change $H_0 : \theta_2 = \theta_1$, Finkelman et al. (2010) defined a Z statistic as a standardized difference between θ estimates from two time points:

$$Z = \frac{\hat{\theta}_2 - \hat{\theta}_1}{\sqrt{I_1(\hat{\theta}_p)^{-1} + I_2(\hat{\theta}_p)^{-1}}} \tag{23}$$

where $\hat{\theta}_1$ and $\hat{\theta}_2$ are MLEs of θ at Time 1 and Time 2, respectively, and $I_t(\theta)$ denotes observed test information at Time *t*. The variance of $\hat{\theta}_2 - \hat{\theta}_1$ was estimated using test information from both time points evaluated at $\hat{\theta}_p$, where $\hat{\theta}_p$ is the MLE of θ using the combined responses and item parameters from Time 1 and Time 2. The Z statistic is compared to a standard normal distribution to make a decision of significance.

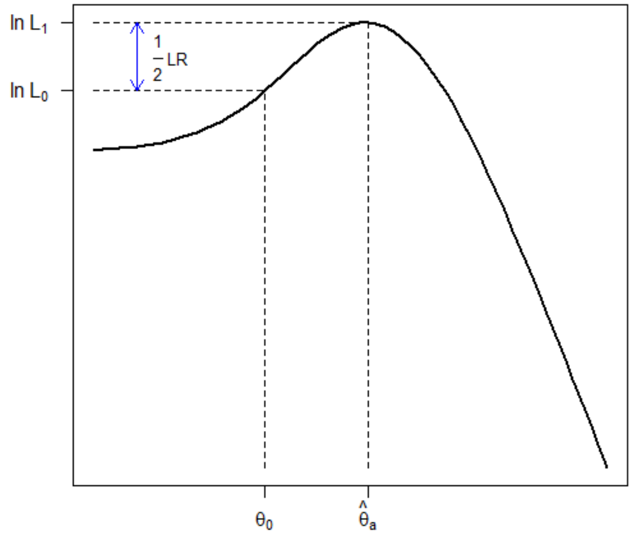
The LR test. A likelihood ratio test is a statistical test based on the ratio of two likelihoods: the maximum of a likelihood function over the parameters with restrictions of the null hypothesis and maximum over the larger set of parameters without the restrictions. The LR statistic is defined as (Neyman & Pearson, 1928):

$$\mathrm{LR} = -2\ln\left[\frac{L(\hat{\theta}_0 \mid \mathbf{u})}{L(\hat{\theta}_a \mid \mathbf{u})}\right] = -2\left[l(\hat{\theta}_0) - l(\hat{\theta}_a)\right], \tag{24}$$

where $\hat{\theta}_0$ is the restricted maximum likelihood (ML) estimate under the null hypothesis, $\hat{\theta}_a$ is the unrestricted ML estimate under the alternative hypothesis, and $l(\cdot)$ is the logarithm of the likelihood function. It has been proven that the LR statistic approximates a chi-square distribution with 1 degree of freedom under the null hypothesis (Wilks, 1938).

Figure 1 shows a plot of a log-likelihood function. The LR statistic compares the log-likelihood value $\ln L_1$ evaluated at $\hat{\theta}_a$ and $\ln L_0$ evaluated at θ_0. The LR statistic is twice the vertical distance between $\ln L_1$ and $\ln L_0$. Assuming the null hypothesis is true, $\hat{\theta}_a$ will be close to θ_0, and the distance between two log-likelihood values will be close to zero. The larger $\hat{\theta}_a - \hat{\theta}_0$ is, the larger a LR statistic will be observed.

Figure 1. The Likelihood Ratio Test



Finkelman et al. (2010) applied the LR statistic in testing the significance of individual change measured with AMC. In testing the null hypothesis of change $H_0 : \theta_2 = \theta_1$, the LR statistic was defined as

$$\mathrm{LR} = -2\ln\left[\frac{L(\hat{\theta}_p \mid \mathbf{u}_{1+2})}{L(\hat{\theta}_1 \mid \mathbf{u}_1) \times L(\hat{\theta}_2 \mid \mathbf{u}_2)}\right], \quad (25)$$

where $\mathbf{u}_1$ and $\mathbf{u}_2$ denote response vectors from Time 1 and Time 2, respectively, and $\mathbf{u}_{1+2}$ is a combined response vector from the two time points. Under the null hypothesis, $\hat{\theta}_p$ is the value that maximizes the likelihood, whereas the likelihood function is maximized when estimating the MLEs separately at each time point, which are denoted as $\hat{\theta}_1$ and $\hat{\theta}_2$. The statistic is compared to a chi-square distribution with 1 degree of freedom to determine the statistical significance of change. Finkelman's two approaches resulted in desirable Type I error rates and power compared to the confidence interval approach.

New statistic 1: The score test. The current study proposes new hypothesis testing methods for AMC as an extension of Finkelman's (2010) approaches. The Score Test (ST) was introduced by Rao (1948) as an alternative way to use the likelihood function to perform large-sample inference. The ST statistic involves the slope and expected curvature of the log-likelihood function instead of considering differences in log-likelihoods. The general form of the ST statistic is defined as

$$\mathrm{ST} = \frac{s(\theta_0 \mid \mathbf{u})^2}{I(\theta_0)}, \tag{26}$$

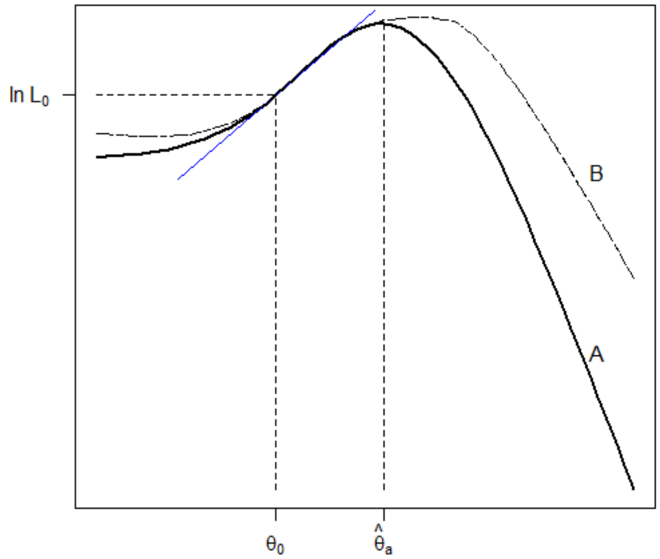
where $s(\theta \mid \mathbf{u}) = \frac{d \ln L(\theta \mid \mathbf{u})}{d\theta}$ is the first derivative of the log-likelihood, which is called a score function, and $I(\theta)$ is test information defined in Equation 13, both evaluated at θ_0. Since $E\left[s(\theta \mid \mathbf{u})\right] = 0$ and $\mathrm{var}\left[s(\theta \mid \mathbf{u})\right] = I(\theta)$, the ratio of the score function to its null standard error has an approximate standard normal distribution, that is,

$$\frac{s(\theta_0 \mid \mathbf{u})}{\sqrt{I(\theta_0)}} \xrightarrow{D} \mathrm{N}(0,1) \tag{27}$$

by the central limit theorem. The ST statistic is the squared value of Equation 27 and follows asymptotically a chi-square distribution with 1 degree of freedom.

Figure 2 illustrates the basic idea behind the ST. The unrestricted ML estimate $\hat{\theta}_a$ maximizes the log-likelihood and thus the slope of the log-likelihood evaluated at $\hat{\theta}_a$ is zero. If the null hypothesis is true, then the restricted ML estimate θ_0 will be close to $\hat{\theta}_a$. The closer the distance between θ_0 and $\hat{\theta}_a$ is, the closer the slope of the log-likelihood at θ_0 is to zero; the slope tends to be larger in absolute value when $\hat{\theta}_a$ is farther from θ_0. In Figure 2, two different log-likelihood functions A and B have the same slope at θ_0 but θ_0 is closer to the maximum of the log-likelihood in function A (i.e., greater curvature) than in function B. The ST statistic weights the slope by the inverse of the curvature of the log-likelihood, so that a greater curvature leads to a smaller test statistic.

Figure 2. The Score Test



In general, the ST statistic is simple to compute as it depends only on estimation of parameters under the null hypothesis, whereas the LR statistic requires estimates both under the null and alternative hypotheses. As the sample size increases to infinity, the ST is asymptotically equivalent to the LR test in the first-order approximation (Rao, 1965; Chandra & Joshi, 1983; Cox & Hinkley, 1974). In finite samples, the two tests tend to generate somewhat different test statistics. For example, when the model is linear, the LR statistic tends to be greater than or equal to the ST statistic (Johnston & DiNardo, 1997, p.150). The second-order powers of the two tests are also different but neither dominates the other (Taniguchi, 1988, 1991).

In testing the no-change hypothesis of $H_0 : \theta_2 = \theta_1$, the ST statistic can be defined as

$$\mathrm{ST} = \frac{s(\hat{\theta}_p \mid \mathbf{u}_1)^2}{I^1(\hat{\theta}_p)} + \frac{s(\hat{\theta}_p \mid \mathbf{u}_2)^2}{I^2(\hat{\theta}_p)} \tag{28}$$

The statistic evaluates the slope and curvature of the log-likelihood functions from each time point at restricted ML estimates and sums over two time points. This is similar to the LR statistic in Finkelman et al. (2010) in that $\ln L(\hat{\theta}_p \mid \mathbf{u}_{1+2}) = \ln L(\hat{\theta}_p \mid \mathbf{u}_1) + \ln L(\hat{\theta}_p \mid \mathbf{u}_2)$ is used under the null hypothesis (i.e., the numerator in Equation 25). If the null hypothesis is true, the MLE at each time point $\hat{\theta}_1$ and $\hat{\theta}_2$ will be close to $\hat{\theta}_p$ and the slope of the log-likelihood at $\hat{\theta}_p$ for each time point will also be close to zero. Hence a smaller ST statistic will be obtained under the null hypothesis. The statistic is then compared to a chi-square distribution with 1 degree of freedom to determine the statistical significance of change.

New statistic 2: The Kullback-Leibler divergence test. The Kullback-Leibler divergence (KLD) is a measure of distance between two distributions. The use of KLD to detect significance of change was first proposed by Wang (2014) for a conventional test using a multidimensional IRT model with a multivariate normal prior. In the current study, KLD was applied for AMC with a unidimensional IRT model using both normal (KLD-N) and uniform priors (KLD-U). Let $\pi_1(\theta \mid \mathbf{u}_1)$ be a posterior distribution at Time 1, and $\pi_2(\theta \mid \mathbf{u}_2)$ be a posterior distribution at Time 2. For convenience, π_1 and π_2 are used interchangeably with $\pi_1(\theta \mid \mathbf{u}_1)$ and $\pi_2(\theta \mid \mathbf{u}_2)$. The KLD (Kullback & Leibler, 1951) between π_1 and π_2 is defined as

$$\begin{aligned} \mathrm{KLD}(\pi_1 \,\|\, \pi_2) &= E_{\pi_1}\left[\ln \frac{\pi_1(\theta \mid \mathbf{u}_1)}{\pi_2(\theta \mid \mathbf{u}_2)}\right] \\ &= \int_{-\infty}^{\infty} \pi_1(\theta \mid \mathbf{u}_1) \ln \frac{\pi_1(\theta \mid \mathbf{u}_1)}{\pi_2(\theta \mid \mathbf{u}_2)} d\theta. \end{aligned} \tag{29}$$

The value of KLD is always nonnegative and is zero if and only if two distributions are identical. Larger KLD indicates that the two distributions differ and, in the present application, that change has occurred between the two time points: if there is no change, KLD will be close to zero.

A numerical approximation to KL divergence takes a finite set of $\{\theta_1, \theta_2, ..., \theta_h\}$ that are evenly spaced between −3 and 3 with 0.1 increments, where h = 61. The posterior distribution of Time 1 is then approximated as follows:

$$\pi_1(\theta_l \mid \mathbf{u}_1) = \frac{L(\theta_l \mid \mathbf{u}_1) p(\theta_l)}{\int_{-\infty}^{\infty} L(\theta \mid \mathbf{u}_1) p(\theta) d\theta} \approx \frac{L(\theta_l \mid \mathbf{u}_1) p(\theta_l)}{\sum_{k=1}^{h} L(\theta_k \mid \mathbf{u}_1) p(\theta_k) \Delta\theta}, \tag{30}$$

where θ_l is one of the 61 θ points used for the approximation, $p(\theta_l)$ is the prior density, and $\Delta\theta = 0.1$. When a uniform prior is used, the prior density is set to 1. The posterior distribution of Time 2 is computed in a similar way based on Time 2 item parameters and responses. Then KL divergence is computed as

$$\mathrm{KLD}(\pi_1 \parallel \pi_2) \approx \sum_{l=1}^{h} \pi_1(\theta_l \mid \mathbf{u}_1) \ln \frac{\pi_1(\theta_l \mid \mathbf{u}_1)}{\pi_2(\theta_l \mid \mathbf{u}_2)} \cdot \Delta\theta, \tag{31}$$

The distribution of KLD is generally unknown. Belov and Armstrong (2011) proved that the distribution of KLD is approximately distributed as a chi-square with 1 degree of freedom under certain assumptions that are common in psychometrics. Assuming that test length is sufficiently long so that the posterior distributions have normal probability density functions:

$$\pi_1(\theta) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left[-\frac{(\theta - \mu_1)^2}{2\sigma_1^2} \right] \approx \mathrm{N}(\mu_1, \sigma_1^2), \tag{32}$$

$$\pi_2(\theta) = \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left[-\frac{(\theta - \mu_2)^2}{2\sigma_2^2} \right] \approx \mathrm{N}(\mu_2, \sigma_2^2). \tag{33}$$

Then KLD in Equation 31 can be simplified to

$$\mathrm{KLD}(\pi_1 \parallel \pi_2) = \frac{1}{2}\left[\ln \frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} + \frac{(\mu_1 - \mu_2)^2}{\sigma_2^2} - 1 \right]. \tag{34}$$

when $\sigma_1^2 = \sigma_2^2$, then Equation 34 transforms to

$$\mathrm{KLD}(\pi_1 \| \pi_2) = \frac{(\mu_1 - \mu_2)^2}{2\sigma_2^2}. \tag{35}$$

Let $Y = \frac{(\mu_1 - \mu_2)}{\sqrt{2}\sigma_2} \sim \mathrm{N}(0,1)$. Then KLD is distributed as a chi-square with one degree of freedom. For each examinee, the KLD statistic is calculated and compared to the chi-square distribution with 1 degree of freedom to determine the significance of change.

Item Selection in AMC

Fisher information. Previous studies in AMC used the maximum information criterion (Kim-Kang & Weiss, 2008; Weiss & Von Minden, 2011; Finkelman at al. 2010). Items that provided the maximum information at the current θ estimate were administered as the next item in the CAT, both at Time 1 and Time 2.

***Kullback-Leibler information*.** The application of Kullback-Leibler information in CAT item selection was first introduced by Chang and Ying (1996). As described in Equation 29, KL information, or KL divergence, is a general measure for the distance between two distributions. When applied to CAT item selection, the larger value of KL information indicates that the item better discriminates between two distributions, or equivalently, between the values of the parameters that index them (Lehmann & Casella, 1998).

Finkelman et al. (2010) applied KL information in the context of detecting individual change by selecting Time 2 items that best differentiated θ_2 from θ_1. Finkelman et al. (2010) used $\hat{\theta}_2$ as the best estimate of θ_2 under the alternative hypothesis, and $\hat{\theta}_p$ as the best estimate of θ_1 under the null hypothesis in computing KL information. The KL measure for the response distributions of item i associated with $\hat{\theta}_2$ and $\hat{\theta}_p$ was then defined as

$$KL_F(\hat{\theta}_2 \,||\, \hat{\theta}_p) = E\left[\ln \frac{L(\hat{\theta}_2 \mid u_i)}{L(\hat{\theta}_p \mid u_i)}\right], \tag{36}$$

where the expectation is taken over response variable u_i. The measure can, therefore, be calculated as

$$KL_F(\hat{\theta}_2 \,||\, \hat{\theta}_p) = P_i(\hat{\theta}_p)\ln\frac{P_i(\hat{\theta}_p)}{P_i(\hat{\theta}_2)} + \left[1 - P_i(\hat{\theta}_p)\right]\ln\frac{1 - P_i(\hat{\theta}_p)}{1 - P_i(\hat{\theta}_2)}. \tag{37}$$

KL information in Finkelman et al. (2010) tended to show lower Type I error rate and more power than Fisher information. However, the use of $\hat{\theta}_p$ can be computationally burdensome in AMC because it requires $\hat{\theta}_p$ estimation based on Time 1 and Time 2 responses given so far for each item selection in a CAT in addition to $\hat{\theta}_2$ estimation.

New item selection method: Modified Kullback-Leibler information. A modified KL information is proposed in the current study to deal with this practical limitation, specifically, $\hat{\theta}_p$ is replaced with the Time 1 final MLE $\hat{\theta}_1$, and items are selected based on a measure of how well an item differentiates between $\hat{\theta}_2$ and $\hat{\theta}_1$. The logic behind this modification is that if the null hypothesis is true, $\hat{\theta}_1$ will be close to $\hat{\theta}_p$. Therefore KL information between $\hat{\theta}_2$ and $\hat{\theta}_p$ will be similar to KL information between $\hat{\theta}_2$ and $\hat{\theta}_1$. Although $\hat{\theta}_1$ is not a better estimate of θ_1 than $\hat{\theta}_p$, because $\hat{\theta}_1$ is based on a smaller number of items than $\hat{\theta}_p$, and $\hat{\theta}_1$ is not affected by examinee status at Time 2, $\hat{\theta}_1$ can still be a good estimate of θ_1, at least the best estimate at Time 1. The modified KL information was defined as

$$KL_M(\hat{\theta}_2 \,||\, \hat{\theta}_1) = E\left[\ln \frac{L(\hat{\theta}_2 \mid u_i)}{L(\hat{\theta}_1 \mid u_i)}\right], \tag{38}$$

Variable-Length Tests

In variable-length AMC, the hypothesis testing method is used as a termination criterion at Time 2, that is, the AMC procedure is terminated before the maximum test length is reached when significant change is observed. Variable-length testing can make AMC more efficient by saving a number of items for examinees with substantial improvement (i.e., change). Variable-length termination was applied only to determine the significance of change at Time 2. The Time 1 assessment was obtained with a fixed-length CAT.

The multiple hypothesis testing problem should be taken into account to terminate AMC based on the significance of change. In a variable-length AMC, a test statistic is computed after each item is administered, including the maximum length. If the statistic exceeds the critical value, the null hypothesis of change is rejected, and test administration is ceased; otherwise, the next item is administered and the statistic is calculated again until the maximum test length is reached. Errors in statistical inference are likely to increase when a set of statistical decisions are made simultaneously based on the observed data (O'Neil & Wetherill, 1971; Ryan, 1959; Tukey, 1953). Therefore, a threshold (i.e., critical value) for individual comparison should not be based on the desired α level. Instead, a higher threshold is generally used to compensate for the multiple comparisons being made to maintain the overall α level, following procedures designed to control “familywise error rates” (Dunn, 1961; Shaffer, 1995; Tukey, 1953).

Bartroff, Finkelman, and Lai (2008) introduced a sequential decision rule to maintain a desirable Type I error rate. Two decision values C_1 and C_2 are used, C_1 for the interim stage, and C_2 for the final stage. For each interim stage (i.e., after each item is administered in the Time 2 CAT), a test statistic, T, is calculated based on administered items and observed responses so far and the null hypothesis is rejected if $T \geq C_1$, where T is a test statistic after k

items are administered, and k is smaller than the maximum test length. The final stage is where the maximum test length is reached without having significant T at any interim stage and the test is forced to terminate. At the final stage, the null hypothesis is rejected if $T \geq C_2$. Lai and Shih (2004) recommended that a proportion of a desired Type I error rate to be observed during the interim stage, and the remaining Type I error rate to be observed at the final stage. Specifically, C_1 is selected such that $\varepsilon\alpha$ is observed during the interim stage, where $1/3 \leq \varepsilon \leq 1/2$, and C_2 is chosen such that the observed Type I error rate at the final stage is $(1-\varepsilon)\alpha$. The decision values C_1 and C_2 are determined through simulations.

Purpose

This study extended the work of Finkelman et al. (2010) in several directions. First, Finkelman's test statistics were more thoroughly examined under extended item bank conditions: the Monte Carlo simulation investigated the effect of item discriminations (i.e., low, medium and high) and bank size in addition to the effect of bank information shape on the performance of the hypothesis tests. The length of AMC was also varied to n = 15, 30, and 50 at both Time 1 and Time 2. Second, two new tests, the LM and KLD tests, were compared to Finkelman's Z and LR tests based on observed Type I error rate and power. Third, the performance of a modified KL CAT item selection method was compared to item selection by Fisher information and Finkelman's KL information. Fourth, variable-length AMC was examined for all combinations of item selection methods and hypothesis testing statistics. The efficiency of variable-length AMC was evaluated compared to fixed-length AMC based on the number of items saved as well as the accuracy of decisions. The results for variable-length AMC can give practical guidelines of threshold values for each bank condition.

Chapter 2:

METHODOLOGY

θ Values

To be comparable to the previous research, the same nine initial θ values, θ_1 = −2.0, −1.5, −1.0, −0.5, 0.0, 0.5, 1.0, 1.5, and 2.0, in Finkelman et al. (2010) were generated. Five levels of positive change were used: true change of $\Delta\theta$ = 0.0, 0.25, 0.5, 1.0, and 1.5, reflecting no change (N), very small change, (VS), small change (S), medium change (M), and large change (L), respectively. Time 2 θ s were generated by adding the specified amount of change to the initial θ values, $\theta_2 = \theta_1 + \Delta\theta$. Table 1 shows how the 45 combinations of θ_1 and θ_2 were generated. 1,000 examinees were generated for each of the 45 combinations of θ_1 and θ_2. The three change levels $\Delta\theta$ = 0.5, 1.0 and 1.5 were also used in Finkelman et al. (2010) and were based on mean change scores in the classical percent-correct metric from alternate forms retest data for students in grades 3 – 5 from the United States, including exams on math, reading, science, and social studies. An additional change level of 0.25 was also examined in the current study.

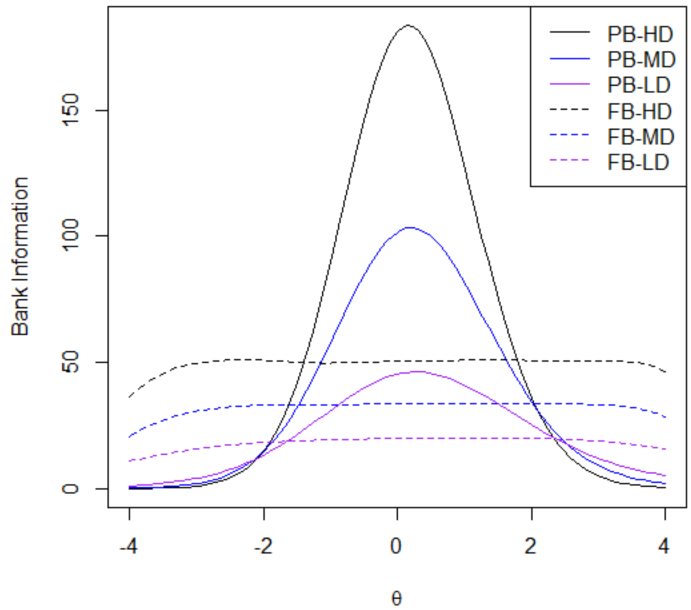
Table 1. Generation of Time 2 θ in Each Simulation Condition

	Time 1 (θ_1)								
Condition	−2.00	−1.50	−1.00	−0.50	0.00	0.50	1.00	1.50	2.00
No Change	−2.00	−1.50	−1.00	−0.50	0.00	0.50	1.00	1.50	2.00
Very Small Change	−1.75	−1.25	−0.75	−0.25	0.25	0.75	1.25	1.75	2.25
Small Change	−1.50	−1.00	−0.50	0.00	0.50	1.00	1.50	2.00	2.50
Medium Change	−1.00	−0.50	0.00	0.50	1.00	1.50	2.00	2.50	3.00
Large Change	−0.50	0.00	0.50	1.00	1.50	2.00	2.50	3.00	3.50

Item Banks

The item banks were generated for the 3PL IRT model in Equation 12. The discrimination parameters (a_i) were simulated from a normal distribution with a mean of 0.6, 1.0, and 1.5 (in the $D = 1.7$ metric), representing low (LD), medium (MD), and high (HD) discrimination conditions, respectively, and a standard deviation of 0.15. Two sets of item difficulty parameters (b_i) were used: a uniform distribution between −4.5 and 4.5 was used as the relatively ideal flat CAT item bank (FB), and a normal distribution with a mean of 0.0 and a standard deviation of 0.8 was used as the relatively realistic peaked item bank (PB). The range of the ideal bank and the mean and the standard deviation of the realistic bank were similar to those in Finkelman at al. (2010). The guessing parameter (c_i) was set to 0.2 for all items. Two bank sizes were used: *n* = 300 items or *n* = 500 items were independently generated (i.e., no overlap) for each time point. The six combinations of bank characteristic conditions, LD-FB, LD-PB, MD-FB, MD-PB, HD-FB, HD-PB, were fully crossed within the two bank sizes. Figure 3 depicts bank information functions for both flat and peaked information conditions at the three discrimination levels for a bank size 300. Average bank information for each item bank over 100 replications is also summarized in Appendix A1. Item parameters were generated using a program written in R (R Development Core Team, 2012).

Figure 3. Test Information Functions of CAT Banks in Three Discrimination Conditions (n = 300)



Data Generation

Item responses were generated using the 3PL model for each condition. The probability of correct response was calculated using Equation 12 and corresponding true θ and item parameters. The probability of correct response for each item was then compared to a randomly generated number from a uniform distribution [0, 1]. If the probability was greater than or equal to the random number, the response to that item was scored as correct (i.e., 1), otherwise scored as incorrect (i.e., 0). 12 item bank conditions, 3 (discrimination) × 2 (bank shape) × 2 (bank size) were fully crossed for each of the 45 combinations of θ_1 and θ_2, 9 (initial θ_1) × 5 (amount of change) for each bank size. The response simulation was done by a program written in R (R Development Core Team, 2012).

AMC Procedures

The initial θ was set to zero for all simulees for Time 1, and the final ML estimate of θ at Time 1 (θ_1) was used as the starting θ estimate for the Time 2 CAT. At Time 1, items were selected based on Fisher information at the current θ estimate. At Time 2, three item selection methods, Fisher information, Finkelman's KL information (KL_F), and modified KL information (KL_M) were used for each bank condition. Five hypothesis testing statistics—Z, LR, ST, KLD with uniform prior (KLD-U), and KLD with normal prior (KLD-N)—were computed for each item selection method at Time 2. The resulting 15 combinations of item selection and hypothesis testing were: FI-Z , FI-LR, FI-ST, FI-KLD-U, FI-KLD-N, KL_F-Z, KL_F-LR, KL_F-ST, KL_F-KLD-U, KL_F-KLD-N, KL_M-Z, KL_M-LR, KL_M-ST, KL_M-KLD-U, and KL_M-KLD-N. The θ estimates were obtained using MLE, with a temporary use of expected a posteriori (EAP) for non-mixed response patterns. The MLE was bounded in the range between −4 and 4.

For a fixed-length AMC, test statistics were calculated at the final stage of Time 2 with the test terminated after 15, 30, or 50 items were administered. The same length was used for both time points (e.g., 30 items in Time 1 and 30 items in Time 2). The AMC simulation was done using a program written in Lazarus (Lazarus and Free Pascal Team, version 2.2, 2008).

Variable-Length AMC Post-Hoc Simulation

For variable-length AMC simulations, a post-hoc simulation based on the results from fixed-length AMCs was used. Test length of the fixed-length AMC was used as the maximum length of the variable-length AMC—15, 30 or 50 items. For example, results for a fixed-length AMC with n = 30 were used for a post-hoc simulation based on the response vectors for a variable-length AMC with the maximum length 30. The minimum length was set to 10 for all variable-length AMCs: results after 10 items were considered to determine the significance of the test statistics. For a post-hoc simulation for variable-length AMC, if the test statistics based on *k* items was significant, $10 < k \le n_{\max}$, where $n_{\max}$, the test was treated as if being terminated at length *k*. The use of post-hoc simulation based on a fixed-length AMC results saved simulation time, compared to a separate run of variable-length AMC based on a separate set of data. The results from variable-length AMC were, therefore, more comparable to those from a fixed-length AMC, as it removed the random error due to a use of different response data.

A separate set of post-hoc simulations was conducted to determine threshold values C_1 and C_2 beforehand. C_1 was found first, then C_2 was found based on C_1. For each test statistic, the test statistic history of the interim stages for each examinee, between item 11 to $n_{\max} - 1$, was compared to candidate C_1 value for each of the 18 combinations of item selection and bank conditions. If any interim statistic exceeded C_1, then the examinee was counted as significant change and early termination. A sequence of candidate C_1 values such that $C_1 \ge T_\alpha$ with an increment 0.01 was evaluated until the proportion of rejecting the null hypothesis at the interim stage was approximately $\varepsilon\alpha = 0.023$. The procedure was repeated for 30 sets of AMC fixed-length results. C_1 was then selected such that the average proportion of incorrectly rejecting the null hypothesis at the interim stage across 30 replications was approximately $\varepsilon\alpha = 0.023$.

C_2 was selected using the similar procedure after finding C_1, based on the same AMC fixed-length results used for C_1. At the interim stage, the test statistic was compared to the selected C_1 to determine significance of change. For examinees who were not terminated at the interim stage, the test statistic at the maximum test length was compared to candidate values of C_2 for each of 18 conditions of item selection and bank. A sequence of candidate C_2 values such that $C_2 \ge T_\alpha$ with an increment 0.01 was evaluated until the proportion of rejecting the null hypothesis at the interim stage was approximately $(1-\varepsilon)\alpha = 0.027$. C_2 was selected such that the average proportion of incorrectly rejecting the null hypothesis at the final stage, after an interim decision was made based on C_1, across 30 replications was approximately the remainder of the desirable Type I error rate, $(1-\varepsilon)\alpha = 0.027$. Tables A2 – A7 in the Appendix summarize values of C_1 and C_2 used for each test statistic and AMC condition, and the average and standard deviation (SD) of observed Type I error rates at both interim and final stages.

Conditions

This study used 9 (Time 1 *θ*) × 5 (level of change) response conditions and 3 (level of discrimination) × 2 (bank information shape) × 2 (bank size) bank conditions. Item responses were generated independently for each of the 45 response conditions. For each of the 45 response sets, 3 item selection methods and 5 hypothesis testing methods were fully crossed, resulted in a 9 (initial *θ*) × 5 (level of change) × 3 (level of discrimination) × 2 (bank information shape) × 2 (bank size) × 3 (item selection) × 5 (hypothesis testing) fully crossed ANOVA design. To ensure stability in the results, all conditions and procedures were repeated 100 times.

Evaluation Criteria

The observed Type I error rate—proportion of incorrect classifications as change in the no-change condition—and the observed power—the proportion of correct classification as change in the four level of change conditions were evaluated conditional on each *θ* level and 12 bank conditions for each hypothesis testing and item selection combination. For variable-length testing, the number of items administered was also evaluated. Means and SDs of these dependent variables were also computed across100 replications.

An ANOVA approach was also used to summarize the effects and two-way interactions on the dependent variables; interactions beyond the two-way were incorporated into the error term of the ANOVAs. Effect size, η^2, was computed for each effect in the ANOVA. η^2 is defined as a ratio of sum of squares,

$$\eta^2 = \frac{SS_{\text{effect}}}{SS_{\text{total}}}. \tag{39}$$

η^2 was multiplied by 100 and reported as a percentage. The Type I error rate was used as the dependent variable in the no-change condition. The proportion of correct classifications as change, i.e., power, was used as the dependent variable in a separate ANOVA analysis in the four change conditions.

Chapter 3

RESULTS

ANOVA

Table 2 summarizes the degrees of freedom, sum of squares and η^2 (as a percentage) from the seven-way ANOVAs for fixed-length and variable-length AMC using observed Type I error rate as the response variable, and Tables 3 – 6 summarize ANOVA results on observed power under four levels of true change.

Table 2 shows that the largest variance in Type I error was accounted for by the test statistic (24.73%) for fixed-length AMC. The other main effects and interactions for fixed-length AMC accounted for less than 5% of variation in Type I error. For variable-length AMC, the largest η^2 was observed for the interaction between test statistic and θ level (8.61%), and the main effect of the test statistic (5.12%). The rest of the main effects and interactions had η^2 for Type I error smaller than 5%.

In Table 3, the largest η^2 for power under very small change ($\Delta\theta = 0.25$) was observed for the main effects of discrimination (41.30%), test length (26.23%), and test statistic (8.41%) for fixed-length AMC. The interaction between discrimination and test length also accounted for 7.29% of variation in power under very small change. For variable-length AMC, the largest η^2 for power was also observed for the main effects of discrimination (38.74%); and test length (31.55%), and the interaction between discrimination and test length (8.20%). The rest of the main effects and interactions each accounted for less than 5% of variation in power under very small change for both fixed- and variable-length AMC.

Table 4 shows ANOVA results for power under small true change ($\Delta\theta = 0.5$). The largest η^2 for power with fixed-length AMC was observed for the main effects of discrimination

(47.23%), test length (32.30%) and the interaction between discrimination and test length (6.76%). Similarly, the largest η^2 for power with variable-length AMC was observed for the main effects of discrimination (45.64%), test length (34.12%) and the interaction between discrimination and test length (7.46%). The rest of the main effects and interactions each accounted for less than 5% of variation in power for both fixed- and variable-length AMC.

In Table 5, the largest η^2 for power under medium change ($\Delta\theta = 1.0$) for fixed-length AMC was observed for the main effects of discrimination (46.48%), test length (36.07%) and θ level (6.04%). For variable-length AMC, the largest η^2 for power was also observed for the main effects of discrimination (46.01%), test length (37.15%) and θ level (6.52%). No interactions accounted for larger than 5% of variation in power under medium true change for both fixed- and variable-length AMC.

Table 6 presents ANOVA results for power under large true change. The largest η^2 for power for fixed-length AMC was observed for the main effects of test length (30.54%), discrimination (26.98%), and θ level (13.39%). The interactions between discrimination and test length (8.86%), and θ and bank information (6.15%), also accounted for relatively large variation in power for fixed-length AMC. The largest η^2 for power for variable-length AMC was also observed for the main effects of test length (31.09%), discrimination (26.89%), and θ level (14.54%). The largest η^2 for interactions was observed for the interaction between discrimination and test length (8.34%), and θ and bank information (6.68%), as they were for fixed-length AMCs.

Table 2. Results of ANOVA on Type I Error Rates

Effect Source		Fixed-Length		Variable-Length	
	df	SS	η^2	SS	η^2
θ	8	1.55	2.09	2.17	4.24
Bank Information	1	.16	.22	.01	.02
Bank Size	1	.00	.00	.00	.00
Discrimination	2	2.35	3.16	.62	1.21
Item Selection	2	.16	.22	.14	.27
Test Length	2	.40	.54	.53	1.03
Test Statistic	4	18.39	**24.73**	2.62	**5.12**
θ × Bank Information	8	.01	.02	.06	.11
θ × Bank Size	8	.00	.01	.01	.02
θ × Discrimination	16	.98	1.32	1.01	1.97
θ × Item Selection	16	.02	.03	.17	.32
θ × Test Length	16	1.18	1.59	1.08	2.11
θ × Test Statistic	32	2.90	3.90	4.42	**8.61**
Bank Information × Bank Size	1	.00	.00	.00	.00
Bank Information × Discrimination	2	.01	.01	.00	.00
Bank Information × Item Selection	2	.00	.00	.01	.01
Bank Information × Test Length	2	.00	.00	.02	.04
Bank Information × Test Statistic	4	1.05	1.42	.03	.05
Bank Size × Discrimination	2	.00	.00	.00	.01
Bank Size × Item Selection	2	.00	.00	.00	.00
Bank Size × Test Length	2	.01	.01	.00	.01
Bank Size × Test Statistic	4	.10	.13	.01	.02
Discrimination × Item Selection	4	.01	.02	.01	.03
Discrimination × Test Length	4	.57	.77	.13	.26
Discrimination × Test Statistic	8	2.75	3.70	1.85	3.61
Item Selection × Test Length	4	.02	.03	.07	.13
Item Selection × Test Statistic	8	2.67	3.59	.18	.35
Test Statistic × Test Length	8	3.48	4.67	.58	1.12
Error	485,826	35.57	47.84	35.55	69.35
Total	485,999	74.36	100.00	51.26	100.00

Table 3. Results of ANOVA on Power Under Very Small True Change

Effect Source	*df*	Fixed-Length SS	Fixed-Length η^2	Variable-Length SS	Variable-Length η^2
θ	8	34.00	2.54	30.40	2.64
Bank Information	1	3.63	.27	1.34	.12
Bank Size	1	6.54	.49	5.48	.48
Discrimination	2	553.30	**41.30**	446.50	**38.74**
Item Selection	2	4.09	.31	5.45	.47
Test Length	2	351.40	**26.23**	363.61	**31.55**
Test Statistic	4	112.63	**8.41**	38.38	3.33
θ × Bank Information	8	24.20	1.81	24.72	2.14
θ × Bank Size	8	.12	.01	.15	.01
θ × Discrimination	16	7.99	.60	6.76	.59
θ × Item Selection	16	.30	.02	.49	.04
θ × Test Length	16	9.20	.69	7.69	.67
θ × Test Statistic	32	6.24	.47	14.13	1.23
Bank Information × Bank Size	1	.05	.00	.03	.00
Bank Information × Discrimination	2	.06	.00	.01	.00
Bank Information × Item Selection	2	.01	.00	.01	.00
Bank Information × Test Length	2	.82	.06	.20	.02
Bank Information × Test Statistic	4	1.66	.12	.31	.03
Bank Size × Discrimination	2	1.92	.14	1.51	.13
Bank Size × Item Selection	2	.03	.00	.02	.00
Bank Size × Test Length	2	2.68	.20	2.31	.20
Bank Size × Test Statistic	4	.10	.01	.06	.01
Discrimination × Item Selection	4	1.13	.08	.92	.08
Discrimination × Test Length	4	97.65	**7.29**	94.49	**8.20**
Discrimination × Test Statistic	8	1.35	.10	1.23	.11
Item Selection × Test Length	4	.18	.01	.12	.01
Item Selection × Test Statistic	8	8.56	.64	.70	.06
Test Statistic × Test Length	8	4.68	.35	3.41	.30
Error	485,826	105.11	7.85	102.08	8.86
Total	485,999	1339.63	100.00	1152.51	100.00

Table 4. Results of ANOVA on Power Under Small True Change

Effect Source		Fixed-Length		Variable-Length	
	df	SS	η^2	SS	η^2
θ	8	482.70	3.23	472.50	3.37
Bank Information	1	18.80	.13	9.90	.07
Bank Size	1	75.30	.50	69.50	.50
Discrimination	2	7059.50	**47.23**	6402.30	**45.64**
Item Selection	2	27.60	.18	34.30	.24
Test Length	2	4827.70	**32.30**	4787.40	**34.12**
Test Statistic	4	380.00	2.54	141.00	1.01
θ × Bank Information	8	305.30	2.04	311.20	2.22
θ × Bank Size	8	1.40	.01	1.30	.01
θ × Discrimination	16	116.40	.78	116.90	.83
θ × Item Selection	16	1.90	.01	1.90	.01
θ × Test Length	16	60.80	.41	60.30	.43
θ × Test Statistic	32	20.70	.14	41.70	.30
Bank Information × Bank Size	1	.60	.00	.40	.00
Bank Information × Discrimination	2	2.00	.01	2.10	.01
Bank Information × Item Selection	2	.10	.00	.10	.00
Bank Information × Test Length	2	2.30	.02	.80	.01
Bank Information × Test Statistic	4	3.60	.02	2.00	.01
Bank Size × Discrimination	2	13.20	.09	12.50	.09
Bank Size × Item Selection	2	.20	.00	.10	.00
Bank Size × Test Length	2	24.90	.17	24.30	.17
Bank Size × Test Statistic	4	.20	.00	.20	.00
Discrimination × Item Selection	4	6.80	.05	5.80	.04
Discrimination × Test Length	4	1009.60	**6.76**	1046.40	**7.46**
Discrimination × Test Statistic	8	4.40	.03	3.30	.02
Item Selection × Test Length	4	3.20	.02	2.40	.02
Item Selection × Test Statistic	8	27.90	.19	3.20	.02
Test Statistic × Test Length	8	19.90	.13	20.10	.14
Error	485,826	448.70	3.00	455.30	3.25
Total	485,999	14945.70	100.00	14029.20	100.00

Table 5. Results of ANOVA on Power Under Medium True Change

Effect Source	*df*	Fixed-Length SS	Fixed-Length η^2	Variable-Length SS	Variable-Length η^2
θ	8	1953.10	**6.04**	2129.00	**6.52**
Bank Information	1	.00	.00	1.70	.01
Bank Size	1	122.90	.38	124.90	.38
Discrimination	2	15030.60	**46.48**	15028.10	**46.01**
Item Selection	2	59.40	.18	66.00	.20
Test Length	2	11664.60	**36.07**	12134.20	**37.15**
Test Statistic	4	625.40	1.93	311.00	.95
θ × Bank Information	8	1027.00	3.18	1121.00	3.43
θ × Bank Size	8	12.70	.04	13.20	.04
θ × Discrimination	16	76.20	.24	89.80	.27
θ × Item Selection	16	8.30	.03	7.50	.02
θ × Test Length	16	22.70	.07	21.60	.07
θ × Test Statistic	32	100.80	.31	116.30	.36
Bank Information × Bank Size	1	.00	.00	.00	.00
Bank Information × Discrimination	2	86.60	.27	83.40	.26
Bank Information × Item Selection	2	.50	.00	.70	.00
Bank Information × Test Length	2	2.30	.01	3.80	.01
Bank Information × Test Statistic	4	2.30	.01	2.90	.01
Bank Size × Discrimination	2	6.50	.02	5.80	.02
Bank Size × Item Selection	2	.10	.00	.10	.00
Bank Size × Test Length	2	4.70	.01	6.10	.02
Bank Size × Test Statistic	4	.50	.00	.20	.00
Discrimination × Item Selection	4	3.10	.01	1.60	.00
Discrimination × Test Length	4	506.30	1.57	483.70	1.48
Discrimination × Test Statistic	8	69.00	.21	32.00	.10
Item Selection × Test Length	4	38.70	.12	37.20	.11
Item Selection × Test Statistic	8	42.70	.13	10.40	.03
Test Statistic × Test Length	8	155.20	.48	98.60	.30
Error	485,826	716.10	2.21	729.90	2.23
Total	485,999	32338.30	100.00	32660.70	100.00

Table 6. Results of ANOVA on Power Under Large True Change

Effect Source	*df*	Fixed-Length		Variable-Length	
		SS	η^2	SS	η^2
θ	8	2567.20	**13.39**	2875.20	**14.54**
Bank Information	1	52.80	.28	70.90	.36
Bank Size	1	63.20	.33	66.40	.34
Discrimination	2	5173.30	**26.98**	5318.10	**26.89**
Item Selection	2	43.20	.23	48.90	.25
Test Length	2	5856.60	**30.54**	6149.30	**31.09**
Test Statistic	4	392.70	2.05	229.00	1.16
θ × Bank Information	8	1179.80	**6.15**	1321.10	**6.68**
θ × Bank Size	8	26.90	.14	29.00	.15
θ × Discrimination	16	83.90	.44	88.80	.45
θ × Item Selection	16	14.30	.07	15.40	.08
θ × Test Length	16	382.80	2.00	403.00	2.04
θ × Test Statistic	32	230.00	1.20	216.30	1.09
Bank Information × Bank Size	1	1.00	.01	1.30	.01
Bank Information × Discrimination	2	56.10	.29	57.90	.29
Bank Information × Item Selection	2	.40	.00	.40	.00
Bank Information × Test Length	2	.30	.00	.80	.00
Bank Information × Test Statistic	4	.50	.00	4.40	.02
Bank Size × Discrimination	2	14.90	.08	15.10	.08
Bank Size × Item Selection	2	.10	.00	.00	.00
Bank Size × Test Length	2	.90	.00	.60	.00
Bank Size × Test Statistic	4	1.20	.01	.60	.00
Discrimination × Item Selection	4	.20	.00	.30	.00
Discrimination × Test Length	4	1698.80	**8.86**	1649.10	**8.34**
Discrimination × Test Statistic	8	177.00	.92	112.60	.57
Item Selection × Test Length	4	43.70	.23	44.40	.22
Item Selection × Test Statistic	8	17.10	.09	7.90	.04
Test Statistic × Test Length	8	168.50	.88	93.90	.47
Error	485,826	926.60	4.83	956.10	4.83
Total	485,999	19174.00	100.00	19776.80	100.00

Effect of θ

Figure 4 depicts marginal means of Type I error and power at each θ level for fixed- and variable-length AMCs. Black solid lines depict marginal means for fixed-length AMC, and red dashed lines depict marginal means for variable-length AMC. The observed η^2 for Type I error for θ level was 2.09% for fixed-length and 4.24% for variable-length AMC (Table 2). Figure 4a shows that marginal means of Type I error rate ranged between 0.048 and 0.054 for fixed-length AMC; and between 0.046 and 0.053 for variable-length AMC, relatively stable across θ. Type I errors were slightly higher for lower θs than those for higher θs, however the differences were negligible. The effect of θ level on power under small change also had η^2 smaller than 5% for both fixed- and variable-length AMCs (Table 4). In Figure 4b, power slightly decreased as θ deviated from 0.0, but differences between θs were trivial relative to the effect of other factors.

Figure 4c displays marginal means for power under medium true change ($\Delta\theta = 1.0$). The observed η^2 for θ level on power under medium change was 6.04% for fixed-length AMC and 6.52% for variable-length AMC. A decrease in power was observed as θ deviated from −0.5 for both fixed- and variable-length AMCs. For fixed-length AMC, marginal mean power was 0.646 at $\theta = -2$; 0.714 at $\theta = -0.5$; and 0.504 at $\theta = 2$. Variable-length AMC had slightly smaller marginal mean power values: observed marginal mean power was 0.637 at $\theta = -2$; 0.705 at $\theta = -0.5$, and 0.485 at $\theta = 2$. A larger decrease in power was observed at high extreme θ than in low extreme θ.

The effect of θ level on power had observed η^2 of 13.39% for fixed-length AMC, and 14.54% for variable-length AMC under large change ($\Delta\theta = 1.5$). As shown in Figure 4d, marginal means for power were relatively stable at $\theta < 1$, with a minor decrease in power as θ deviated from 0; and a larger decrease in power at $\theta \geq 1$. For fixed-length AMC, marginal means of power were 0.863 at $\theta = -2$; 0.893 at $\theta = -1$; 0.898 at $\theta = 0$; 0.868 at $\theta = 1$; and 0.659 at $\theta = 2$. Variable-length AMC had slightly smaller marginal mean values of power than those for fixed-length AMC. For variable-length AMC, marginal means of power were 0.858 at $\theta = -2$; 0.889 at $\theta = -1$; 0.895 at $\theta = 0$; 0.862 at $\theta = 1$; and 0.641 at $\theta = 2$. Although results for power under very small change ($\Delta\theta = 0.25$) are not displayed in Figure 4, small but relatively constant level of power between 0.091 and 0.116 was observed for fixed-length AMC, and between 0.085 and 0.110 was observed for variable-length AMC. Marginal means and SDs at each θ level for fixed- and variable-length AMCs are summarized in Appendix Table A8.

Variable-length AMC used smaller number of items on average than fixed-length AMC, although the two termination methods had similar observed η^2 s for the effect of θ level, as well as similar marginal means for Type I error and power under all true change conditions. Figure 5 compares means for the number of items used to obtain the marginal effects of θ shown in Figure 4. In Figure 5, black dashed lines indicate average number of items used for fixed-length AMC, and blue dashed lines indicate marginal means of the number of items administered for variable-length AMC. Fixed-length AMC had an average of 31.67 items constantly across θ, which was the average of the three test lengths 15, 30 and 50. Under the no-change condition, variable-length AMC administered approximately 31.33 items, which was relatively stable across θ and was similar to fixed-length AMC (Figure 5a). AMC terminated earlier (i.e., smaller number of items were used compared to fixed-length AMC with the same maximum test length) as the amount of change increased.

Under small change ($\Delta\theta = 0.5$), variable-length AMC used about 28.99 items on average (about 8.43% saved), which was relatively constant but slightly more items were used at $\theta = 2$: the marginal mean number of items used was 29.27 at $\theta = -2$; 28.74 at $\theta = -1$; 28.65 at $\theta = 0$; 28.82 at $\theta = 1$; and 29.97 at $\theta = 2$ (Figure 5b). Under medium true change ($\Delta\theta = 1.0$), variable-length AMC administered about 21.84 items on average (about 31.01% saved), and more items were used at $\theta \geq 1.0$ (minimum 17.81% saved). The marginal mean number of items used was 21.98 at $\theta = -2$; 20.86 at $\theta = -1$; 20.63 at $\theta = 0$; 21.41 at $\theta = 1$; and 26.03 at $\theta = 2$ (Figure 5c). Under large true change in Figure 5d ($\Delta\theta = 1.5$), the number of items used was similar between 15 and 16 items (about 46.78% saved) at $\theta < 1.0$, and the number of items increased at $\theta \geq 1.0$ (minimum 26.65% saved): the marginal mean number of items used was 16.27 at $\theta = -2$; 15.61 at $\theta = -1$; 15.18 at $\theta = 0$; 16.27 at $\theta = 1$; and 23.23 at $\theta = 2$. The average and SD of the number of items used at each θ level are summarized in Appendix Table A9.

Figure 4. Effect of θ Level on Type I Error and Power for Fixed- and Variable-Length AMC

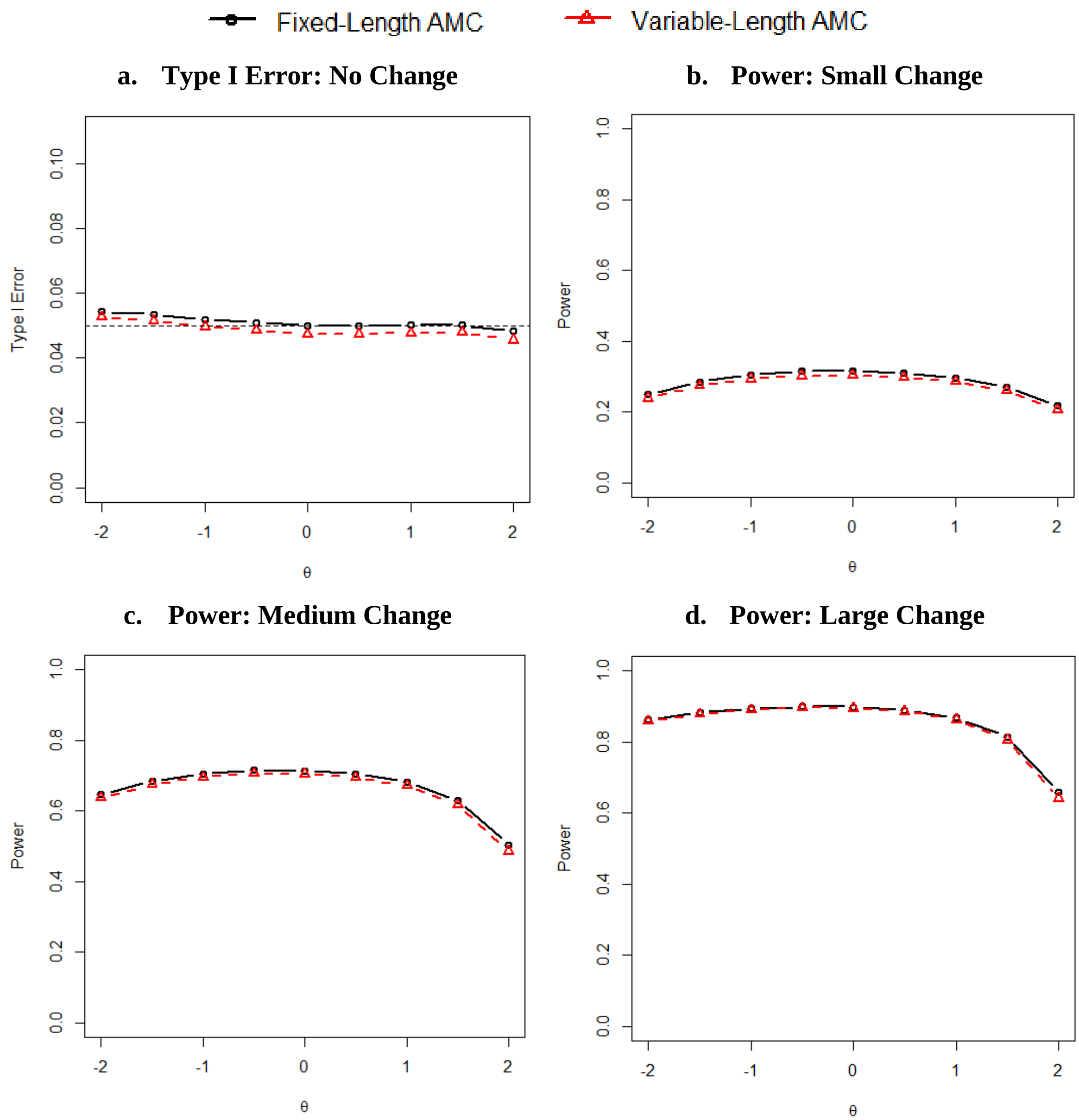
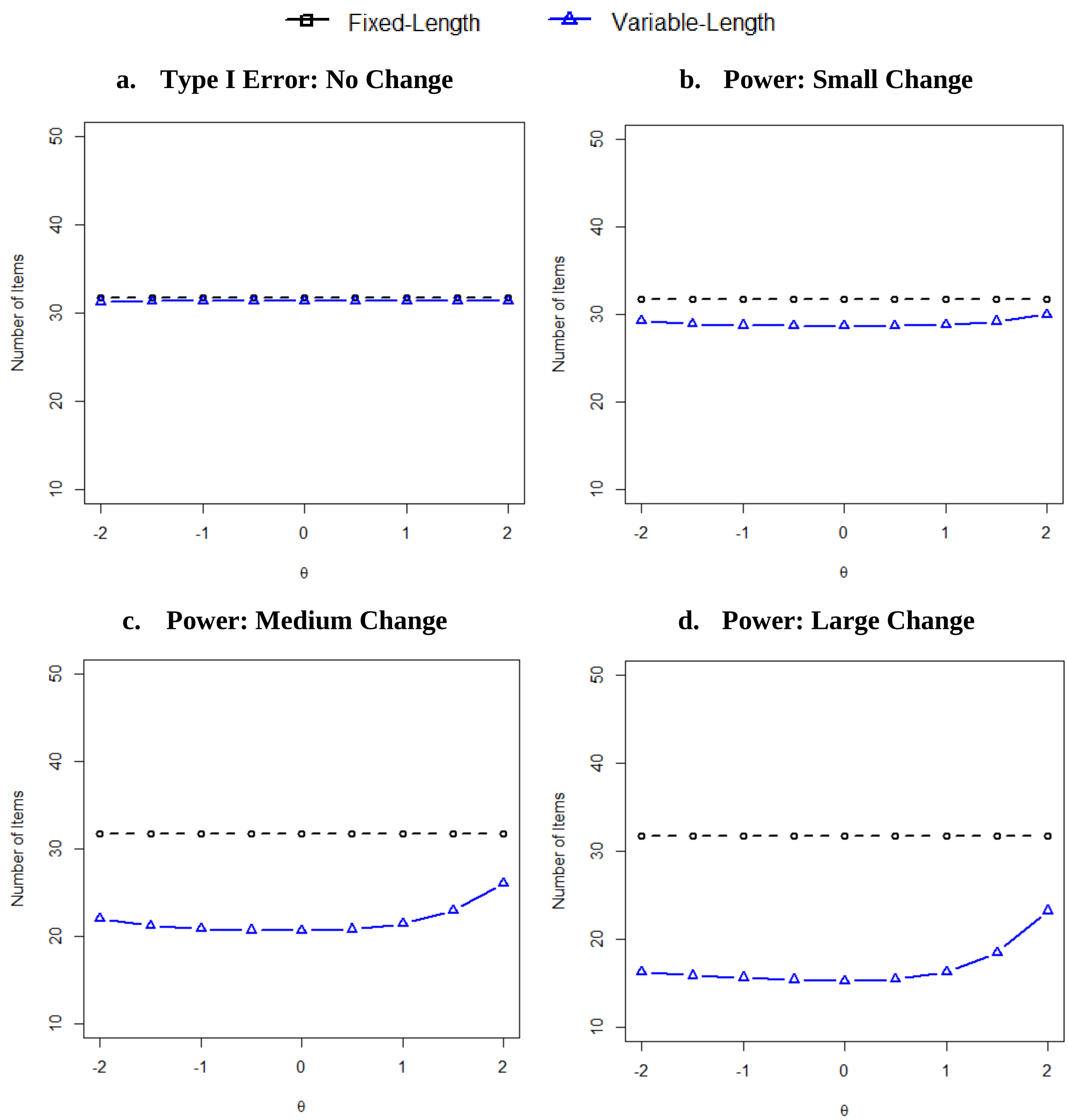


Figure 5. Mean Number of Items Used Conditional on θ for Fixed- and Variable-Length AMC



Effect of Bank Information

The effect of bank information was small both for fixed-length and variable-length AMC, with similar effect sizes: the main effects of bank information, interactions between the bank information and bank size, discrimination, item selection, test length, and test statistic accounted for less than 1% of variance for Type I error and power under all change conditions. The interactions between the bank information and θ had η^2 for Type I error smaller than 1%, and η^2 for power smaller than 5% under very small to medium conditions. Under large change, the interaction between the bank information and θ had $\eta^2 = 6.15\%$ for fixed-length AMC and 6.68% for variable-length AMC.

Figure 6 shows marginal means of Type I error and power for peaked and flat banks conditional on θ. Solid lines indicate results from a peaked bank and dashed lines indicate results from a flat bank. Fixed- and variable-length AMCs are distinguished by color, black and red, respectively. In Figure 6a, Type I errors between peaked and flat banks were indistinguishable across θ for both fixed- and variable-length AMCs. In Figures 6b – 6d, flat AMC banks had stable power functions across θ, and peaked AMC had a decrease in power as θ deviated from 0.0. Larger decrease in power at high θ range was observed as the amount of true change increased. Note that the effect sizes of the interactions under small and medium change were about 2% for small change (Figure 6b), and about 3% for medium change (Figure 6c) for both fixed-length and variable-length AMC.

In Figure 6d under large true change ($\Delta\theta = 1.5$), a similar pattern of interaction was observed, however differences between the two bank information structures conditional on θ were larger. Constant levels of power around 0.86 were observed, with a slight decrease at $\theta = 2$, for AMC with flat bank information under large true change ($\Delta\theta = 1.5$). Marginal means for the flat bank were 0.86 at $\theta = -2$; 0.87 at $\theta = -1$; 0.88 at $\theta = 0$; 0.87 at $\theta = 1$; and 0.80 at $\theta = 2$.

Peaked and flat banks had identical marginal mean power level at $\theta = -2$, and a peaked bank had slightly higher power than a flat bank at $-2 < \theta < 1$. A large decrease in power at $\theta > 1.0$ was observed for AMC with peaked bank information, whereas power decreased slightly in the low θ range. Marginal means for the peaked bank were 0.86 at $\theta = -2$; 0.89 at $\theta = -1$; 0.92 at $\theta = 0$; 0.87 at $\theta = 1$; and 0.52 at $\theta = 2$. Results for power under very small change ($\Delta\theta = 0.25$) are not displayed in Figure 6, but power functions for both fixed-length and variable-length AMC, and for both peaked and flat bank information were small and relatively stable between 0.074 and 0.126 across θ. Marginal means and SDs of Type I error and power conditional on bank information and θ are summarized in Appendix Table A10.

Variable-length AMC used smaller numbers of items than fixed-length AMC of the same maximum test length condition. Figure 7 compares marginal mean number of items used for fixed-length and variable-length with peaked and flat bank information. Fixed-length AMC administered a constant 31.67 items, on average. Under the no-change condition, variable-length AMC used 31.33 items on average, which was similar to that of fixed-length AMC, for both fixed and flat banks (Figure 7a). In Figures 7b – 7d, variable-length AMC used smaller numbers of items than fixed-length AMC as the amount of change increased. Peaked banks used more items than flat banks at $\theta > 1$ under medium and large change.

Under small change, peaked and flat banks both used a mean of approximately 29 items (8.43% saved) across θ (Figure 7b). Under medium true change in Figure 7c, variable-length AMC with a flat bank used 21.72 items on average (31.42% saved), relatively constantly across θ. Peaked banks with medium change used about 21.09 items on average at $\theta \leq 1$ (33.41% saved) and more items at $\theta > 1$, resulting in a minimal item reduction in test length of 8.46% at $\theta = 2$. Under large true change in Figure 7d, flat banks used about 16.02 items at $\theta \leq 1$ (49.42% saved) and used more items at $\theta > 1$: 16.92 items at $\theta = 1.5$ (46.54% saved) and 18.83 items at $\theta = 2.0$

(40.54% saved). Peaked banks used 15.41 items (51.34% saved) on average at $\theta \leq 1$, which was smaller than that of flat banks, and used 19.90 items at $\theta = 1.5$ (37.16% saved); and 27.62 items at $\theta = 2$ (12.78% saved). The average number of items used and SD conditional on θ level and bank information are summarized in Appendix Table A11.

Figure 6. Effect of Bank Information on Type I Error and Power Conditional on θ for Fixed- and Variable-Length AMC

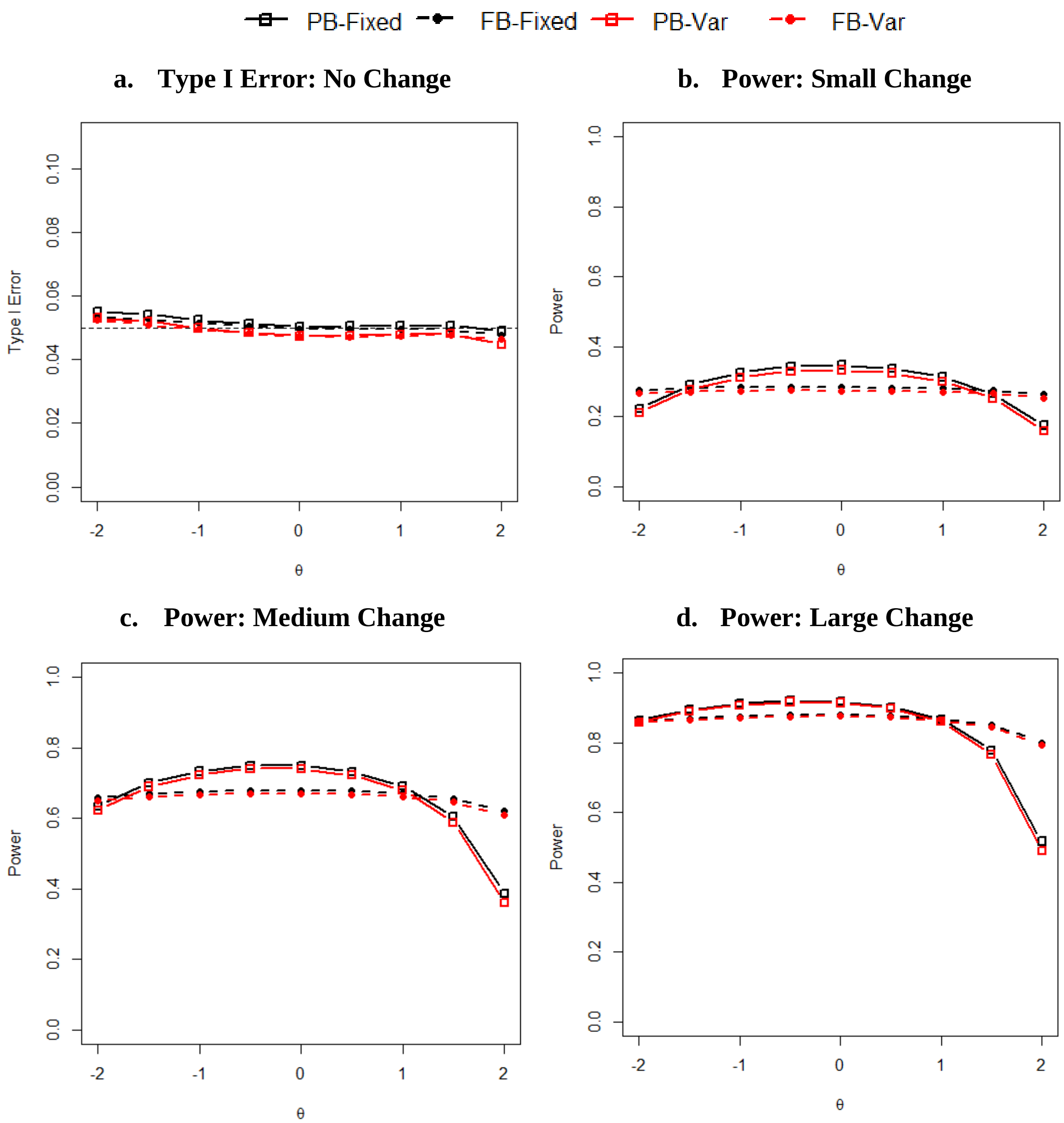
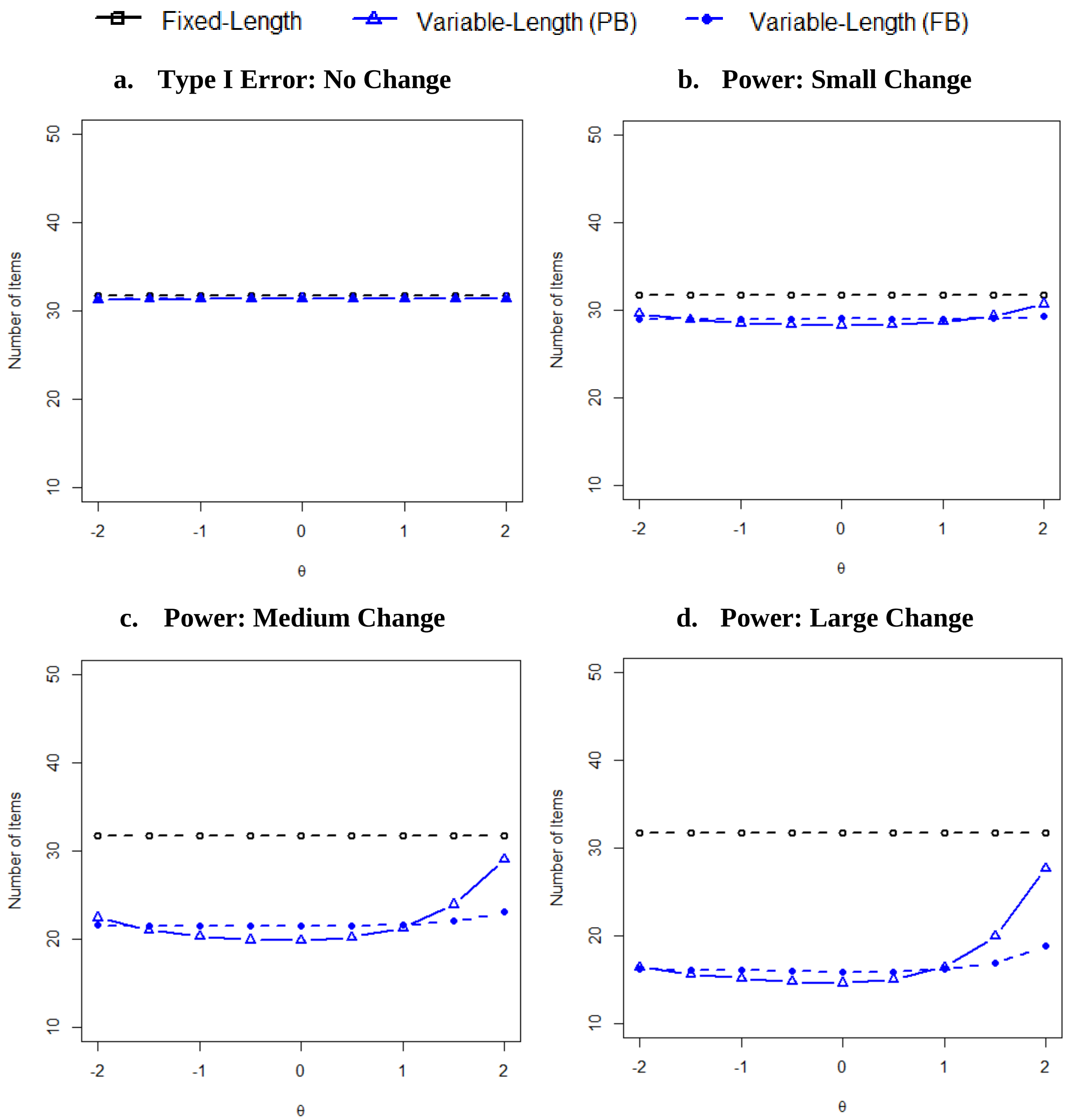


Figure 7. Mean Number of Items Used Conditional on Bank Information and θ for Fixed- and Variable-Length AMC



Effect of Bank Size

The effect of the bank size on Type I error and power was negligible. The main effect of bank size and interactions between the bank size and θ, bank information, discrimination, item selection, test length, test statistic all accounted for less than 1% of variance in Type I error and power under all change conditions for both fixed-length and variable-length AMCs (Tables 2 – 6). The two bank sizes, 300-item and 500-item banks, had similar Type I error, but the 500-item bank had slightly higher power than the 300-item bank for both fixed-length and variable-length AMC. Differences in power between the 500-item and the 300-item banks averaged 0.025 under small change, 0.032 under medium change, and 0.023 under large change. Marginal effects of the bank size are summarized in Appendix Table A12.

Effect of Discrimination

The main effect of item discrimination for Type I error had observed η^2 smaller than 5% for both fixed-length and variable-length AMCs. The main effect of discrimination on power was similar for fixed- and variable-length AMC: the effect was slightly larger for fixed-length AMC. The observed η^2 for power for the main effect of discrimination was 41.3% for fixed-length AMC and 38.74% for variable-length AMC under very small change. The observed η^2 for power for the main effect of discrimination was 47.23% for fixed-length AMC and 45.64% for variable-length AMC under small change. Under medium change, observed η^2 for power for the main effect of discrimination was similar at 46.48% for fixed-length AMC and 46.01% for variable-length AMC. The main effect of discrimination on power was somewhat lower under large change accounting for 26.98% of variance for fixed-length AMC and 26.89% of variance for variable-length AMC.

Type I error slightly increased from 0.048 to 0.054 for fixed-length and from 0.047 to 0.050 for variable-length AMC as discrimination increased from low to high (Figure 8a). In Figures 8b – 8d, higher power was observed for higher discrimination conditions for all change levels. Fixed-length and variable-length AMC resulted in almost identical power at each discrimination level. Marginal means for power under small change were 0.147, 0.266, and 0.441 for low, moderate, and high discrimination conditions, respectively for fixed-length AMC; and 0.142, 0.255, and 0.422 for variable-length AMC (Figure 8b). Under medium change, observed marginal means of power were 0.435, 0.697 and 0.862 for low, moderate, and high discrimination conditions for fixed-length AMC; and 0.425, 0.685, and 0.852 for variable-length AMC (Figure 8c). Under large change, marginal means of power were 0.710, 0.893, and 0.952 for fixed-length AMC, and 0.702, 0.887, and 0.948 for variable-length AMC for low, moderate and high discrimination conditions, respectively (Figure 8d). Although not displayed in Figure 8, a similar pattern of power was observed under very small change ($\Delta\theta = 0.25$): observed marginal means of power were 0.071, 0.101 and 0.153 for low, moderate, and high discrimination conditions for fixed-length AMC; and 0.069, 0.096, and 0.142 for variable-length AMC. Marginal means and SDs for the discrimination conditions are summarized in Appendix Table A13.

Figure 9 compares the average number of items used conditional on discrimination for four change conditions. Fixed-length AMC used 31.67 items on average constantly for all discrimination conditions. Variable-length AMC used smaller number of items on average than fixed-length AMC: differences between fixed- and variable-length AMC increased with larger change, and/or as discrimination increased. The averages for the number of items used were 30.51 (3.66% saved), 29.35 (7.33% saved), and 27.09 (14.46% saved) for low, moderate and high discrimination, respectively under small change (Figure 9b). Under medium change, 26.94 (14.94% saved), 21.80 (31.17% saved) and 16.78 (47.02% saved) items were used on average for low, moderate and high discrimination conditions, respectively (Figure 9c). The mean test lengths for variable-length AMC were 21.28, 15.68, and 13.58 for low, moderate, and high discrimination conditions, respectively, under large true change (Figure 9d) resulting in a minimum item reduction in test length of 32.81% and a maximum of 57.12% from that of the fixed-length test. The means and SDs for the number of items used for each discrimination and change condition are summarized in Appendix Table A14.

Figure 8. Effect of Discrimination on Type I Error and Power for Fixed- and Variable-Length AMC

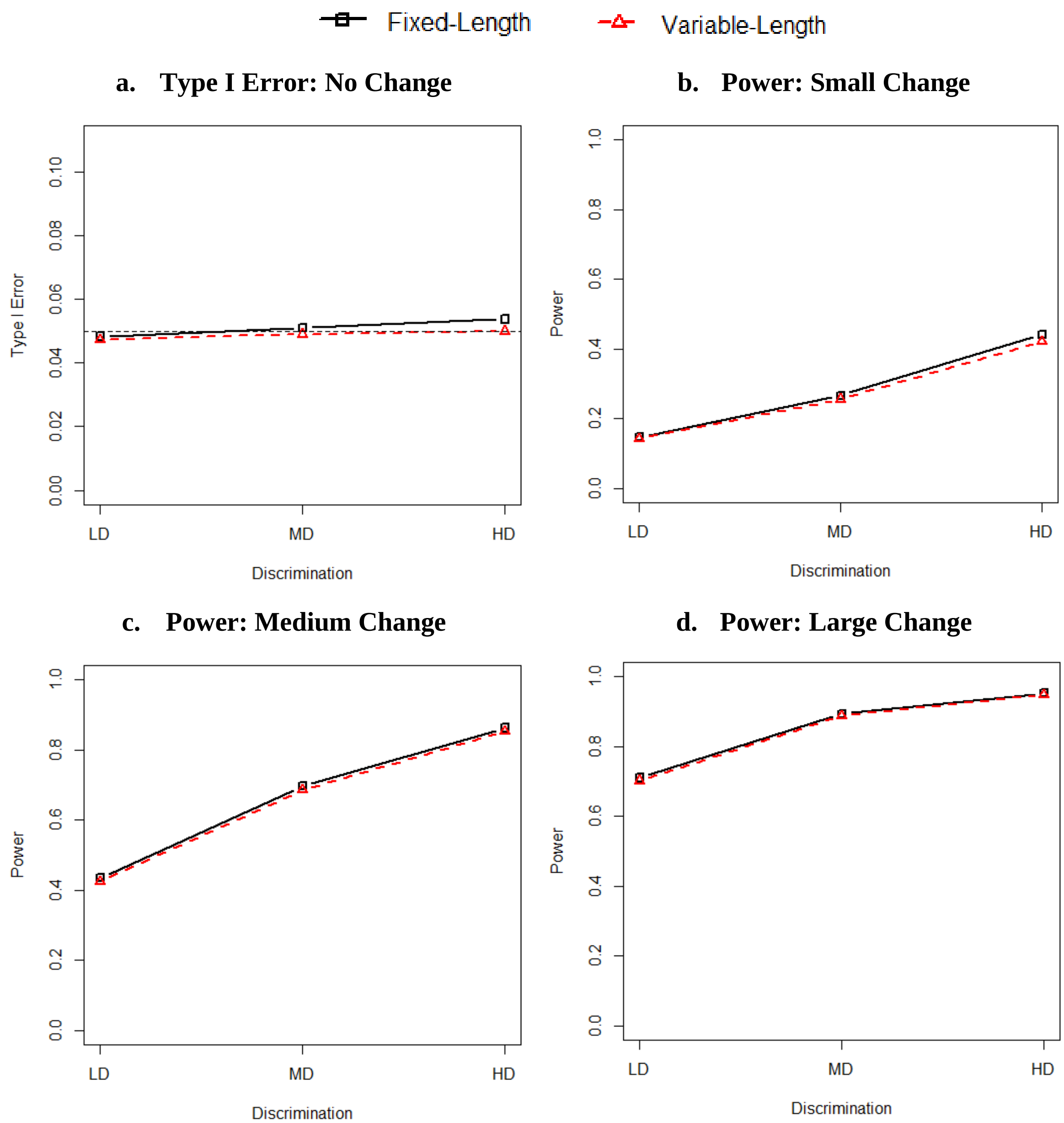
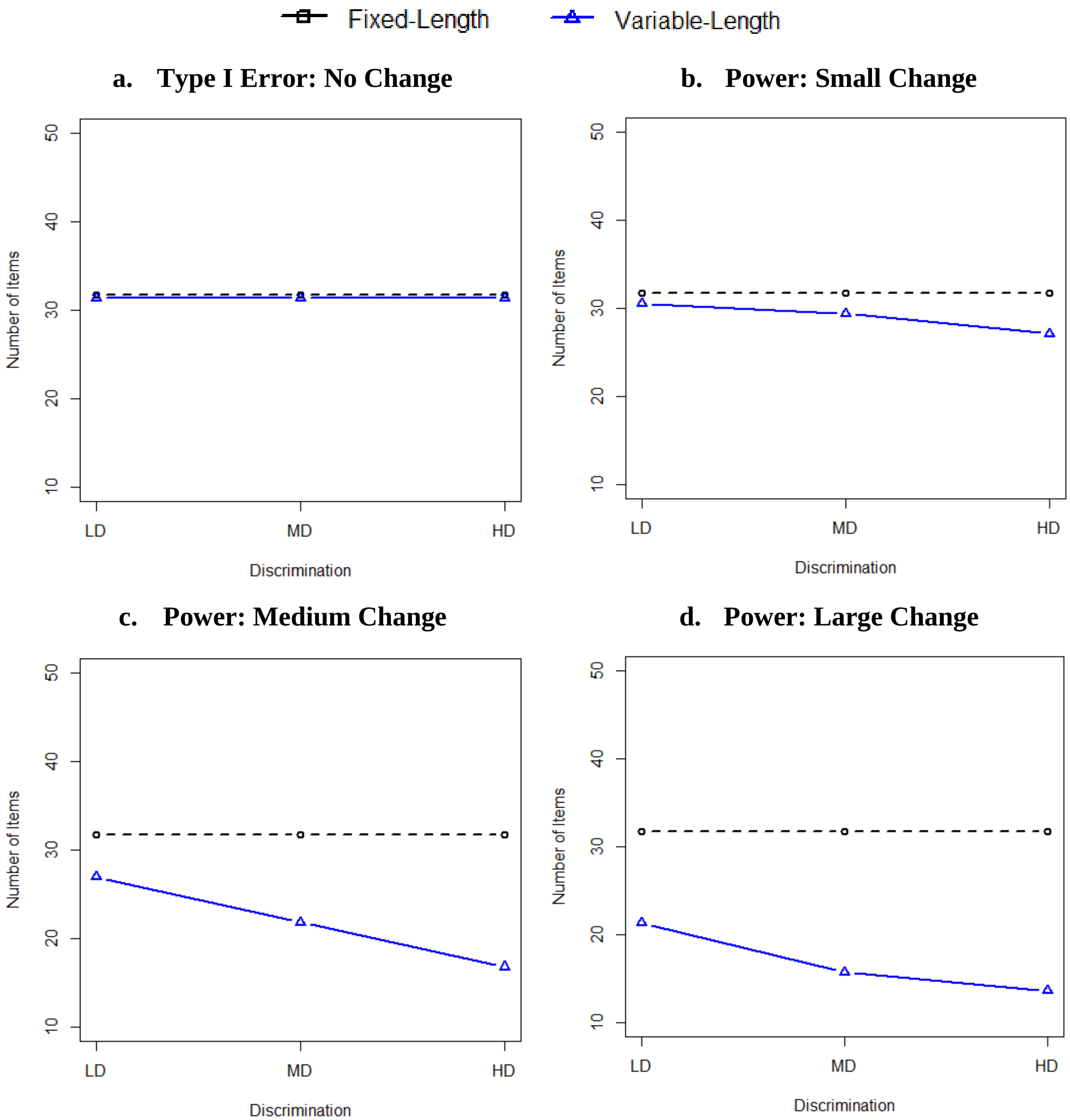


Figure 9. Mean Number of Items Used Conditional on Discrimination for Fixed- and Variable-Length AMC



Effect of Item Selection Method

The effect of the item selection method on Type I error and power was small. The main effect of the item selection method, and the interactions between item selection and θ, bank information, discrimination, item selection, and test length accounted for less than 1% of variance in Type I error and power for both fixed-length and variable-length AMCs. The interaction between item selection and test statistic on Type I error had observed $\eta^2 = 3.59\%$ for fixed-length AMC, and 0.35% for variable-length AMC (Tables 2). KL-F and KL-M item selection methods had slightly higher Type I error and power than the FI method for both fixed- and variable-length AMC, but the differences were observed only in the second or third decimal places. Marginal means and SDs are summarized in Appendix Table A15.

Effect of Test Length

Test length had a small main effect on Type I error: observed η^2 was 0.54% for fixed-length AMC and 1.03% for variable-length AMC (Table 2). The main effects for test length on power were larger for all level of change conditions. The main effect of test length had observed η^2 for power of 26.23% for fixed-length AMC and 31.55 % for variable-length AMC under very small change (Table 3). The observed η^2 for power for the main effect of test length was 32.30% for fixed-length AMC and 34.12 % for variable-length AMC under small change (Table 4). The main effect of test length accounted for 36.07% of variance in power for fixed-length AMC, and 37.15% of variance for variable-length AMC under medium change (Table 5). Under large change, the main effect of test length on power had observed $\eta^2 = 30.54\%$ for fixed-length AMC, and 31.09% for variable-length AMC (Table 6).

Figure 10 shows Type I error and power for test length. In Figure 10a, a slightly better adherence to the nominal error rate was observed as test length increased. Marginal means of Type I error for each test length were 0.052, 0.051 and 0.050 for 15-item, 30-item and 50-item fixed-length AMC, respectively. Variable-length AMC had slightly lower Type I error than fixed-length AMC for 15-item and 30-item tests: observed marginal means of Type I error for variable-length AMC were 0.047, 0.049 and 0.050 for the 15-item, 30-item, and 50-item maximum test lengths.

For both fixed- and variable-length AMC, higher level of power was observed as test length increased, with a larger increase between 15-item and 30-item tests than between 30-item and 50-item tests (Figure 10b – 10d). Under small change, power increased from 0.161 to 0.406 for fixed-length AMC, and from 0.150 to 0.393 for variable-length AMC on average as test length increased from 15 items to 50 items (Figure 10b). Power increased from 0.458 to 0.830 for fixed-length AMC, and from 0.443 to 0.823 for variable-length AMC as test length increased from 15 items to 50 items with medium true change (Figure 10c). Under large change, marginal means of power increased from 0.701 to 0.959 for fixed-length, and from 0.691 to 0.956 for variable-length as test length increased from 15 items to 50 items (Figure 10d). Although not presented in Figure 10, similar but small power was observed under very small change ($\Delta\theta = 0.25$): power increased from 0.076 to 0.142 for fixed-length AMC and from 0.069 to 0.136 for variable-length AMC as test length increased from 15 items to 50 items. Marginal means and SDs of Type I error and power conditional on test length are summarized in Appendix Table A16.

Figure 11 shows that variable-length AMC used smaller number of items on average than fixed-length AMC of the same test length, but yielded similar levels of Type I error and power. Larger saving in the number of items was observed for tests with longer length. Under no change, variable-length used an average of 14.93 items for the maximum 15-item test, 29.70 items for the maximum 30-item test, and 49.38 items for the maximum 50-item test on average (Figure 11a).

The average number of items used decreased under larger true change conditions. Under small change, variable-length AMC used 14.75 items (1.66% saved), 28.09 items (6.38% saved), and 44.12 items (11.76% saved) for maximum 15-item, 30-item and 50-item tests, re spectively (Figure 11b). Variable-length AMC with medium change used 14.01 items (6.60% saved), 23.30 items (22.65% saved), and 29.21 items (41.58% saved) for maximum 15-item, 30-item and 50-item tests, respectively (Figure 11c). Under large change, variable-length AMC used 13.06 items (12.94% saved), 17.42 items (41.93% saved) and 20.06 items (59.88% saved) for maximum 15-item, 30-item and 50-item tests (Figure 11d). Marginal means and SDs of the number of items used conditional on the maximum test length are summarized in Appendix Table A17.

Figure 10. Effect of Test Length on Type I Error and Power for Fixed- and Variable-Length AMC

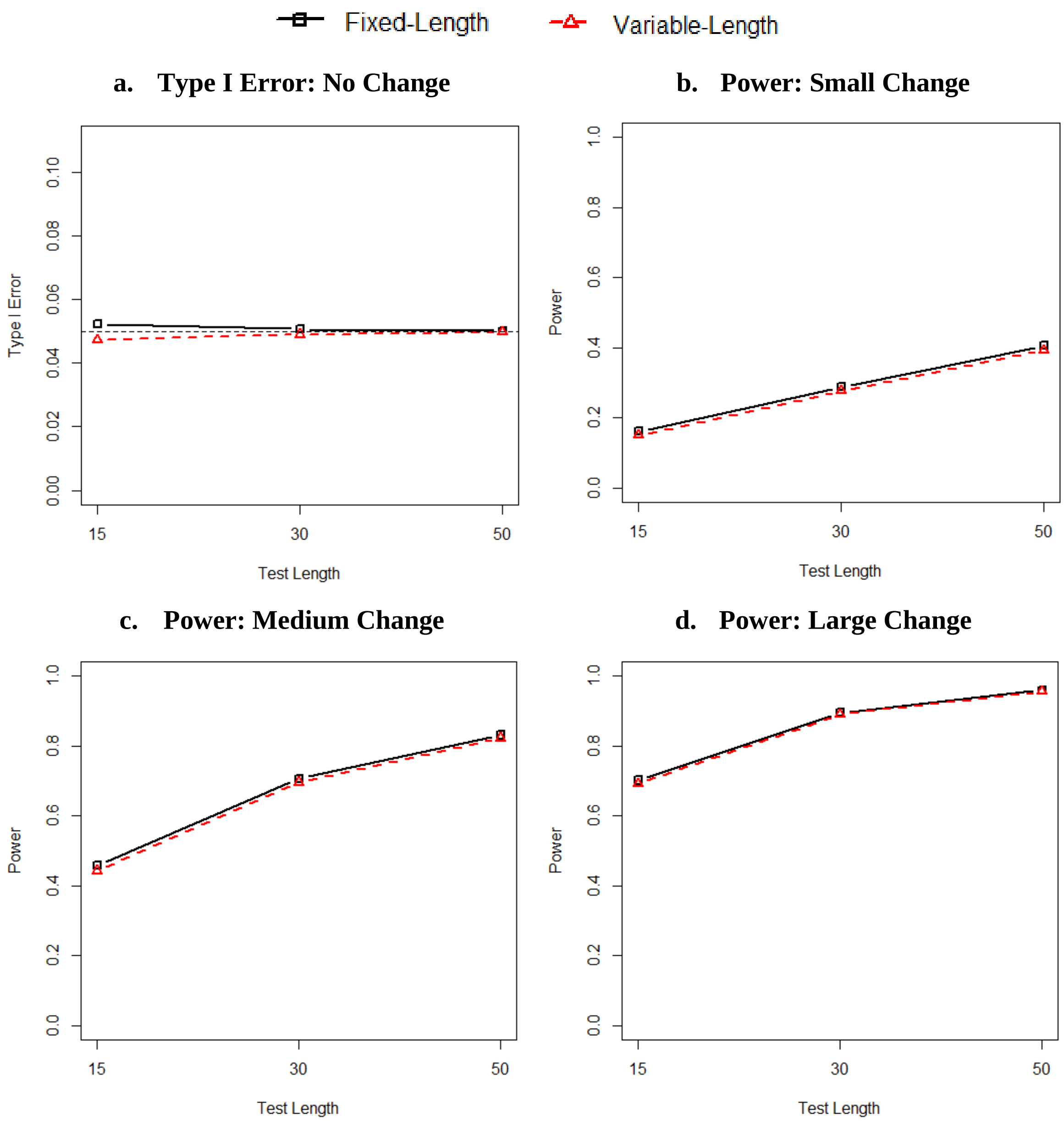
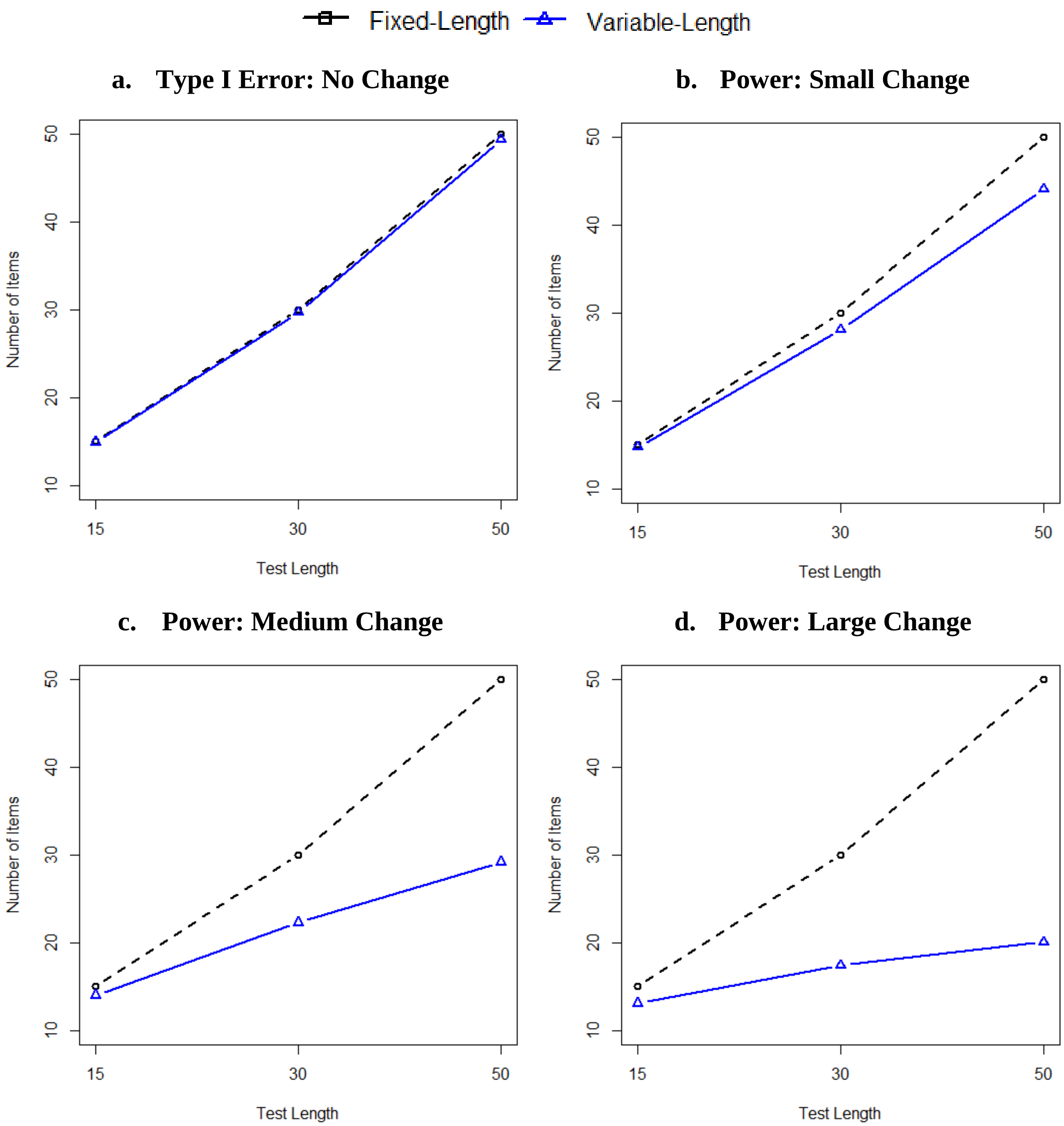


Figure 11. Mean Number of Items Used Conditional on Test Length for Fixed- and Variable-Length AMC



The interaction between test length and discrimination also accounted for greater than 5% of variance for power under very small, small, and large true change. The observed η^2 for power under very small change was 7.29% for fixed-length AMC and 8.2% for variable-length AMC ; power under small change was 6.76% for fixed-length AMC and 7.46% for variable-length AMC; and observed η^2 for power under large change was 8.86% for fixed-length AMC and 8.34% for variable-length AMC. Under medium change, the interaction between test length and discrimination had η^2 smaller than 2% for both fixed- and variable-length AMC.

Figure 12 displays the interaction plot of test length and discrimination on Type I error and power. Test lengths are depicted on the x-axis and discrimination conditions are depicted by different colors of black, blue, and red, representing low, moderate and high discrimination conditions, respectively. As shown in the small η^2 for the main effect of test length and the interaction between test length and discrimination, negligible differences in Type I error were observed across test lengths or across test length × discrimination combinations for both fixed-length and variable-length AMC (Figure 12a, 12e). Power increased in longer tests and higher discriminations; however, the increase rate between test lengths was smaller under larger change conditions (Figures 12b – d, 12f – h).

Under small change, the difference in power between discrimination conditions increased as test length increased (Figure 12b, 12f). For the 15-item test, power for HD was higher than LD by 0.150 for fixed-length AMC and by 0.134 for variable-length AMC. For the 50-item test, power for HD was higher than LD by 0.422 for fixed-length and by 0.410 for variable-length AMC. Larger increase in power was observed for higher discrimination as test length increased. For LD, power increased from 0.093 to 0.204 for fixed-length, and from 0.088 to 0.198 for variable-length as test length increased from 15 items to 50 items. For MD, power increased from 0.149 to 0.386 for fixed-length, and from 0.139 to 0.373 for variable-length as test length increased from 15 items to 50 items. For HD, power increased from 0.243 to 0.626 for fixed-length, and from 0.222 to 0.608 for variable-length as test length increased from 15 items to 50 items.

Under large true change, increase in power for longer test was smaller for higher discrimination conditions (Figures 12d, 12h). For LD, power increased from 0.465 to 0.905 for fixed-length, and from 0.456 to 0.899 for variable-length as test length increased from 15 items to 50 items. For MD, power increased from 0.743 to 0.985 for fixed-length, and from 0.732 to 0.984 for variable-length as test length increased from 15 items to 50 items. For HD, power increased from 0.895 to 0.987 for fixed-length, and from 0.885 to 0.985 for variable-length as test length increased from 15 items to 50 items. Differences between the discrimination conditions decreased as test length increased. For fixed-length AMC, differences in power between HD and LD were 0.429 for the 15-item test, 0.217 for the 30-item test, and 0.081 for the 50-item test. For variable-length AMC, HD had higher power than LD by 0.429 for the 15-item test, 0.223 for the 30-item test, and 0.085 for the 50-item test.

A similar pattern of power was observed under very small change, although not presented in Figure 12. For LD, power increased from 0.057 to 0.085 for fixed-length, and from 0.054 to 0.083 for variable-length as test length increased from 15 items to 50 items. For MD, power increased from 0.073to 0.131 for fixed-length, and from 0.067 to 0.127 for variable-length as test length increased from 15 items to 50 items. For HD, power increased from 0.098 to 0.210 for fixed-length, and from 0.087 to 0.198 for variable-length as test length increased from 15 items to 50 items. Means and standard deviations for Type I error and power conditional on test length and discrimination conditions are summarized in Appendix Table A18.

Figure 13 shows the number of items used for variable-length AMC compared to that of fixed-length AMC of the same test length × discrimination condition. Variable-length AMC used almost identical number of items as fixed-length AMC, with no difference between three discrimination conditions under no change (Figure 13a). In Figures 13b – 13d, variable-length AMC used smaller numbers of items as the amount of change increased, and larger saving was observed for higher discrimination conditions within the same maximum test length. Differences between discrimination conditions were larger for tests with longer length. Under small change, the number of items used with HD and LD ranged from 14.63 to 14.85 items (2.47% to 1% saved) for 15-item tests; from 26.80 to 29.12 items (10.67% to 2.93% saved) for 30-item tests; and from 39.84 to 47.56 items (20.32% to 4.88% saved) for 50-item tests (Figure 13b). Under medium change, the number of items used with HD and LD ranged from 13.39 to 14.55 items (10.73% to 3.02% saved) for 15-item tests; from 17.66 to 26.62 items (41.12% to 11.28% saved) for 30-item tests; and from 19.30 to 39.65 items (61.40% to 20.70% saved) for 50-item tests (Figure 13c). Under large change, the number of items used with HD and LD ranged from 12.18 to 13.96 items (18.77% to 6.91% saved) for 15-item tests; from 13.72 to 22.12 items (54.27% to 26.28% saved) for 30-item tests; and from 14.83 to 27.76 items (70.34% to 44.47% saved) for 50-item tests (Figure 13d). Marginal means and SDs of the number of items used conditional on the maximum test length × discrimination are also summarized in Appendix Table A19.

Figure 12. Effect of Test Length on Type I Error and Power Conditional on Discrimination for Fixed- and Variable-Length AMC

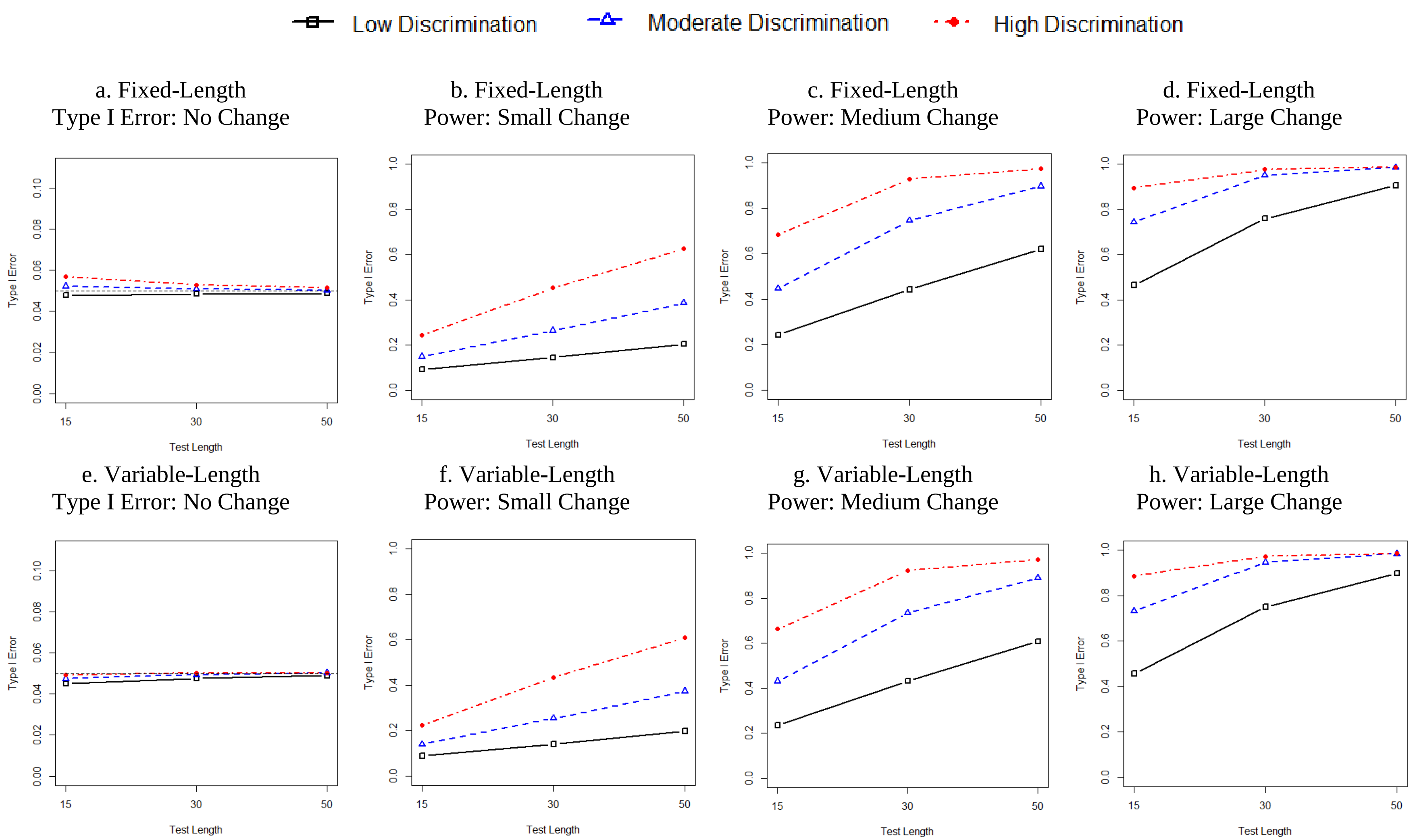
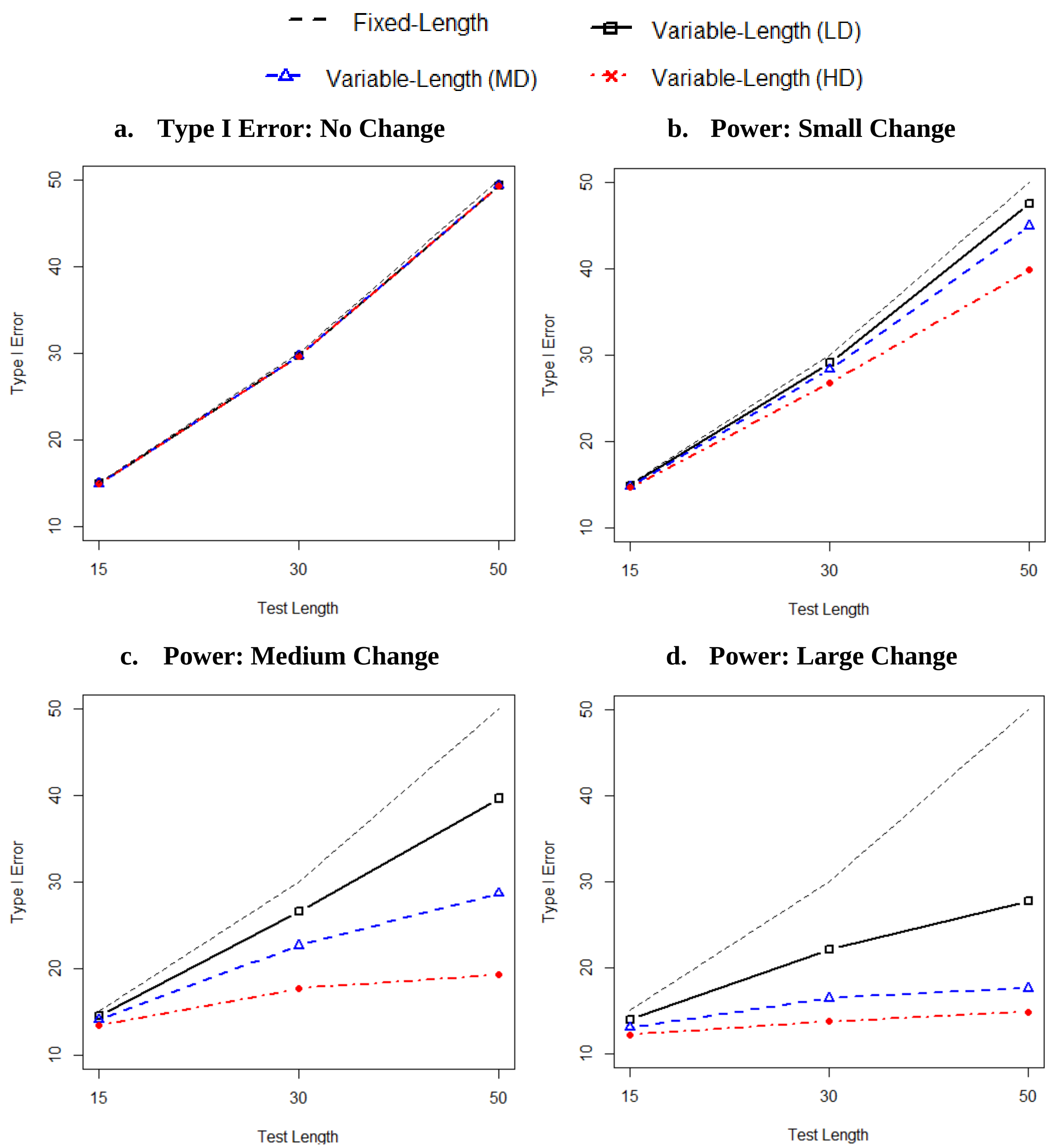


Figure 13. Mean Number of Items Used Conditional on Test Length and Discrimination for Fixed- and Variable-Length AMC



Effect of Test Statistic

The main effects for test statistic on Type I error had observed η^2 = 24.73% for fixed-length AMC and 5.12% for variable-length AMC. The observed η^2 for the main effects of test statistic on power was 8.41% for fixed-length AMC and 3.33% for variable-length AMC under very small change. The main effects for test statistic on power had observed η^2 around 2% for fixed-length AMC and around 1% for variable-length AMC for small to large change conditions.

Figure 14 shows marginal means of Type I error and power for each statistic. Figure 14a shows that larger differences in Type I error were observed for fixed-length than variable-length AMC. For fixed-length AMC, Z and KLD-U had Type I error of 0.05 on average, while the LR and ST statistics had higher Type I error of 0.056 and 0.059, respectively, and the KLD-N statistic had lower Type I error rate of 0.041. For variable-length AMC, marginal means of Type I error were 0.05 for the Z, LR, and ST statistics; 0.049 for the KLD-U; and 0.044 for the KLD-N statistic. Figures 14b – 14d show that higher power was observed for the test statistics in the order of ST, LR, Z, KLD-U and KLD-N, however the effect sizes for test statistic on power were smaller than 5%. Marginal means and SDs for Type I error for each statistic are summarized in Appendix Table A20.

Variable-length AMCs used smaller numbers of items than fixed-length AMC. Figure 15 shows that the five test statistics used similar numbers of items, with LR using slightly less items and KLD-N using slightly more items on average under medium and large change. At the marginal level, fixed-length AMC used 31.67 items on average for each statistic. Variable-length AMC used 31.30 items on average for Z, LR and ST; and 31.39 items on average for KLD-U and KLD-N under no change (Figure 15a). Under small change, the minimum length of 28.53 items (9.92% saved) was observed for LR and the maximum length of 29.41 items (7.15% saved) was observed for KLD-N (Figure 15b). Under medium change in Figure 15c, LR had the minimum average test length of 20.85 (34.17% saved) and KLD-N had the maximum average test length of 22.82 (27.95% saved). Under large change, the LR statistic had the shortest average length of 15.69 items (50.47% saved) and KLD-N had the longest average test length of 17.84 items (43.67% saved), as shown in Figure 15d. Means and SDs for the number of items used are summarized in Appendix Table A21.

Figure 14. Effect of Test Statistic on Type I Error and Power for Fixed- and Variable-Length AMC

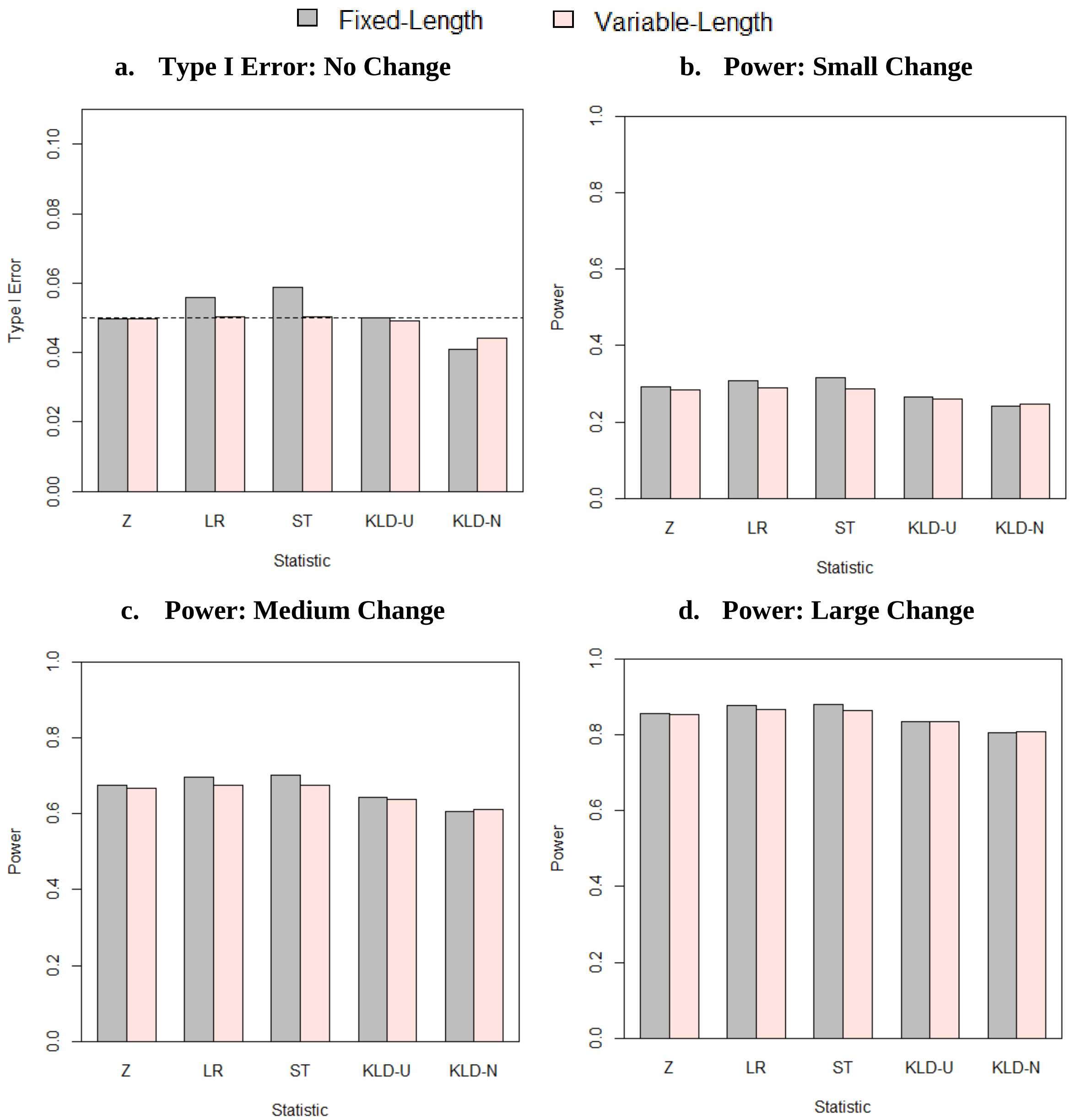
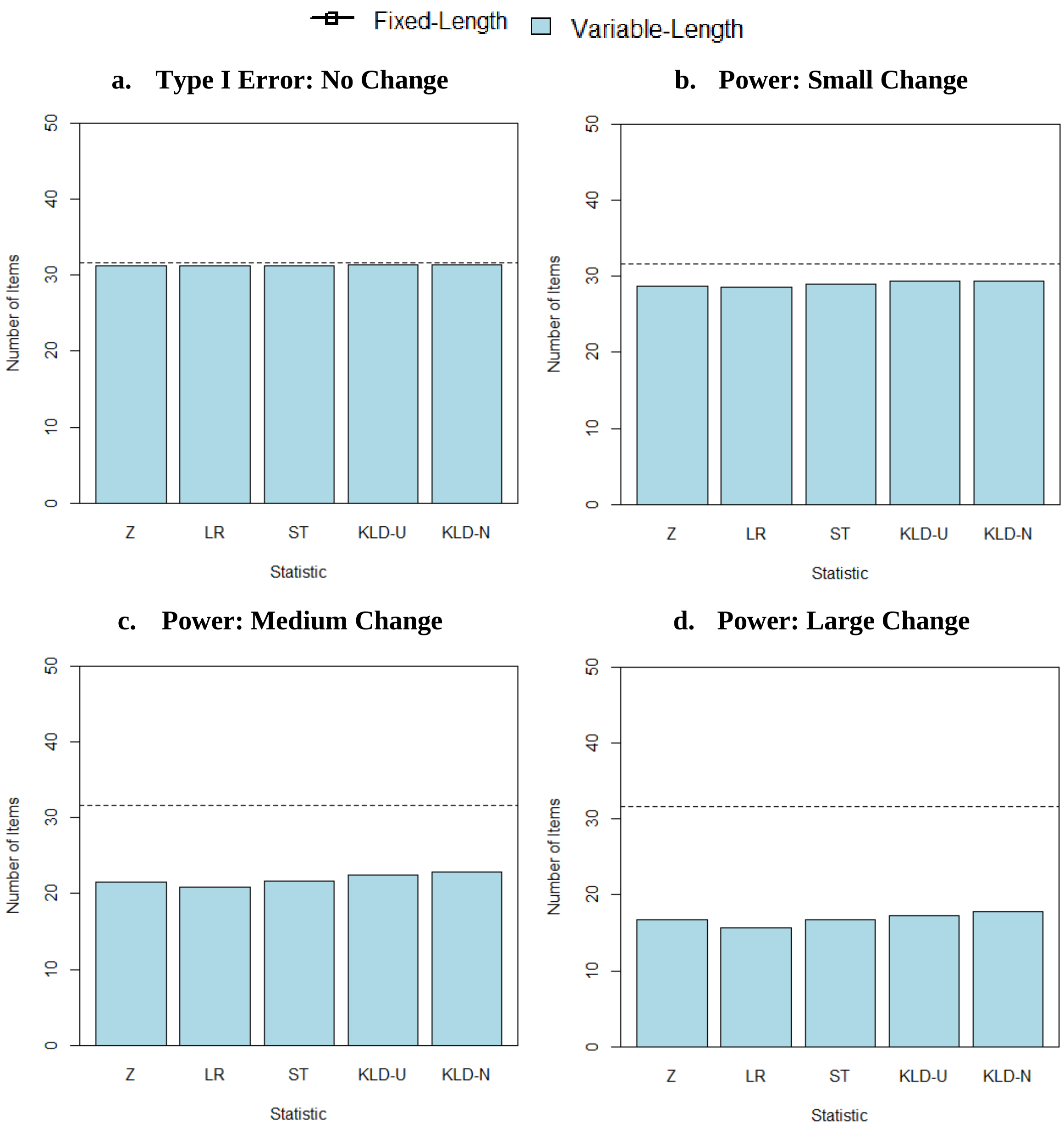


Figure 15. Mean Number of Items Used Conditional on Test Statistic for Fixed- and Variable-Length AMC



In the ANOVA, the interaction between test statistic and θ had observed η^2 for Type I error of 3.9% for fixed-length and 8.61% for variable-length. For fixed-length, Type I error rate for the Z statistic slightly increased as θ deviated from 0.0: average Type I error ranged between 0.046 and 0.051 at $-1 \le \theta \le 1$; 0.059 at $\theta = -2$, and 0.054 at $\theta = 2$ (Figure 16a). A slight decrease in Type I error was observed for LR and ST statistics as θ increased: average Type I error decreased from 0.059 to 0.054 for LR, and from 0.061 to 0.056 for the ST statistic as θ increased from $\theta = -2$ to $\theta = 2$. A better adherence to the nominal error rate was observed for the KLD-U statistic: observed Type I error ranged between .049 and 0.053 at $-1 \le \theta \le 1$; 0.045 at $\theta = -2$, and 0.053at $\theta = 2$. The Type I error for the KLD-N statistic decreased from 0.047 to 0.032 as θ increased from $\theta = -2$ to $\theta = 2$.

Figure 16e shows that variable-length AMC had better adherence to the nominal level than fixed-length AMC. Type I error for the Z statistic ranged between 0.044 and 0.05 at $-1 \le \theta \le 1$, 0.061 at $\theta = -2$, and 0.055 at $\theta = 2$. Average Type I error rates ranged between 0.047 and 0.053 for LR; and between 0.046 and 0.052 for ST, both slightly decreased from $\theta = -2$ to $\theta = 2$. The KLD-U statistic had average Type I error that ranged between 0.049 and 0.052 at $-1 \le \theta \le 1$, 0.046 at $\theta = -2$, and 0.044 at $\theta = 2$.

The interactions between test statistic and θ on power had observed η^2 smaller than 2% for all levels of change for both fixed-length and variable-length AMC. For both fixed- and variable-length AMCs, power increased as the amount of true change increased (Figures 16b – 16d, 16f – 16h). A decrease in power was observed as θ deviated from 0.0, with a larger decrease in power at the high θ range than the low θ range for all test statistics. The KLD-U and KLD-N statistics had lower power than Z, LR and ST with a larger decrease in power as θ increased. Marginal means and SDs of Type I error and power for test statistic × θ for fixed- and variable-length AMC are summarized in Appendix Tables A22 – A23.

Figure 17 shows the mean number of items used for variable-length AMC for test statistic × θ conditions. Fixed-length AMC used 31.67 items on average constantly across θ. Under no change, trivial differences were found between fixed- and variable-length AMC, and between the test statistics: Z, LR and LM used 31.30 items (1.17% saved), and KLD-U and KLD-N used 31.39 items (0.88% saved) on average across θ (Figure 17a). In Figures 17b – 17d, the LR statistic generally used less items and the KLD-N statistic generally used more items than the other statistics. More items were used at higher θ ranges for LR, ST, KLD-U, and KLD-N. Z used more items in different θ ranges depending on the amount of change.

Figure 17b shows that Z under small true change used slightly less items as θ increased with minimum 28.04 items at $\theta = 1.5$ (11.46% saved) and maximum 29.41 items at $\theta = -2$ (7.12% saved). The LR, ST, and KLD-U statistics used slightly more items at the high θ range. The LR statistic used a minimum of 28.06 items at $\theta = 0$ (11.39% saved) and maximum 29.61 items at $\theta = 2$ (6.51% saved). The ST statistic used a minimum of 28.39 items at $\theta = 0$ (10.36% saved) and maximum 30.09 items at $\theta = 2$ (5% saved). The KLD-U statistic used a minimum of 28.95 items at $\theta = 0.5$ (8.59% saved) and maximum 30.44 items at $\theta = 2$ (3.87% saved). The KLD-N statistic used slightly more items as θ increased, with minimum 28.67 items at $\theta = -1.5$ (9.48% saved) and maximum 30.87 items at $\theta = 2$ (2.52% saved).

More items were generally used at the high θ range under medium change, with the largest saving observed for LR and the smallest saving observed for KLD-N (Figure 17c). The Z statistic used a minimum of 19.94 items at $\theta = 1.0$ (37.02% saved) and used more items as θ deviated from 1.0, with the maximum 24 items at $\theta = 2$ (23.22% saved). The LR statistic used a minimum of 19.59 items at $\theta = 0$ (38.14% saved) and a maximum of 24.86 items at $\theta = 2$ (21.52% saved). The ST statistic used a minimum of 20.22 items at $\theta = -0.5$ (36.14 % saved) and a maximum of 25.79 items at $\theta = 2$ (18.58% saved). The KLD-U statistic used a minimum of

21.01 items at $\theta = 0$ (33.66% saved) and a maximum of 27.24 items at $\theta = 2$ (13.98 % saved). The KLD-N statistic used more items as θ increased, with a minimum of 20.75 items at $\theta = -1.5$ (34.47 % saved) and a maximum of 28.28 items at $\theta = 2$ (10.72% saved).

Similar patterns were observed under large change, with the largest saving observed for LR and the smallest saving observed for KLD-N (Figure 17d). The Z statistic used a minimum of 14.56 items at $\theta = 0.5$ (54.02% saved) and used more items as θ deviated from 1.0, with the maximum 21.42 items at $\theta = 2$ (32.38 % saved). The LR statistic used a minimum of 14.29 items at $\theta = -0.5$ (54.86% saved) and a maximum of 21.38 items at $\theta = 2$ (32.48% saved). The ST statistic used a minimum of 15.18 items at $\theta = 0$ (52.07 % saved) and a maximum of 22.25 items at $\theta = 2$ (29.76 % saved). The KLD-U statistic used a minimum of 15.28 items at $\theta = 0$ (51.76 % saved) and a maximum of 24.93 items at $\theta = 2$ (21.27 % saved). The KLD-N statistic used a minimum of 15.12 items at $\theta = -1.5$ (52.26 % saved) and used more items as θ increased, with a maximum of 26.16 items at $\theta = 2$ (17.39 % saved). Means and SDs of the number of items used with variable-length AMC for test statistic × θ are summarized in Appendix Table A24. Average Type I error and power for each simulation condition at each θ level are also summarized in Appendix Tables A25 – A96, and at aggregated θ levels in Appendix Tables A97 – A120.

Figure 16. Effect of Test Statistic on Type I Error and Power Conditional on θ for Fixed- and Variable-Length AMC

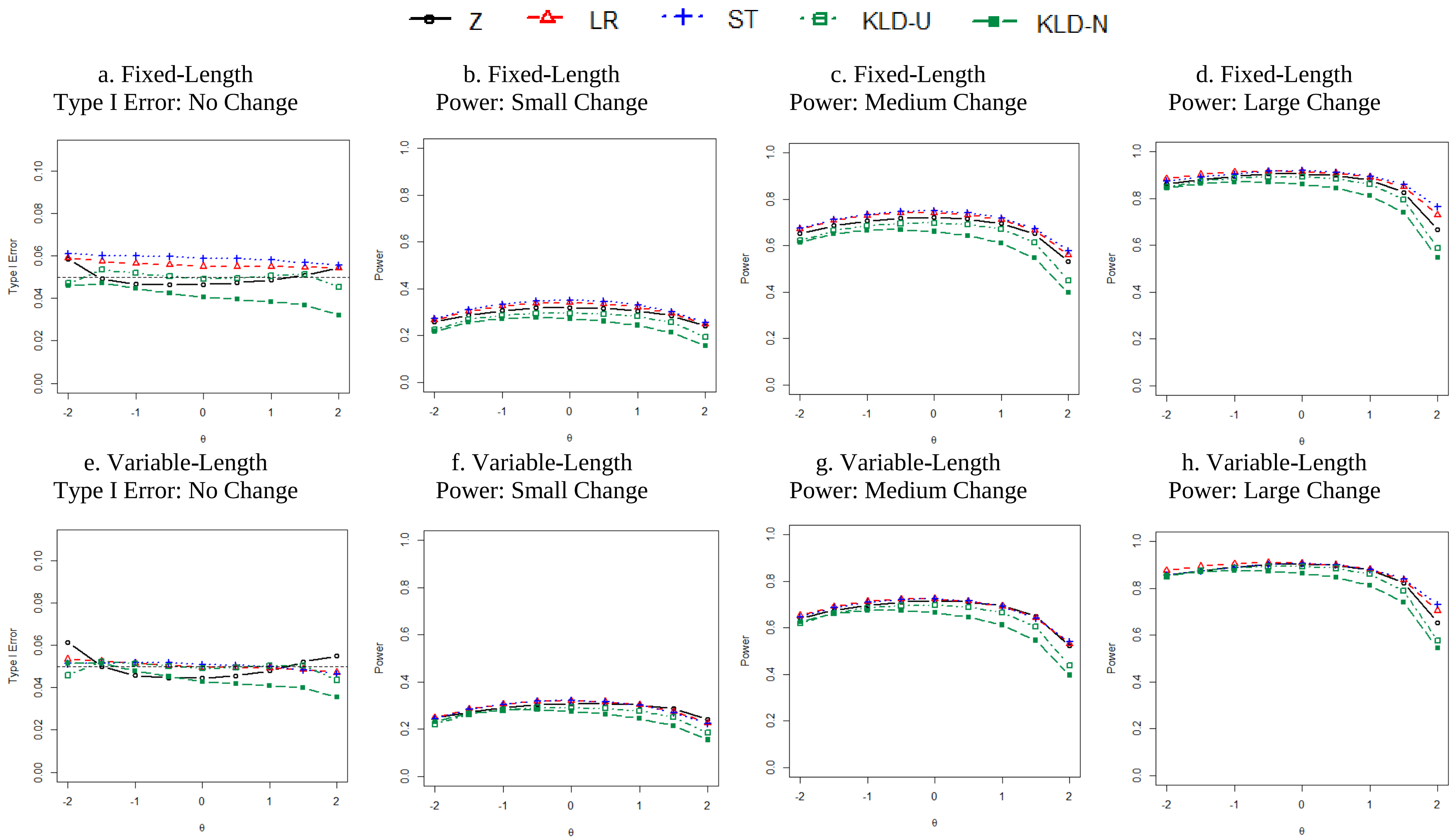
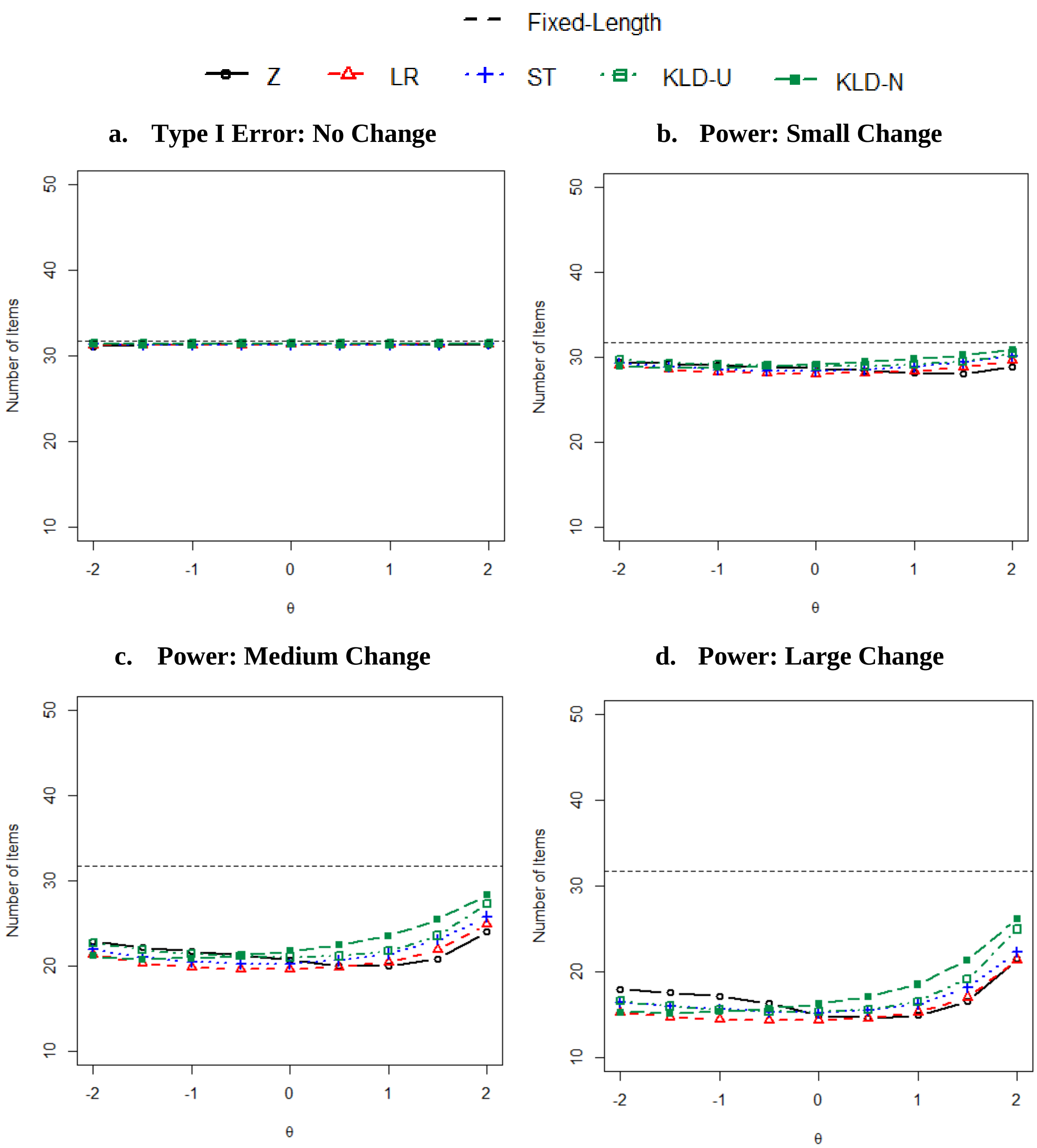


Figure 17. Mean Number of Items Used Conditional on Test Statistic and θ for Fixed- and Variable-Length AMC



Chapter 4

DISCUSSION AND CONCLUSIONS

As an extension of Finkelman et al. (2010), this study introduced three new test statistics: ST, KLD-U and KLD-N. The new item selection method of modified KL information was also compared to the two existing item selection methods of Fisher information and KL information in Finkelman's study. The effect of item discrimination, test length, bank information shape, and bank size were further examined to better attribute findings in Finkelman et al. (2010). In addition, the study extended to a variable-length termination for all combinations of item selection methods and test statistics.

Major Effects on Type I Error and Power

Item discrimination. Type I error slightly increased as discrimination increased: Type I error increased from 0.048 to 0.054 for fixed-length, and from 0.047 to 0.050 for variable-length AMC. However, item discrimination had a small effect on Type I error: observed η^2 for discrimination on Type I error was smaller than 5% for both fixed- and variable-length AMC. For fixed-length AMC, moderate discrimination had a better adherence to the nominal level, and for variable-length AMC, high discrimination ($\bar{a} = 1.5$) yielded a better adherence to the nominal level. However, differences between the discrimination conditions were found in the third decimal places, especially smaller for variable-length AMC.

Larger effects of discrimination on power were observed: observed η^2 for discrimination for power ranged between 26.98% and 47.23% for fixed-length AMC and between 26.89% and 46.01% for variable-length AMC. Power increased in higher discrimination conditions on average for both fixed- and variable-length AMC. High discrimination was related to an increase in the bank information. The results suggested that better quality of an AMC bank is related to better performance of AMC in detecting individual change, while maintaining Type I error around the nominal level.

Under large change ($\Delta\theta = 1.5$), differences in power between discrimination conditions decreased due to a ceiling effect. High discrimination had mean observed power of 0.862 under medium true change ($\Delta\theta = 1.0$), and power increased to 0.952 under large change of $\Delta\theta = 1.5$ (i.e., power increased by 0.09). Since observed power under medium true change was high, there was not much gain in power for the high discrimination condition when the amount of change increased in Time 2. For the low discrimination condition ($\bar{a} = 0.6$), on the other hand, power increased from 0.696 under medium change to 0.893 (i.e., power increased by 0.275) under large true change. The results imply that the benefits of using a high-quality AMC bank (i.e., high discrimination) were larger when change in θ between two occasions was smaller. However, a medium discriminating bank ($\bar{a} = 1.0$) yielded comparable performance to a high discriminating bank under large true change.

Test length. A slightly better adherence to the nominal error rate was observed as test length increased; however, the effect sizes for test length on Type I error were trivial for both fixed- ($\eta^2 = 0.54$) and variable-length AMC ($\eta^2 = 1.03$). Longer tests yielded higher power on average, and test length accounted for greater than 30% of variance in power under all change conditions for both fixed- and variable-length AMC. These results also suggested that the power of hypothesis testing in AMC was related to the magnitude of the bank information function, since longer tests had higher information functions.

The increase in power as a function of test length varied depending on the level of discrimination. Under very small ($\Delta\theta = 0.25$) and small change ($\Delta\theta = 0.5$), power increased at a constant rate as test length increased for all discrimination conditions. Under medium and large true change, differences in power between test lengths were also smaller for higher discrimination conditions. Also, the increase in power was larger between 15 items and 30 items than between 30 items and 50 items. This implied that the advantage of administering more items in detecting individual change decreased when change between the two occasions was medium to large, especially with a high discriminating bank.

Test statistic. Differences in Type I errors between the test statistics were observed, with larger effects of test statistic for fixed-length AMC ($\eta^2 = 24.75\%$) than variable-length AMC ($\eta^2 = 5.12\%$). For fixed-length AMC, Z and KLD-U had the best adherence to the nominal level with the average Type I error equivalent to 0.05, and KLD-N had the largest deviation from the nominal level with the average Type I error equivalent to 0.041. LR and ST had Type I errors that were higher than the nominal level, 0.056 and 0.059, respectively. For variable-length AMC, only KLD-N showed a deviation from the nominal level, with the average Type I error equivalent to 0.044. The Z, LR, and ST statistics had average Type I error of 0.05 and KLD-U had average Type I error rate of 0.049.

Variable-length AMC had a better adherence to the nominal level on average than fixed-length AMC as well as smaller differences between the statistics. Two critical values C_1 and C_2 were used, which were greater than or equal to the nominal level critical value, to have the proportion of rejecting the null hypothesis to be about 0.023 at the interim stage and 0.027 at the final stage of variable-length AMC. This correction for familywise error did not yield large differences in observed Type I error for Z and KLD-U, which already had a good adherence to the nominal level in fixed-length AMC. For LR and ST, which exceeded the nominal level in fixed-length AMC, the use of the two critical values in variable-length AMC reduced the observed Type I error rate to 0.05. For the KLD-N statistics, the variable-length AMC had an increase in observed Type I error from 0.041 to 0.044, which was closer to the nominal level as a result of using the two critical values. Similar patterns were observed for all item selection methods.

The KLD-N statistic showed the smallest correction to the nominal level in variable-length AMC because of its low Type I error in fixed-length AMC. The use of the two critical values in variable-length AMC was intended to reduce to the nominal level inflated Type I error due to multiple comparisons. However, KLD-N had observed Type I error at the final stage lower than 0.027 after selecting C_1 to have interim stage Type I error to be 0.023. Therefore, C_2 was selected at the nominal level (i.e., the minimum possible value) for most conditions, and observed Type I errors were lower than 0.05 as a result. See Appendix Tables A2 – A7 for C_1 and C_2.

The interaction between test statistic and θ on Type I error further clarified differences between the test statistics ($\eta^2 = 3.9\%$ for fixed-length and $\eta^2 = 8.61\%$ for variable-length AMC). For fixed-length AMC, the KLD-U and Z statistics showed small deviation from the nominal level across θ with a slightly larger deviation at extreme θ ranges. The KLD-U had Type I errors about 0.001 above 0.05 at $-1.5 \le \theta \le 1.5$ and a slight decrease at extreme θ ranges: 0.003 below 0.05 at $\theta = -2$ and 0.005 below 0.05 at $\theta = 2$. The Z statistic had Type I errors about 0.002 lower than 0.05 at $-1.5 \le \theta \le 1.5$ and a slight increase at extreme θ ranges: 0.009 higher than 0.05 at $\theta = -2$ and 0.004 higher at $\theta = 2$. The LR and ST statistics had Type I errors constantly above the nominal level by 0.006 or 0.009, respectively. KLD-N had a relatively better performance at $\theta \le -1$ with 0.004 lower than the nominal level but had a increase in deviation as θ increased with a maximum 0.018 lower than the nominal level at $\theta = 2$.

For variable-length AMC, deviations from the nominal level were generally reduced except for the Z statistic. KLD-U had Type I error slightly closer to the nominal level at $-1.5 \le \theta \le 1.5$ and a slightly larger deviation at extreme θ ranges compared to those for the fixed-length AMC. The LR and ST statistics had a decrease in Type I error with 0.001 and 0.002 above the nominal level on average, respectively. KLD-N also had a relatively better performance at $\theta \le -1$ with 0.002 lower than the nominal level and had larger deviation as θ increased to the maximum deviance of 0.014 below the nominal level at $\theta = 2$. The Z statistic, on the other hand, showed larger deviation from the nominal level across all θ levels: average 0.003 deviation below the nominal level was observed at $-1.5 \le \theta \le 1.5$, Type I error was 0.011 above the nominal level at $\theta = -2$ and 0.005 above the nominal level at $\theta = 2$.

Differences in deviation were due to the way variable-length AMC corrected for familywise error. Critical values for the variable-length AMC were determined to have the overall Type I error close to 0.05, rather than controlling Type I error at each θ level. Therefore, a better adherence to the nominal level for variable-length AMC across θ was observed for statistics if their Type I errors for fixed-length AMC were constantly higher or lower than the nominal level (e.g., LR, ST, and KLD-N). However, if Type I errors were lower than the nominal level at one θ range and higher than the nominal level at another θ range, such as Z or KLD-U, variable-length AMC could yield larger deviation from the nominal level for some θ levels while achieving a better adherence to the nominal level for other θ levels.

Five test statistics yielded different levels of power functions; however, the effect sizes were small for the main effect ($0.95\% \le \eta^2 \le 8.41\%$) and the interactions ($0\% \le \eta^2 \le 1.23\%$) for all change conditions. At the marginal level, five statistics had similar power at $\theta = -2$ and larger differences in power were observed as θ increased. The LR and ST statistics had slightly higher power than the other statistics and smaller decrease in power at high θ ranges. For the Z, LR and ST statistics, a larger decrease in power was observed at $\theta > 1.5$; however, power decreased from $\theta \geq -1.5$ for the KLD-U and KLD-N statistics. Similar patterns were observed across different change conditions. Despite the observed differences in power, differences in Type I error described more meaningful evaluation on the performance of the test statistics due to small effect sizes of the test statistics for power.

KLD with both prior distributions had lower performance than the Z, LR and ST statistics in detecting individual change. KLD with normal prior had Type I error around 0.04 on average and KLD with both prior distributions had lower power than the other three statistics. Results for KLD-N indicated the effect of a normal prior, which regressed the test statistic values toward the mean and led to a smaller rejection rate than the other statistics. Low power for KLD-U and KLD-N implied that the distribution of KLD might not follow a chi-square distribution with one degree of freedom when measurements are taken with AMC. Normal posterior distributions with equal means and variance were assumed to approximate KLD with a chi-square distribution with one degree of freedom (Belov and Armstrong, 2011). These assumptions are justifiable when parallel tests are administered to a population. However, the assumption of equal population distributions in repeated measurement is violated under change conditions, and AMC also administers items with difficulties adapted to the abilities of candidates.

Moderate Effects on Type I Error and Power

***θ*.** The main effects of the θ level on Type I error were small: the observed η^2 was 2.09% for fixed-length and 4.24% for variable-length AMC. Marginal means of Type I error slightly decreased from 0.054 to 0.048 for fixed-length; and from 0.053 to 0.046 for variable-length AMC. The results suggested that AMC did not have noticeable differences in Type I error rate depending on Time 1 θ level in detecting individual change: all θ levels yielded Type I error close to the nominal level, although AMC had a slight better adherence to the nominal level when Time 1 θ was close to 0.0 compared to θs at extreme ranges.

The main effects of θ were larger on power under medium and large change conditions: the observed η^2 for power was about 6% under medium true change and about 14% under large change. Power decreased as θ deviated from 0.0: a larger decrease in power was observed at high θ range than at low θ range. These results imply that the power of AMC is related to the difficulties of the items in the AMC bank. Item banks used in the current study had difficulties that ranged between −4.5 and 4.5. θ values in extreme high or low ranges, e.g, $\theta = -2$ or $\theta = 2$ did not have a sufficient number of items at Time 1 measurement, especially using an item bank with peaked information. When change was made at Time 2, Time 1 low θ values then moved toward the middle, for example from $\theta_1 = -2$ to $\theta_2 = -0.5$, and had more items at Time 2 measurement. A smaller decrease in power at low θ range reflected this better measurement achieved at Time 2. Time 1 high θ values, on the other hand, had smaller numbers of item with appropriate difficulties as larger change was made at Time 2, for example from $\theta_1 = 2$ to $\theta_2 = 3.5$. A large decrease in power at high θ ranges reflected this lack of items with appropriate difficulties for Time 1 simulees with high θ. This lower power for high θ ranges could be a practical issue of applying AMC in detecting change, as items with extreme difficulties might be difficult to develop for some traits.

Bank information. Effects of the distribution of bank information across θ were small for both fixed-length and variable-length AMC: main effects of the bank information on Type I error and power had observed η^2 smaller than 1%. Only the interaction between the bank information and θ on power under large change had η^2 larger than 5%. Consistent with Finkelman et al. (2010), power functions were affected by the shape of the test or bank information function. The power functions decreased as θ deviated from 0 in AMC with a peaked bank in ranges of θ, while constant levels of power were observed for AMC with flat bank information. A large decrease in power at positive extreme levels of θ under large true change in a peaked bank was attributable to lack of test information at Time 2 in that range. The shape of the bank information was related to the number of items available at each difficulty level.

No Effect on Type I Error and Power

Bank size. Effect of the bank size on Type I error and power was negligible. Main effects of the bank size and interactions between the bank size and other factors all had observed η^2 smaller than 1%. Although the 500-item bank had slightly higher power than the 300-item bank for both fixed- and variable-length AMC, the results implied that the improvement obtained by an additional 200 items would not be practically meaningful. The results further led to the question of whether similar performance of AMC could have been obtained with a bank size smaller than 300 items. Since the performance of AMC was affected by the distribution of item difficulties, i.e., bank information shape, a decrease in the bank size might decrease the performance of AMC. However, if a decrease in performance is relatively small, further investigation on an optimal bank size could give a cost efficient way of implementing AMC without sacrificing the performance of detecting individual change.

Item selection. Although the Z statistic combined with KL information item selection was advocated by Finkelman et al. (2010), the results from the current study showed that no item selection method dominated another. The FI method had a slightly better adherence to the nominal level: the marginal mean of Type I error was 0.05 for FI; 0.051 for KL-F; and 0.052 for KL-M. However, the KL-F and KL-M item selection methods had slightly higher power than the FI method on average for both fixed- and variable-length AMC. However, the differences in Type I error and power were in the second or third decimal places. ANOVA analysis further revealed that the main effects for the item selection method had η^2 for Type I error and power smaller than 1% for both fixed- and variable-length AMC. Thus, it can be concluded that the effect of the item selection method had a negligible effect on the performance of AMC under conditions studied, which was consistent with previous findings that differences in item selection methods in CAT disappeared after 10 items had been administered (e.g., Chen & Ankenmann, 2004)

Comparisons to Previous Findings

The current study analyzed the results of the simulations in a series of ANOVAs, which helped highlight which effects were important. In Finkelman et al. (2010), results were presented focusing on describing observed differences between the simulation conditions. This was an informative approach as an initial investigation of hypothesis testing, since no information on how the proposed methods worked under certain circumstances had been available. However, their study did not evaluate the magnitude of the observed difference (i.e., effect size) in any systematic way. Observed differences are negligible if the magnitude of the difference is small. Evaluating effect size is especially meaningful in simulation studies, since differences between the factors are induced by the study design. Thus, reporting effect sizes and interpreting differences based on that helped to focus on the substantive factors as opposed to describing all observed findings.

Consistent with Finkelman et al.'s (2010) advocacy of the *Z* statistic, the results in this study suggested that *Z* generally had the best balance between Type I error and power. The KLD-U statistic had similarly good adherence to the nominal level; however, power was lower than the other statistics. Type I errors for the LR and ST statistics were consistently higher than the nominal level for fixed-length AMC, 0.056 and 0.059, respectively, on average. The main effects for test statistic on Type I error had observed η^2 = 24.73% for fixed-length AMC. The *Z* statistic, however, had an increase in Type I error at $\theta = \pm 2$, with an observed Type I error exceeding 0.1 in the higher discriminating bank. LR and ST, on the other hand, were rarely affected by the discrimination conditions. The results could imply that the performance of the *Z* statistic could be less controllable as a better quality test could reduce the performance of identifying significant individual change. However, the effect sizes of the interaction between test statistic and discrimination were relatively small ($\eta^2 = 3.7\%$ for fixed-length AMC and $\eta^2 = 3.6\%$ for variable-length AMC).

In the current study, modified KL information was introduced as a new item selection method; however, the main effects for item selection method had η^2 for Type I error and power that were smaller than 1% for both fixed- and variable-length AMC. In addition, no one item selection method dominated another at the marginal level, which was inconsistent with the advocacy of KL information in Finkelman et al. (2010). In the current findings, the FI method had a slightly better adherence to the nominal level but the KL-F and KL-M item selection methods had slightly higher power than the FI method, on average. Given the fact that observed differences between item selection methods were negligible, FI would have practical computational advantages over KL information methods. Using Fisher information, an information table can be pre-constructed before the administration of AMC, which will facilitate item selection, instead of calculating information based on current $\hat{\theta}$ at each time of item selection, which is required by the KL method.

The findings for the effect of discrimination and test length further supported the level of power in relation to bank information. Power increased in higher discrimination conditions and with longer test lengths for both fixed- and variable-length AMC. However, differences between 30-item and 50-item AMCs decreased as the amount of change increased. Also, test length had smaller effects on power for high discriminating banks. The current findings suggested several implications for AMC banks: (1) tests need to be longer than 15 items to achieve sufficient levels of power for all discrimination conditions and change conditions; (2) when change between the two occasions is small, it is critical to obtain high quality banks or to administer enough number of items to accurately detect significant individual change; (3) however, when change between the two occasions is medium to large, the effect of discrimination and test lengths decreased so that cost-efficient AMC banks (e.g., medium discrimination with 30-item AMCs) can yield comparable results to banks with higher quality.

In the current study, variable-length AMC was applied to each simulation condition, whereas one AMC procedure, i.e., Z statistic with KL information in 50-item test, was examined in Finkelman et al. (2010) for brevity. Consistent with the previous findings, variable-length showed comparable performance relative to fixed-length AMC with important reductions on test length under many conditions. The effect sizes of factors examined were also similar between fixed-length and variable-length AMC. Variable-length used fewer items than fixed-length AMC and the mean saving increased as the level of change increased: about 8.43% items were saved with small change; 31.01% items were saved on average with the minimum 17.81% saving at $\theta \geq 1.0$ under medium change; and 46.78% items were saved with a minimum 26.65% saving at

$\theta \geq 1.0$ under large change. The results from the current study also implied that the efficiency of variable-length AMC was affected by bank information. More items were administered at high θ levels, especially under larger true change, because few items with target difficulties were available. Tests with low discrimination also used more items than higher discrimination conditions.

The maximum saving in the number of items largely depended on the minimum 10-item length, especially under a large change of $\Delta\theta = 1.5$. SDs of the number of items saved were smaller under large change, which implied that the final test lengths in variable-length AMC tended to be the minimum test length. The minimum test length of 10 items was arbitrarily chosen in the current study. However, increasing test efficiency by lowering the minimum length is not always desirable since it could sacrifice accuracy of final θ estimation and significance decisions. Further investigation would be necessary to determine optimal minimum test length to achieve both efficiency and accuracy of variable-length AMC.

Bias in MLE and Type I Error in the Z Statistic

A post-hoc analysis was conducted to clarify large Type I error rates for the Z statistic at extreme θ ranges. Figure 18 shows Time 1 θ estimates on the x-axis against Time 2 θ estimates on the y-axis for $\theta = -2.0, -1.0, 0.0, 1.0$, and 2.0 under the no-change condition. Five θ levels are displayed in Figure 18 for brevity: the remaining four θ values have plots that are similar to those of $\theta = -1.0, 0.0$, and 1.0. Estimates for simulees who were falsely detected as significant change are colored in red. Time 1 true θ value is indicated by a vertical dashed line and Time 2 true θ value is indicated by a horizontal line in each figure: estimates are closer to the intersection of the two lines if estimates from both time points are precise, and further away from the center as estimates have larger differences from the true value. Results from the 30-item AMC with high discriminating peaked bank are used to demonstrate the findings.

Figure 18 shows that MLE is biased outward in AMC at extreme θ ranges. A larger spread around the true θ is shown in Figure 18a ($\theta = -2$) and 18e ($\theta = 2$) than in figures with less extreme θ values (Figures 18b – 18d). In Figure 18a, more θ estimates are around the lower left corner of the plot, which indicates that biased θs are smaller than $\theta = -2$ at Time 1 or at Time 2 or at both occasions. In Figure 18e, more θ estimates are in the upper right corner of the plot, which indicates that biased θs are larger than $\theta = 2$ at Time 1 or at Time 2 or at both occasions. In Figure 18b – 18d, θ estimates are closely spread around the true θ values.

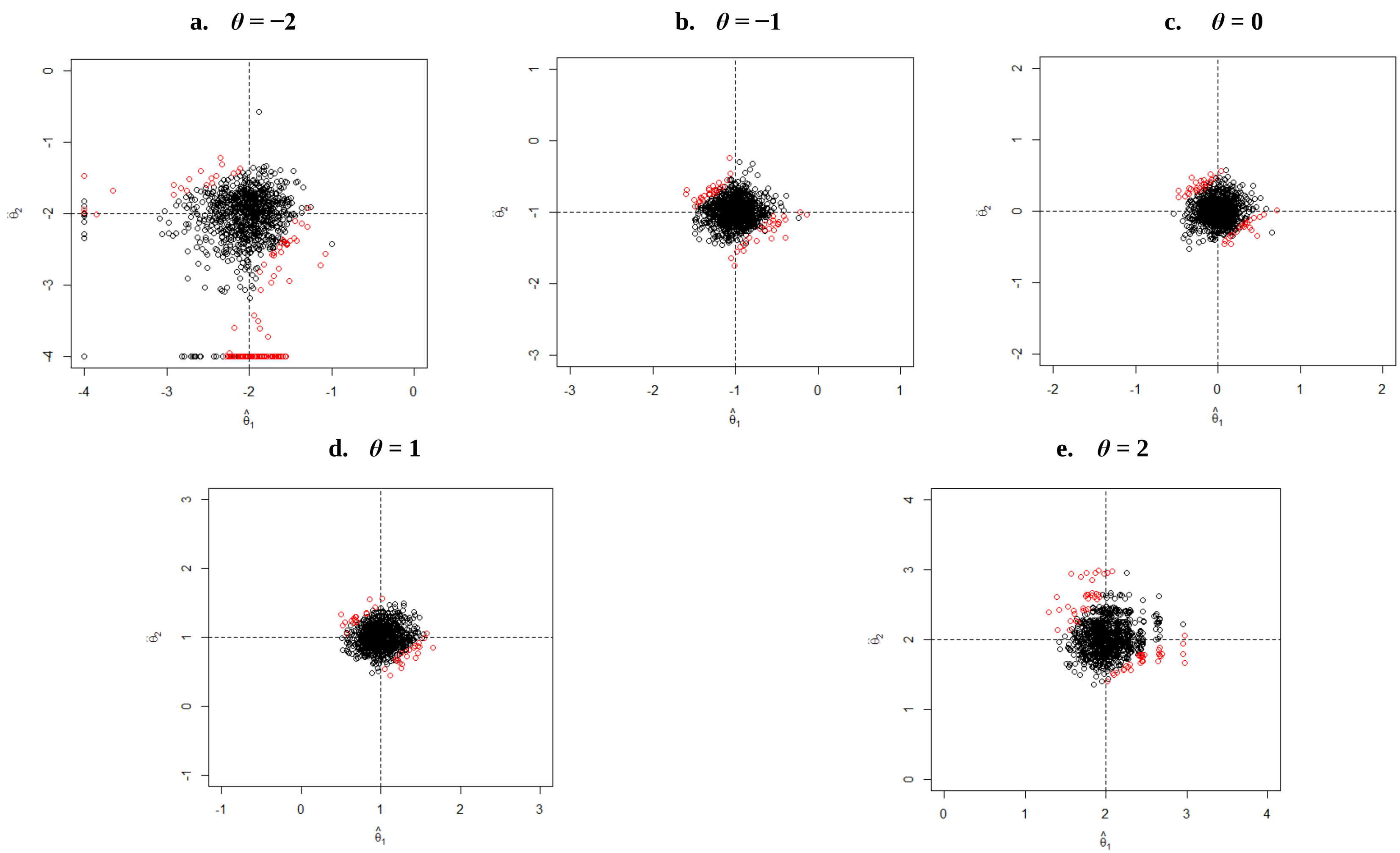
The patterns of bias in MLE in Figure 18 have also been reported in previous studies on the behavior of MLE in CAT (Wang & Vispoel, 1998; Wang, Hanson & Lai, 1999; Warm 1989). Lord (1983) derived the bias function of MLE as

$$\text{Bias}(\text{MLE}(\theta)) \approx \frac{D}{I^2}\sum_{i=1}^{n} a_i I_i (\phi_i - 1/2), \tag{40}$$

where $\phi = (P_i - c_i)/(1 - c_i)$. Equation 40 suggests that the bias will be close to zero if all items have difficulties the same as a simulee's θ: replacing $b_i = \theta$ in Equation 12, P_i becomes $P_i = c_i + \frac{1}{2}(1 - c_i)$ and $\phi = \frac{1}{2}$, which leads bias to be zero. The bias will be negative if θ is smaller than difficulties: $b_i < \theta$ makes $P_i < c_i + \frac{1}{2}(1 - c_i)$, $\phi < \frac{1}{2}$ and thus bias to be negative. The bias will be positive if θ is greater than difficulties (i.e., biased outward). In CAT, this implies that MLE is theoretically unbiased or biased in a small degree as CAT administers items based on each examinee's θ level. CATs in practice, however, will have a slightly larger bias at extreme θ ranges because a sufficient number of items with extreme difficulties are not available. Equation 40 also implies that bias increases with high discriminating items.

Figure 18 also shows that simulees were detected as significant by the Z statistic (i.e., red points) had bias in θ estimates in opposite directions, that is one estimate was biased outward and the other estimate was biased inward. The upper left part of the figure is the region where $\hat{\theta}_1$ was smaller than the true θ and $\hat{\theta}_2$ was larger than the true θ. In the lower right part of the figure, $\hat{\theta}_1$ was larger than the true θ and $\hat{\theta}_2$ was smaller than the true θ. In all five sub-figures, most of the red points are located in these two regions of the plots. The change estimate $\hat{\theta}_2 - \hat{\theta}_1$ (i.e., the numerator of the *Z* statistic) will be larger when θ estimates from the two occasions are biased in opposite directions, which makes the *Z* statistic more likely to be rejected. More simulees were rejected at $\theta = -2$ and at $\theta = 2$ because larger bias in those regions made more $\hat{\theta}_2 - \hat{\theta}_1$ larger, and made it easier for the *Z* statistic to be rejected.

Figure 18. MLE Estimates of θs at Time 1 and Time 2 and Significant Detection of the Z Statistic for 30-item AMC with High Discriminating Peaked Bank



Recommendations for the Design of AMC

The performance of AMC largely depends on the quality of the item bank from which items are drawn. When developing an item bank for AMC, (1) the size and (2) test information function of the bank should be determined adequately. The number of items in a bank should be sufficient to provide information throughout a test. A bank of 500 items is considered as adequate for CAT (Thompson & Weiss, 2011); however, the current study showed that a 300-item AMC bank performed as well as a 500-item AMC bank in terms of Type I error and power. A bank with smaller size would reduce the cost of developing an AMC bank, while obtaining comparable decision accuracy to that of a larger bank. In practice, additional factors can also be considered in determining the bank size: if AMC is used for high-stakes testing, for example, a larger bank might be preferred to reduce item exposure and enhance test security; if content constraints are imposed, it will be important to make sure that there is a sufficient number of items in each content domain.

The test information of an AMC bank should provide adequate information at all θ levels of interest; that is, the item bank should have sufficient numbers of items available at corresponding difficulty levels. In the current study, the power of AMC decreased at the high θ range when not many items with high difficulties were available. Lack of sufficient items in a bank also decreased the efficiency of variable-length AMC at the high θ range (i.e., more items were administered). A bank with a similar number of high discriminating items at each θ (i.e., flat test information) is ideal for AMC to have equivalent level of Type I error and power across θ. At Time 2 or later time points, more difficult items are necessary than easy items since many examinees might be in a higher θ range after positive change (the reverse would also be true if negative change is expected). However, if it is a problem to add difficult items in practice, then the use of AMC needs to make sure to allow enough items to be administered before determining a change decision by setting the appropriate minimum so that a correct decision of significant change for high θ examinees can be made.

Once an item bank is developed, extensive monte-carlo simulations should be made in order to determine optimal specifications for AMC with a given item bank. Performance of the test statistic given the item bank under various expected change conditions should be examined to ensure that the selected test statistic performs better than the other test statistics to determine significant individual change. Since the performance of AMC was affected by the number of items administered in the current study, the minimum and maximum for numbers of items for variable-length AMC should be carefully determined to obtain adequate levels of power and Type I error (or if a fixed-length test is necessary, the best test length of the AMC should be identified in the simulations). In practice, precise θ estimation or correct classification of pass/fail can be another main purpose of a test in addition to detecting significance of change. In such a case, a combined termination criterion, for example, minimum standard error combined with significant change, should be carefully examined to satisfy multiple test purposes. Note that these recommendations are limited to the conditions examined in this study, and by the factors that are described below.

Limitations and Implications for Future Research

Generalizability of findings. The generalizability of the results of the study is limited to model-fitting data. The use of simulated data was considered important, since the current study was only a second attempt to apply hypothesis testing methods to determine significant individual change. Thus, the behavior of the test statistics under various conditions had not yet been investigated. Simulated data has the advantages that true values are known and data fit the model.

Therefore, it helped to clarify previous findings obtained by Finkelman et al. (2010), as well as to introduce new methods for determining significant individual change.

In addition, the advantage of using real data over simulated data is not clear. A live-testing study would require real items administered by CAT at two points in time to the same examinees. Because there would be no "true" criterion to evaluate change in real data, there would be no means of knowing whether the observed "significant" change actually captured "true" change (e.g., Brouwer et al., 2013). Using real data it would then be difficult to demonstrate anything other than the amount of significant change observed, which would make it difficult to compare performance of test statistics in terms of decision accuracy. Thus, the information gained from the use of real data would be small, and such a study would be very expensive.

The use of simulated item banks could also limit the generalizability of this study. However, item banks should be both realistic and relatively optimal, especially for an initial investigation. Thus, simulated item banks with two bank information structures were used in the current study. If a real item bank were used, which likely would have an information function somewhat like the peaked bank, it would be difficult to attribute whether some of the results might have been due to idiosyncrasies of the item bank, not the hypothesis testing methods or other variables manipulated in the study. With a realistic but relatively "pure" item bank, the results obtained in relation to the factors studied can be more confidently interpreted.

Some applied CATs are barely adaptive in practice (Thissen & Mislevy, 2000; Drasgow & Chuah, 2006). Practical constraints often imposed on CATs include content specifications or item exposure controls. As an initial investigation, the current study focused on a constraint-free ideal condition to better focus on the factors of interest. Constraints on CATs would generally be expected to reduce the effectiveness and efficiency of AMC. If additional constraints were imposed on AMC, therefore, it might be difficult to attribute results due to the effect of factors, separated from the effect of constraints on AMC.

After a better idea of how hypothesis testing methods work under ideal conditions is clarified, future studies should further investigate how hypothesis testing methods function under more realistic conditions. Item parameter estimation error, the use of more realistic item banks, possibly real data, item and person misfit, item exposure or content constraints on AMC, and a variety of other conditions would be worth investigating.

Statistical significance and scientific or practical importance. Statistical significance does not necessarily mean scientific importance. Setting aside the problem of an arbitrary use of the *p*-value, a small change that is statistically significant could be not large enough to make a practical difference. For example, if an observed change in depression is significant, it may indicate that the change is real, but it does not inform whether the observed change is clinically meaningful so that treatments can be reduced or terminated. The issue of scientific importance also applies to variable-length AMC with respect to whether it is appropriate to terminate an AMC based on statistical significance, or if practical importance should also be incorporated when making early termination. Thus, significance of change may not be informative by itself in some applications.

Since the current study already had many conditions, it was difficult to have additional criteria to apply to the effect size measure within a single study. However, the practical usefulness of AMC in determining individual change would largely increase if it can provide scientific or practical importance of change in addition to a statistical decision of significant change. Future study should investigate methods of defining effect size or confidence intervals of the observed change at the individual level. A non-parametric approach would be appealing, which does not assume normality of the distributions used in the hypothesis tests.

Higher-order interaction in ANOVA. The current study only analyzed the main effects and two-way interactions in the ANOVAs, leaving all higher-order interactions in the error variance. Higher-order interactions were not included mainly due to difficulties in interpretation. Based on the ANOVA results, however, higher-order interactions would explain only a small portion of the variance in power if they were included. Error variances for power, which would be explained by higher-order interactions, were around 3% (3.25%), 2.21% (2.23%) and 4.83% (4.83%) under small, medium and large change, respectively, for fixed-length (variable-length) AMC. On the other hand, higher-order interactions could possibly explain larger portions of the variance for Type I error. Error variance for Type I error was 47.84% for fixed-length AMC and 69.35% for variable-length AMC. Inclusion of higher-order interactions could further clarify the performance of hypothesis testing due to combinations of factors of interest.

Effect of θ estimation. Findings in the current study were based on MLE. The *Z* statistic showed an increase in Type I error at extreme $\theta = \pm 2$ and it was found in a follow-up analysis that high Type I errors at extreme θ were related to bias in MLE. At non-extreme θ ranges, a better adherence to the nominal level was observed for the *Z* statistic. The LR and ST statistics showed relatively stable but slightly higher Type I error than the nominal level across θ, and smaller decrease in power at $\theta = 2$ under medium and large true change. The KLD-U and KLD-N statistics had lower Type I error than the nominal level and lowest power across conditions examined, yet the statistics were not based on point estimation of θ.

Future studies can further investigate the effect of θ estimation methods in identifying the significance of individual change. There are four θ estimation techniques that have been primarily investigated in the CAT literature: ML, weighted maximum likelihood, and two Bayesian methods — EAP and MAP. The four θ estimation techniques have shown differences in bias and standard error (SE) in the implementation of CAT (e.g., Wang, Hanson & Lau, 1999; Wang & Vispoel, 1998; Weiss, 1982; Yi et al., 2001). MLE has shown smaller bias than EAP and MAP, and WLE was derived to further reduce the bias in MLE. The Bayesian methods have had the smallest SEs across θ, while MLE had the largest SEs and WLE had SE slightly lower than MLE. Future research is needed that examines whether θ estimation methods in relation to difference in bias and SE can affect the performance of test statistics and item selection methods in detecting individual change under various item banks.

Multiple time points. The pre-posttest design in the current study can be generalized to multiple time points. In a clinical setting, patients with chronic disease rarely have only one follow-up visit. Instead, several follow-up measurements are made with pre-defined schedules, for example nine measurements are taken biweekly for posttraumatic stress (e.g., Foa et al., 1991; Frison & Pocock, 1992). It is a loss of information if only two observations are included in testing the significance of change, which might limit the practical usefulness of AMC as operationalized to date in various fields where longitudinal measurements are often used. Therefore, future study should generalize the test statistics to include multiple occasions. To track change at multiple time points, future studies can also examine various growth patterns with different growth rates between occasions in addition to linear growth, since the methods studied here do not assume any a priori pattern of change.

Generalization of IRT models. Future studies can also extend the IRT models that are used for testing individual change. Although it is common to assume items on a test as measures of a unidimensional latent trait, it is not always justifiable. Many educational and psychological variables have been described as inherently multidimensional, and many personality and health assessments measure multiple dimensions (Ackerman, 1994; Reckase, 1985; Reckase, 2009). It is also difficult to assume unidimensionality for achievement tests with subdomains that are not highly inter-correlated. A generalization of test statistics examined here to multidimensional situations will then better reflect characteristics of change, and may deal with situations where one trait changes and others do not.

Both unidimensional and multidimensional models can be further extended to polytomous items. Although dichotomous IRT models have major implications in educational settings with multiple-choice items, polytomous items are also commonly used in applied psychological measurement, such as Likert-type attitude items and partial-credit cognitive items. Thus, further study is necessary to clarify the nature of test statistics in individual change using polytomous IRT models. These developments will extend the practical applicability of AMC in various contexts of measuring individual change.

References

Ackerman, T. A. (1994). Using multidimensional item response theory to understand what items and tests are measuring. *Applied Measurement in Education, 7* (4), 255-278.

Ali, B., Rahbar, M., Naeem, S., Gul, A., Mubeen, S., & Iqbal, A. (2003). The effectiveness of counseling on anxiety and depression by minimally trained counselors: A randomized control trial. *American Journal of Psychotherapy, 57*, 324–336.

Bartroff, J., Finkelman, M., & Lai, T. L. (2008). Modern sequential analysis and its applications to computerized adaptive testing. *Psychometrika, 73*, 473-486.

Belov, D. I., & Armstrong, R. D. (2011). Distributions of the Kullback-Leibler divergence with applications. *British Journal of Mathematical Statistical Psychology, 64*, 291-309.

Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C. Harris (Ed.), *Problems in measuring change* (pp. 3–20). Madison, WI: University of Wisconsin Press.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In Lord, F.M. & Novick, M.R. (Eds.), *Statistical theories of mental test scores*. Reading, MA: Addison–Wesley.

Brouwer, D., Meijer, R. R., & Zevalkink, J. (2013). Measuring individual significant change on the Beck Depression Inventory-II through IRT-based statistics. *Psychotherapy Research, 23* (5), 489-501.

Brown, C. A., Smith, M. S., & Stein, M. K. (1995). *Linking teacher support to enhanced classroom instruction*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Bryk, A. S., & Randenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*, 147-158.

Burr, J. A., & Nesselroade, J. R. (1990). Change measurement. In A. von Eye (Ed.), *Statistical methods in longitudinal research* (vol. 1) (pp. 3-34). Boston: Academic Press.

Chang, H. H., & Ying, Z. L. (1996). A global information approach to computerized adaptive testing. *Applied Psychological Measurement*, *20*(3), 213-229.

Chen, S-Y., & Ankenmann, R. D. (2004). Effects of practical constraints of item selection rules at the early stage of computerized adaptive testing. *Journal of Educational Measurement*, *41*, 149-174.

Cardon, L. R., Fulker, D. W., DeFries, J. C. and Plomin, R. (1992). Continuity and change in general cognitive ability From 1 to 7 years of age. *Dev. Psychol*, *28*, 1–10.

Chandra, T. K. & Joshi, S. N. (1983). Comparison of likelihood ratio, Wald's and Rao's test, *Sankhyā, Series A, 45*, 226-246.

Chapman, O. (2002). Belief structure and inservice high school mathematics teacher growth. In G. Leader, E. Pehkonen and G. Torner (Eds.), Beliefs: *A Hidden Variable in Mathematics Education*, Kluwer Academic Publishers, Dordrecht, 117-193.

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical models, design, and statistical model. *Annual Review of Psychology*, *57*, 505-528.

Cox, D. R., & Hinkley, D. V. (1974). *Theoretical statistics*. New York: Wiley.

Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory*. Mason, OH: Thomson Wadsworth.

Cronbach, L.J., & Furby, L. (1970). How we should measure “change” – or should we? *Psychological Bulletin, 74*, 68–80.

Drasgow, F., & Chuah, S. C. (2006). Computer-based testing. In Eid, M. & Diener, E. (Eds). *Handbook of multimethod measurement in psychology* (pp. 87-100). Washington, DC: American Psychological Association

DuBois, P. H. (1957). *Multivariate correlational analysis*. Harper: New York.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association, 56,* 52–64.

Embretson, S. E. (1991). Implications of a multidimensional latent trait model for measuring change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 184-197). Washington, DC: American Psychological Association.

Embretson, S. E. (1996). Item response theory models and spurious interaction effects in factorial ANOVA designs. *Applied Psychological Measurement, 20,* 201-212.

Falloon, I. H., Boyd, J. L., McGill, C. W., Williamson, M., Razani, J., Moss, H. B., Gilderman, A. M., & Simpson, G. M. (1985). Family management in the prevention of morbidity of schizophrenia: Clinical outcome of a two-year longitudinal study. *Arch Gen Psychiatry, 42*(9), 887-896.

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in Mathematics instruction. *Journal for Research in Mathematics Education , 27* (4), 403-434.

Fliege, H., Becker, J., Walter, O. B., Bjorner, J. B., Klapp, B. F., & Rose, M. (2005). Development of a computer-adaptive test for depression (D-CAT). *Quality of Life Research, 14,* 2277–2291.

Finkelman, M. D., Weiss, D. J., & Kim-Kang, G. (2010). Item selection and hypothesis testing for the adaptive measurement of change. *Applied Psychological Measurement, 34,* 238-254.

Fischer, G. H. (1976). Some probabilistic models for measuring change. In D. N. M. de Gruijter & L. J. T. van der Kamp (Eds). *Advances in psychological and educational measurement* (pp. 97-110). New York: John Wiley & Sons.

Fischer, G. H. (1983). Logistic latent trait models with linear constraints. *Psychometrika, 48,* 3-26.

Fischer, G. H. (2001). Gain scores revisited under an IRT perspective. In A. Boomsma, M. van Duijn, T. Snijders (Eds.) *Essays on item response theory* (pp. 43-68). Springer-Verlag New York, Inc.

Fischer, G. H. (2003). The precision of gain scores under an item response theory perspective: A comparison of asymptotic and exact conditional inference about change. *Applied Psychological Measurement,27,* 3-26.

Foa, E. B., Rothbaum, B. O., Riggs, D. S., & Murdock, T. B. (1991). Treatment of posttraumatic stress disorder in rape victims: A comparison between cognitive-behavioral procedures and counseling. *Journal of Consulting and Clinical Psychology, 59*, 715-723.

Frison, L., & Pocock, S. J. (1992). Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Statistics in Medicine, 11*, 1865 – 1704.

Gagne, D., & Toye, R. C. (1994). The effects of therapeutic touch and relaxation therapy in reducing anxiety. *Archives of Psychiatric Nursing, 8*, 184-189.

Gialluca, K. A., & Weiss, D. J. (1979). *Efficiency of an adaptive inter-subset branching strategy in the measurement of classroom achievement* (Research Report 79-6). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program. Available from CAT Central (http://www.psych.umn.edu/psylabs/catcentral/).

Guo, J., & Drasgow, F. (2010). Identifying cheating on unproctored internet test: The Z-test and the likelihood ratio test. *International Journal of Selection and Assessment*, *18* (4), 351-364.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66* , 361-396.

Hart, D. L., Cook, K. F., Mioduski, J. E., Teal, C. R., & Crane, P. K. (2006). Simulated computerized adaptive test for patients with shoulder impairments was efficient and produced valid measures of function. *Journal of Clinical Epidemiology, 59*, 290-298.

Heppner, M. J., & Heppner, P. P. (2003). Identifying process variables in career counseling: A research agenda. *Journal of Vocational Behavior, 62*, 429-452.

Hummel-Rossi, B., & Weinberg, S. L. (1975). Practical guidelines in applying current theories to the measurement of change. I. Problems in measuring change and recommended procedures. *JSAS Catalog of Selected Documents in Psychology, 5*, 226 (Ms. No. 916).

Johnston, J. & DiNardo, J. (1997) *Econometric methods*. New York, NY: The McGraw-Hill Companies, Inc.

Kajander, A., & Mason, R. (2007). Examining teacher growth in professional learning groups for in-service teachers of mathematics. *Canadian Journal of Science, Mathematics and Technology Education, 7,* 417-438.

Kang, S.-M., & Waller, N. G. (2005). Moderated multiple regression, spurious interaction effects, and IRT. *Applied Psychological Measurement, 29,* 87-105.

Kim-Kang, G., & Weiss, D. J. (2007). Comparison of computerized adaptive testing and classical methods for measuring individual change. In D. J. Weiss (Ed.), *Proceedings of the 2007 GMAC Conference on Computerized Adaptive Testing*. Available from www.psych.umn.edu/psylabs/CATCentral

Kim-Kang, G., & Weiss, D. J. (2008). Adaptive measurement of individual change. *Zeitschrift fur Psychologie, 216*, 49-58.

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. 257–283). New York: Academic Press.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics, 22*, 79-86.

Lai, T. L., & Shih, M. C. (2004). Power, sample size and adaptation considerations in the design of group sequential clinical trials. *Biometrika, 91*, 507–528.

Lazarus and Free Pascal Team (2008). Lazarus: The professional Free Pascal RAD IDE.

Lehmann, E. L. & Casella, G. (1998). *Theory of point estimation*. New York: Springer-Verlag.

Lei, J., & Zhao, Y. (2007). Technology uses and student achievement: A longitudinal study. *Computers & Education, 49*, 284-296.

Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 21–38). Madison: WI: University of Wisconsin Press.

Lord, F. M. (1983). Unbiased estimators of ability parameters, of their variance, and of their parallel-forms reliability. *Psychometrika, 48*, 233-245.

Markus, G. (1980). *Models for the analysis of panel data.* Beverly Hills: Sage.

Manning, W. H., & DuBois, P. H. (1962). Correlation methods in research on human learning. *Perceptual and Motor Skills, 15*, 287-321.

Maurelli, V., & Weiss, D. J. (1981). *Factors influencing the psychometric characteristics of an adaptive testing strategy for test batteries*. (Research Rep. No. 81-4). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, Computerized Adaptive Testing Laboratory. Available from CAT Central (http://www.psych.umn.edu/psylabs/catcentral/).

May, K., & Nicewander, W. A. (1998). Measuring change conventionally and adaptively. *Educational and Psychological Measurement, 58,* 882-897.

McArdle, J. J., & Epstein, D. B. (1987). Latent growth curves within developmental structural equation models. *Child Development, 57*, 110-133.

Mellenbergh G. J., & van den Brink, W. P. (1998). The measurement of individual change. *Psychological Methods, 3*(4), 470-485.

Nation, K., Cooksey, J., Taylor, J., & Bishop, D. (2010). A longitudinal investigation of the early language and reading skills in children with reading comprehension impairment. *Journal of Child Psychology & Psychiatry, 51*, 1031–1039.

Neyman, J. & Pearson, E. S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika, 20*, 175-240, 263-294.

O'Neil, R., & Wetherill, G. B (1971). The present state of multiple comparison methods. *Journal of Royal Satistical Society, 33,* 218-250.

Perdrixa, S., Stauffera, S., Masdonatib, J., Massoudia, K., & Rossiera, J. (2012). Effectiveness of career counseling: A one-year follow-up. Journal of *Vocational Behavior, 80*, 565–578.

R Development Core Team (2012), *R: A language and environment for statistical computing.* Vienna, Austria: the R Foundation for Statistical Computing.

Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with application to problems of estimation. *Proceedings of the Cambridge Philosophical Society, 44,* 50-57.

Rao, C. R. (1965). *Linear statistical inference and its applications*, New York: Willey.

Reckase, M. D. (1985). The difficulty of test items that measure more than one ability. *Applied Psychological Measurement, 9* (4), 401-412.

Reckase, M. D. (2009). *Multidimensional item response theory*. New York: Springer.

Rogosa, D. R., Brand, D., & Zimowski, M. (1982). A growth curve approach to the scores in the measurement of change. *Journal of Educational Measurement, 20*, 333-343.

Rogosa, D. R., & Willet, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. *Psychometrika*, *50*, 203-228.

Ryan, T. A. (1959). Multiple comparisons in psychological research, *Psychological Bulletin, 56*, 26-47.

Shaffer, J. P. (1995). Multiple hypothesis testing, *Annual Review of Psychology*, *46*, 561–584.

Simms, L. J., & Clark, L. A. (2005). Validation of a computerized adaptive version of the Schedule for Nonadaptive and Adaptive Personality (SNAP). *Psychological Assessment*, *17*, 28–43.

Smits, J. A. J., Berry, A. C., Powers, M. B., Behar, E., & Otto, M. W (2008). Reducing anxiety sensitivity with exercise. *Depression and Anxiety*, *25*, 689-699.

Taniguchi, M. (1988). Asymptotic expansions of the distribution of some statistics for Gaussian ARMA process. *Journal of Multivariate Analysis, 27*, 494-511.

Taniguchi, M. (1991). *Higher order asymptotic theory for time series analysis, Lecture Notes in Statistics, 68,* Springer-Verlag, New York.

Thissen, D., & Mislevy, R.J. (2000). Testing algorithms. In Wainer, H. (Ed.) *Computerized Adaptive Testing: A Primer*. Mahwah, NJ: Lawrence Erlbaum Associates.

Thompson, N. A., & Weiss, D. J. (2011). A framework for the development of computerized adaptive tests. *Practical Assessment, Research & Evaluation, 16*(1). Available online: http://pareonline.net/getvn.asp?v=16&n=1.

Thompson, R., & Haddock, G. (2012). Sometimes stories sell: When are narrative appeals most likely to work? *European Journal of Social Psychology, 42*, 92–102.

Thorndike, R. L. (1963). *The concepts of over- and underachievement*. New York: Columbia University, Teachers College, Bureau of Publications.

Tukey, J. W. (1953). The problem of multiple comparisons. Unpublished manuscript, Princeton University.

van der Linden, W. J., & Pashley, P. J. (2010). Item selection and ability estimation in adaptive testing. In van der Linden, W. J., & Glas, C. A. W. (Eds), *Elements of Adaptive Testing*. (pp. 3 – 30). New York: Springer.

Von Eye, A. (1990). *Statistical methods in longitudinal research: Time series and categorical longitudinal data.* Boston, MA: Academic

Von Minden, S. (2011). *Measuring individual change: A comparison of conventional and adaptive tests.* Unpublished Master's thesis, Department of Psychology, University of Minnesota

Walthera, E., Gawronskib, B., Blankc, H., & Langera, T. (2009). Changing likes and dislikes through the back door: The US-revaluation effect. *Cognition & Emotion, 23*, 889-917.

Warm, T. A. (1989). Weighted likelihood estimation of ability in the item response theory. *Psychometrika, 54*, 427-450.

Wang, C. (2014). Reporting reliable change in students' overall and domain abilities across two time points. *Paper presented at the 76th meeting of the National Council on Measurement in Education, Philadelphia, PA* .

Wang, T., Hanson, B. A. & Lau, C.-M. A. (1999). Reducing bias in CAT trait estimation: A comparison of approaches. *Applied Psychological Measurement*, *23*, 263-278.

Wang, T., & Vispoel, W. P. (1998). Properties of ability estimation methods in computerized adaptive testing. *Journal of Educational Measurement*, *35*, 109-135.

Weiss, D. J. (1982). Improving measurement quality and efficiency with adaptive testing. *Applied Psychological Measurement, 4,* 473-285.

Weiss, D. J. (2011). Better data from better measurements using computerized adaptive testing. *Journal of Methods and Measurement in the Social Sciences, 2*(1), 1-23.

Weiss, D. J., & Kingsbury, G. G. (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement, 21,* 361-375.

Weiss, D. J., & Von Minden, S. (2011). Measuring individual growth with conventional and adaptive tests. *Journal of Methods and Measurement in the Social Sciences, 2*(1), 80-101.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics 9,* 60–62.

Willett, J. B. (1988). Questions and answers in the measurement of change. *Review of research in education, 15,* 345-422.

Willett, J. B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth, *Educational and Psychological Measurement, 49*, 587-602.

Willett, J. B. (1997). Measuring change: What individual growth modeling buys you. In E. Arnsel & K. A. Reninger (Eds.), *Change and development* (pp. 213-243). Maywah, NJ: Erlbaum.

Williams, R. H., & Zimmerman, D. W. (1996). Are simple gain scores obsolete? *Applied Psychological Measurement*, *20*, 59–69.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*, 57-67

Yi, Q., Wang, T., & Ban, J-C. (2001). Effects of scale transformation and test-termination rule on the precision of ability estimation in computerized adaptive testing. *Journal of Educational Measurement, 38*, 267-292.

Zimmerman, D. W., & Williams, R. H. (1982). Gain scores in research can be highly reliable. *Journal of Educational Measurement, 19*, 149-154.

Appendices

Table A1. Average Bank Information of CAT Banks in Three Discrimination Conditions

Bank	HD	MD	LD
300-Item Bank			
Peaked Bank	56.528	37.491	21.777
Flat Bank	49.579	32.174	18.19
500-Item Bank			
Peaked Bank	94.262	62.569	36.272
Flat Bank	82.881	53.839	30.362

Table A2. Critical Values for 15-Item Variable-Length AMC With 300-Item Bank

Item Selection	Bank	Z		LR		ST		KLD-U		KLD-N	
		C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2
FI	HD-PB	2.44	1.97	6.88	4.49	7.00	4.63	5.52	3.84	5.31	3.84
	MD-PB	2.44	1.98	6.6	4.33	6.86	4.51	5.16	3.84	4.72	3.84
	LD-PB	2.47	2.00	6.38	4.21	6.67	4.42	4.89	3.84	4.11	3.84
	HD-FB	2.34	1.96	7.05	4.58	7.17	4.75	5.44	3.84	5.25	3.84
	MD-FB	2.33	1.96	6.76	4.43	7.00	4.64	5.13	3.84	4.71	3.84
	LD-FB	2.33	1.96	6.53	4.30	6.78	4.52	4.81	3.84	4.08	3.84
KL-F	HD-PB	2.54	1.99	6.42	4.24	6.69	4.42	6.06	4.17	5.66	3.87
	MD-PB	2.53	2.00	6.25	4.17	6.49	4.31	5.62	3.92	4.99	3.84
	LD-PB	2.55	2.02	6.18	4.12	6.38	4.26	5.25	3.84	4.31	3.84
	HD-FB	2.41	1.96	6.54	4.28	6.86	4.50	6.05	4.16	5.67	3.86
	MD-FB	2.40	1.96	6.34	4.18	6.60	4.38	5.61	3.90	5.00	3.84
	LD-FB	2.41	1.96	6.29	4.13	6.51	4.29	5.21	3.84	4.30	3.84
KL-M	HD-PB	2.58	2.01	6.33	4.19	6.53	4.34	6.36	4.38	5.87	4.03
	MD-PB	2.56	2.01	6.20	4.09	6.39	4.21	5.88	4.07	5.19	3.84
	LD-PB	2.57	2.03	6.12	4.08	6.30	4.19	5.46	3.84	4.46	3.84
	HD-FB	2.44	1.96	6.44	4.19	6.66	4.35	6.44	4.41	5.97	4.07
	MD-FB	2.44	1.96	6.30	4.14	6.50	4.29	5.91	4.09	5.25	3.84
	LD-FB	2.45	1.96	6.19	4.05	6.35	4.19	5.42	3.84	4.44	3.84

Table A3. Critical Values for 30-Item Variable-Length AMC With 300-Item Bank

Item Selection	Bank	Z		LR		ST		KLD-U		KLD-N	
		C1	C2	C1	C2	C1	C2	C1	C2	C1	C2
FI	HD-PB	2.81	2.09	7.93	4.41	7.98	4.43	6.06	3.92	5.88	3.84
	MD-PB	2.8	2.08	7.73	4.35	7.88	4.41	5.88	3.84	5.56	3.84
	LD-PB	2.83	2.09	7.59	4.31	7.77	4.37	5.74	3.84	5.14	3.84
	HD-FB	2.61	2.01	8.06	4.44	8.21	4.53	5.95	3.87	5.85	3.84
	MD-FB	2.59	2.00	7.76	4.36	8.03	4.49	5.79	3.84	5.52	3.84
	LD-FB	2.60	2.00	7.59	4.32	7.83	4.44	5.64	3.84	5.09	3.84
KL-F	HD-PB	2.87	2.08	7.32	4.20	7.41	4.24	6.26	4.02	5.95	3.84
	MD-PB	2.86	2.08	7.35	4.20	7.43	4.23	6.12	3.94	5.68	3.84
	LD-PB	2.88	2.08	7.34	4.20	7.38	4.24	5.98	3.84	5.28	3.84
	HD-FB	2.68	2.00	7.46	4.22	7.62	4.34	6.26	4.02	6.04	3.87
	MD-FB	2.69	2.00	7.40	4.17	7.58	4.29	6.13	3.92	5.74	3.84
	LD-FB	2.68	2.00	7.30	4.18	7.44	4.27	5.91	3.84	5.28	3.84
KL-M	HD-PB	2.91	2.09	7.29	4.15	7.33	4.18	6.45	4.10	6.12	3.86
	MD-PB	2.88	2.08	7.31	4.16	7.31	4.19	6.29	4.01	5.81	3.84
	LD-PB	2.90	2.08	7.29	4.16	7.28	4.19	6.10	3.92	5.39	3.84
	HD-FB	2.73	2.01	7.41	4.19	7.49	4.26	6.46	4.14	6.22	3.98
	MD-FB	2.73	2.01	7.34	4.16	7.45	4.24	6.29	4.03	5.89	3.84
	LD-FB	2.72	2.01	7.25	4.14	7.32	4.21	6.05	3.91	5.38	3.84

Table A4. Critical Values for 50-Item Variable-Length AMC With 300-Item Bank

Item Selection	Bank	Z		LR		ST		KLD-U		KLD-N	
		C1	C2	C1	C2	C1	C2	C1	C2	C1	C2
FI	HD-PB	2.95	2.12	8.35	4.4	8.48	4.39	6.36	3.99	6.19	3.84
	MD-PB	2.92	2.09	8.2	4.3	8.38	4.33	6.23	3.91	5.97	3.84
	LD-PB	2.95	2.1	8.11	4.3	8.3	4.33	6.13	3.86	5.65	3.84
	HD-FB	2.71	2.03	8.4	4.37	8.57	4.41	6.18	3.9	6.12	3.84
	MD-FB	2.71	2.03	8.29	4.35	8.54	4.42	6.13	3.88	5.97	3.84
	LD-FB	2.71	2.03	8.07	4.31	8.36	4.39	5.97	3.84	5.58	3.84
KL-F	HD-PB	2.99	2.12	7.81	4.25	7.86	4.25	6.49	4.05	6.23	3.87
	MD-PB	2.96	2.09	7.84	4.19	7.92	4.22	6.37	4	6.04	3.84
	LD-PB	2.99	2.09	7.85	4.2	7.91	4.22	6.25	3.91	5.73	3.84
	HD-FB	2.79	2.02	7.86	4.18	7.97	4.24	6.31	3.96	6.17	3.86
	MD-FB	2.8	2.03	7.89	4.19	8.04	4.25	6.3	3.94	6.06	3.84
	LD-FB	2.79	2.03	7.81	4.2	7.94	4.26	6.17	3.91	5.7	3.84
KL-M	HD-PB	3.03	2.13	7.79	4.21	7.81	4.21	6.57	4.08	6.3	3.88
	MD-PB	3	2.09	7.8	4.17	7.82	4.18	6.45	4.04	6.12	3.84
	LD-PB	3.02	2.09	7.83	4.15	7.83	4.15	6.34	3.91	5.8	3.84
	HD-FB	2.83	2.03	7.83	4.16	7.86	4.2	6.44	4.02	6.27	3.91
	MD-FB	2.83	2.03	7.84	4.17	7.91	4.21	6.42	3.99	6.16	3.84
	LD-FB	2.83	2.04	7.75	4.18	7.83	4.23	6.28	3.96	5.79	3.84

Table A5. Critical Values for 15-Item Variable-Length AMC With 500-Item Bank

Item Selection	Bank	Z		LR		ST		KLD-U		KLD-N	
		C1	C2	C1	C2	C1	C2	C1	C2	C1	C2
FI	HD-PB	2.42	1.96	6.98	4.55	7.11	4.66	5.58	3.88	5.41	3.84
	MD-PB	2.42	1.97	6.70	4.38	6.94	4.57	5.20	3.84	4.79	3.84
	LD-PB	2.46	1.99	6.41	4.26	6.73	4.46	4.91	3.84	4.19	3.84
	HD-FB	2.33	1.96	7.01	4.54	7.16	4.75	5.42	3.84	5.28	3.84
	MD-FB	2.32	1.96	6.76	4.42	7.08	4.67	5.19	3.84	4.78	3.84
	LD-FB	2.33	1.96	6.56	4.32	6.86	4.55	4.87	3.84	4.14	3.84
KL-F	HD-PB	2.49	1.97	6.42	4.25	6.72	4.45	6.11	4.21	5.77	3.94
	MD-PB	2.50	1.99	6.33	4.21	6.58	4.37	5.69	3.96	5.10	3.84
	LD-PB	2.53	2.01	6.21	4.11	6.44	4.25	5.28	3.84	4.39	3.84
	HD-FB	2.39	1.96	6.48	4.24	6.83	4.49	6.06	4.20	5.73	3.92
	MD-FB	2.41	1.96	6.41	4.20	6.71	4.42	5.72	3.96	5.12	3.84
	LD-FB	2.41	1.96	6.30	4.15	6.54	4.33	5.27	3.84	4.38	3.84
KL-M	HD-PB	2.52	1.98	6.36	4.18	6.57	4.33	6.44	4.42	6.03	4.10
	MD-PB	2.53	2.00	6.28	4.15	6.48	4.28	5.99	4.14	5.30	3.84
	LD-PB	2.56	2.02	6.13	4.06	6.32	4.19	5.51	3.84	4.53	3.84
	HD-FB	2.43	1.96	6.36	4.18	6.65	4.38	6.49	4.44	6.02	4.13
	MD-FB	2.45	1.96	6.34	4.14	6.58	4.30	6.08	4.17	5.40	3.84
	LD-FB	2.46	1.96	6.23	4.10	6.43	4.24	5.54	3.84	4.56	3.84

Table A6. Critical Values for 30-Item Variable-Length AMC With 500-Item Bank

Item Selection	Bank	Z		LR		ST		KLD-U		KLD-N	
		C1	C2	C1	C2	C1	C2	C1	C2	C1	C2
FI	HD-PB	2.73	2.05	7.98	4.40	8.10	4.45	6.04	3.93	5.92	3.84
	MD-PB	2.73	2.06	7.80	4.37	7.99	4.45	5.93	3.87	5.63	3.84
	LD-PB	2.79	2.07	7.64	4.30	7.84	4.38	5.75	3.84	5.50	3.84
	HD-FB	2.60	2.00	8.00	4.44	8.24	4.58	5.96	3.87	5.86	3.84
	MD-FB	2.59	2.00	7.85	4.37	8.11	4.51	5.83	3.84	5.59	3.84
	LD-FB	2.60	1.99	7.64	4.30	7.92	4.44	5.62	3.84	5.50	3.84
KL-F	HD-PB	2.79	2.05	7.42	4.22	7.54	4.27	6.33	4.08	6.05	3.88
	MD-PB	2.81	2.05	7.45	4.22	7.54	4.28	6.23	4.01	5.84	3.84
	LD-PB	2.84	2.06	7.33	4.16	7.39	4.21	5.94	3.84	5.50	3.84
	HD-FB	2.69	2.00	7.50	4.24	7.72	4.37	6.32	4.07	6.10	3.93
	MD-FB	2.69	2.00	7.43	4.20	7.59	4.33	6.16	3.98	5.83	3.84
	LD-FB	2.69	2.00	7.35	4.18	7.49	4.28	5.93	3.84	5.50	3.84
KL-M	HD-PB	2.83	2.05	7.41	4.18	7.46	4.21	6.50	4.15	6.15	3.93
	MD-PB	2.84	2.06	7.36	4.17	7.40	4.21	6.39	4.09	5.94	3.84
	LD-PB	2.86	2.07	7.28	4.15	7.30	4.18	6.09	3.93	5.50	3.84
	HD-FB	2.73	2.01	7.47	4.21	7.59	4.31	6.50	4.18	6.15	3.98
	MD-FB	2.73	2.01	7.38	4.15	7.49	4.25	6.33	4.06	5.96	3.84
	LD-FB	2.73	2.01	7.30	4.12	7.38	4.20	6.10	3.92	5.50	3.84

Table A7. Critical Values for 50-Item Variable-Length AMC With 500-Item Bank

Item Selection	Bank	Z		LR		ST		KLD-U		KLD-N	
		C1	C2	C1	C2	C1	C2	C1	C2	C1	C2
FI	HD-PB	2.87	2.09	8.49	4.41	8.63	4.43	6.41	4.03	6.28	3.90
	MD-PB	2.88	2.08	8.30	4.33	8.54	4.39	6.26	3.96	6.05	3.84
	LD-PB	2.93	2.10	8.13	4.31	8.38	4.37	6.14	3.89	5.73	3.84
	HD-FB	2.72	2.03	8.57	4.42	8.82	4.49	6.22	3.92	6.19	3.86
	MD-FB	2.71	2.02	8.34	4.37	8.66	4.46	6.12	3.86	5.97	3.84
	LD-FB	2.71	2.02	8.14	4.30	8.47	4.39	6.00	3.84	5.65	3.84
KL-F	HD-PB	2.94	2.09	7.96	4.24	8.07	4.28	6.58	4.10	6.34	3.92
	MD-PB	2.94	2.07	7.91	4.20	8.00	4.25	6.42	4.03	6.14	3.84
	LD-PB	2.97	2.09	7.91	4.22	7.98	4.25	6.31	3.95	5.83	3.84
	HD-FB	2.79	2.03	7.98	4.26	8.16	4.34	6.39	4.01	6.26	3.93
	MD-FB	2.79	2.02	7.91	4.21	8.11	4.28	6.30	3.95	6.08	3.84
	LD-FB	2.80	2.03	7.83	4.18	7.99	4.25	6.16	3.88	5.75	3.84
KL-M	HD-PB	2.97	2.09	7.93	4.22	7.98	4.24	6.68	4.14	6.43	3.96
	MD-PB	2.96	2.08	7.90	4.19	7.95	4.21	6.53	4.08	6.21	3.85
	LD-PB	3.00	2.09	7.85	4.20	7.88	4.21	6.39	3.99	5.91	3.84
	HD-FB	2.84	2.04	7.98	4.22	8.09	4.28	6.53	4.06	6.40	3.98
	MD-FB	2.84	2.03	7.90	4.17	8.02	4.23	6.44	4.01	6.21	3.86
	LD-FB	2.84	2.04	7.81	4.19	7.89	4.23	6.27	3.95	5.86	3.84

Table A8. Marginal Mean (SD) for Type I Error and Power at Each θ Level for Fixed- and Variable-Length AMC

Termination Method	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
Fixed	.054	.053	.052	.051	.050	.050	.050	.050	.048
	(.017)	(.012)	(.011)	(.011)	(.011)	(.011)	(.011)	(.012)	(.014)
Variable	.053	.052	.050	.048	.047	.047	.048	.048	.046
	(.015)	(.009)	(.008)	(.008)	(.008)	(.009)	(.009)	(.009)	(.012)
Power: Very Small Change $\Delta\theta$ = 0.25									
Fixed	.097	.109	.114	.116	.116	.114	.111	.105	.091
	(.043)	(.048)	(.053)	(.056)	(.058)	(.057)	(.054)	(.049)	(.046)
Variable	.093	.104	.108	.110	.109	.108	.105	.100	.085
	(.039)	(.044)	(.049)	(.052)	(.053)	(.053)	(.050)	(.046)	(.044)
Power: Small Change $\Delta\theta$ = 0.5									
Fixed	.248	.285	.304	.314	.315	.310	.297	.270	.218
	(.147)	(.168)	(.181)	(.188)	(.189)	(.187)	(.179)	(.162)	(.147)
Variable	.239	.274	.292	.302	.303	.298	.285	.259	.206
	(.141)	(.163)	(.175)	(.182)	(.183)	(.181)	(.173)	(.157)	(.142)
Power: Medium Change $\Delta\theta$ = 1.0									
Fixed	.646	.684	.704	.714	.714	.704	.681	.630	.504
	(.252)	(.245)	(.242)	(.241)	(.243)	(.245)	(.250)	(.258)	(.273)
Variable	.637	.675	.695	.705	.705	.695	.671	.617	.485
	(.252)	(.246)	(.242)	(.242)	(.244)	(.246)	(.251)	(.259)	(.273)
Power: Large Change $\Delta\theta$ = 1.5									
Fixed	.863	.882	.893	.899	.898	.890	.868	.814	.659
	(.170)	(.157)	(.151)	(.149)	(.153)	(.160)	(.179)	(.217)	(.284)
Variable	.858	.878	.889	.896	.895	.886	.862	.805	.641
	(.170)	(.158)	(.152)	(.150)	(.154)	(.162)	(.181)	(.220)	(.287)

Table A9. Marginal Mean (SD) for Number of Items Used at Each θ Level for Fixed- and Variable-Length AMC

Change	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Fixed-Length									
	31.67	31.67	31.67	31.67	31.67	31.67	31.67	31.67	31.67
	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)
Variable-Length									
No	31.28	31.31	31.34	31.35	31.36	31.35	31.34	31.33	31.35
	(14.09)	(14.11)	(14.12)	(14.12)	(14.13)	(14.13)	(14.12)	(14.11)	(14.10)
Very Small	30.91	30.87	30.86	30.86	30.86	30.87	30.88	30.91	31.09
	(13.76)	(13.71)	(13.70)	(13.70)	(13.70)	(13.70)	(13.71)	(13.74)	(13.87)
Small	29.27	28.92	28.74	28.65	28.65	28.68	28.82	29.18	29.97
	(12.43)	(12.13)	(12.01)	(11.96)	(11.97)	(12.00)	(12.09)	(12.37)	(12.98)
Medium	21.98	21.22	20.86	20.69	20.63	20.80	21.41	22.94	26.03
	(8.45)	(8.11)	(8.00)	(7.97)	(7.99)	(8.11)	(8.42)	(9.12)	(11.16)
Large	16.27	15.84	15.61	15.38	15.18	15.44	16.27	18.41	23.23
	(5.28)	(5.07)	(4.96)	(4.73)	(4.59)	(4.88)	(5.58)	(6.88)	(10.23)

Table A10. Marginal Mean (SD) for Type I Error and Power Conditional on Bank Information and θ Level for Fixed- and Variable-Length AMC

Termination	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
Fixed-Length									
PB	.055	.054	.052	.051	.050	.051	.051	.051	.049
	(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
FB	.053	.053	.051	.051	.050	.049	.050	.049	.048
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Variable-Length									
PB	.053	.052	.050	.049	.047	.048	.048	.048	.045
	(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
FB	.052	.051	.049	.048	.047	.047	.047	.048	.046
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Power: Very Small Change, $\Delta\theta$ = 0.25									
Fixed-Length									
PB	.088	.109	.119	.125	.126	.124	.117	.106	.082
	(.03)	(.05)	(.06)	(.06)	(.06)	(.06)	(.06)	(.05)	(.04)
FB	.107	.108	.108	.107	.106	.105	.105	.103	.100
	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)
Variable-Length									
PB	.083	.104	.112	.118	.119	.116	.111	.100	.074
	(.03)	(.04)	(.05)	(.06)	(.06)	(.06)	(.05)	(.05)	(.03)
FB	.103	.104	.103	.102	.100	.100	.100	.099	.096
	(.04)	(.04)	(.04)	(.04)	(.05)	(.05)	(.05)	(.05)	(.05)
Power: Small Change, $\Delta\theta$ = 0.5									
Fixed-Length									
PB	.222	.290	.325	.344	.347	.338	.313	.265	.175
	(.12)	(.17)	(.19)	(.20)	(.20)	(.20)	(.18)	(.15)	(.10)
FB	.274	.280	.284	.285	.283	.282	.281	.275	.262
	(.17)	(.17)	(.17)	(.17)	(.17)	(.17)	(.17)	(.17)	(.17)
Variable-Length									
PB	.211	.278	.311	.329	.333	.324	.300	.252	.159
	(.11)	(.16)	(.19)	(.19)	(.20)	(.19)	(.18)	(.15)	(.09)
FB	.266	.271	.273	.275	.272	.271	.271	.266	.253
	(.16)	(.16)	(.16)	(.16)	(.16)	(.16)	(.17)	(.17)	(.17)
Power: Medium Change, $\Delta\theta$ = 1.0									
Fixed-Length									
PB	.634	.700	.733	.750	.750	.731	.691	.604	.386
	(.25)	(.24)	(.23)	(.23)	(.23)	(.23)	(.24)	(.25)	(.21)
FB	.658	.669	.676	.678	.678	.677	.671	.655	.621
	(.25)	(.25)	(.25)	(.25)	(.25)	(.25)	(.25)	(.26)	(.28)
Variable-Length									
PB	.623	.690	.723	.740	.740	.722	.680	.588	.361
	(.25)	(.24)	(.23)	(.23)	(.23)	(.24)	(.25)	(.25)	(.20)
FB	.651	.660	.667	.669	.669	.668	.662	.646	.609
	(.25)	(.25)	(.25)	(.25)	(.25)	(.25)	(.26)	(.26)	(.28)

Continued on the next page

Table A10 (continued). Marginal Mean (SD) for Type I Error and Power Conditional on Bank Information and θ Level for Fixed- and Variable-Length AMC

Termination	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Power: Large Change, $\Delta\theta$ = 1.5									
Fixed-Length									
PB	.864	.894	.912	.920	.917	.902	.867	.778	.517
	(.17)	(.15)	(.13)	(.13)	(.13)	(.15)	(.18)	(.23)	(.25)
FB	.862	.869	.874	.879	.879	.877	.868	.850	.800
	(.17)	(.17)	(.17)	(.17)	(.17)	(.17)	(.18)	(.20)	(.24)
Variable-Length									
PB	.859	.890	.908	.917	.914	.898	.861	.765	.491
	(.17)	(.15)	(.13)	(.13)	(.13)	(.15)	(.18)	(.23)	(.25)
FB	.858	.865	.871	.875	.875	.873	.864	.845	.792
	(.17)	(.17)	(.17)	(.17)	(.17)	(.17)	(.18)	(.20)	(.24)

Table A11. Marginal Mean (SD) for Number of Items Used for Peaked and Flat Bank at Each θ Level for Fixed- and Variable-Length AMC

	θ								
Change	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Fixed-Length									
	31.67	31.67	31.67	31.67	31.67	31.67	31.67	31.67	31.67
	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)
Variable-Length									
No Change									
PB	31.26	31.31	31.33	31.35	31.36	31.35	31.34	31.33	31.37
	(14.08)	(14.10)	(14.12)	(14.13)	(14.14)	(14.13)	(14.12)	(14.11)	(14.11)
FB	31.31	31.32	31.34	31.34	31.35	31.35	31.34	31.33	31.33
	(14.10)	(14.11)	(14.12)	(14.12)	(14.13)	(14.12)	(14.12)	(14.11)	(14.10)
Very Small Change, $\Delta\theta$ = 0.25									
PB	30.98	30.87	30.83	30.81	30.80	30.80	30.83	30.92	31.25
	(13.83)	(13.71)	(13.67)	(13.65)	(13.64)	(13.64)	(13.68)	(13.75)	(14.00)
FB	30.85	30.87	30.90	30.92	30.93	30.93	30.92	30.90	30.93
	(13.69)	(13.71)	(13.73)	(13.75)	(13.75)	(13.76)	(13.75)	(13.74)	(13.74)
Small Change, $\Delta\theta$ = 0.5									
PB	29.59	28.89	28.52	28.32	28.27	28.36	28.65	29.32	30.69
	(12.72)	(12.12)	(11.84)	(11.70)	(11.67)	(11.74)	(11.95)	(12.47)	(13.51)
FB	28.94	28.95	28.96	28.98	29.02	29.00	28.98	29.04	29.24
	(12.13)	(12.15)	(12.17)	(12.20)	(12.25)	(12.24)	(12.23)	(12.26)	(12.38)
Medium Change, $\Delta\theta$ = 1.0									
PB	22.42	20.97	20.27	19.89	19.81	20.18	21.26	23.88	28.99
	(8.69)	(7.96)	(7.60)	(7.38)	(7.35)	(7.65)	(8.28)	(9.38)	(12.12)
FB	21.53	21.46	21.45	21.49	21.45	21.43	21.56	22.00	23.07
	(8.17)	(8.25)	(8.33)	(8.45)	(8.50)	(8.51)	(8.55)	(8.76)	(9.19)
Large Change, $\Delta\theta$ = 1.5									
PB	16.39	15.59	15.13	14.78	14.56	15.02	16.38	19.90	27.62
	(5.43)	(4.99)	(4.64)	(4.16)	(3.81)	(4.39)	(5.63)	(7.35)	(11.04)
FB	16.15	16.09	16.09	15.98	15.81	15.85	16.15	16.92	18.83
	(5.12)	(5.13)	(5.22)	(5.17)	(5.19)	(5.30)	(5.52)	(6.01)	(6.98)

Table A12. Marginal Mean (SD) for Type I Error and Power Conditional on Bank Size for Fixed- and Variable-Length AMC

Termination	Fixed-Length AMC		Variable-Length AMC	
	300-Item	500-Item	300-Item	500-Item
Type I Error	.051	.051	.049	.049
	(.013)	(.012)	(.011)	(.010)
Power: Very Small Change $\Delta\theta = 0.25$	.104	.112	.099	.106
	(.050)	(.055)	(.046)	(.051)
Power: Small Change $\Delta\theta = 0.5$	.272	.297	.261	.285
	(.168)	(.182)	(.162)	(.177)
Power: Medium Change $\Delta\theta = 1.0$	.649	.680	.638	.670
	(.260)	(.254)	(.262)	(.256)
Power: Large Change $\Delta\theta = 1.5$	.840	.863	.834	.857
	(.206)	(.190)	(.209)	(.193)

Table A13. Marginal Mean (SD) for Type I Error and Power Conditional on Discrimination for Fixed- and Variable-Length AMC

Termination	Fixed-Length AMC			Variable-Length AMC		
	LD	MD	HD	LD	MD	HD
Type I Error	.048	.051	.054	.047	.049	.050
	(.013)	(.011)	(.012)	(.010)	(.009)	(.011)
Power: Very Small Change $\Delta\theta = 0.25$	.071	.101	.153	.069	.096	.142
	(.022)	(.033)	(.057)	(.020)	(.031)	(.055)
Power: Small Change $\Delta\theta = 0.5$	.147	.266	.441	.142	.255	.422
	(.060)	(.112)	(.180)	(.056)	(.108)	(.180)
Power: Medium Change $\Delta\theta = 1.0$	.435	.697	.862	.425	.685	.852
	(.180)	(.214)	(.169)	(.175)	(.216)	(.177)
Power: Large Change $\Delta\theta = 1.5$	.710	.893	.952	.702	.887	.948
	(.221)	(.156)	(.116)	(.220)	(.160)	(.123)

Table A14. Marginal Mean (SD) for Number of Items Used for Discrimination for Fixed- and Variable-Length AMC

Termination	Discrimination		
	LD	MD	HD
Fixed-Length AMC	31.67	31.67	31.67
	(14.34)	(14.34)	(14.34)
Variable-Length AMC			
No Change	31.34	31.33	31.33
	(14.12)	(14.11)	(14.11)
Very Small Change $\Delta\theta = 0.25$	31.17	30.96	30.57
	(13.97)	(13.78)	(13.43)
Small Change $\Delta\theta = 0.5$	30.51	29.35	27.09
	(13.40)	(12.38)	(10.46)
Medium Change $\Delta\theta = 1.0$	26.94	21.80	16.78
	(10.45)	(6.83)	(5.00)
Large Change $\Delta\theta = 1.5$	21.28	15.68	13.58
	(6.83)	(4.77)	(5.12)

Table A15. Marginal Mean (SD) for Type I Error and Power Conditional on Item Selection for Fixed- and Variable-Length AMC

Termination	Fixed-Length AMC			Variable-Length AMC		
	FI	KL-F	KL-M	FI	KL-F	KL-M
Type I Error	.050	.051	.052	.048	.049	.049
	(.014)	(.012)	(.011)	(.011)	(.010)	(.010)
Power: Very Small Change $\Delta\theta = 0.25$	.104	.109	.111	.098	.104	.105
	(.053)	(.052)	(.052)	(.048)	(.049)	(.049)
Power: Small Change $\Delta\theta = 0.5$	.274	.289	.291	.261	.278	.280
	(.176)	(.176)	(.174)	(.169)	(.171)	(.170)
Power: Medium Change $\Delta\theta = 1.0$	.649	.672	.673	.637	.662	.662
	(.265)	(.255)	(.253)	(.266)	(.256)	(.254)
Power: Large Change $\Delta\theta = 1.5$	.838	.858	.858	.831	.853	.852
	(.208)	(.194)	(.193)	(.211)	(.197)	(.197)

Table A16. Marginal Mean (SD) for Type I Error and Power Conditional on Test Length for Fixed- and Variable-Length AMC

Termination	Fixed-Length AMC			Variable-Length AMC		
	15	30	50	15	30	50
Type I Error	.052	.051	.050	.047	.049	.050
	(.016)	(.011)	(.009)	(.012)	(.010)	(.009)
Power: Very Small Change $\Delta\theta = 0.25$	.076	.107	.142	.069	.102	.136
	(.029)	(.041)	(.060)	(.023)	(.036)	(.055)
Power: Small Change $\Delta\theta = 0.5$	.161	.287	.406	.150	.276	.393
	(.079)	(.140)	(.191)	(.069)	(.133)	(.186)
Power: Medium Change $\Delta\theta = 1.0$	.458	.706	.830	.443	.696	.823
	(.214)	(.225)	(.177)	(.206)	(.227)	(.182)
Power: Large Change $\Delta\theta = 1.5$	.701	.895	.959	.691	.890	.956
	(.236)	(.142)	(.080)	(.235)	(.147)	(.085)

Table A17. Marginal Mean (SD) for Number of Items

Used Conditional on Test Length for Variable-Length AMC

Termination	Maximum Test Length		
	15	30	50
No Change	14.93	29.70	49.38
	(.03)	(.11)	(.23)
Very Small Change $\Delta\theta = 0.25$	14.89	29.40	48.41
	(.05)	(.30)	(.82)
Small Change $\Delta\theta = 0.5$	14.75	28.09	44.12
	(.15)	(1.23)	(3.80)
Medium Change $\Delta\theta = 1.0$	14.01	22.30	29.21
	(.64)	(4.40)	(9.84)

Table A18. Marginal Mean (SD) for Type I Error and Power Conditional on Test Length and Discrimination for Fixed- and Variable-Length AMC

Termination	Fixed-Length AMC			Variable-Length AMC		
	15	30	50	15	30	50
Type I Error						
High Discrimination (a = 1.5)	.057 (.015)	.053 (.011)	.051 (.010)	.049 (.012)	.050 (.010)	.050 (.010)
Moderate Discrimination (a = 1.0)	.052 (.014)	.051 (.010)	.050 (.009)	.047 (.010)	.049 (.009)	.050 (.009)
Low Discrimination (a = 0.6)	.048 (.016)	.048 (.012)	.049 (.010)	.045 (.012)	.048 (.010)	.049 (.009)
Power: Very Small Change $\Delta\theta$ = 0.25						
High Discrimination (a = 1.5)	.098 (.027)	.150 (.030)	.210 (.044)	.087 (.021)	.142 (.026)	.198 (.041)
Moderate Discrimination (a = 1.0)	.073 (.023)	.099 (.021)	.131 (.024)	.067 (.017)	.095 (.018)	.127 (.022)
Low Discrimination (a = 0.6)	.057 (.022)	.070 (.019)	.085 (.018)	.054 (.017)	.069 (.016)	.083 (.015)
Power: Small Change $\Delta\theta$ = 0.5						
High Discrimination (a = 1.5)	.243 (.064)	.453 (.081)	.626 (.115)	.222 (.054)	.435 (.077)	.608 (.115)
Moderate Discrimination (a = 1.0)	.149 (.045)	.264 (.052)	.386 (.070)	.139 (.037)	.254 (.046)	.373 (.066)
Low Discrimination (a = 0.6)	.093 (.035)	.145 (.037)	.204 (.043)	.088 (.029)	.140 (.032)	.198 (.038)
Power: Medium Change $\Delta\theta$ = 1.0						
High Discrimination (a = 1.5)	.684 (.137)	.929 (.106)	.973 (.087)	.662 (.135)	.923 (.112)	.970 (.094)
Moderate Discrimination (a = 1.0)	.446 (.115)	.746 (.104)	.897 (.088)	.431 (.107)	.734 (.104)	.890 (.091)
Low Discrimination (a = 0.6)	.242 (.085)	.442 (.094)	.620 (.099)	.235 (.075)	.431 (.088)	.607 (.097)
Power: Large Change $\Delta\theta$ = 1.5						
High Discrimination (a = 1.5)	.895 (.151)	.976 (.090)	.987 (.064)	.885 (.159)	.973 (.098)	.985 (.071)
Moderate Discrimination (a = 1.0)	.743 (.163)	.950 (.094)	.985 (.056)	.732 (.163)	.946 (.099)	.984 (.059)
Low Discrimination (a = 0.6)	.465 (.148)	.759 (.125)	.905 (.088)	.456 (.141)	.750 (.125)	.899 (.091)

Table A19. Marginal Mean (SD) for Number of Items Used Conditional on Test Length and Discrimination for Variable-Length AMC

Termination/Discrimination	Maximum Test Length		
	15	30	50
No Change			
High Discrimination (*a* = 1.5)	14.93	29.69	49.38
	(.03)	(.13)	(.27)
Moderate Discrimination (*a* = 1.0)	14.93	29.70	49.38
	(.02)	(.10)	(.22)
Low Discrimination (*a* = 0.6)	14.93	29.70	49.38
	(.02)	(.10)	(.20)
Very Small Change $\Delta\theta = 0.25$			
High Discrimination (*a* = 1.5)	14.87	29.18	47.66
	(.06)	(.32)	(.79)
Moderate Discrimination (*a* = 1.0)	14.90	29.44	48.55
	(.04)	(.22)	(.52)
Low Discrimination (*a* = 0.6)	14.91	29.58	49.02
	(.03)	(.17)	(.39)
Small Change $\Delta\theta = 0.5$			
High Discrimination (*a* = 1.5)	14.63	26.80	39.84
	(.17)	(1.11)	(3.04)
Moderate Discrimination (*a* = 1.0)	14.77	28.34	44.95
	(.10)	(.61)	(1.58)
Low Discrimination (*a* = 0.6)	14.85	29.12	47.56
	(.06)	(.36)	(.90)
Medium Change $\Delta\theta = 1.0$			
High Discrimination (*a* = 1.5)	13.39	17.66	19.30
	(.62)	(3.33)	(6.71)
Moderate Discrimination (*a* = 1.0)	14.09	22.63	28.67
	(.37)	(2.29)	(5.20)
Low Discrimination (*a* = 0.6)	14.55	26.62	39.65
	(.20)	(1.28)	(3.26)
Large Change $\Delta\theta = 1.5$			
High Discrimination (*a* = 1.5)	12.18	13.72	14.83
	(.96)	(4.14)	(7.55)
Moderate Discrimination (*a* = 1.0)	13.03	16.43	17.59
	(.77)	(3.68)	(6.54)
Low Discrimination (*a* = 0.6)	13.96	22.12	27.76
	(.46)	(2.76)	(5.99)

Table A20. Marginal Mean (SD) for Type I Error and Power Conditional on Test Statistic for Fixed- and Variable-Length AMC

	Test Statistic				
	Z	LR	ST	KLD-U	KLD-N
Type I Error					
Fixed-Length	.050	.056	.059	.050	.041
	(.011)	(.009)	(.010)	(.011)	(.013)
Variable-Length	.050	.050	.050	.049	.044
	(.012)	(.008)	(.007)	(.010)	(.012)
Power: Very Small Change $\Delta\theta = 0.25$					
Fixed-Length	.112	.121	.126	.098	.084
	(.050)	(.049)	(.049)	(.050)	(.053)
Variable-Length	.109	.110	.109	.096	.088
	(.045)	(.047)	(.047)	(.048)	(.053)
Power: Small Change $\Delta\theta = 0.5$					
Fixed-Length	.293	.308	.316	.266	.241
	(.172)	(.170)	(.169)	(.175)	(.180)
Variable-Length	.284	.288	.287	.261	.246
	(.163)	(.166)	(.166)	(.171)	(.179)
Power: Medium Change $\Delta\theta = 1.0$					
Fixed-Length	.675	.696	.703	.643	.606
	(.247)	(.235)	(.228)	(.269)	(.292)
Variable-Length	.668	.676	.675	.639	.611
	(.246)	(.243)	(.241)	(.267)	(.289)
Power: Large Change $\Delta\theta = 1.5$					
Fixed-Length	.857	.879	.881	.836	.805
	(.190)	(.166)	(.152)	(.213)	(.247)
Variable-Length	.853	.867	.864	.834	.809
	(.192)	(.176)	(.169)	(.213)	(.243)

Table A21. Marginal Mean (SD) for Number of Items Used Conditional on Test Statistic for Variable-Length AMC

Termination	Test Statistic				
	Z	LR	ST	KLD-U	KLD-N
Fixed-Length AMC	31.67	31.67	31.67	31.67	31.67
	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)
Variable-Length AMC					
No Change	31.30	31.30	31.30	31.39	31.39
	(14.08)	(14.09)	(14.08)	(14.16)	(14.16)
Very Small Change $\Delta\theta = 0.25$	30.77	30.73	30.89	31.06	31.06
	(13.64)	(13.60)	(13.74)	(13.84)	(13.84)
Small Change $\Delta\theta = 0.5$	28.74	28.53	28.92	29.33	29.41
	(12.06)	(11.92)	(12.24)	(12.40)	(12.48)
Medium Change $\Delta\theta = 1.0$	21.48	20.85	21.64	22.42	22.82
	(8.44)	(8.32)	(8.75)	(8.97)	(9.35)
Large Change $\Delta\theta = 1.5$	16.79	15.69	16.72	17.20	17.84
	(6.29)	(5.62)	(6.08)	(6.80)	(7.45)

Table A22. Marginal Mean (SD) for Type I Error and Power Conditional on Test Statistic and θ for Fixed-Length AMC

Test		θ								
		−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error										
	Z	.059	.049	.047	.046	.046	.047	.049	.051	.054
		(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
	LR	.059	.058	.056	.056	.055	.055	.055	.055	.054
		(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
	ST	.061	.060	.060	.060	.059	.059	.058	.057	.056
		(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
	KLD-U	.047	.053	.052	.050	.049	.049	.051	.051	.045
		(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
	KLD-N	.046	.047	.045	.043	.040	.039	.038	.037	.032
		(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Power: Very Small Change, $\Delta\theta = 0.25$										
	Z	.106	.109	.113	.116	.117	.116	.115	.112	.103
		(.04)	(.05)	(.05)	(.06)	(.06)	(.06)	(.05)	(.05)	(.04)
	LR	.108	.118	.125	.129	.130	.129	.125	.118	.106
		(.04)	(.04)	(.05)	(.05)	(.06)	(.06)	(.05)	(.05)	(.04)
	ST	.111	.123	.130	.134	.136	.135	.131	.122	.110
		(.04)	(.04)	(.05)	(.05)	(.06)	(.06)	(.05)	(.05)	(.04)
	KLD-U	.082	.101	.105	.106	.105	.104	.102	.098	.079
		(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.04)
	KLD-N	.079	.092	.096	.095	.092	.088	.083	.075	.058
		(.04)	(.05)	(.05)	(.06)	(.06)	(.06)	(.05)	(.05)	(.04)
Power: Small Change, $\Delta\theta = 0.5$										
	Z	.259	.287	.306	.317	.320	.316	.305	.284	.241
		(.14)	(.17)	(.18)	(.19)	(.19)	(.18)	(.18)	(.16)	(.14)
	LR	.267	.304	.325	.338	.340	.336	.322	.295	.248
		(.14)	(.16)	(.18)	(.18)	(.18)	(.18)	(.17)	(.16)	(.14)
	ST	.272	.311	.334	.347	.350	.346	.331	.301	.255
		(.14)	(.16)	(.18)	(.18)	(.18)	(.18)	(.17)	(.15)	(.14)
	KLD-U	.225	.269	.286	.294	.295	.292	.281	.257	.194
		(.16)	(.17)	(.18)	(.19)	(.19)	(.18)	(.17)	(.16)	(.15)
	KLD-N	.218	.256	.271	.276	.271	.261	.244	.213	.156
		(.15)	(.17)	(.18)	(.19)	(.19)	(.19)	(.19)	(.17)	(.15)
Power: Medium Change, $\Delta\theta = 1.0$										
	Z	.652	.685	.706	.719	.722	.715	.695	.651	.532
		(.24)	(.24)	(.24)	(.23)	(.23)	(.23)	(.24)	(.24)	(.27)
	LR	.672	.709	.730	.741	.741	.733	.711	.665	.560
		(.23)	(.23)	(.22)	(.22)	(.22)	(.22)	(.23)	(.23)	(.25)
	ST	.674	.712	.734	.747	.749	.741	.719	.673	.578
		(.23)	(.22)	(.22)	(.22)	(.22)	(.22)	(.22)	(.23)	(.23)
	KLD-U	.621	.664	.685	.695	.697	.691	.670	.613	.450
		(.27)	(.26)	(.25)	(.25)	(.25)	(.25)	(.25)	(.27)	(.28)
	KLD-N	.613	.650	.666	.669	.661	.643	.610	.546	.398
		(.27)	(.27)	(.27)	(.27)	(.28)	(.28)	(.29)	(.30)	(.29)

Continued on the next page

Table A22 (continued). Marginal Mean (SD) for Type I Error and Power Conditional on Test Statistic and θ for Fixed-Length AMC

Test		θ								
		−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Power: Large Change, $\Delta\theta = 1.5$										
	Z	.862	.880	.894	.905	.905	.898	.879	.826	.667
		(.16)	(.15)	(.14)	(.14)	(.14)	(.14)	(.16)	(.20)	(.29)
	LR	.884	.902	.912	.916	.915	.908	.891	.850	.729
		(.15)	(.13)	(.13)	(.13)	(.13)	(.13)	(.15)	(.18)	(.25)
	ST	.873	.890	.903	.916	.918	.912	.895	.859	.763
		(.15)	(.14)	(.13)	(.12)	(.12)	(.13)	(.14)	(.16)	(.21)
	KLD-U	.849	.873	.886	.892	.892	.885	.862	.795	.588
		(.19)	(.17)	(.16)	(.16)	(.16)	(.16)	(.18)	(.23)	(.29)
	KLD-N	.845	.863	.870	.868	.860	.844	.811	.739	.546
		(.19)	(.18)	(.18)	(.19)	(.20)	(.22)	(.24)	(.28)	(.31)

Table A23. Marginal Mean (SD) for Type I Error and Power Conditional on Test Statistic and θ for Variable-Length AMC

Test		θ								
		−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error										
	Z	.061	.050	.046	.045	.044	.046	.048	.052	.055
		(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
	LR	.053	.053	.052	.051	.050	.049	.049	.049	.047
		(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
	ST	.051	.051	.052	.052	.051	.051	.050	.048	.046
		(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
	KLD-U	.046	.052	.051	.050	.049	.049	.050	.050	.044
		(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
	KLD-N	.051	.051	.048	.045	.043	.042	.041	.040	.036
		(.02)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Power: Very Small Change, $\Delta\theta$ = 0.25										
	Z	.103	.103	.106	.109	.110	.112	.114	.116	.107
		(.04)	(.04)	(.05)	(.05)	(.05)	(.05)	(.05)	(.04)	(.04)
	LR	.099	.109	.115	.118	.119	.118	.114	.108	.094
		(.04)	(.04)	(.05)	(.05)	(.05)	(.05)	(.05)	(.04)	(.04)
	ST	.096	.107	.114	.118	.119	.118	.113	.105	.092
		(.03)	(.04)	(.05)	(.05)	(.05)	(.05)	(.05)	(.04)	(.04)
	KLD-U	.080	.099	.103	.104	.104	.103	.101	.095	.074
		(.04)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.04)	(.04)
	KLD-N	.089	.100	.101	.099	.095	.090	.084	.076	.058
		(.04)	(.05)	(.05)	(.06)	(.06)	(.06)	(.05)	(.05)	(.04)
Power: Small Change, $\Delta\theta$ = 0.5										
	Z	.247	.273	.291	.303	.307	.307	.302	.287	.240
		(.14)	(.16)	(.17)	(.18)	(.18)	(.17)	(.17)	(.15)	(.14)
	LR	.249	.285	.306	.318	.320	.315	.302	.274	.225
		(.14)	(.16)	(.17)	(.18)	(.18)	(.18)	(.17)	(.15)	(.14)
	ST	.245	.283	.305	.318	.321	.316	.300	.271	.224
		(.14)	(.16)	(.17)	(.18)	(.18)	(.18)	(.17)	(.15)	(.13)
	KLD-U	.221	.264	.281	.290	.291	.287	.277	.250	.185
		(.15)	(.17)	(.18)	(.18)	(.18)	(.18)	(.17)	(.16)	(.14)
	KLD-N	.232	.267	.279	.282	.275	.263	.245	.213	.155
		(.15)	(.17)	(.18)	(.19)	(.19)	(.19)	(.18)	(.17)	(.14)
Power: Medium Change, $\Delta\theta$ = 1.0										
	Z	.641	.675	.696	.710	.715	.711	.694	.650	.523
		(.24)	(.24)	(.24)	(.23)	(.23)	(.23)	(.23)	(.24)	(.27)
	LR	.653	.691	.712	.723	.724	.715	.692	.643	.531
		(.24)	(.24)	(.23)	(.23)	(.23)	(.23)	(.23)	(.24)	(.25)
	ST	.646	.686	.709	.722	.724	.715	.691	.641	.539
		(.24)	(.24)	(.23)	(.23)	(.23)	(.23)	(.23)	(.24)	(.24)
	KLD-U	.619	.663	.684	.694	.696	.688	.666	.604	.437
		(.27)	(.26)	(.25)	(.25)	(.25)	(.25)	(.25)	(.26)	(.28)
	KLD-N	.626	.661	.674	.675	.665	.646	.612	.546	.396
		(.26)	(.26)	(.26)	(.27)	(.27)	(.28)	(.29)	(.30)	(.29)

Continued on the next page

Table A23 (continued). Marginal Mean (SD) for Type I Error and Power Conditional on Test Statistic and θ for Variable-Length AMC

Test		θ								
		−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Power: Large Change, $\Delta\theta$ = 1.5										
	Z	.856	.874	.890	.903	.904	.898	.879	.823	.653
		(.17)	(.15)	(.14)	(.14)	(.14)	(.14)	(.16)	(.21)	(.29)
	LR	.875	.894	.904	.908	.907	.900	.880	.834	.704
		(.15)	(.14)	(.14)	(.13)	(.14)	(.14)	(.15)	(.19)	(.26)
	ST	.856	.875	.889	.902	.906	.899	.880	.839	.730
		(.16)	(.15)	(.14)	(.13)	(.14)	(.14)	(.15)	(.18)	(.23)
	KLD-U	.851	.875	.887	.893	.893	.885	.860	.790	.576
		(.19)	(.17)	(.16)	(.15)	(.15)	(.16)	(.18)	(.23)	(.29)
	KLD-N	.853	.870	.875	.873	.864	.847	.812	.739	.544
		(.18)	(.17)	(.18)	(.18)	(.20)	(.21)	(.24)	(.27)	(.31)

Table A24. Marginal Mean (SD) for Number of Items Used Conditional on Test Statistic and θ for Variable-Length AMC

Test	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Fixed-Length AMC									
	31.67	31.67	31.67	31.67	31.67	31.67	31.67	31.67	31.67
	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)	(14.34)
Variable-Length AMC									
Type I Error: No Change									
Z	31.08	31.28	31.36	31.38	31.39	31.37	31.33	31.25	31.23
	(13.93)	(14.07)	(14.13)	(14.15)	(14.15)	(14.14)	(14.11)	(14.05)	(14.02)
LR	31.28	31.27	31.28	31.29	31.30	31.30	31.30	31.32	31.35
	(14.08)	(14.07)	(14.07)	(14.08)	(14.09)	(14.09)	(14.09)	(14.10)	(14.11)
ST	31.32	31.30	31.28	31.27	31.28	31.28	31.29	31.31	31.34
	(14.11)	(14.09)	(14.07)	(14.06)	(14.07)	(14.07)	(14.07)	(14.09)	(14.10)
KLD-U	31.43	31.38	31.38	31.38	31.39	31.38	31.38	31.38	31.43
	(14.21)	(14.16)	(14.16)	(14.16)	(14.16)	(14.16)	(14.16)	(14.15)	(14.17)
KLD-N	31.31	31.34	31.38	31.41	31.42	31.42	31.41	31.39	31.38
	(14.13)	(14.14)	(14.16)	(14.18)	(14.19)	(14.18)	(14.17)	(14.15)	(14.12)
Power: Very Small Change, $\Delta\theta = 0.25$									
Z	30.86	30.97	30.97	30.92	30.87	30.76	30.59	30.43	30.56
	(13.76)	(13.81)	(13.80)	(13.76)	(13.72)	(13.63)	(13.50)	(13.36)	(13.42)
LR	30.82	30.72	30.67	30.64	30.63	30.64	30.67	30.76	30.98
	(13.69)	(13.60)	(13.55)	(13.53)	(13.52)	(13.53)	(13.55)	(13.63)	(13.78)
ST	30.96	30.86	30.79	30.76	30.76	30.80	30.88	31.00	31.16
	(13.81)	(13.72)	(13.66)	(13.63)	(13.64)	(13.67)	(13.74)	(13.83)	(13.97)
KLD-U	31.15	31.02	31.02	31.01	31.00	30.99	31.00	31.07	31.31
	(13.92)	(13.80)	(13.80)	(13.79)	(13.78)	(13.78)	(13.79)	(13.85)	(14.03)
KLD-N	30.78	30.78	30.88	30.97	31.06	31.14	31.22	31.31	31.43
	(13.64)	(13.62)	(13.69)	(13.77)	(13.83)	(13.89)	(13.96)	(14.03)	(14.13)
Power: Small Change, $\Delta\theta = 0.5$									
Z	29.41	29.25	29.05	28.85	28.67	28.39	28.12	28.04	28.83
	(12.64)	(12.45)	(12.28)	(12.14)	(12.01)	(11.81)	(11.59)	(11.49)	(12.03)
LR	29.02	28.54	28.25	28.09	28.06	28.11	28.30	28.77	29.61
	(12.27)	(11.89)	(11.69)	(11.60)	(11.59)	(11.62)	(11.73)	(12.05)	(12.70)
ST	29.33	28.85	28.55	28.39	28.39	28.53	28.83	29.37	30.09
	(12.50)	(12.12)	(11.92)	(11.84)	(11.86)	(11.95)	(12.15)	(12.55)	(13.13)
KLD-U	29.68	29.27	29.11	29.00	28.96	28.95	29.08	29.50	30.44
	(12.65)	(12.31)	(12.19)	(12.11)	(12.09)	(12.10)	(12.20)	(12.53)	(13.28)
KLD-N	28.88	28.67	28.74	28.91	29.16	29.42	29.76	30.23	30.87
	(12.07)	(11.88)	(11.94)	(12.08)	(12.26)	(12.45)	(12.69)	(13.06)	(13.60)

Continued on the next page

Table A24 (continued). Marginal Mean (SD) for Number of Items Used Conditional on Test Statistic and θ for Variable-Length AMC

Test	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Power: Medium Change, $\Delta\theta = 1.0$									
Z	22.85	22.19	21.70	21.26	20.61	20.03	19.94	20.78	24.00
	(9.12)	(8.75)	(8.47)	(8.21)	(7.86)	(7.59)	(7.43)	(7.58)	(9.75)
LR	21.28	20.33	19.84	19.61	19.59	19.81	20.42	21.89	24.86
	(8.15)	(7.78)	(7.61)	(7.55)	(7.56)	(7.63)	(7.86)	(8.44)	(10.46)
ST	21.93	20.97	20.48	20.22	20.23	20.61	21.41	23.08	25.79
	(8.37)	(8.01)	(7.85)	(7.82)	(7.90)	(8.09)	(8.43)	(9.14)	(11.12)
KLD-U	22.69	21.84	21.40	21.14	21.01	21.12	21.76	23.54	27.24
	(8.64)	(8.18)	(8.03)	(7.97)	(7.97)	(8.07)	(8.42)	(9.37)	(11.66)
KLD-N	21.14	20.75	20.87	21.22	21.72	22.45	23.52	25.42	28.28
	(7.75)	(7.65)	(7.85)	(8.14)	(8.48)	(8.85)	(9.35)	(10.20)	(12.09)
Power: Large Change, $\Delta\theta = 1.5$									
Z	17.92	17.52	17.14	16.28	14.91	14.56	14.89	16.48	21.42
	(6.76)	(6.67)	(6.43)	(5.49)	(4.19)	(4.05)	(4.24)	(5.00)	(9.01)
LR	15.21	14.68	14.40	14.29	14.35	14.59	15.27	17.01	21.38
	(4.50)	(4.13)	(3.99)	(3.95)	(4.00)	(4.16)	(4.68)	(5.78)	(9.17)
ST	16.41	15.94	15.62	15.29	15.18	15.44	16.22	18.13	22.25
	(4.90)	(4.65)	(4.48)	(4.30)	(4.36)	(4.67)	(5.32)	(6.47)	(9.84)
KLD-U	16.58	15.95	15.58	15.33	15.28	15.54	16.49	19.12	24.93
	(5.13)	(4.73)	(4.60)	(4.56)	(4.61)	(4.83)	(5.63)	(7.34)	(10.79)
KLD-N	15.23	15.12	15.33	15.67	16.21	17.05	18.47	21.33	26.16
	(4.24)	(4.25)	(4.55)	(4.97)	(5.46)	(6.03)	(6.93)	(8.22)	(11.20)

Table A25 . Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.097	.044	.035	.036	.04	.042	.044	.05	.05
LR	.084	.074	.071	.068	.066	.064	.062	.059	.055
ST	.077	.073	.076	.076	.075	.073	.07	.064	.059
KLD-U	.064	.058	.048	.043	.044	.043	.046	.052	.048
KLD-N	.077	.062	.046	.039	.038	.035	.034	.035	.031
Moderate Discrimination (a = 1.0)									
Z	.074	.05	.039	.04	.041	.045	.046	.054	.054
LR	.076	.071	.065	.063	.061	.061	.059	.059	.052
ST	.076	.073	.072	.072	.069	.068	.065	.062	.059
KLD-U	.042	.046	.044	.042	.042	.044	.046	.048	.032
KLD-N	.046	.044	.038	.035	.032	.031	.029	.027	.019
Low Discrimination (a = 0.6)									
Z	.059	.058	.051	.046	.046	.049	.053	.059	.051
LR	.069	.069	.066	.061	.059	.059	.058	.057	.045
ST	.075	.073	.071	.068	.065	.064	.061	.059	.056
KLD-U	.026	.036	.043	.043	.044	.046	.046	.037	.019
KLD-N	.023	.027	.03	.027	.026	.025	.022	.017	.009
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.114	.074	.071	.079	.086	.091	.090	.096	.067
LR	.115	.117	.123	.130	.130	.128	.119	.110	.083
ST	.106	.117	.130	.141	.145	.144	.131	.118	.094
KLD-U	.058	.061	.058	.064	.065	.068	.070	.080	.059
KLD-N	.077	.071	.061	.059	.054	.050	.046	.049	.036
Moderate Discrimination (a = 1.0)									
Z	.086	.066	.062	.065	.069	.072	.076	.081	.062
LR	.093	.094	.096	.097	.097	.095	.091	.087	.066
ST	.093	.098	.104	.107	.108	.105	.099	.091	.080
KLD-U	.037	.046	.047	.049	.051	.055	.059	.058	.036
KLD-N	.047	.050	.045	.041	.037	.034	.031	.028	.018
Low Discrimination (a = 0.6)									
Z	.069	.069	.062	.059	.061	.063	.070	.071	.055
LR	.078	.082	.079	.076	.077	.075	.075	.069	.050
ST	.083	.086	.085	.083	.083	.081	.079	.072	.068
KLD-U	.024	.036	.042	.044	.047	.049	.051	.039	.021
KLD-N	.024	.030	.032	.029	.026	.024	.020	.015	.008
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.196	.172	.194	.222	.250	.254	.246	.233	.109
LR	.230	.262	.295	.321	.334	.322	.296	.253	.152
ST	.218	.263	.302	.339	.356	.346	.316	.265	.176
KLD-U	.114	.134	.154	.178	.193	.198	.202	.199	.110
KLD-N	.146	.157	.164	.169	.165	.154	.142	.136	.070

Continued on the next page

Table A25 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.140	.126	.134	.152	.159	.166	.162	.154	.091
LR	.162	.178	.191	.206	.206	.203	.185	.164	.105
ST	.161	.184	.204	.222	.223	.219	.196	.171	.135
KLD-U	.061	.087	.098	.110	.116	.124	.126	.113	.060
KLD-N	.082	.095	.096	.096	.087	.081	.071	.058	.031
Low Discrimination (a = 0.6)									
Z	.098	.104	.102	.105	.110	.112	.117	.109	.071
LR	.111	.124	.127	.130	.132	.127	.123	.108	.067
ST	.116	.130	.136	.140	.141	.134	.128	.114	.096
KLD-U	.031	.054	.067	.073	.080	.082	.081	.061	.029
KLD-N	.035	.050	.054	.049	.046	.040	.032	.023	.011
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.496	.534	.589	.686	.771	.760	.712	.589	.167
LR	.653	.731	.792	.826	.839	.815	.760	.633	.297
ST	.604	.673	.737	.812	.850	.830	.773	.645	.358
KLD-U	.439	.512	.589	.656	.703	.704	.675	.567	.221
KLD-N	.517	.574	.613	.637	.649	.622	.576	.460	.154
Moderate Discrimination (a = 1.0)									
Z	.348	.385	.443	.502	.527	.516	.482	.380	.134
LR	.443	.509	.564	.593	.594	.566	.510	.413	.202
ST	.438	.512	.574	.613	.614	.584	.523	.431	.291
KLD-U	.248	.324	.376	.418	.444	.445	.415	.312	.128
KLD-N	.295	.348	.371	.375	.363	.332	.285	.200	.075
Low Discrimination (a = 0.6)									
Z	.227	.258	.281	.301	.308	.310	.292	.218	.103
LR	.264	.308	.333	.346	.345	.332	.301	.229	.112
ST	.273	.319	.348	.363	.360	.344	.307	.255	.193
KLD-U	.108	.174	.208	.234	.246	.246	.214	.138	.056
KLD-N	.119	.161	.172	.170	.155	.134	.104	.062	.023
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.686	.727	.803	.951	.976	.960	.928	.731	.175
LR	.917	.954	.977	.986	.986	.980	.955	.830	.372
ST	.702	.674	.719	.890	.983	.982	.958	.848	.459
KLD-U	.761	.829	.904	.954	.967	.962	.935	.779	.281
KLD-N	.849	.889	.918	.944	.949	.936	.894	.690	.203
Moderate Discrimination (a = 1.0)									
Z	.599	.659	.770	.854	.861	.836	.767	.510	.132
LR	.785	.844	.881	.899	.893	.861	.791	.621	.272
ST	.735	.796	.864	.904	.901	.870	.799	.668	.428
KLD-U	.591	.678	.752	.800	.813	.794	.718	.508	.184
KLD-N	.651	.707	.740	.754	.735	.686	.582	.375	.117
Low Discrimination (a = 0.6)									
Z	.451	.513	.568	.593	.593	.571	.492	.304	.110
LR	.526	.590	.627	.638	.627	.589	.519	.361	.147
ST	.532	.602	.641	.653	.640	.597	.530	.433	.302
KLD-U	.321	.422	.481	.512	.520	.490	.399	.239	.087
KLD-N	.330	.392	.415	.403	.372	.316	.232	.124	.039

Table A26. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.095	.047	.040	.040	.044	.045	.047	.051	.052
LR	.070	.063	.061	.060	.059	.059	.058	.056	.053
ST	.072	.068	.069	.067	.066	.065	.063	.059	.055
KLD-U	.065	.067	.059	.056	.056	.054	.057	.062	.054
KLD-N	.072	.064	.054	.048	.047	.044	.043	.042	.034
Moderate Discrimination (a = 1.0)									
Z	.074	.053	.044	.043	.044	.047	.049	.056	.055
LR	.066	.063	.059	.058	.055	.057	.056	.056	.051
ST	.069	.067	.065	.065	.061	.062	.060	.058	.056
KLD-U	.044	.054	.055	.052	.051	.054	.056	.056	.036
KLD-N	.045	.048	.045	.042	.040	.039	.036	.031	.021
Low Discrimination (a = 0.6)									
Z	.061	.060	.052	.048	.048	.050	.054	.060	.053
LR	.063	.062	.060	.057	.056	.056	.056	.056	.045
ST	.069	.067	.065	.062	.061	.060	.058	.056	.053
KLD-U	.030	.041	.050	.051	.053	.054	.053	.041	.020
KLD-N	.025	.030	.033	.032	.031	.029	.025	.018	.010
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.127	.098	.098	.104	.103	.102	.096	.098	.069
LR	.103	.109	.119	.125	.126	.124	.117	.109	.084
ST	.104	.117	.131	.137	.138	.137	.130	.118	.096
KLD-U	.074	.094	.094	.098	.096	.093	.090	.092	.066
KLD-N	.077	.087	.086	.087	.082	.076	.068	.062	.042
Moderate Discrimination (a = 1.0)									
Z	.098	.082	.078	.076	.077	.077	.079	.082	.065
LR	.085	.088	.090	.090	.093	.091	.090	.085	.068
ST	.087	.093	.098	.099	.102	.100	.097	.090	.081
KLD-U	.047	.067	.073	.071	.071	.069	.071	.066	.041
KLD-N	.047	.058	.060	.057	.054	.049	.043	.036	.022
Low Discrimination (a = 0.6)									
Z	.074	.076	.070	.065	.065	.067	.071	.072	.058
LR	.070	.076	.075	.073	.073	.074	.074	.070	.051
ST	.076	.080	.080	.078	.079	.079	.076	.072	.067
KLD-U	.028	.045	.056	.058	.059	.059	.057	.044	.022
KLD-N	.025	.034	.039	.037	.035	.031	.026	.018	.009
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.245	.250	.272	.291	.293	.277	.258	.238	.113
LR	.228	.270	.306	.327	.333	.320	.299	.258	.157
ST	.232	.284	.325	.348	.355	.342	.320	.272	.183
KLD-U	.179	.229	.251	.267	.270	.256	.242	.220	.122
KLD-N	.174	.216	.237	.247	.243	.219	.195	.162	.083

Continued on the next page

Table A26 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.170	.167	.172	.179	.177	.178	.166	.158	.095
LR	.156	.176	.191	.202	.202	.202	.184	.166	.110
ST	.159	.186	.203	.216	.218	.217	.197	.174	.139
KLD-U	.092	.139	.152	.158	.157	.156	.145	.126	.068
KLD-N	.092	.123	.131	.134	.125	.114	.095	.072	.039
Low Discrimination (a = 0.6)									
Z	.113	.120	.117	.116	.116	.116	.118	.110	.075
LR	.104	.120	.124	.126	.128	.126	.122	.109	.072
ST	.109	.124	.130	.134	.137	.133	.126	.116	.100
KLD-U	.041	.074	.094	.099	.099	.098	.090	.066	.032
KLD-N	.039	.059	.068	.068	.062	.054	.042	.028	.013
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.681	.740	.786	.808	.806	.775	.719	.596	.176
LR	.687	.763	.811	.835	.838	.815	.762	.640	.309
ST	.694	.775	.822	.847	.852	.831	.779	.655	.372
KLD-U	.644	.713	.763	.786	.789	.762	.713	.591	.242
KLD-N	.627	.701	.749	.767	.758	.718	.646	.502	.181
Moderate Discrimination (a = 1.0)									
Z	.459	.505	.542	.560	.554	.530	.486	.386	.146
LR	.458	.526	.571	.594	.592	.565	.512	.421	.220
ST	.466	.540	.588	.612	.611	.584	.529	.441	.309
KLD-U	.374	.465	.504	.523	.518	.496	.445	.336	.144
KLD-N	.356	.434	.469	.478	.456	.410	.334	.233	.094
Low Discrimination (a = 0.6)									
Z	.273	.303	.316	.322	.323	.317	.294	.224	.116
LR	.263	.308	.333	.343	.344	.331	.302	.237	.126
ST	.271	.318	.346	.357	.358	.342	.310	.260	.207
KLD-U	.153	.239	.276	.287	.286	.273	.232	.148	.062
KLD-N	.140	.196	.218	.216	.200	.167	.125	.073	.028
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.937	.964	.977	.983	.977	.961	.929	.740	.186
LR	.937	.964	.979	.986	.985	.979	.956	.837	.388
ST	.936	.964	.980	.987	.987	.982	.960	.856	.478
KLD-U	.931	.961	.975	.981	.979	.971	.944	.797	.307
KLD-N	.924	.956	.971	.977	.973	.959	.916	.726	.238
Moderate Discrimination (a = 1.0)									
Z	.799	.842	.868	.880	.871	.839	.772	.526	.145
LR	.806	.854	.882	.897	.889	.860	.796	.637	.296
ST	.811	.860	.889	.904	.898	.869	.806	.682	.458
KLD-U	.761	.818	.848	.861	.852	.820	.743	.536	.209
KLD-N	.738	.796	.824	.830	.804	.745	.632	.420	.148
Low Discrimination (a = 0.6)									
Z	.532	.582	.605	.614	.606	.577	.497	.316	.127
LR	.530	.595	.626	.637	.625	.589	.525	.379	.168
ST	.538	.606	.639	.651	.638	.599	.538	.447	.331
KLD-U	.419	.525	.563	.574	.562	.516	.419	.254	.096
KLD-N	.379	.458	.484	.475	.435	.362	.264	.142	.047

Table A27. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.095	.049	.042	.043	.045	.046	.048	.053	.052
LR	.068	.062	.060	.058	.057	.056	.056	.055	.053
ST	.070	.065	.065	.064	.062	.061	.060	.057	.055
KLD-U	.066	.073	.068	.064	.061	.060	.062	.067	.056
KLD-N	.073	.070	.060	.055	.051	.049	.048	.046	.036
Moderate Discrimination (a = 1.0)									
Z	.076	.054	.044	.045	.046	.049	.051	.056	.054
LR	.065	.061	.057	.056	.055	.056	.055	.054	.050
ST	.067	.064	.061	.061	.060	.060	.058	.055	.055
KLD-U	.048	.059	.060	.059	.057	.059	.060	.060	.038
KLD-N	.049	.051	.048	.046	.044	.043	.040	.034	.023
Low Discrimination (a = 0.6)									
Z	.063	.059	.053	.050	.050	.051	.056	.060	.050
LR	.060	.059	.058	.056	.055	.054	.054	.055	.042
ST	.065	.062	.062	.059	.058	.057	.055	.054	.052
KLD-U	.032	.044	.054	.056	.057	.057	.056	.044	.022
KLD-N	.027	.032	.036	.036	.034	.031	.028	.021	.011
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.128	.102	.102	.107	.109	.106	.101	.100	.067
LR	.102	.106	.115	.120	.124	.121	.115	.108	.083
ST	.102	.112	.124	.130	.134	.132	.126	.116	.094
KLD-U	.080	.111	.112	.112	.109	.104	.100	.098	.067
KLD-N	.081	.096	.097	.097	.094	.087	.079	.069	.044
Moderate Discrimination (a = 1.0)									
Z	.100	.084	.081	.082	.082	.080	.081	.083	.061
LR	.083	.085	.087	.089	.092	.089	.086	.084	.065
ST	.084	.088	.093	.095	.099	.096	.092	.088	.079
KLD-U	.052	.077	.083	.083	.082	.077	.076	.071	.042
KLD-N	.051	.063	.066	.065	.063	.055	.049	.040	.024
Low Discrimination (a = 0.6)									
Z	.077	.077	.071	.069	.068	.068	.073	.072	.052
LR	.068	.072	.072	.072	.072	.072	.072	.068	.047
ST	.072	.075	.076	.076	.076	.076	.074	.070	.066
KLD-U	.032	.051	.063	.067	.066	.064	.061	.046	.023
KLD-N	.028	.037	.042	.043	.039	.034	.029	.020	.010
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.250	.257	.281	.303	.301	.286	.265	.239	.110
LR	.228	.265	.300	.323	.327	.314	.291	.253	.156
ST	.229	.276	.315	.339	.346	.334	.311	.267	.181
KLD-U	.196	.260	.281	.296	.291	.276	.258	.227	.122
KLD-N	.184	.233	.258	.272	.263	.242	.217	.172	.085

Continued on the next page

Table A27 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.172	.173	.178	.187	.185	.184	.172	.157	.087
LR	.151	.173	.186	.198	.199	.199	.182	.163	.107
ST	.153	.179	.196	.209	.212	.211	.192	.170	.138
KLD-U	.103	.159	.171	.178	.176	.170	.155	.130	.070
KLD-N	.098	.133	.144	.149	.141	.128	.107	.080	.042
Low Discrimination (a = 0.6)									
Z	.114	.124	.120	.121	.121	.120	.119	.104	.064
LR	.100	.116	.120	.125	.127	.124	.119	.106	.064
ST	.103	.119	.124	.130	.134	.130	.123	.113	.099
KLD-U	.047	.084	.104	.111	.112	.106	.096	.070	.033
KLD-N	.041	.063	.073	.074	.069	.060	.047	.032	.014
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.686	.742	.789	.810	.807	.775	.720	.587	.168
LR	.680	.754	.801	.826	.829	.805	.753	.633	.305
ST	.684	.762	.809	.837	.841	.820	.769	.649	.368
KLD-U	.677	.742	.785	.802	.802	.773	.723	.592	.241
KLD-N	.643	.720	.766	.782	.776	.735	.667	.505	.182
Moderate Discrimination (a = 1.0)									
Z	.466	.513	.550	.565	.561	.533	.491	.370	.125
LR	.453	.518	.563	.585	.584	.555	.507	.413	.210
ST	.458	.528	.576	.600	.600	.572	.522	.434	.305
KLD-U	.406	.496	.532	.547	.542	.509	.459	.339	.145
KLD-N	.373	.452	.490	.499	.478	.433	.356	.243	.096
Low Discrimination (a = 0.6)									
Z	.278	.310	.323	.329	.327	.321	.292	.199	.090
LR	.258	.304	.326	.338	.338	.326	.301	.226	.108
ST	.262	.310	.335	.349	.350	.335	.307	.258	.205
KLD-U	.170	.260	.297	.308	.302	.286	.241	.152	.062
KLD-N	.147	.206	.231	.232	.214	.184	.139	.079	.028
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.936	.962	.977	.982	.977	.960	.925	.730	.178
LR	.932	.960	.976	.984	.983	.976	.952	.834	.382
ST	.930	.957	.976	.984	.985	.979	.957	.853	.470
KLD-U	.940	.964	.977	.982	.979	.971	.944	.796	.304
KLD-N	.928	.957	.973	.978	.974	.961	.917	.727	.237
Moderate Discrimination (a = 1.0)									
Z	.800	.841	.869	.879	.869	.837	.762	.490	.120
LR	.797	.844	.875	.889	.881	.851	.790	.626	.281
ST	.799	.849	.881	.895	.889	.860	.799	.676	.450
KLD-U	.781	.833	.860	.870	.858	.823	.746	.536	.204
KLD-N	.746	.804	.834	.839	.815	.759	.645	.427	.145
Low Discrimination (a = 0.6)									
Z	.536	.587	.610	.618	.607	.574	.465	.264	.092
LR	.522	.585	.616	.628	.618	.584	.514	.354	.140
ST	.526	.593	.626	.640	.629	.591	.529	.442	.324
KLD-U	.444	.546	.584	.592	.575	.529	.425	.255	.093
KLD-N	.390	.471	.499	.493	.458	.384	.278	.145	.044

Table A28. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.038	.036	.038	.039	.039	.041	.042	.044	.046
LR	.080	.075	.071	.068	.065	.063	.064	.065	.062
ST	.080	.078	.076	.076	.074	.072	.072	.073	.069
KLD-U	.064	.054	.048	.046	.043	.042	.044	.047	.043
KLD-N	.080	.058	.045	.040	.035	.034	.032	.032	.027
Moderate Discrimination (a = 1.0)									
Z	.035	.036	.037	.040	.041	.042	.045	.047	.048
LR	.072	.069	.065	.064	.064	.062	.062	.063	.062
ST	.074	.074	.071	.072	.072	.070	.070	.070	.067
KLD-U	.046	.046	.042	.042	.041	.042	.043	.045	.032
KLD-N	.052	.044	.036	.033	.030	.028	.027	.025	.018
Low Discrimination (a = 0.6)									
Z	.038	.040	.040	.042	.041	.044	.046	.049	.051
LR	.065	.066	.063	.063	.060	.061	.061	.061	.061
ST	.071	.072	.071	.071	.068	.069	.068	.066	.064
KLD-U	.035	.040	.041	.041	.039	.041	.040	.034	.020
KLD-N	.033	.031	.027	.025	.021	.019	.017	.014	.010
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.066	.068	.072	.077	.079	.084	.084	.080	.079
LR	.119	.117	.118	.121	.121	.121	.119	.113	.110
ST	.120	.121	.124	.130	.134	.135	.133	.128	.124
KLD-U	.063	.060	.060	.061	.058	.062	.065	.064	.056
KLD-N	.084	.069	.061	.055	.047	.044	.040	.032	.024
Moderate Discrimination (a = 1.0)									
Z	.053	.055	.057	.062	.063	.065	.065	.066	.068
LR	.094	.093	.091	.094	.091	.092	.090	.088	.086
ST	.100	.099	.100	.104	.102	.104	.101	.097	.095
KLD-U	.043	.045	.045	.047	.047	.049	.051	.050	.036
KLD-N	.056	.049	.042	.038	.033	.029	.024	.020	.013
Low Discrimination (a = 0.6)									
Z	.049	.049	.051	.051	.053	.054	.056	.061	.059
LR	.078	.077	.077	.075	.074	.075	.073	.075	.070
ST	.084	.083	.085	.083	.083	.084	.081	.081	.075
KLD-U	.031	.036	.040	.039	.040	.041	.040	.035	.021
KLD-N	.035	.032	.028	.024	.020	.016	.013	.010	.007
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.152	.171	.181	.202	.215	.224	.221	.209	.188
LR	.246	.260	.270	.287	.294	.292	.285	.270	.246
ST	.249	.264	.279	.301	.314	.315	.310	.296	.271
KLD-U	.126	.138	.142	.158	.162	.170	.177	.169	.135
KLD-N	.161	.158	.148	.148	.136	.124	.113	.088	.059

Continued on the next page

Table A28 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.104	.113	.120	.128	.132	.138	.139	.137	.130
LR	.164	.172	.176	.179	.181	.183	.180	.174	.162
ST	.172	.181	.188	.194	.198	.200	.196	.189	.175
KLD-U	.074	.084	.088	.092	.094	.101	.106	.101	.070
KLD-N	.096	.093	.086	.077	.067	.059	.051	.038	.025
Low Discrimination (a = 0.6)									
Z	.078	.081	.084	.087	.086	.089	.093	.095	.092
LR	.116	.116	.117	.119	.116	.116	.118	.116	.110
ST	.123	.125	.127	.130	.128	.127	.129	.124	.117
KLD-U	.042	.053	.057	.060	.060	.063	.062	.053	.033
KLD-N	.051	.050	.044	.038	.029	.023	.018	.013	.010
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.498	.547	.584	.643	.702	.715	.691	.641	.572
LR	.663	.705	.745	.780	.786	.784	.763	.721	.664
ST	.620	.659	.707	.767	.799	.799	.783	.744	.693
KLD-U	.455	.500	.544	.593	.624	.644	.637	.589	.431
KLD-N	.532	.548	.561	.569	.559	.543	.494	.406	.267
Moderate Discrimination (a = 1.0)									
Z	.331	.359	.390	.426	.439	.448	.438	.417	.369
LR	.451	.481	.502	.521	.521	.521	.505	.479	.436
ST	.450	.483	.513	.541	.545	.545	.530	.501	.460
KLD-U	.264	.295	.320	.344	.356	.375	.372	.327	.209
KLD-N	.313	.316	.312	.298	.272	.250	.212	.162	.107
Low Discrimination (a = 0.6)									
Z	.200	.214	.222	.229	.235	.237	.249	.242	.210
LR	.270	.282	.288	.287	.289	.285	.291	.278	.249
ST	.279	.295	.304	.306	.308	.304	.306	.291	.270
KLD-U	.126	.150	.160	.167	.176	.178	.178	.138	.080
KLD-N	.143	.140	.127	.108	.091	.073	.061	.045	.031
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.788	.793	.826	.925	.964	.955	.935	.905	.804
LR	.910	.942	.965	.975	.979	.975	.967	.951	.910
ST	.700	.713	.759	.888	.972	.977	.970	.957	.925
KLD-U	.773	.823	.881	.925	.945	.946	.934	.891	.680
KLD-N	.847	.871	.891	.907	.914	.901	.863	.777	.548
Moderate Discrimination (a = 1.0)									
Z	.631	.661	.723	.777	.797	.792	.767	.724	.560
LR	.776	.807	.830	.842	.847	.840	.817	.784	.700
ST	.720	.761	.815	.850	.859	.853	.831	.798	.739
KLD-U	.586	.628	.672	.707	.733	.740	.711	.609	.372
KLD-N	.647	.653	.652	.639	.623	.585	.510	.400	.256
Low Discrimination (a = 0.6)									
Z	.417	.435	.457	.471	.481	.489	.482	.449	.333
LR	.517	.529	.540	.544	.545	.545	.533	.505	.423
ST	.520	.540	.559	.566	.566	.563	.549	.521	.483
KLD-U	.321	.348	.368	.383	.399	.402	.365	.273	.152
KLD-N	.334	.319	.297	.262	.232	.197	.158	.119	.081

Table A29. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.036	.038	.041	.043	.043	.043	.044	.046	.046
LR	.067	.064	.062	.062	.059	.057	.058	.059	.058
ST	.076	.071	.069	.068	.066	.064	.065	.065	.063
KLD-U	.075	.066	.061	.057	.053	.051	.054	.058	.055
KLD-N	.084	.065	.054	.049	.044	.041	.041	.041	.036
Moderate Discrimination (a = 1.0)									
Z	.038	.039	.041	.042	.044	.046	.045	.047	.050
LR	.063	.060	.059	.058	.058	.058	.057	.057	.058
ST	.069	.066	.065	.065	.064	.064	.063	.062	.061
KLD-U	.054	.055	.053	.051	.050	.052	.052	.053	.041
KLD-N	.056	.049	.044	.040	.038	.036	.033	.032	.024
Low Discrimination (a = 0.6)									
Z	.041	.042	.043	.044	.044	.046	.047	.050	.051
LR	.059	.060	.057	.058	.056	.058	.057	.057	.057
ST	.065	.066	.063	.064	.062	.063	.062	.061	.060
KLD-U	.040	.047	.048	.050	.048	.048	.047	.040	.024
KLD-N	.036	.035	.032	.030	.026	.025	.022	.018	.012
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.081	.086	.093	.095	.095	.092	.089	.084	.080
LR	.110	.110	.114	.115	.116	.114	.113	.108	.106
ST	.125	.124	.126	.126	.127	.127	.127	.122	.122
KLD-U	.094	.093	.093	.090	.086	.083	.084	.084	.077
KLD-N	.095	.086	.083	.078	.072	.066	.062	.057	.048
Moderate Discrimination (a = 1.0)									
Z	.066	.068	.070	.072	.071	.070	.068	.067	.066
LR	.088	.088	.087	.086	.087	.088	.084	.084	.082
ST	.097	.097	.096	.095	.095	.097	.094	.094	.092
KLD-U	.061	.067	.066	.066	.064	.064	.063	.062	.048
KLD-N	.062	.059	.055	.052	.047	.042	.038	.033	.024
Low Discrimination (a = 0.6)									
Z	.057	.056	.057	.056	.056	.058	.058	.060	.059
LR	.073	.071	.071	.069	.070	.071	.070	.071	.068
ST	.080	.078	.078	.076	.077	.079	.077	.077	.073
KLD-U	.040	.048	.052	.051	.051	.052	.049	.042	.027
KLD-N	.038	.037	.035	.030	.027	.025	.020	.016	.010
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.208	.229	.244	.256	.253	.248	.240	.222	.194
LR	.252	.267	.278	.288	.292	.290	.284	.271	.248
ST	.277	.288	.297	.306	.312	.312	.308	.297	.279
KLD-U	.212	.223	.228	.233	.228	.225	.223	.217	.182
KLD-N	.204	.210	.212	.212	.200	.188	.176	.159	.125

Continued on the next page

Table A29 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.136	.145	.151	.153	.152	.149	.145	.142	.129
LR	.164	.172	.172	.177	.178	.178	.175	.171	.161
ST	.178	.186	.186	.190	.193	.194	.190	.188	.179
KLD-U	.117	.130	.133	.133	.132	.129	.130	.126	.094
KLD-N	.114	.118	.114	.109	.101	.090	.083	.070	.052
Low Discrimination (a = 0.6)									
Z	.094	.096	.097	.096	.094	.093	.097	.095	.090
LR	.111	.113	.113	.114	.111	.112	.115	.112	.106
ST	.121	.122	.122	.124	.122	.122	.125	.121	.116
KLD-U	.059	.074	.080	.080	.079	.079	.078	.065	.042
KLD-N	.057	.059	.056	.051	.044	.038	.033	.024	.016
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.652	.700	.733	.751	.743	.738	.706	.656	.582
LR	.696	.733	.761	.783	.784	.782	.761	.726	.680
ST	.714	.746	.774	.797	.800	.799	.783	.754	.718
KLD-U	.650	.684	.707	.721	.718	.718	.701	.661	.541
KLD-N	.636	.670	.689	.697	.679	.664	.632	.568	.457
Moderate Discrimination (a = 1.0)									
Z	.433	.456	.469	.480	.471	.470	.452	.425	.383
LR	.474	.492	.507	.520	.519	.519	.504	.481	.452
ST	.493	.511	.527	.542	.541	.542	.529	.508	.484
KLD-U	.395	.420	.430	.440	.434	.436	.427	.384	.271
KLD-N	.379	.394	.394	.389	.365	.344	.312	.263	.186
Low Discrimination (a = 0.6)									
Z	.246	.255	.256	.253	.252	.248	.255	.244	.216
LR	.272	.282	.286	.285	.285	.281	.287	.277	.262
ST	.288	.298	.302	.302	.303	.299	.304	.294	.281
KLD-U	.181	.210	.218	.219	.220	.213	.209	.166	.103
KLD-N	.167	.171	.164	.150	.137	.118	.100	.075	.048
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.925	.950	.963	.969	.967	.955	.936	.906	.813
LR	.932	.953	.968	.975	.977	.974	.966	.953	.927
ST	.931	.953	.969	.977	.980	.977	.971	.961	.944
KLD-U	.926	.946	.958	.963	.966	.962	.953	.926	.785
KLD-N	.917	.939	.953	.956	.956	.946	.927	.887	.746
Moderate Discrimination (a = 1.0)									
Z	.773	.793	.806	.812	.809	.802	.773	.733	.615
LR	.797	.816	.830	.839	.840	.837	.816	.791	.745
ST	.807	.825	.840	.851	.853	.850	.832	.812	.777
KLD-U	.746	.766	.778	.784	.786	.786	.758	.681	.464
KLD-N	.727	.741	.745	.738	.720	.699	.641	.557	.385
Low Discrimination (a = 0.6)									
Z	.491	.497	.500	.500	.501	.503	.490	.459	.372
LR	.522	.530	.537	.541	.543	.544	.532	.514	.478
ST	.539	.547	.555	.560	.563	.563	.551	.535	.515
KLD-U	.422	.443	.452	.455	.458	.454	.411	.320	.190
KLD-N	.382	.379	.366	.343	.320	.286	.238	.177	.112

Table A30. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.038	.039	.042	.045	.045	.045	.046	.046	.049
LR	.064	.061	.060	.059	.057	.056	.056	.055	.056
ST	.072	.067	.065	.064	.062	.061	.062	.061	.060
KLD-U	.084	.076	.069	.064	.060	.057	.059	.062	.062
KLD-N	.092	.072	.060	.055	.049	.046	.045	.044	.043
Moderate Discrimination (a = 1.0)									
Z	.039	.041	.043	.044	.046	.048	.047	.049	.052
LR	.060	.058	.057	.055	.056	.056	.055	.055	.055
ST	.065	.063	.062	.060	.061	.061	.060	.059	.057
KLD-U	.059	.061	.059	.057	.056	.056	.057	.057	.046
KLD-N	.061	.054	.048	.043	.041	.040	.037	.035	.029
Low Discrimination (a = 0.6)									
Z	.042	.044	.044	.046	.046	.047	.049	.051	.054
LR	.057	.058	.056	.056	.054	.054	.055	.056	.056
ST	.061	.062	.060	.060	.058	.059	.059	.059	.057
KLD-U	.042	.050	.053	.054	.053	.053	.052	.043	.027
KLD-N	.040	.038	.034	.032	.028	.026	.024	.021	.015
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.084	.090	.097	.100	.100	.098	.095	.088	.086
LR	.108	.108	.111	.111	.112	.112	.111	.104	.104
ST	.121	.119	.120	.120	.122	.123	.123	.116	.118
KLD-U	.110	.110	.109	.104	.098	.095	.094	.091	.088
KLD-N	.104	.097	.093	.088	.083	.077	.073	.068	.062
Moderate Discrimination (a = 1.0)									
Z	.070	.072	.074	.075	.075	.074	.072	.072	.072
LR	.085	.085	.085	.084	.084	.084	.084	.082	.081
ST	.094	.092	.092	.091	.091	.092	.092	.090	.088
KLD-U	.069	.078	.077	.076	.073	.072	.071	.069	.054
KLD-N	.068	.065	.061	.058	.054	.049	.045	.042	.032
Low Discrimination (a = 0.6)									
Z	.059	.060	.060	.059	.059	.061	.060	.062	.061
LR	.069	.069	.069	.067	.068	.069	.068	.068	.067
ST	.074	.074	.074	.073	.073	.075	.074	.073	.071
KLD-U	.044	.054	.060	.058	.058	.059	.054	.046	.031
KLD-N	.040	.040	.038	.034	.030	.028	.024	.020	.015
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.216	.236	.252	.264	.264	.256	.246	.232	.202
LR	.246	.262	.273	.282	.289	.285	.278	.265	.242
ST	.268	.280	.289	.297	.305	.303	.298	.287	.270
KLD-U	.240	.252	.255	.256	.252	.246	.241	.234	.198
KLD-N	.220	.229	.231	.231	.223	.211	.197	.190	.149

Continued on the next page

Table A30 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.143	.153	.157	.158	.158	.155	.150	.148	.137
LR	.160	.168	.170	.172	.174	.174	.170	.167	.160
ST	.173	.179	.181	.183	.186	.188	.184	.180	.174
KLD-U	.133	.150	.151	.149	.148	.145	.141	.136	.105
KLD-N	.123	.129	.126	.120	.113	.105	.097	.090	.067
Low Discrimination (a = 0.6)									
Z	.099	.100	.102	.101	.098	.098	.101	.098	.093
LR	.108	.110	.110	.110	.109	.110	.112	.110	.106
ST	.115	.116	.117	.118	.118	.118	.120	.117	.114
KLD-U	.067	.085	.091	.092	.090	.089	.086	.071	.048
KLD-N	.062	.063	.061	.057	.050	.045	.039	.032	.023
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.655	.704	.736	.755	.746	.738	.709	.663	.574
LR	.685	.722	.750	.774	.773	.770	.749	.715	.661
ST	.700	.733	.760	.785	.787	.787	.768	.739	.698
KLD-U	.680	.710	.729	.741	.734	.732	.712	.679	.558
KLD-N	.653	.687	.705	.715	.700	.688	.658	.609	.507
Moderate Discrimination (a = 1.0)									
Z	.439	.462	.479	.490	.480	.476	.459	.435	.364
LR	.462	.482	.498	.512	.509	.508	.494	.472	.440
ST	.479	.497	.514	.528	.528	.528	.517	.494	.471
KLD-U	.423	.447	.461	.467	.459	.458	.442	.404	.289
KLD-N	.394	.409	.415	.412	.391	.371	.351	.301	.231
Low Discrimination (a = 0.6)									
Z	.254	.263	.263	.261	.260	.256	.260	.244	.203
LR	.267	.276	.279	.279	.279	.277	.283	.272	.253
ST	.278	.287	.292	.293	.294	.292	.296	.286	.276
KLD-U	.201	.232	.239	.240	.239	.230	.223	.181	.113
KLD-N									
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.922	.947	.962	.967	.966	.955	.936	.902	.777
LR	.925	.946	.963	.970	.974	.969	.961	.944	.914
ST	.924	.944	.963	.972	.976	.972	.966	.954	.932
KLD-U	.935	.951	.961	.965	.966	.962	.953	.926	.796
KLD-N	.920	.941	.955	.957	.957	.949	.934	.896	.790
Moderate Discrimination (a = 1.0)									
Z	.775	.794	.807	.814	.807	.802	.774	.720	.523
LR	.788	.806	.820	.830	.828	.826	.803	.779	.717
ST	.795	.814	.827	.839	.840	.838	.818	.798	.755
KLD-U	.765	.782	.793	.798	.795	.792	.764	.690	.479
KLD-N	.736	.751	.756	.753	.736	.720	.677	.596	.443
Low Discrimination (a = 0.6)									
Z	.500	.506	.506	.506	.508	.506	.486	.430	.305
LR	.514	.522	.527	.531	.534	.534	.522	.504	.440
ST	.527	.535	.542	.547	.550	.550	.539	.524	.501
KLD-U	.449	.472	.476	.477	.479	.468	.427	.338	.201
KLD-N	.395	.392	.383	.364	.343	.319	.276	.215	.141

Table A31. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.105	.057	.046	.045	.046	.047	.047	.052	.070
LR	.057	.060	.059	.057	.058	.058	.056	.053	.056
ST	.056	.060	.062	.061	.062	.062	.060	.054	.052
KLD-U	.039	.060	.048	.045	.045	.046	.046	.053	.062
KLD-N	.037	.054	.046	.042	.042	.042	.041	.043	.042
Moderate Discrimination (a = 1.0)									
Z	.079	.056	.048	.046	.047	.047	.047	.052	.064
LR	.060	.059	.056	.056	.057	.055	.053	.054	.055
ST	.060	.060	.059	.060	.061	.058	.056	.055	.052
KLD-U	.037	.052	.047	.045	.044	.045	.045	.052	.046
KLD-N	.034	.044	.042	.040	.039	.038	.036	.038	.032
Low Discrimination (a = 0.6)									
Z	.065	.059	.053	.049	.050	.049	.053	.057	.062
LR	.060	.059	.058	.055	.057	.054	.056	.056	.055
ST	.062	.062	.060	.058	.060	.057	.057	.056	.054
KLD-U	.026	.042	.049	.046	.048	.047	.051	.050	.031
KLD-N	.021	.032	.036	.036	.036	.034	.034	.030	.019
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.145	.143	.149	.160	.165	.162	.159	.156	.132
LR	.113	.154	.173	.185	.191	.186	.175	.156	.116
ST	.110	.156	.182	.195	.201	.195	.182	.157	.110
KLD-U	.079	.128	.125	.131	.134	.134	.133	.137	.093
KLD-N	.075	.120	.124	.128	.126	.121	.116	.115	.065
Moderate Discrimination (a = 1.0)									
Z	.107	.099	.102	.106	.108	.111	.107	.105	.108
LR	.092	.106	.116	.124	.124	.125	.116	.105	.095
ST	.092	.109	.121	.130	.131	.132	.120	.106	.090
KLD-U	.054	.083	.085	.087	.088	.091	.088	.086	.065
KLD-N	.053	.075	.079	.080	.078	.076	.069	.061	.044
Low Discrimination (a = 0.6)									
Z	.080	.080	.078	.078	.079	.081	.082	.082	.085
LR	.076	.082	.085	.087	.088	.087	.086	.080	.076
ST	.078	.085	.089	.091	.092	.091	.088	.080	.073
KLD-U	.032	.056	.066	.066	.066	.067	.068	.060	.036
KLD-N	.030	.047	.052	.053	.052	.048	.043	.035	.021
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.344	.426	.478	.508	.521	.513	.484	.427	.269
LR	.327	.455	.521	.550	.561	.547	.506	.419	.256
ST	.324	.462	.532	.562	.575	.560	.516	.417	.248
KLD-U	.281	.384	.426	.450	.464	.460	.440	.388	.201
KLD-N	.265	.382	.427	.443	.447	.432	.404	.343	.149

Continued on the next page

Table A31 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.224	.252	.284	.299	.312	.307	.287	.264	.220
LR	.214	.269	.311	.330	.340	.331	.300	.260	.198
ST	.216	.275	.321	.342	.352	.341	.307	.261	.192
KLD-U	.153	.218	.244	.254	.266	.264	.248	.224	.132
KLD-N	.147	.206	.235	.241	.243	.231	.206	.173	.093
Low Discrimination (a = 0.6)									
Z	.141	.153	.166	.171	.176	.175	.168	.162	.145
LR	.139	.159	.179	.187	.191	.186	.171	.157	.131
ST	.142	.163	.184	.194	.198	.191	.174	.156	.127
KLD-U	.068	.118	.141	.145	.148	.148	.139	.119	.064
KLD-N	.068	.101	.119	.123	.119	.111	.093	.072	.037
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.882	.926	.953	.975	.980	.974	.956	.884	.465
LR	.905	.960	.977	.983	.984	.978	.958	.871	.528
ST	.898	.953	.973	.983	.985	.979	.959	.865	.542
KLD-U	.871	.929	.955	.968	.971	.965	.946	.851	.416
KLD-N	.874	.935	.956	.966	.966	.958	.935	.810	.336
Moderate Discrimination (a = 1.0)									
Z	.676	.764	.808	.839	.839	.821	.773	.687	.465
LR	.690	.789	.831	.858	.857	.835	.778	.673	.464
ST	.692	.794	.837	.864	.863	.841	.781	.668	.468
KLD-U	.624	.723	.766	.800	.803	.786	.735	.618	.308
KLD-N	.612	.718	.758	.784	.777	.748	.679	.538	.249
Low Discrimination (a = 0.6)									
Z	.406	.472	.516	.541	.545	.524	.485	.428	.335
LR	.411	.489	.538	.564	.565	.536	.486	.414	.320
ST	.417	.498	.547	.572	.573	.542	.488	.411	.318
KLD-U	.312	.423	.470	.494	.498	.476	.428	.320	.150
KLD-N	.295	.387	.432	.447	.437	.396	.329	.224	.102
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.918	.935	.981	.999	1.000	1.000	.999	.977	.510
LR	.996	.999	1.000	1.000	1.000	1.000	.999	.976	.648
ST	.946	.924	.955	.995	1.000	1.000	.999	.973	.685
KLD-U	.989	.996	.999	1.000	1.000	1.000	.998	.966	.516
KLD-N	.994	.998	.999	1.000	1.000	1.000	.998	.948	.431
Moderate Discrimination (a = 1.0)									
Z	.946	.968	.986	.992	.991	.985	.968	.905	.564
LR	.969	.986	.991	.993	.993	.987	.967	.899	.646
ST	.966	.984	.991	.994	.993	.987	.967	.895	.690
KLD-U	.950	.974	.983	.988	.987	.980	.958	.851	.443
KLD-N	.952	.975	.982	.985	.983	.972	.939	.798	.386
Low Discrimination (a = 0.6)									
Z	.767	.831	.861	.870	.868	.836	.783	.692	.492
LR	.778	.843	.873	.881	.877	.841	.779	.678	.510
ST	.781	.847	.877	.884	.880	.843	.778	.673	.531
KLD-U	.714	.797	.832	.842	.840	.804	.724	.536	.248
KLD-N	.688	.770	.806	.808	.792	.732	.623	.429	.194

Table A32. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.103	.057	.047	.047	.046	.047	.049	.052	.070
LR	.055	.055	.055	.055	.054	.054	.054	.052	.056
ST	.054	.056	.058	.058	.058	.058	.056	.052	.051
KLD-U	.041	.065	.056	.053	.051	.051	.052	.056	.064
KLD-N	.038	.056	.052	.049	.047	.047	.046	.045	.044
Moderate Discrimination (a = 1.0)									
Z	.077	.056	.048	.048	.047	.048	.048	.054	.064
LR	.057	.056	.053	.055	.053	.053	.052	.053	.054
ST	.057	.057	.055	.058	.056	.056	.054	.054	.051
KLD-U	.038	.057	.052	.051	.050	.050	.051	.056	.048
KLD-N	.035	.046	.045	.045	.043	.043	.041	.041	.033
Low Discrimination (a = 0.6)									
Z	.064	.058	.052	.049	.050	.050	.053	.057	.062
LR	.057	.056	.055	.053	.055	.053	.054	.055	.055
ST	.060	.058	.057	.056	.057	.055	.056	.055	.053
KLD-U	.028	.046	.053	.051	.052	.051	.054	.053	.033
KLD-N	.024	.034	.039	.040	.040	.038	.036	.033	.021
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.146	.154	.162	.174	.175	.170	.163	.157	.132
LR	.109	.152	.173	.188	.190	.185	.175	.157	.116
ST	.108	.155	.181	.196	.198	.193	.183	.159	.110
KLD-U	.084	.155	.156	.162	.161	.155	.148	.142	.095
KLD-N	.076	.137	.148	.156	.153	.144	.134	.120	.068
Moderate Discrimination (a = 1.0)									
Z	.108	.104	.106	.112	.114	.114	.110	.105	.108
LR	.089	.103	.112	.121	.124	.123	.116	.104	.094
ST	.089	.106	.116	.126	.129	.129	.120	.106	.091
KLD-U	.060	.099	.101	.104	.103	.102	.097	.090	.068
KLD-N	.054	.083	.091	.095	.093	.089	.079	.067	.046
Low Discrimination (a = 0.6)									
Z	.081	.081	.080	.080	.082	.081	.083	.083	.084
LR	.073	.079	.082	.085	.088	.086	.085	.080	.075
ST	.076	.081	.085	.088	.091	.089	.086	.080	.073
KLD-U	.036	.065	.075	.075	.075	.073	.072	.065	.039
KLD-N	.032	.050	.058	.060	.058	.054	.048	.038	.023
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.358	.459	.512	.537	.541	.530	.494	.430	.271
LR	.329	.465	.530	.556	.564	.550	.511	.422	.258
ST	.327	.471	.541	.567	.576	.562	.520	.420	.251
KLD-U	.310	.452	.497	.515	.517	.504	.468	.396	.206
KLD-N	.281	.432	.487	.505	.503	.484	.440	.354	.156

Continued on the next page

Table A32 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.235	.268	.301	.316	.323	.316	.293	.266	.220
LR	.214	.270	.312	.332	.340	.332	.303	.262	.198
ST	.214	.274	.321	.342	.350	.341	.309	.262	.193
KLD-U	.178	.258	.285	.296	.299	.290	.265	.232	.137
KLD-N	.159	.229	.267	.278	.277	.260	.227	.183	.101
Low Discrimination (a = 0.6)									
Z	.146	.158	.174	.179	.182	.179	.168	.161	.144
LR	.136	.156	.178	.186	.190	.185	.170	.155	.131
ST	.139	.160	.183	.191	.195	.190	.172	.155	.127
KLD-U	.079	.136	.162	.165	.164	.160	.145	.122	.068
KLD-N	.073	.109	.132	.137	.134	.124	.103	.078	.043
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.918	.963	.978	.982	.982	.975	.957	.885	.470
LR	.915	.965	.979	.984	.984	.978	.958	.873	.533
ST	.914	.966	.980	.985	.985	.979	.960	.868	.547
KLD-U	.916	.962	.976	.980	.980	.971	.950	.856	.425
KLD-N	.904	.959	.974	.978	.977	.967	.942	.817	.347
Moderate Discrimination (a = 1.0)									
Z	.705	.790	.825	.848	.846	.825	.777	.689	.471
LR	.699	.797	.835	.858	.856	.835	.781	.675	.470
ST	.701	.801	.840	.864	.862	.840	.784	.671	.476
KLD-U	.684	.778	.812	.832	.829	.805	.748	.627	.321
KLD-N	.650	.760	.799	.818	.810	.776	.700	.555	.267
Low Discrimination (a = 0.6)									
Z	.419	.490	.530	.550	.552	.530	.486	.429	.337
LR	.412	.493	.540	.562	.564	.538	.485	.416	.324
ST	.417	.498	.546	.569	.571	.543	.488	.413	.322
KLD-U	.348	.467	.509	.524	.522	.494	.439	.330	.160
KLD-N	.314	.413	.464	.478	.468	.421	.348	.242	.116
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	1.000	1.000	1.000	1.000	.999	.977	.516
LR	.998	.999	1.000	1.000	1.000	1.000	.999	.976	.653
ST	.997	.999	1.000	1.000	1.000	1.000	.999	.974	.692
KLD-U	.998	.999	1.000	1.000	1.000	1.000	.998	.967	.525
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.998	.950	.444
Moderate Discrimination (a = 1.0)									
Z	.972	.986	.991	.992	.991	.985	.968	.907	.572
LR	.972	.986	.991	.993	.992	.986	.968	.901	.655
ST	.971	.987	.992	.994	.992	.987	.968	.897	.699
KLD-U	.969	.984	.989	.991	.989	.982	.960	.858	.458
KLD-N	.963	.982	.988	.989	.987	.976	.944	.812	.409
Low Discrimination (a = 0.6)									
Z	.781	.839	.868	.874	.871	.839	.783	.694	.498
LR	.780	.843	.873	.880	.876	.841	.778	.681	.520
ST	.783	.846	.876	.883	.879	.843	.778	.677	.538
KLD-U	.754	.826	.854	.858	.852	.813	.732	.550	.263
KLD-N	.712	.793	.826	.830	.812	.752	.641	.456	.217

Table A33. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.103	.057	.048	.048	.048	.049	.049	.054	.070
LR	.054	.055	.054	.054	.053	.054	.052	.052	.056
ST	.054	.055	.056	.056	.056	.056	.054	.052	.051
KLD-U	.040	.067	.058	.056	.053	.054	.055	.059	.065
KLD-N	.037	.057	.054	.052	.049	.050	.048	.048	.044
Moderate Discrimination (a = 1.0)									
Z	.077	.057	.048	.049	.048	.049	.050	.055	.065
LR	.056	.055	.051	.053	.052	.053	.052	.053	.055
ST	.056	.056	.053	.056	.054	.055	.054	.053	.051
KLD-U	.039	.059	.055	.055	.052	.053	.054	.058	.049
KLD-N	.035	.048	.047	.048	.046	.045	.044	.043	.033
Low Discrimination (a = 0.6)									
Z	.064	.058	.053	.050	.051	.051	.054	.058	.063
LR	.054	.055	.054	.052	.053	.052	.054	.055	.054
ST	.057	.056	.056	.054	.055	.054	.054	.054	.051
KLD-U	.030	.048	.056	.054	.054	.054	.056	.055	.033
KLD-N	.026	.036	.041	.041	.041	.040	.038	.034	.021
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.147	.157	.166	.178	.178	.173	.167	.159	.132
LR	.109	.152	.172	.186	.187	.182	.174	.157	.116
ST	.107	.153	.177	.193	.194	.188	.180	.158	.110
KLD-U	.085	.164	.166	.173	.170	.164	.156	.145	.096
KLD-N	.076	.143	.157	.165	.162	.153	.142	.124	.069
Moderate Discrimination (a = 1.0)									
Z	.108	.107	.108	.116	.116	.116	.113	.105	.108
LR	.088	.103	.111	.120	.122	.121	.116	.103	.094
ST	.087	.104	.115	.124	.126	.126	.120	.104	.090
KLD-U	.062	.106	.108	.111	.109	.108	.102	.093	.068
KLD-N	.055	.087	.096	.100	.098	.094	.085	.070	.047
Low Discrimination (a = 0.6)									
Z	.083	.082	.082	.082	.083	.083	.083	.083	.084
LR	.072	.078	.081	.084	.086	.085	.083	.079	.075
ST	.074	.079	.084	.086	.088	.086	.084	.079	.072
KLD-U	.038	.068	.080	.080	.079	.077	.074	.065	.040
KLD-N	.034	.052	.060	.062	.062	.057	.050	.040	.024
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.358	.464	.515	.540	.546	.533	.496	.429	.271
LR	.328	.462	.525	.551	.560	.546	.507	.420	.257
ST	.325	.466	.534	.560	.570	.556	.515	.419	.251
KLD-U	.312	.467	.511	.528	.530	.515	.477	.397	.206
KLD-N	.282	.443	.498	.516	.517	.497	.451	.356	.156

Continued on the next page

Table A33 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.236	.272	.305	.320	.327	.319	.296	.267	.218
LR	.213	.269	.310	.328	.337	.328	.301	.261	.198
ST	.212	.271	.316	.336	.345	.336	.306	.261	.192
KLD-U	.184	.271	.299	.308	.311	.300	.273	.234	.138
KLD-N	.163	.239	.277	.289	.290	.273	.238	.188	.103
Low Discrimination (a = 0.6)									
Z	.146	.161	.177	.182	.185	.181	.170	.161	.143
LR	.134	.156	.177	.185	.188	.183	.170	.154	.130
ST	.136	.158	.180	.188	.192	.186	.172	.153	.127
KLD-U	.084	.144	.171	.173	.172	.165	.151	.124	.069
KLD-N	.075	.113	.139	.144	.141	.130	.108	.081	.045
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.918	.962	.977	.981	.981	.975	.957	.885	.470
LR	.912	.963	.978	.982	.982	.976	.957	.872	.532
ST	.911	.964	.978	.983	.983	.977	.958	.867	.547
KLD-U	.919	.963	.976	.980	.980	.972	.951	.856	.425
KLD-N	.905	.960	.974	.978	.978	.968	.943	.817	.347
Moderate Discrimination (a = 1.0)									
Z	.707	.790	.826	.848	.846	.826	.776	.689	.469
LR	.695	.791	.831	.854	.852	.831	.777	.674	.469
ST	.695	.794	.835	.858	.856	.835	.780	.670	.475
KLD-U	.695	.786	.819	.838	.834	.810	.751	.629	.322
KLD-N	.655	.764	.804	.823	.816	.784	.706	.558	.270
Low Discrimination (a = 0.6)									
Z	.423	.491	.533	.551	.554	.530	.487	.430	.332
LR	.409	.488	.536	.556	.559	.533	.482	.414	.320
ST	.412	.491	.541	.562	.565	.537	.484	.410	.320
KLD-U	.359	.479	.521	.533	.531	.501	.443	.335	.164
KLD-N	.318	.422	.471	.487	.477	.431	.356	.250	.123
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	1.000	1.000	1.000	1.000	.999	.977	.515
LR	.997	.999	1.000	1.000	1.000	1.000	.999	.976	.652
ST	.997	.999	1.000	1.000	1.000	1.000	.999	.974	.691
KLD-U	.998	.999	1.000	1.000	1.000	1.000	.998	.967	.525
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.998	.950	.444
Moderate Discrimination (a = 1.0)									
Z	.971	.986	.990	.992	.991	.985	.968	.906	.565
LR	.970	.986	.990	.993	.991	.985	.967	.900	.652
ST	.969	.986	.990	.993	.992	.986	.967	.896	.697
KLD-U	.971	.985	.989	.991	.989	.982	.960	.858	.459
KLD-N	.964	.983	.987	.990	.987	.977	.944	.814	.412
Low Discrimination (a = 0.6)									
Z	.781	.839	.866	.874	.870	.838	.783	.690	.480
LR	.775	.839	.868	.877	.872	.838	.776	.678	.510
ST	.777	.841	.870	.879	.874	.839	.775	.674	.535
KLD-U	.760	.832	.858	.863	.856	.816	.736	.554	.267
KLD-N	.715	.796	.830	.834	.818	.758	.650	.465	.224

Table A34. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.046	.046	.047	.047	.046	.048	.047	.048	.048
LR	.058	.059	.058	.058	.056	.058	.057	.058	.059
ST	.062	.062	.062	.062	.060	.062	.061	.062	.062
KLD-U	.051	.049	.047	.045	.044	.046	.045	.048	.048
KLD-N	.052	.049	.045	.042	.040	.041	.039	.040	.038
Moderate Discrimination (a = 1.0)									
Z	.047	.046	.047	.046	.048	.046	.047	.047	.047
LR	.058	.057	.057	.056	.057	.056	.057	.056	.056
ST	.062	.061	.061	.060	.061	.060	.060	.060	.060
KLD-U	.045	.045	.046	.044	.044	.044	.045	.046	.043
KLD-N	.045	.042	.042	.039	.038	.035	.035	.034	.030
Low Discrimination (a = 0.6)									
Z	.046	.046	.046	.045	.046	.046	.047	.046	.048
LR	.056	.056	.056	.055	.056	.056	.055	.054	.056
ST	.059	.059	.059	.059	.060	.060	.059	.058	.059
KLD-U	.038	.042	.044	.043	.043	.044	.044	.041	.032
KLD-N	.036	.035	.034	.031	.030	.029	.027	.024	.021
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.130	.134	.136	.138	.140	.141	.140	.136	.136
LR	.154	.157	.157	.158	.162	.161	.160	.157	.156
ST	.162	.164	.164	.165	.170	.170	.169	.166	.165
KLD-U	.112	.112	.111	.110	.112	.114	.115	.113	.112
KLD-N	.118	.115	.110	.106	.103	.101	.097	.090	.083
Moderate Discrimination (a = 1.0)									
Z	.089	.091	.095	.094	.092	.093	.094	.094	.092
LR	.106	.107	.110	.110	.108	.109	.110	.108	.107
ST	.112	.114	.116	.116	.114	.116	.116	.114	.114
KLD-U	.071	.074	.077	.074	.072	.074	.077	.078	.071
KLD-N	.076	.074	.073	.067	.062	.059	.057	.054	.044
Low Discrimination (a = 0.6)									
Z	.067	.068	.068	.068	.066	.067	.068	.070	.068
LR	.079	.080	.080	.080	.079	.080	.080	.081	.078
ST	.083	.084	.084	.084	.084	.085	.086	.085	.082
KLD-U	.046	.054	.055	.055	.053	.054	.056	.054	.040
KLD-N	.048	.049	.046	.043	.038	.035	.031	.028	.019
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.393	.407	.421	.434	.433	.436	.435	.426	.404
LR	.436	.447	.458	.470	.471	.472	.469	.462	.441
ST	.448	.458	.468	.480	.484	.484	.482	.476	.454
KLD-U	.350	.358	.368	.373	.378	.384	.384	.377	.356
KLD-N	.364	.367	.368	.365	.359	.354	.343	.323	.289

Continued on the next page

Table A34 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.232	.241	.245	.248	.246	.242	.250	.242	.238
LR	.262	.271	.273	.277	.276	.272	.277	.270	.263
ST	.272	.280	.282	.287	.287	.283	.288	.279	.274
KLD-U	.191	.201	.204	.205	.203	.203	.213	.207	.186
KLD-N	.201	.202	.197	.190	.178	.169	.167	.151	.124
Low Discrimination (a = 0.6)									
Z	.133	.134	.134	.135	.133	.135	.135	.132	.135
LR	.152	.154	.153	.154	.153	.154	.154	.149	.152
ST	.158	.160	.160	.162	.161	.162	.161	.157	.158
KLD-U	.096	.107	.108	.108	.107	.110	.109	.102	.080
KLD-N	.100	.099	.093	.086	.078	.071	.063	.052	.039
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.919	.933	.943	.950	.953	.952	.948	.941	.930
LR	.939	.948	.955	.960	.961	.959	.956	.951	.942
ST	.936	.945	.953	.960	.963	.961	.959	.954	.945
KLD-U	.901	.913	.923	.930	.936	.936	.932	.925	.894
KLD-N	.913	.919	.924	.925	.927	.921	.912	.896	.851
Moderate Discrimination (a = 1.0)									
Z	.713	.725	.728	.738	.739	.741	.736	.724	.702
LR	.746	.755	.758	.768	.769	.769	.762	.750	.731
ST	.751	.761	.764	.776	.778	.778	.771	.760	.741
KLD-U	.658	.672	.674	.684	.688	.695	.694	.677	.579
KLD-N	.672	.673	.663	.659	.647	.640	.621	.580	.473
Low Discrimination (a = 0.6)									
Z	.410	.411	.418	.415	.416	.419	.417	.416	.409
LR	.443	.443	.450	.449	.450	.453	.448	.446	.438
ST	.453	.453	.461	.462	.462	.465	.458	.456	.447
KLD-U	.346	.355	.361	.359	.361	.367	.362	.338	.238
KLD-N	.343	.333	.325	.305	.288	.272	.247	.212	.149
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	1.000	1.000	1.000	1.000	.999	.977	.516
LR	.998	.999	1.000	1.000	1.000	1.000	.999	.976	.653
ST	.997	.999	1.000	1.000	1.000	1.000	.999	.974	.692
KLD-U	.998	.999	1.000	1.000	1.000	1.000	.998	.967	.525
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.998	.950	.444
Moderate Discrimination (a = 1.0)									
Z	.972	.986	.991	.992	.991	.985	.968	.907	.572
LR	.972	.986	.991	.993	.992	.986	.968	.901	.655
ST	.971	.987	.992	.994	.992	.987	.968	.897	.699
KLD-U	.969	.984	.989	.991	.989	.982	.960	.858	.458
KLD-N	.963	.982	.988	.989	.987	.976	.944	.812	.409
Low Discrimination (a = 0.6)									
Z	.781	.839	.868	.874	.871	.839	.783	.694	.498
LR	.780	.843	.873	.880	.876	.841	.778	.681	.520
ST	.783	.846	.876	.883	.879	.843	.778	.677	.538
KLD-U	.754	.826	.854	.858	.852	.813	.732	.550	.263
KLD-N	.712	.793	.826	.830	.812	.752	.641	.456	.217

Table A35. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.046	.047	.047	.048	.047	.048	.047	.049	.049
LR	.054	.056	.055	.055	.054	.055	.054	.056	.055
ST	.058	.059	.058	.058	.057	.058	.058	.059	.058
KLD-U	.058	.057	.054	.051	.048	.050	.050	.053	.054
KLD-N	.058	.054	.051	.048	.044	.045	.045	.046	.044
Moderate Discrimination (a = 1.0)									
Z	.046	.047	.048	.047	.048	.046	.048	.048	.047
LR	.054	.054	.055	.054	.055	.053	.055	.054	.054
ST	.057	.057	.058	.057	.059	.056	.057	.057	.057
KLD-U	.051	.051	.052	.049	.050	.047	.050	.051	.049
KLD-N	.049	.047	.046	.043	.042	.039	.040	.039	.036
Low Discrimination (a = 0.6)									
Z	.047	.047	.047	.045	.047	.047	.047	.047	.048
LR	.054	.053	.054	.052	.053	.054	.054	.053	.053
ST	.057	.056	.056	.055	.057	.057	.056	.055	.056
KLD-U	.042	.048	.049	.046	.047	.048	.048	.045	.036
KLD-N	.038	.039	.038	.035	.034	.033	.031	.028	.024
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.140	.143	.145	.147	.147	.145	.145	.142	.139
LR	.154	.158	.159	.160	.161	.159	.158	.156	.154
ST	.163	.165	.166	.166	.167	.167	.165	.164	.162
KLD-U	.140	.138	.135	.133	.130	.131	.131	.130	.130
KLD-N	.135	.133	.130	.126	.122	.119	.117	.112	.109
Moderate Discrimination (a = 1.0)									
Z	.094	.096	.099	.098	.095	.096	.096	.096	.094
LR	.104	.106	.108	.107	.105	.107	.107	.106	.104
ST	.110	.112	.114	.112	.111	.112	.113	.112	.110
KLD-U	.087	.090	.091	.087	.084	.085	.087	.088	.083
KLD-N	.084	.083	.082	.077	.073	.071	.069	.068	.062
Low Discrimination (a = 0.6)									
Z	.069	.071	.069	.070	.068	.068	.070	.070	.068
LR	.076	.078	.076	.077	.076	.076	.078	.078	.076
ST	.079	.082	.080	.081	.080	.081	.082	.082	.080
KLD-U	.053	.062	.062	.062	.061	.061	.062	.060	.047
KLD-N	.051	.052	.049	.047	.043	.042	.039	.036	.028
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.422	.433	.442	.453	.448	.449	.447	.440	.418
LR	.446	.457	.462	.473	.471	.473	.470	.466	.444
ST	.459	.468	.473	.482	.482	.483	.482	.478	.458
KLD-U	.414	.419	.419	.420	.416	.420	.421	.418	.399
KLD-N	.406	.411	.410	.408	.399	.398	.393	.382	.358

Continued on the next page

Table A35 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.246	.255	.256	.257	.256	.251	.256	.248	.245
LR	.264	.272	.272	.274	.274	.271	.276	.268	.263
ST	.274	.282	.280	.283	.283	.281	.286	.278	.275
KLD-U	.229	.237	.236	.233	.230	.226	.235	.230	.215
KLD-N	.220	.225	.222	.215	.207	.197	.198	.188	.171
Low Discrimination (a = 0.6)									
Z	.140	.141	.141	.140	.138	.140	.139	.135	.139
LR	.151	.152	.152	.152	.151	.153	.152	.147	.152
ST	.158	.158	.159	.158	.158	.160	.159	.154	.158
KLD-U	.113	.126	.125	.124	.121	.124	.122	.115	.094
KLD-N	.108	.107	.103	.098	.091	.088	.081	.071	.060
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.940	.948	.953	.955	.956	.955	.951	.945	.936
LR	.945	.952	.956	.959	.960	.960	.956	.952	.944
ST	.947	.953	.958	.961	.962	.961	.959	.955	.948
KLD-U	.939	.944	.946	.947	.949	.948	.944	.939	.922
KLD-N	.936	.941	.943	.943	.943	.940	.934	.926	.907
Moderate Discrimination (a = 1.0)									
Z	.737	.744	.745	.751	.749	.749	.746	.734	.716
LR	.754	.759	.760	.767	.768	.767	.763	.753	.738
ST	.762	.766	.766	.774	.776	.775	.772	.762	.748
KLD-U	.719	.724	.721	.723	.722	.724	.723	.712	.641
KLD-N	.708	.710	.703	.700	.691	.685	.676	.654	.594
Low Discrimination (a = 0.6)									
Z	.429	.427	.430	.427	.427	.429	.426	.425	.415
LR	.446	.445	.449	.448	.450	.451	.449	.446	.438
ST	.456	.453	.458	.459	.460	.462	.459	.457	.449
KLD-U	.391	.396	.397	.393	.395	.398	.391	.370	.280
KLD-N	.367	.359	.352	.338	.327	.317	.298	.276	.219
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	.999	1.000	1.000	1.000	.999	.999	.999
LR	.998	.999	.999	1.000	1.000	1.000	1.000	.999	.999
ST	.998	.999	.999	1.000	1.000	1.000	1.000	.999	.999
KLD-U	.999	.999	.999	.999	.999	1.000	.999	.999	.994
KLD-N	.998	.999	.999	.999	.999	.999	.999	.999	.995
Moderate Discrimination (a = 1.0)									
Z	.970	.971	.973	.974	.975	.975	.972	.968	.963
LR	.972	.973	.976	.977	.978	.977	.975	.972	.968
ST	.973	.974	.977	.978	.979	.979	.976	.974	.970
KLD-U	.967	.966	.968	.969	.971	.970	.967	.960	.890
KLD-N	.964	.963	.964	.963	.963	.961	.955	.945	.886
Low Discrimination (a = 0.6)									
Z	.759	.757	.766	.760	.764	.763	.762	.753	.747
LR	.773	.771	.781	.777	.780	.779	.778	.770	.767
ST	.779	.777	.787	.783	.788	.786	.786	.778	.775
KLD-U	.733	.732	.738	.733	.736	.736	.728	.674	.513
KLD-N	.705	.695	.693	.678	.668	.654	.631	.583	.472

Table A36. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.046	.048	.048	.049	.048	.048	.049	.050	.049
LR	.054	.054	.053	.054	.053	.053	.054	.055	.054
ST	.056	.057	.056	.056	.055	.056	.056	.058	.056
KLD-U	.062	.060	.056	.054	.052	.053	.054	.056	.057
KLD-N	.061	.057	.053	.051	.047	.048	.048	.048	.048
Moderate Discrimination (a = 1.0)									
Z	.047	.048	.049	.048	.049	.048	.048	.049	.049
LR	.052	.053	.055	.053	.054	.052	.053	.053	.053
ST	.055	.056	.057	.056	.056	.054	.055	.055	.055
KLD-U	.053	.055	.056	.053	.052	.050	.053	.054	.052
KLD-N	.051	.050	.049	.046	.044	.042	.043	.042	.039
Low Discrimination (a = 0.6)									
Z	.048	.047	.048	.047	.047	.048	.049	.047	.049
LR	.053	.052	.052	.051	.052	.052	.053	.051	.052
ST	.055	.054	.054	.053	.054	.055	.055	.053	.054
KLD-U	.044	.050	.051	.050	.050	.050	.051	.048	.037
KLD-N	.042	.040	.039	.036	.035	.035	.033	.031	.026
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.143	.145	.149	.150	.149	.148	.147	.146	.142
LR	.152	.154	.157	.156	.157	.157	.156	.155	.152
ST	.159	.160	.162	.161	.162	.162	.162	.161	.159
KLD-U	.150	.146	.145	.140	.137	.137	.138	.139	.137
KLD-N	.141	.138	.137	.133	.128	.127	.125	.123	.119
Moderate Discrimination (a = 1.0)									
Z	.097	.099	.101	.100	.097	.098	.099	.098	.096
LR	.103	.106	.107	.106	.103	.106	.106	.105	.103
ST	.108	.109	.111	.109	.107	.110	.111	.109	.108
KLD-U	.094	.097	.097	.093	.089	.091	.093	.093	.088
KLD-N	.088	.088	.088	.082	.076	.077	.076	.074	.069
Low Discrimination (a = 0.6)									
Z	.071	.072	.071	.072	.070	.070	.072	.071	.070
LR	.076	.076	.076	.076	.075	.075	.077	.076	.076
ST	.078	.079	.078	.079	.078	.078	.080	.079	.079
KLD-U	.057	.068	.068	.066	.065	.065	.067	.064	.051
KLD-N	.053	.055	.052	.050	.046	.045	.043	.040	.034
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.426	.438	.447	.457	.453	.454	.450	.444	.421
LR	.442	.452	.459	.468	.466	.469	.465	.461	.438
ST	.453	.461	.467	.476	.474	.478	.474	.471	.450
KLD-U	.431	.435	.434	.435	.428	.434	.432	.431	.409
KLD-N	.418	.423	.423	.422	.412	.412	.408	.401	.378

Continued on the next page

Table A36 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.251	.257	.259	.263	.260	.256	.261	.253	.248
LR	.262	.267	.268	.272	.271	.268	.274	.265	.260
ST	.270	.275	.275	.278	.278	.276	.282	.273	.268
KLD-U	.243	.247	.247	.245	.241	.239	.246	.240	.224
KLD-N	.230	.231	.230	.225	.217	.210	.212	.202	.189
Low Discrimination (a = 0.6)									
Z	.144	.144	.143	.143	.141	.143	.143	.138	.141
LR	.150	.150	.149	.150	.149	.152	.152	.146	.149
ST	.155	.154	.154	.154	.154	.157	.156	.151	.154
KLD-U	.120	.133	.132	.131	.128	.131	.130	.121	.101
KLD-N	.111	.112	.107	.104	.096	.095	.089	.081	.072
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.940	.946	.951	.954	.954	.953	.950	.944	.934
LR	.942	.949	.953	.957	.957	.956	.953	.949	.939
ST	.944	.950	.954	.958	.959	.958	.956	.951	.943
KLD-U	.941	.945	.947	.948	.949	.949	.946	.940	.923
KLD-N	.937	.941	.944	.944	.944	.943	.937	.930	.914
Moderate Discrimination (a = 1.0)									
Z	.739	.744	.744	.751	.750	.750	.746	.734	.716
LR	.748	.753	.753	.761	.761	.762	.758	.747	.730
ST	.755	.760	.758	.766	.768	.769	.766	.754	.739
KLD-U	.729	.733	.728	.731	.730	.733	.731	.719	.653
KLD-N	.713	.716	.708	.707	.701	.698	.690	.672	.623
Low Discrimination (a = 0.6)									
Z	.432	.431	.434	.431	.431	.433	.430	.429	.420
LR	.442	.440	.444	.444	.445	.447	.443	.442	.434
ST	.449	.447	.451	.453	.453	.456	.451	.451	.443
KLD-U	.403	.410	.411	.408	.407	.409	.403	.380	.294
KLD-N	.372	.368	.362	.351	.342	.334	.317	.301	.250
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	.999	1.000	.999	.999	.999	.999	.998
LR	.998	.999	.999	1.000	.999	1.000	.999	.999	.999
ST	.997	.999	.999	1.000	.999	1.000	.999	.999	.999
KLD-U	.999	.999	.999	.999	.999	.999	.999	.999	.995
KLD-N	.998	.999	.999	.999	.999	.999	.999	.999	.996
Moderate Discrimination (a = 1.0)									
Z	.969	.971	.973	.973	.974	.973	.971	.967	.961
LR	.970	.972	.974	.974	.976	.975	.973	.969	.964
ST	.971	.972	.974	.975	.977	.976	.975	.970	.967
KLD-U	.967	.968	.970	.969	.971	.970	.968	.960	.901
KLD-N	.964	.964	.965	.963	.964	.962	.958	.948	.905
Low Discrimination (a = 0.6)									
Z	.759	.756	.765	.760	.762	.762	.760	.754	.747
LR	.766	.764	.772	.770	.773	.772	.770	.765	.761
ST	.771	.768	.777	.775	.780	.778	.776	.771	.767
KLD-U	.741	.738	.744	.739	.743	.742	.732	.685	.534
KLD-N	.708	.698	.700	.686	.679	.669	.652	.614	.521

Table A37. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.099	.056	.048	.047	.046	.047	.050	.055	.071
LR	.054	.052	.053	.053	.052	.053	.053	.052	.056
ST	.052	.052	.055	.055	.055	.056	.055	.051	.051
KLD-U	.037	.058	.048	.046	.044	.045	.049	.055	.063
KLD-N	.032	.049	.046	.044	.043	.043	.045	.048	.045
Moderate Discrimination (a = 1.0)									
Z	.072	.056	.052	.049	.047	.049	.049	.052	.065
LR	.053	.054	.055	.054	.052	.053	.052	.050	.056
ST	.052	.054	.057	.057	.054	.056	.054	.050	.054
KLD-U	.037	.056	.051	.047	.045	.047	.048	.052	.056
KLD-N	.032	.046	.046	.044	.042	.043	.042	.042	.040
Low Discrimination (a = 0.6)									
Z	.063	.056	.053	.051	.050	.051	.052	.054	.059
LR	.057	.055	.055	.054	.053	.054	.053	.053	.053
ST	.058	.056	.056	.056	.055	.056	.054	.053	.052
KLD-U	.030	.047	.051	.050	.048	.050	.050	.052	.039
KLD-N	.024	.036	.040	.041	.040	.041	.039	.037	.027
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.154	.196	.228	.247	.253	.252	.237	.208	.157
LR	.121	.193	.242	.264	.270	.268	.244	.199	.132
ST	.119	.194	.248	.271	.278	.275	.248	.195	.122
KLD-U	.095	.192	.205	.220	.224	.224	.212	.181	.108
KLD-N	.084	.176	.203	.217	.218	.214	.200	.162	.077
Moderate Discrimination (a = 1.0)									
Z	.116	.129	.141	.148	.153	.150	.147	.140	.122
LR	.100	.128	.148	.159	.164	.160	.152	.136	.107
ST	.099	.129	.152	.164	.170	.164	.155	.135	.102
KLD-U	.076	.123	.127	.130	.134	.132	.130	.122	.084
KLD-N	.067	.107	.120	.124	.126	.122	.115	.101	.061
Low Discrimination (a = 0.6)									
Z	.087	.089	.096	.097	.097	.098	.097	.096	.094
LR	.082	.089	.099	.102	.103	.102	.099	.093	.086
ST	.082	.091	.102	.104	.105	.104	.100	.092	.083
KLD-U	.045	.078	.087	.086	.086	.086	.085	.081	.054
KLD-N	.042	.063	.074	.075	.074	.071	.065	.055	.036
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.412	.619	.700	.739	.746	.733	.684	.560	.341
LR	.386	.622	.717	.757	.762	.748	.689	.542	.308
ST	.384	.624	.724	.763	.768	.754	.691	.534	.291
KLD-U	.371	.601	.668	.703	.713	.702	.654	.513	.245
KLD-N	.340	.589	.666	.699	.705	.689	.636	.476	.186

Continued on the next page

Table A37 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.290	.372	.425	.458	.471	.453	.426	.365	.283
LR	.273	.376	.441	.477	.489	.467	.431	.354	.256
ST	.273	.379	.447	.484	.497	.473	.434	.352	.244
KLD-U	.248	.356	.395	.422	.437	.419	.394	.327	.195
KLD-N	.223	.334	.385	.411	.420	.396	.363	.284	.152
Low Discrimination (a = 0.6)									
Z	.170	.205	.235	.250	.255	.248	.233	.212	.182
LR	.165	.207	.242	.260	.264	.255	.233	.205	.168
ST	.167	.210	.245	.264	.268	.258	.234	.203	.164
KLD-U	.114	.188	.216	.227	.230	.224	.207	.177	.105
KLD-N	.106	.161	.194	.206	.206	.193	.169	.130	.074
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.975	.996	.999	.999	.999	.999	.994	.948	.591
LR	.974	.997	.999	1.000	.999	.999	.994	.940	.622
ST	.973	.997	.999	.999	.999	.999	.994	.935	.618
KLD-U	.973	.995	.998	.999	.999	.998	.993	.929	.497
KLD-N	.970	.996	.998	.999	.999	.998	.992	.906	.407
Moderate Discrimination (a = 1.0)									
Z	.844	.927	.954	.963	.963	.951	.917	.826	.630
LR	.841	.930	.958	.967	.966	.953	.917	.814	.603
ST	.841	.931	.959	.968	.967	.954	.916	.809	.589
KLD-U	.832	.915	.944	.955	.955	.942	.903	.788	.454
KLD-N	.812	.910	.942	.951	.949	.934	.886	.741	.393
Low Discrimination (a = 0.6)									
Z	.550	.656	.707	.739	.741	.710	.652	.573	.466
LR	.548	.661	.716	.749	.748	.714	.650	.559	.442
ST	.552	.665	.720	.753	.752	.716	.649	.555	.433
KLD-U	.505	.634	.682	.712	.713	.681	.615	.497	.269
KLD-N	.468	.600	.656	.685	.680	.634	.548	.414	.217
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.991	.996	1.000	1.000	1.000	1.000	1.000	.994	.647
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.993	.753
ST	.996	.995	.999	1.000	1.000	1.000	1.000	.992	.770
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.990	.611
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.983	.520
Moderate Discrimination (a = 1.0)									
Z	.997	.999	1.000	1.000	1.000	.999	.996	.972	.783
LR	.997	.999	1.000	1.000	1.000	1.000	.996	.968	.809
ST	.997	.999	1.000	1.000	1.000	1.000	.996	.966	.812
KLD-U	.996	.999	1.000	1.000	1.000	.999	.995	.954	.620
KLD-N	.996	.999	1.000	1.000	1.000	.999	.993	.934	.573
Low Discrimination (a = 0.6)									
Z	.911	.951	.968	.971	.968	.954	.917	.844	.710
LR	.912	.953	.970	.973	.969	.954	.914	.832	.699
ST	.913	.954	.971	.974	.970	.955	.913	.828	.692
KLD-U	.899	.944	.962	.966	.961	.945	.896	.760	.430
KLD-N	.881	.935	.955	.959	.952	.927	.856	.692	.386

Table A38. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.098	.056	.050	.048	.046	.048	.050	.054	.070
LR	.054	.051	.053	.052	.051	.052	.052	.051	.056
ST	.053	.051	.054	.054	.053	.054	.053	.050	.051
KLD-U	.038	.060	.053	.050	.048	.049	.052	.056	.063
KLD-N	.033	.050	.050	.048	.046	.046	.048	.048	.046
Moderate Discrimination (a = 1.0)									
Z	.071	.056	.051	.049	.047	.049	.050	.052	.065
LR	.053	.053	.053	.052	.051	.052	.051	.050	.056
ST	.052	.054	.055	.054	.053	.054	.052	.050	.053
KLD-U	.038	.058	.053	.050	.048	.050	.050	.054	.057
KLD-N	.033	.047	.048	.047	.045	.046	.044	.043	.041
Low Discrimination (a = 0.6)									
Z	.062	.055	.052	.051	.050	.051	.052	.054	.059
LR	.055	.053	.053	.053	.052	.052	.052	.052	.053
ST	.056	.054	.055	.054	.053	.054	.053	.052	.051
KLD-U	.032	.050	.054	.052	.050	.052	.052	.053	.040
KLD-N	.026	.038	.042	.043	.043	.043	.041	.037	.028
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.154	.198	.235	.256	.260	.258	.239	.208	.157
LR	.120	.192	.242	.266	.271	.268	.244	.199	.132
ST	.118	.192	.246	.271	.276	.274	.248	.196	.122
KLD-U	.096	.202	.227	.244	.246	.242	.222	.183	.108
KLD-N	.085	.182	.222	.239	.240	.234	.210	.164	.078
Moderate Discrimination (a = 1.0)									
Z	.116	.132	.144	.154	.157	.152	.148	.140	.122
LR	.098	.128	.147	.161	.164	.159	.151	.136	.107
ST	.098	.128	.150	.165	.168	.162	.153	.136	.102
KLD-U	.077	.133	.138	.145	.145	.141	.136	.124	.085
KLD-N	.067	.113	.129	.137	.138	.131	.122	.104	.063
Low Discrimination (a = 0.6)									
Z	.086	.091	.098	.098	.099	.098	.098	.096	.093
LR	.079	.088	.098	.101	.102	.101	.098	.092	.084
ST	.080	.090	.100	.103	.104	.103	.099	.092	.082
KLD-U	.048	.083	.094	.092	.091	.090	.088	.082	.056
KLD-N	.043	.065	.079	.080	.079	.075	.069	.058	.038
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.413	.630	.715	.752	.754	.739	.687	.561	.342
LR	.388	.626	.723	.762	.764	.749	.690	.542	.309
ST	.385	.627	.728	.767	.769	.754	.692	.535	.292
KLD-U	.379	.627	.704	.738	.739	.721	.665	.517	.247
KLD-N	.344	.611	.698	.733	.732	.711	.649	.481	.189

Continued on the next page

Table A38 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.294	.380	.436	.467	.479	.457	.430	.366	.284
LR	.274	.378	.443	.478	.490	.467	.433	.356	.256
ST	.273	.380	.447	.483	.496	.472	.436	.353	.246
KLD-U	.260	.378	.423	.448	.459	.436	.404	.331	.198
KLD-N	.229	.351	.408	.435	.444	.415	.376	.289	.156
Low Discrimination (a = 0.6)									
Z	.172	.208	.237	.255	.257	.251	.234	.212	.182
LR	.164	.206	.239	.260	.263	.254	.233	.205	.169
ST	.166	.208	.242	.263	.267	.257	.233	.204	.164
KLD-U	.123	.199	.228	.240	.239	.232	.211	.180	.109
KLD-N	.108	.168	.203	.218	.215	.203	.174	.135	.078
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.978	.997	.999	1.000	.999	.999	.994	.949	.592
LR	.976	.997	.999	1.000	.999	.999	.994	.940	.623
ST	.975	.997	.999	1.000	.999	.999	.994	.935	.620
KLD-U	.978	.997	.999	.999	.999	.999	.993	.930	.500
KLD-N	.974	.996	.999	.999	.999	.998	.992	.907	.412
Moderate Discrimination (a = 1.0)									
Z	.850	.932	.958	.965	.964	.952	.918	.828	.633
LR	.844	.933	.959	.967	.966	.954	.917	.815	.606
ST	.843	.934	.960	.968	.967	.954	.917	.810	.592
KLD-U	.848	.929	.954	.961	.960	.945	.906	.791	.461
KLD-N	.824	.922	.950	.958	.956	.938	.890	.746	.405
Low Discrimination (a = 0.6)									
Z	.557	.663	.714	.743	.744	.711	.654	.573	.467
LR	.550	.662	.718	.749	.748	.713	.649	.559	.443
ST	.552	.665	.721	.752	.752	.714	.648	.556	.434
KLD-U	.524	.653	.699	.726	.724	.686	.619	.503	.279
KLD-N	.480	.615	.672	.700	.694	.645	.556	.426	.233
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.994	.649
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.993	.754
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.992	.772
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.990	.614
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.984	.525
Moderate Discrimination (a = 1.0)									
Z	.997	.999	1.000	1.000	1.000	.999	.996	.973	.785
LR	.997	.999	1.000	1.000	1.000	.999	.996	.968	.811
ST	.997	.999	1.000	1.000	1.000	.999	.996	.966	.814
KLD-U	.997	.999	1.000	1.000	1.000	.999	.995	.955	.629
KLD-N	.997	.999	1.000	1.000	1.000	.999	.994	.938	.587
Low Discrimination (a = 0.6)									
Z	.913	.953	.969	.972	.968	.955	.918	.845	.711
LR	.912	.954	.970	.973	.968	.955	.914	.833	.701
ST	.912	.954	.970	.973	.969	.955	.913	.829	.694
KLD-U	.908	.950	.965	.968	.963	.947	.898	.766	.444
KLD-N	.888	.940	.959	.962	.954	.929	.862	.706	.411

Table A39. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.098	.056	.050	.047	.048	.048	.051	.054	.071
LR	.054	.052	.052	.050	.051	.051	.052	.051	.056
ST	.052	.051	.053	.052	.052	.052	.053	.050	.051
KLD-U	.038	.060	.055	.051	.051	.051	.053	.057	.064
KLD-N	.033	.051	.052	.049	.049	.048	.049	.049	.046
Moderate Discrimination (a = 1.0)									
Z	.071	.056	.052	.050	.049	.050	.050	.052	.065
LR	.053	.052	.053	.052	.051	.052	.051	.050	.056
ST	.052	.052	.054	.053	.052	.053	.052	.049	.053
KLD-U	.039	.059	.055	.052	.051	.052	.052	.055	.057
KLD-N	.033	.047	.050	.048	.047	.047	.045	.044	.041
Low Discrimination (a = 0.6)									
Z	.062	.056	.052	.052	.050	.051	.052	.055	.059
LR	.054	.053	.052	.053	.051	.052	.051	.053	.053
ST	.055	.054	.053	.054	.052	.052	.051	.052	.051
KLD-U	.032	.051	.054	.053	.052	.052	.053	.055	.040
KLD-N	.027	.038	.043	.045	.043	.043	.041	.039	.028
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.154	.199	.237	.258	.262	.261	.240	.208	.157
LR	.121	.191	.240	.263	.268	.267	.244	.198	.133
ST	.119	.191	.243	.268	.273	.272	.246	.195	.123
KLD-U	.096	.204	.234	.249	.251	.249	.226	.184	.108
KLD-N	.085	.184	.227	.244	.245	.241	.215	.165	.079
Moderate Discrimination (a = 1.0)									
Z	.116	.133	.146	.156	.159	.155	.149	.141	.122
LR	.098	.128	.147	.159	.164	.159	.150	.136	.107
ST	.097	.128	.149	.162	.167	.161	.152	.135	.102
KLD-U	.078	.136	.144	.149	.151	.146	.138	.125	.085
KLD-N	.068	.116	.134	.141	.143	.136	.124	.105	.063
Low Discrimination (a = 0.6)									
Z	.086	.092	.099	.099	.100	.100	.098	.097	.093
LR	.078	.088	.098	.100	.101	.101	.097	.093	.084
ST	.079	.089	.099	.101	.102	.102	.097	.091	.081
KLD-U	.050	.086	.097	.095	.094	.093	.089	.083	.057
KLD-N	.044	.067	.080	.082	.082	.078	.070	.059	.039
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.414	.632	.716	.752	.755	.739	.686	.560	.342
LR	.388	.626	.721	.757	.761	.744	.688	.542	.309
ST	.385	.625	.724	.761	.764	.748	.690	.535	.292
KLD-U	.379	.632	.711	.742	.743	.725	.667	.518	.247
KLD-N	.344	.615	.705	.736	.737	.716	.652	.481	.189

Continued on the next page

Table A39 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.295	.383	.438	.468	.481	.459	.430	.366	.284
LR	.274	.377	.440	.474	.488	.464	.431	.355	.256
ST	.273	.377	.443	.478	.492	.468	.433	.352	.245
KLD-U	.263	.385	.431	.454	.466	.441	.408	.332	.198
KLD-N	.230	.356	.415	.441	.451	.422	.381	.290	.157
Low Discrimination (a = 0.6)									
Z	.173	.210	.240	.256	.262	.253	.235	.213	.181
LR	.162	.206	.239	.258	.264	.254	.233	.205	.167
ST	.164	.207	.241	.261	.267	.256	.233	.203	.162
KLD-U	.126	.205	.234	.246	.247	.236	.214	.181	.110
KLD-N	.110	.170	.208	.223	.222	.207	.178	.139	.081
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.978	.997	.999	.999	.999	.999	.994	.949	.592
LR	.976	.997	.999	.999	.999	.999	.994	.940	.623
ST	.975	.997	.999	.999	.999	.999	.994	.935	.620
KLD-U	.978	.997	.999	.999	.999	.999	.993	.930	.501
KLD-N	.974	.997	.999	.999	.999	.999	.992	.907	.413
Moderate Discrimination (a = 1.0)									
Z	.850	.932	.957	.964	.963	.952	.917	.827	.632
LR	.842	.931	.957	.965	.964	.952	.916	.814	.605
ST	.841	.931	.958	.966	.965	.953	.916	.809	.592
KLD-U	.850	.930	.954	.961	.960	.946	.906	.791	.462
KLD-N	.826	.924	.951	.958	.956	.940	.890	.747	.406
Low Discrimination (a = 0.6)									
Z	.558	.661	.712	.745	.744	.710	.655	.573	.467
LR	.548	.658	.713	.747	.746	.709	.649	.558	.442
ST	.548	.659	.715	.749	.748	.710	.648	.554	.433
KLD-U	.530	.657	.704	.732	.727	.688	.623	.505	.282
KLD-N	.483	.617	.675	.706	.698	.648	.563	.430	.239
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.994	.648
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.993	.754
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.992	.772
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.990	.614
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.984	.525
Moderate Discrimination (a = 1.0)									
Z	.997	.999	1.000	1.000	1.000	.999	.996	.973	.783
LR	.997	.999	1.000	1.000	1.000	.999	.996	.968	.811
ST	.997	.999	1.000	1.000	1.000	.999	.996	.966	.814
KLD-U	.997	.999	1.000	1.000	1.000	.999	.995	.955	.629
KLD-N	.996	.999	1.000	1.000	1.000	.999	.993	.938	.589
Low Discrimination (a = 0.6)									
Z	.912	.953	.968	.971	.967	.954	.917	.844	.709
LR	.910	.953	.969	.971	.967	.953	.913	.832	.699
ST	.910	.953	.969	.972	.968	.953	.912	.827	.692
KLD-U	.909	.951	.966	.968	.963	.947	.899	.768	.448
KLD-N	.888	.941	.960	.962	.955	.931	.864	.710	.419

Table A40. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.049	.050	.050	.049	.049	.050	.050	.050	.048
LR	.054	.054	.054	.053	.054	.055	.054	.055	.053
ST	.055	.056	.056	.054	.055	.057	.056	.057	.055
KLD-U	.049	.049	.047	.048	.048	.048	.048	.050	.048
KLD-N	.050	.048	.046	.045	.045	.045	.044	.045	.042
Moderate Discrimination (a = 1.0)									
Z	.048	.049	.049	.050	.048	.048	.048	.049	.048
LR	.054	.054	.054	.055	.054	.053	.054	.054	.053
ST	.056	.056	.056	.057	.055	.055	.056	.056	.056
KLD-U	.047	.047	.046	.047	.046	.046	.047	.048	.047
KLD-N	.045	.044	.043	.043	.041	.040	.040	.040	.038
Low Discrimination (a = 0.6)									
Z	.050	.049	.049	.049	.048	.048	.048	.045	.048
LR	.055	.054	.054	.055	.053	.053	.053	.050	.053
ST	.057	.056	.056	.057	.055	.056	.056	.053	.055
KLD-U	.044	.046	.047	.047	.046	.046	.046	.043	.040
KLD-N	.041	.039	.039	.038	.037	.035	.034	.030	.030
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.186	.193	.194	.192	.192	.190	.193	.191	.194
LR	.197	.204	.207	.203	.203	.201	.204	.202	.205
ST	.202	.207	.209	.206	.207	.205	.209	.207	.209
KLD-U	.166	.169	.168	.168	.167	.167	.170	.171	.173
KLD-N	.168	.170	.167	.164	.160	.157	.157	.155	.153
Moderate Discrimination (a = 1.0)									
Z	.121	.123	.123	.121	.123	.123	.123	.123	.120
LR	.131	.132	.133	.131	.133	.133	.133	.133	.129
ST	.135	.136	.136	.134	.137	.137	.138	.138	.133
KLD-U	.106	.105	.105	.104	.105	.106	.107	.109	.104
KLD-N	.106	.104	.102	.096	.096	.093	.092	.090	.082
Low Discrimination (a = 0.6)									
Z	.081	.083	.085	.082	.080	.079	.083	.082	.080
LR	.089	.090	.092	.088	.088	.087	.091	.090	.087
ST	.092	.093	.095	.092	.092	.091	.094	.093	.090
KLD-U	.064	.070	.073	.070	.068	.069	.072	.071	.061
KLD-N	.063	.064	.064	.059	.055	.052	.052	.048	.041
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.580	.587	.592	.597	.589	.592	.596	.586	.582
LR	.596	.602	.606	.612	.604	.608	.611	.602	.598
ST	.601	.606	.610	.615	.608	.612	.616	.607	.603
KLD-U	.545	.548	.552	.560	.551	.555	.561	.555	.548
KLD-N	.550	.549	.550	.552	.539	.539	.539	.528	.512

Continued on the next page

Table A40 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.342	.353	.360	.353	.357	.351	.357	.357	.349
LR	.360	.370	.377	.371	.375	.369	.376	.374	.366
ST	.367	.376	.383	.377	.382	.376	.382	.381	.373
KLD-U	.310	.317	.324	.318	.322	.317	.326	.327	.313
KLD-N	.311	.314	.317	.305	.302	.290	.292	.288	.268
Low Discrimination (a = 0.6)									
Z	.186	.189	.190	.186	.190	.185	.189	.190	.190
LR	.199	.201	.203	.200	.204	.198	.201	.202	.202
ST	.204	.205	.206	.204	.209	.203	.206	.206	.206
KLD-U	.156	.164	.165	.161	.166	.162	.166	.166	.146
KLD-N	.153	.153	.148	.140	.138	.130	.127	.121	.106
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.991	.993	.992	.991	.992	.992	.992	.991	.990
LR	.992	.993	.993	.992	.993	.993	.993	.992	.991
ST	.991	.993	.993	.992	.993	.993	.993	.992	.991
KLD-U	.988	.990	.990	.989	.990	.990	.990	.989	.986
KLD-N	.988	.990	.989	.988	.988	.988	.988	.986	.983
Moderate Discrimination (a = 1.0)									
Z	.887	.891	.894	.890	.891	.890	.894	.887	.886
LR	.895	.898	.902	.898	.900	.898	.902	.896	.895
ST	.897	.900	.904	.901	.902	.901	.905	.898	.897
KLD-U	.866	.870	.873	.870	.872	.871	.877	.870	.844
KLD-N	.866	.868	.867	.860	.857	.852	.854	.840	.813
Low Discrimination (a = 0.6)									
Z	.578	.583	.585	.575	.580	.582	.583	.578	.578
LR	.596	.600	.604	.594	.599	.599	.600	.596	.594
ST	.601	.605	.610	.600	.606	.606	.606	.602	.600
KLD-U	.538	.542	.545	.535	.542	.546	.547	.532	.449
KLD-N	.526	.521	.516	.495	.490	.483	.471	.448	.391
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.997	.997	.997	.997	.998	.997	.997	.997	.997
LR	.997	.998	.998	.998	.998	.998	.998	.997	.997
ST	.997	.998	.998	.998	.998	.998	.998	.998	.997
KLD-U	.996	.996	.996	.996	.996	.996	.996	.996	.983
KLD-N	.996	.996	.996	.996	.996	.995	.995	.994	.983
Low Discrimination (a = 0.6)									
Z	.899	.904	.904	.903	.902	.900	.901	.900	.895
LR	.906	.911	.911	.911	.910	.908	.908	.907	.902
ST	.907	.913	.913	.913	.912	.910	.911	.909	.904
KLD-U	.878	.884	.884	.884	.885	.883	.881	.862	.736
KLD-N	.870	.871	.866	.861	.853	.845	.833	.813	.720

Table A41. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.049	.050	.050	.049	.050	.050	.050	.050	.049
LR	.053	.053	.053	.051	.053	.054	.053	.053	.053
ST	.054	.054	.054	.052	.054	.055	.054	.055	.054
KLD-U	.053	.052	.050	.049	.050	.050	.050	.052	.051
KLD-N	.052	.050	.049	.046	.047	.047	.046	.047	.046
Moderate Discrimination (a = 1.0)									
Z	.048	.049	.048	.050	.049	.048	.049	.049	.048
LR	.053	.053	.052	.054	.052	.052	.053	.052	.052
ST	.055	.055	.054	.055	.054	.053	.054	.054	.053
KLD-U	.051	.051	.049	.050	.049	.048	.050	.050	.050
KLD-N	.048	.048	.045	.046	.044	.043	.043	.043	.041
Low Discrimination (a = 0.6)									
Z	.050	.049	.049	.050	.048	.048	.049	.045	.049
LR	.053	.053	.052	.053	.051	.052	.053	.049	.052
ST	.055	.054	.054	.055	.053	.053	.055	.051	.053
KLD-U	.046	.049	.049	.049	.048	.048	.050	.046	.043
KLD-N	.042	.042	.041	.040	.038	.037	.036	.032	.032
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.191	.197	.199	.195	.194	.193	.195	.192	.197
LR	.198	.204	.207	.202	.202	.200	.203	.201	.205
ST	.202	.207	.210	.205	.205	.203	.207	.205	.208
KLD-U	.181	.183	.182	.178	.177	.175	.179	.180	.184
KLD-N	.178	.181	.179	.173	.169	.166	.168	.167	.168
Moderate Discrimination (a = 1.0)									
Z	.124	.125	.126	.124	.124	.125	.124	.124	.122
LR	.130	.131	.132	.130	.131	.132	.131	.131	.128
ST	.134	.135	.135	.133	.134	.135	.135	.135	.132
KLD-U	.116	.116	.114	.112	.113	.113	.114	.115	.112
KLD-N	.112	.111	.108	.104	.104	.101	.100	.099	.095
Low Discrimination (a = 0.6)									
Z	.084	.084	.085	.083	.081	.079	.084	.083	.080
LR	.088	.089	.090	.088	.087	.085	.089	.089	.085
ST	.091	.091	.092	.091	.090	.088	.092	.091	.088
KLD-U	.071	.076	.077	.075	.073	.072	.076	.076	.065
KLD-N	.068	.067	.066	.063	.059	.057	.057	.055	.048
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.589	.596	.599	.602	.593	.597	.599	.592	.589
LR	.600	.606	.609	.611	.604	.607	.611	.603	.599
ST	.604	.609	.613	.614	.607	.611	.615	.608	.604
KLD-U	.572	.572	.572	.576	.566	.570	.575	.572	.567
KLD-N	.569	.570	.568	.568	.554	.555	.557	.551	.541

Continued on the next page

Table A41 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.352	.361	.367	.360	.360	.356	.361	.361	.355
LR	.363	.371	.377	.372	.372	.368	.374	.373	.367
ST	.369	.376	.382	.377	.378	.374	.380	.380	.373
KLD-U	.335	.340	.344	.337	.337	.333	.340	.344	.332
KLD-N	.327	.330	.332	.322	.318	.309	.311	.312	.298
Low Discrimination (a = 0.6)									
Z	.192	.193	.194	.189	.192	.186	.191	.192	.192
LR	.200	.200	.202	.198	.201	.196	.200	.199	.200
ST	.204	.204	.205	.202	.206	.201	.204	.204	.205
KLD-U	.170	.176	.177	.172	.176	.172	.176	.174	.158
KLD-N	.160	.159	.157	.150	.149	.140	.140	.136	.126
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.992	.993	.993	.992	.992	.992	.992	.991	.991
LR	.992	.993	.993	.992	.993	.993	.993	.991	.991
ST	.992	.993	.993	.992	.993	.992	.993	.992	.991
KLD-U	.990	.992	.991	.989	.990	.990	.991	.990	.988
KLD-N	.990	.991	.991	.989	.989	.989	.990	.988	.986
Moderate Discrimination (a = 1.0)									
Z	.892	.896	.895	.893	.894	.892	.896	.890	.890
LR	.896	.900	.900	.899	.899	.898	.902	.895	.896
ST	.898	.902	.902	.900	.901	.900	.904	.898	.898
KLD-U	.882	.883	.883	.880	.881	.879	.886	.880	.863
KLD-N	.878	.878	.876	.871	.869	.864	.869	.860	.845
Low Discrimination (a = 0.6)									
Z	.587	.589	.593	.580	.585	.586	.586	.583	.582
LR	.598	.599	.604	.592	.599	.598	.598	.594	.595
ST	.603	.604	.609	.598	.604	.604	.604	.600	.601
KLD-U	.562	.562	.566	.554	.561	.561	.561	.550	.475
KLD-N	.540	.535	.533	.514	.513	.507	.499	.483	.440
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.997	.998	.997	.997	.997	.997	.997	.997	.997
LR	.997	.998	.997	.998	.998	.998	.998	.998	.997
ST	.997	.998	.997	.998	.998	.998	.998	.998	.997
KLD-U	.996	.997	.997	.997	.997	.997	.997	.997	.988
KLD-N	.996	.997	.996	.996	.996	.996	.996	.996	.990
Low Discrimination (a = 0.6)									
Z	.901	.906	.905	.905	.903	.902	.903	.902	.899
LR	.906	.910	.910	.910	.908	.907	.907	.907	.904
ST	.907	.911	.911	.912	.910	.909	.910	.909	.906
KLD-U	.889	.893	.893	.894	.892	.891	.890	.874	.769
KLD-N	.877	.879	.876	.871	.865	.860	.854	.841	.774

Table A42. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.050	.050	.051	.049	.050	.051	.050	.051	.049
LR	.052	.052	.052	.051	.052	.053	.053	.053	.051
ST	.053	.053	.053	.052	.053	.054	.053	.054	.052
KLD-U	.054	.053	.052	.051	.051	.052	.052	.054	.052
KLD-N	.054	.052	.050	.048	.048	.048	.048	.049	.046
Moderate Discrimination (a = 1.0)									
Z	.049	.050	.049	.051	.049	.049	.049	.049	.048
LR	.052	.052	.051	.053	.052	.051	.052	.052	.051
ST	.053	.054	.052	.054	.053	.052	.053	.053	.052
KLD-U	.053	.053	.051	.052	.051	.050	.052	.051	.051
KLD-N	.050	.049	.047	.047	.045	.044	.045	.044	.042
Low Discrimination (a = 0.6)									
Z	.050	.050	.050	.050	.048	.048	.049	.046	.049
LR	.052	.052	.052	.053	.051	.051	.052	.049	.051
ST	.053	.054	.053	.054	.052	.052	.053	.050	.052
KLD-U	.047	.051	.051	.051	.049	.050	.051	.047	.044
KLD-N	.044	.044	.042	.042	.039	.038	.038	.034	.034
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.192	.199	.201	.196	.196	.193	.197	.194	.198
LR	.197	.202	.205	.200	.200	.198	.202	.200	.203
ST	.200	.205	.207	.202	.202	.200	.204	.203	.206
KLD-U	.185	.187	.186	.182	.181	.180	.183	.184	.187
KLD-N	.182	.184	.183	.177	.174	.171	.172	.172	.174
Moderate Discrimination (a = 1.0)									
Z	.126	.127	.126	.125	.125	.126	.125	.125	.123
LR	.130	.130	.130	.128	.130	.130	.129	.130	.127
ST	.133	.133	.133	.131	.132	.133	.132	.133	.129
KLD-U	.121	.120	.117	.116	.116	.117	.117	.119	.116
KLD-N	.115	.114	.110	.108	.107	.106	.104	.104	.101
Low Discrimination (a = 0.6)									
Z	.084	.085	.086	.084	.083	.081	.085	.084	.082
LR	.086	.088	.089	.087	.087	.085	.088	.087	.085
ST	.088	.090	.090	.088	.089	.087	.090	.089	.087
KLD-U	.073	.079	.080	.078	.078	.076	.080	.079	.069
KLD-N	.068	.069	.069	.065	.063	.060	.061	.058	.052
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.589	.597	.600	.602	.593	.596	.601	.593	.589
LR	.595	.602	.606	.608	.600	.603	.608	.601	.596
ST	.599	.604	.609	.610	.603	.606	.611	.604	.599
KLD-U	.577	.578	.578	.581	.572	.575	.581	.578	.572
KLD-N	.572	.574	.572	.573	.560	.562	.564	.558	.550

Continued on the next page

Table A42 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.354	.363	.370	.362	.363	.357	.362	.362	.356
LR	.360	.369	.375	.369	.370	.364	.370	.370	.363
ST	.365	.373	.379	.372	.374	.369	.375	.376	.368
KLD-U	.342	.348	.351	.342	.344	.339	.345	.349	.338
KLD-N	.331	.336	.338	.328	.324	.317	.319	.321	.310
Low Discrimination (a = 0.6)									
Z	.194	.195	.196	.191	.193	.190	.193	.193	.194
LR	.198	.199	.200	.196	.199	.195	.198	.198	.199
ST	.202	.201	.203	.200	.202	.198	.202	.202	.202
KLD-U	.175	.182	.182	.178	.180	.177	.181	.179	.162
KLD-N	.164	.163	.161	.154	.154	.147	.147	.144	.134
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.991	.993	.992	.991	.992	.992	.992	.990	.990
LR	.991	.993	.992	.991	.992	.992	.992	.991	.990
ST	.991	.993	.992	.991	.992	.992	.992	.991	.991
KLD-U	.990	.991	.991	.989	.990	.990	.991	.990	.988
KLD-N	.990	.991	.990	.989	.990	.989	.990	.988	.986
Moderate Discrimination (a = 1.0)									
Z	.891	.894	.896	.892	.892	.892	.895	.891	.889
LR	.894	.896	.898	.895	.896	.896	.899	.894	.892
ST	.894	.898	.899	.897	.898	.897	.901	.896	.894
KLD-U	.884	.885	.885	.881	.883	.882	.888	.883	.864
KLD-N	.878	.879	.877	.872	.871	.868	.873	.865	.853
Low Discrimination (a = 0.6)									
Z	.588	.590	.593	.581	.586	.586	.587	.584	.583
LR	.594	.596	.600	.589	.595	.594	.594	.590	.591
ST	.598	.599	.603	.593	.600	.598	.599	.595	.596
KLD-U	.568	.570	.572	.560	.567	.566	.568	.556	.484
KLD-N	.543	.541	.537	.522	.521	.514	.509	.498	.459
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.997	.998	.997	.997	.997	.997	.997	.997	.996
LR	.997	.998	.997	.997	.997	.997	.997	.997	.997
ST	.997	.998	.997	.997	.997	.997	.997	.997	.997
KLD-U	.997	.997	.996	.997	.997	.997	.997	.996	.989
KLD-N	.996	.997	.996	.996	.996	.996	.996	.995	.991
Low Discrimination (a = 0.6)									
Z	.900	.906	.904	.904	.901	.901	.902	.902	.899
LR	.902	.908	.907	.907	.904	.904	.905	.905	.902
ST	.903	.909	.908	.908	.906	.905	.906	.906	.904
KLD-U	.891	.896	.894	.894	.892	.892	.892	.877	.777
KLD-N	.877	.880	.876	.873	.868	.864	.860	.850	.791

Table A43. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.106	.049	.038	.037	.040	.042	.045	.052	.050
LR	.067	.059	.055	.051	.050	.047	.045	.042	.038
ST	.055	.052	.053	.054	.054	.051	.049	.044	.038
KLD-U	.064	.059	.049	.044	.045	.044	.047	.054	.049
KLD-N	.080	.064	.048	.040	.038	.035	.034	.036	.032
Moderate Discrimination (a = 1.0)									
Z	.081	.054	.041	.039	.041	.045	.046	.056	.053
LR	.064	.058	.054	.051	.048	.049	.046	.046	.040
ST	.057	.055	.053	.053	.051	.051	.047	.045	.042
KLD-U	.043	.048	.046	.044	.044	.046	.049	.049	.033
KLD-N	.052	.049	.042	.037	.034	.033	.032	.029	.022
Low Discrimination (a = 0.6)									
Z	.060	.059	.050	.044	.044	.046	.051	.058	.048
LR	.060	.060	.057	.052	.050	.050	.048	.048	.035
ST	.059	.058	.055	.052	.050	.049	.046	.045	.041
KLD-U	.028	.039	.046	.046	.048	.050	.049	.039	.020
KLD-N	.030	.035	.037	.034	.032	.031	.028	.022	.013
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.120	.076	.072	.079	.087	.093	.095	.102	.067
LR	.095	.095	.098	.103	.103	.101	.094	.085	.059
ST	.077	.086	.096	.106	.109	.108	.100	.087	.063
KLD-U	.059	.063	.060	.065	.068	.070	.074	.083	.060
KLD-N	.082	.074	.064	.061	.056	.051	.047	.050	.036
Moderate Discrimination (a = 1.0)									
Z	.090	.068	.061	.064	.069	.073	.079	.086	.062
LR	.079	.081	.080	.080	.080	.078	.076	.070	.050
ST	.070	.075	.080	.082	.083	.080	.076	.069	.058
KLD-U	.039	.048	.050	.052	.054	.059	.063	.061	.037
KLD-N	.054	.057	.050	.045	.040	.037	.033	.030	.020
Low Discrimination (a = 0.6)									
Z	.068	.068	.060	.056	.058	.062	.070	.071	.052
LR	.068	.071	.069	.067	.067	.064	.065	.058	.039
ST	.065	.068	.068	.066	.066	.063	.061	.055	.049
KLD-U	.025	.038	.046	.049	.051	.054	.054	.041	.021
KLD-N	.034	.041	.042	.038	.032	.029	.025	.018	.010
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.199	.172	.193	.220	.250	.258	.254	.241	.108
LR	.193	.220	.250	.273	.284	.273	.250	.210	.114
ST	.168	.207	.244	.276	.295	.286	.259	.212	.126
KLD-U	.116	.137	.157	.181	.198	.204	.209	.204	.111
KLD-N	.154	.164	.169	.173	.168	.157	.145	.138	.070

Continued on the next page

Table A43 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.139	.124	.130	.148	.157	.167	.168	.161	.089
LR	.142	.153	.166	.179	.178	.176	.160	.138	.082
ST	.128	.147	.165	.182	.183	.179	.159	.136	.101
KLD-U	.064	.090	.102	.116	.122	.131	.132	.117	.061
KLD-N	.094	.106	.105	.104	.093	.086	.074	.060	.031
Low Discrimination (a = 0.6)									
Z	.095	.100	.096	.100	.106	.111	.119	.109	.066
LR	.099	.111	.113	.115	.116	.113	.108	.092	.053
ST	.094	.107	.111	.114	.116	.110	.102	.090	.072
KLD-U	.034	.058	.072	.080	.087	.088	.085	.063	.030
KLD-N	.051	.067	.069	.062	.056	.046	.038	.026	.012
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.493	.531	.584	.682	.773	.763	.719	.592	.164
LR	.607	.689	.753	.791	.802	.776	.714	.572	.233
ST	.534	.611	.676	.755	.805	.784	.719	.575	.272
KLD-U	.445	.520	.597	.666	.712	.713	.683	.571	.222
KLD-N	.531	.585	.624	.647	.657	.629	.582	.462	.154
Moderate Discrimination (a = 1.0)									
Z	.342	.378	.434	.497	.526	.518	.491	.382	.131
LR	.406	.472	.523	.553	.553	.525	.470	.363	.163
ST	.380	.451	.514	.557	.560	.527	.466	.370	.228
KLD-U	.257	.334	.388	.432	.458	.458	.424	.316	.128
KLD-N	.319	.371	.392	.394	.378	.344	.292	.204	.075
Low Discrimination (a = 0.6)									
Z	.217	.247	.269	.292	.302	.310	.293	.213	.097
LR	.243	.284	.308	.320	.319	.306	.273	.200	.091
ST	.234	.278	.305	.318	.316	.297	.263	.212	.151
KLD-U	.116	.185	.221	.247	.259	.257	.220	.140	.056
KLD-N	.153	.197	.204	.197	.176	.149	.114	.066	.024
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.681	.720	.794	.950	.976	.962	.929	.728	.172
LR	.904	.946	.971	.982	.982	.973	.940	.781	.298
ST	.646	.627	.663	.838	.972	.974	.940	.791	.358
KLD-U	.771	.840	.914	.959	.970	.964	.937	.780	.281
KLD-N	.860	.898	.927	.950	.953	.938	.896	.691	.203
Moderate Discrimination (a = 1.0)									
Z	.590	.648	.761	.851	.861	.838	.770	.505	.127
LR	.760	.823	.862	.880	.871	.837	.759	.567	.223
ST	.682	.748	.826	.877	.873	.836	.754	.604	.344
KLD-U	.606	.694	.767	.814	.824	.802	.723	.510	.184
KLD-N	.677	.731	.761	.772	.749	.696	.589	.377	.118
Low Discrimination (a = 0.6)									
Z	.435	.497	.555	.584	.588	.570	.489	.294	.102
LR	.500	.562	.600	.611	.598	.559	.484	.321	.121
ST	.484	.554	.597	.607	.591	.547	.477	.375	.242
KLD-U	.339	.442	.501	.530	.536	.501	.404	.240	.087
KLD-N	.382	.442	.459	.441	.401	.338	.245	.129	.040

Table A44. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.098	.048	.040	.039	.042	.043	.045	.050	.050
LR	.058	.053	.052	.050	.050	.049	.048	.046	.042
ST	.056	.052	.054	.052	.052	.050	.048	.044	.040
KLD-U	.057	.057	.050	.046	.047	.046	.048	.053	.045
KLD-N	.072	.064	.054	.048	.047	.044	.043	.041	.034
Moderate Discrimination (a = 1.0)									
Z	.077	.054	.042	.041	.042	.045	.047	.054	.051
LR	.057	.055	.052	.050	.048	.050	.049	.048	.042
ST	.056	.054	.053	.052	.050	.050	.049	.046	.043
KLD-U	.044	.053	.054	.051	.050	.053	.055	.054	.035
KLD-N	.047	.050	.047	.044	.042	.041	.038	.033	.023
Low Discrimination (a = 0.6)									
Z	.058	.057	.049	.044	.044	.046	.051	.055	.047
LR	.056	.056	.054	.051	.050	.050	.049	.048	.037
ST	.057	.056	.053	.051	.050	.049	.047	.045	.042
KLD-U	.031	.043	.052	.053	.055	.056	.055	.042	.021
KLD-N	.029	.036	.039	.037	.036	.034	.030	.023	.013
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.126	.096	.094	.100	.101	.099	.095	.097	.065
LR	.089	.096	.104	.108	.109	.108	.102	.096	.069
ST	.082	.095	.106	.112	.113	.113	.106	.096	.073
KLD-U	.063	.080	.080	.084	.082	.079	.078	.079	.054
KLD-N	.078	.088	.086	.088	.083	.076	.068	.062	.041
Moderate Discrimination (a = 1.0)									
Z	.098	.080	.075	.073	.074	.074	.078	.081	.060
LR	.074	.078	.080	.081	.083	.081	.080	.076	.056
ST	.072	.077	.081	.083	.086	.084	.082	.075	.065
KLD-U	.045	.066	.071	.070	.070	.068	.070	.064	.040
KLD-N	.051	.063	.064	.061	.056	.051	.045	.037	.023
Low Discrimination (a = 0.6)									
Z	.070	.072	.064	.060	.060	.062	.068	.068	.051
LR	.062	.068	.068	.066	.066	.067	.067	.061	.043
ST	.063	.067	.068	.067	.066	.066	.064	.060	.054
KLD-U	.029	.047	.059	.062	.062	.062	.059	.045	.022
KLD-N	.031	.043	.047	.044	.040	.035	.029	.021	.011
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.241	.243	.265	.284	.287	.272	.256	.236	.106
LR	.205	.245	.277	.298	.305	.293	.271	.232	.133
ST	.196	.246	.283	.304	.312	.301	.278	.233	.146
KLD-U	.155	.202	.224	.240	.242	.227	.218	.195	.101
KLD-N	.176	.217	.239	.249	.243	.218	.194	.161	.082

Continued on the next page

Table A44 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.166	.161	.166	.173	.171	.173	.165	.155	.089
LR	.141	.160	.174	.185	.185	.185	.169	.149	.094
ST	.136	.161	.177	.190	.191	.189	.171	.150	.116
KLD-U	.090	.137	.150	.157	.155	.155	.143	.122	.064
KLD-N	.099	.131	.139	.141	.130	.119	.097	.074	.039
Low Discrimination (a = 0.6)									
Z	.105	.113	.108	.108	.108	.111	.115	.104	.067
LR	.094	.109	.114	.117	.118	.116	.112	.098	.061
ST	.092	.108	.113	.118	.119	.116	.109	.098	.082
KLD-U	.043	.078	.098	.104	.104	.102	.093	.067	.032
KLD-N	.050	.072	.080	.078	.070	.059	.046	.031	.014
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.675	.733	.780	.802	.800	.769	.716	.586	.166
LR	.660	.738	.788	.814	.817	.791	.736	.605	.268
ST	.653	.739	.789	.818	.822	.798	.741	.606	.310
KLD-U	.609	.684	.738	.762	.761	.735	.684	.555	.207
KLD-N	.629	.703	.751	.769	.758	.718	.644	.498	.178
Moderate Discrimination (a = 1.0)									
Z	.448	.494	.530	.548	.543	.523	.485	.376	.135
LR	.433	.501	.546	.569	.566	.541	.487	.390	.191
ST	.429	.504	.550	.576	.574	.547	.489	.401	.265
KLD-U	.369	.461	.502	.521	.515	.494	.441	.328	.139
KLD-N	.370	.446	.482	.488	.464	.417	.339	.235	.094
Low Discrimination (a = 0.6)									
Z	.259	.288	.301	.306	.308	.307	.286	.209	.103
LR	.246	.290	.314	.325	.325	.312	.283	.215	.108
ST	.242	.287	.315	.327	.326	.310	.279	.231	.175
KLD-U	.158	.246	.284	.295	.294	.279	.235	.148	.062
KLD-N	.164	.222	.241	.236	.215	.179	.132	.076	.028
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.935	.962	.976	.982	.977	.961	.927	.729	.175
LR	.930	.959	.976	.984	.982	.975	.948	.810	.341
ST	.925	.956	.975	.984	.982	.976	.948	.819	.405
KLD-U	.922	.955	.971	.978	.975	.965	.934	.766	.265
KLD-N	.924	.956	.972	.977	.973	.959	.916	.723	.234
Moderate Discrimination (a = 1.0)									
Z	.792	.835	.862	.874	.864	.836	.766	.507	.133
LR	.788	.839	.870	.885	.876	.845	.776	.602	.261
ST	.785	.839	.871	.887	.879	.848	.778	.641	.400
KLD-U	.758	.816	.847	.859	.850	.817	.737	.525	.201
KLD-N	.749	.804	.831	.835	.810	.749	.635	.422	.148
Low Discrimination (a = 0.6)									
Z	.514	.563	.588	.597	.591	.566	.480	.295	.113
LR	.509	.574	.606	.617	.605	.568	.500	.349	.145
ST	.504	.572	.607	.618	.604	.564	.501	.407	.287
KLD-U	.428	.535	.573	.583	.570	.522	.422	.255	.096
KLD-N	.414	.488	.510	.498	.453	.376	.272	.145	.048

Table A45. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.096	.047	.040	.040	.041	.042	.045	.050	.047
LR	.058	.054	.051	.050	.049	.048	.048	.047	.043
ST	.056	.052	.053	.051	.050	.049	.048	.045	.041
KLD-U	.054	.058	.053	.049	.047	.046	.047	.052	.041
KLD-N	.069	.065	.056	.050	.047	.044	.043	.042	.032
Moderate Discrimination (a = 1.0)									
Z	.077	.053	.042	.042	.043	.046	.049	.053	.049
LR	.058	.055	.051	.051	.049	.051	.050	.049	.043
ST	.057	.054	.052	.052	.050	.051	.049	.046	.045
KLD-U	.043	.053	.054	.053	.052	.054	.055	.053	.033
KLD-N	.050	.053	.050	.047	.045	.044	.041	.035	.024
Low Discrimination (a = 0.6)									
Z	.058	.056	.050	.044	.044	.046	.051	.053	.043
LR	.054	.053	.053	.050	.050	.050	.050	.049	.036
ST	.055	.053	.053	.050	.050	.049	.047	.046	.043
KLD-U	.032	.045	.056	.058	.059	.059	.057	.044	.022
KLD-N	.030	.036	.041	.040	.038	.035	.031	.024	.013
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.126	.096	.097	.101	.103	.100	.098	.095	.060
LR	.090	.094	.103	.107	.110	.109	.103	.096	.070
ST	.083	.093	.105	.110	.114	.112	.106	.097	.076
KLD-U	.062	.086	.087	.088	.086	.081	.079	.076	.047
KLD-N	.076	.089	.090	.090	.087	.080	.072	.062	.038
Moderate Discrimination (a = 1.0)									
Z	.098	.081	.076	.077	.078	.078	.080	.079	.054
LR	.075	.078	.081	.082	.085	.083	.081	.077	.056
ST	.072	.077	.081	.083	.087	.084	.081	.077	.067
KLD-U	.047	.070	.076	.076	.075	.070	.070	.063	.036
KLD-N	.055	.067	.070	.068	.066	.056	.051	.041	.024
Low Discrimination (a = 0.6)									
Z	.072	.072	.066	.063	.062	.065	.069	.064	.044
LR	.062	.066	.067	.066	.067	.067	.067	.061	.040
ST	.062	.065	.066	.066	.067	.066	.065	.061	.056
KLD-U	.033	.052	.066	.070	.069	.066	.063	.047	.023
KLD-N	.033	.044	.049	.048	.044	.038	.033	.023	.011
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.242	.244	.268	.288	.288	.274	.261	.227	.098
LR	.207	.243	.278	.299	.302	.291	.270	.232	.134
ST	.196	.242	.281	.302	.308	.298	.276	.234	.149
KLD-U	.159	.214	.237	.250	.246	.230	.219	.185	.090
KLD-N	.175	.221	.246	.257	.249	.226	.203	.158	.075

Continued on the next page

Table A45 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.166	.164	.169	.178	.178	.180	.170	.148	.077
LR	.141	.162	.174	.187	.188	.188	.172	.150	.094
ST	.135	.161	.176	.189	.192	.191	.173	.153	.119
KLD-U	.094	.146	.159	.167	.164	.157	.144	.117	.060
KLD-N	.104	.140	.150	.155	.145	.131	.110	.080	.042
Low Discrimination (a = 0.6)									
Z	.106	.115	.110	.111	.114	.116	.113	.094	.054
LR	.092	.108	.112	.117	.120	.116	.111	.095	.056
ST	.090	.105	.110	.116	.120	.116	.110	.099	.084
KLD-U	.049	.086	.107	.116	.116	.109	.098	.070	.033
KLD-N	.050	.074	.084	.082	.077	.066	.050	.034	.014
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.674	.730	.778	.799	.796	.766	.710	.566	.151
LR	.656	.733	.782	.808	.810	.785	.730	.602	.268
ST	.647	.731	.781	.810	.814	.791	.735	.607	.313
KLD-U	.624	.695	.742	.762	.760	.729	.676	.528	.186
KLD-N	.630	.706	.753	.770	.761	.719	.646	.479	.163
Moderate Discrimination (a = 1.0)									
Z	.453	.498	.535	.552	.549	.527	.482	.349	.112
LR	.436	.500	.545	.567	.566	.538	.489	.389	.187
ST	.429	.499	.547	.571	.571	.542	.492	.404	.269
KLD-U	.384	.475	.513	.528	.520	.489	.436	.314	.126
KLD-N	.384	.462	.499	.507	.484	.437	.359	.244	.097
Low Discrimination (a = 0.6)									
Z	.261	.292	.304	.312	.314	.312	.274	.178	.076
LR	.243	.289	.311	.324	.324	.313	.284	.206	.094
ST	.239	.284	.310	.324	.324	.310	.284	.234	.177
KLD-U	.175	.265	.303	.316	.309	.291	.244	.152	.062
KLD-N	.168	.229	.250	.247	.227	.194	.144	.081	.029
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.932	.958	.975	.980	.974	.958	.919	.710	.160
LR	.926	.955	.973	.981	.980	.972	.944	.809	.340
ST	.920	.950	.971	.980	.980	.973	.946	.821	.408
KLD-U	.924	.954	.970	.975	.971	.961	.924	.743	.238
KLD-N	.923	.954	.971	.976	.970	.957	.907	.702	.213
Moderate Discrimination (a = 1.0)									
Z	.790	.833	.860	.871	.862	.832	.747	.462	.107
LR	.783	.833	.866	.879	.872	.840	.773	.598	.254
ST	.778	.831	.865	.881	.874	.842	.777	.644	.403
KLD-U	.765	.820	.849	.859	.844	.810	.724	.506	.180
KLD-N	.754	.811	.839	.843	.818	.762	.646	.428	.145
Low Discrimination (a = 0.6)									
Z	.516	.565	.589	.599	.594	.558	.436	.236	.078
LR	.504	.568	.600	.612	.601	.566	.490	.326	.123
ST	.497	.564	.598	.612	.600	.562	.500	.408	.283
KLD-U	.450	.552	.591	.599	.581	.534	.427	.256	.093
KLD-N	.418	.496	.519	.509	.471	.395	.284	.147	.044

Table A46. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.046	.042	.042	.043	.043	.044	.045	.048	.050
LR	.062	.056	.052	.049	.046	.044	.044	.045	.044
ST	.054	.052	.050	.050	.049	.048	.048	.049	.046
KLD-U	.065	.055	.048	.046	.044	.043	.045	.048	.044
KLD-N	.083	.060	.047	.041	.036	.034	.033	.033	.028
Moderate Discrimination (a = 1.0)									
Z	.040	.041	.041	.043	.044	.045	.048	.051	.054
LR	.057	.054	.050	.050	.048	.047	.046	.047	.045
ST	.052	.052	.050	.051	.051	.049	.048	.048	.045
KLD-U	.048	.047	.044	.044	.043	.044	.045	.046	.033
KLD-N	.059	.048	.039	.036	.032	.030	.029	.028	.021
Low Discrimination (a = 0.6)									
Z	.041	.043	.043	.045	.045	.046	.049	.054	.057
LR	.054	.054	.051	.050	.048	.049	.049	.049	.048
ST	.052	.053	.052	.052	.050	.050	.049	.048	.046
KLD-U	.037	.042	.044	.044	.042	.044	.043	.036	.021
KLD-N	.043	.040	.034	.030	.026	.024	.022	.019	.014
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.073	.073	.076	.080	.084	.089	.091	.090	.092
LR	.093	.091	.091	.092	.092	.091	.089	.084	.081
ST	.086	.085	.088	.092	.096	.097	.094	.090	.087
KLD-U	.065	.062	.062	.062	.060	.064	.067	.066	.057
KLD-N	.090	.073	.064	.057	.048	.045	.040	.032	.024
Moderate Discrimination (a = 1.0)									
Z	.057	.059	.060	.066	.067	.070	.071	.075	.078
LR	.077	.075	.074	.075	.072	.072	.070	.068	.065
ST	.072	.071	.072	.077	.075	.074	.073	.069	.066
KLD-U	.045	.048	.047	.050	.050	.052	.053	.051	.036
KLD-N	.065	.056	.047	.042	.036	.030	.025	.021	.014
Low Discrimination (a = 0.6)									
Z	.052	.052	.054	.055	.056	.059	.061	.069	.068
LR	.066	.064	.064	.062	.061	.061	.059	.061	.056
ST	.063	.063	.064	.063	.062	.062	.059	.059	.054
KLD-U	.034	.040	.044	.042	.044	.045	.042	.036	.021
KLD-N	.049	.043	.037	.030	.024	.020	.016	.013	.009
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.161	.177	.186	.207	.222	.234	.234	.226	.208
LR	.200	.213	.219	.235	.241	.238	.233	.218	.194
ST	.189	.204	.213	.234	.249	.248	.244	.229	.207
KLD-U	.129	.142	.146	.161	.166	.174	.181	.172	.136
KLD-N	.172	.165	.154	.152	.138	.126	.114	.089	.059

Continued on the next page

Table A46 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.108	.117	.125	.133	.139	.146	.150	.153	.146
LR	.139	.144	.146	.148	.149	.149	.148	.142	.129
ST	.131	.138	.145	.150	.152	.155	.151	.144	.131
KLD-U	.077	.088	.092	.096	.099	.106	.110	.102	.071
KLD-N	.111	.104	.094	.082	.071	.063	.052	.040	.026
Low Discrimination (a = 0.6)									
Z	.081	.085	.088	.091	.091	.096	.103	.108	.104
LR	.099	.101	.100	.101	.097	.098	.098	.096	.089
ST	.095	.098	.100	.102	.099	.098	.098	.094	.088
KLD-U	.047	.058	.061	.065	.065	.067	.066	.054	.033
KLD-N	.071	.067	.056	.047	.035	.028	.021	.016	.012
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.509	.555	.590	.651	.713	.726	.705	.662	.588
LR	.609	.650	.693	.730	.734	.732	.708	.661	.595
ST	.543	.580	.630	.693	.734	.740	.719	.674	.616
KLD-U	.464	.508	.554	.604	.634	.653	.644	.592	.432
KLD-N	.548	.562	.573	.578	.567	.549	.497	.407	.267
Moderate Discrimination (a = 1.0)									
Z	.337	.366	.398	.436	.451	.462	.456	.445	.388
LR	.406	.432	.453	.469	.468	.468	.451	.424	.378
ST	.380	.410	.442	.470	.475	.473	.456	.426	.384
KLD-U	.273	.305	.331	.355	.368	.385	.379	.330	.209
KLD-N	.338	.338	.329	.311	.282	.257	.216	.163	.107
Low Discrimination (a = 0.6)									
Z	.205	.220	.230	.238	.245	.250	.267	.266	.227
LR	.240	.251	.256	.255	.256	.252	.257	.244	.213
ST	.232	.248	.256	.257	.258	.250	.253	.238	.218
KLD-U	.136	.161	.171	.178	.186	.186	.182	.139	.080
KLD-N	.180	.171	.151	.127	.105	.082	.067	.049	.033
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.796	.798	.830	.930	.966	.957	.940	.911	.807
LR	.894	.929	.955	.967	.970	.964	.954	.930	.876
ST	.641	.658	.699	.828	.949	.966	.956	.934	.891
KLD-U	.784	.835	.892	.932	.950	.948	.936	.891	.680
KLD-N	.859	.882	.901	.915	.919	.905	.864	.777	.548
Moderate Discrimination (a = 1.0)									
Z	.637	.668	.731	.786	.805	.802	.783	.743	.567
LR	.744	.774	.798	.810	.815	.806	.779	.738	.642
ST	.657	.699	.760	.804	.817	.808	.780	.740	.670
KLD-U	.602	.646	.687	.720	.744	.747	.715	.610	.372
KLD-N	.673	.677	.672	.656	.635	.593	.513	.400	.256
Low Discrimination (a = 0.6)									
Z	.423	.443	.467	.482	.494	.507	.507	.474	.346
LR	.481	.491	.501	.504	.505	.505	.490	.461	.374
ST	.460	.482	.500	.506	.505	.501	.484	.456	.416
KLD-U	.339	.366	.386	.400	.413	.412	.369	.274	.152
KLD-N	.384	.362	.334	.292	.253	.211	.167	.125	.083

Table A47. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.040	.042	.044	.046	.046	.045	.047	.048	.050
LR	.056	.053	.051	.050	.048	.046	.047	.048	.047
ST	.057	.054	.052	.051	.048	.046	.047	.048	.045
KLD-U	.065	.057	.052	.049	.045	.043	.046	.049	.046
KLD-N	.085	.065	.054	.050	.044	.041	.041	.041	.037
Moderate Discrimination (a = 1.0)									
Z	.041	.042	.045	.045	.046	.048	.047	.051	.054
LR	.054	.052	.051	.049	.050	.050	.048	.049	.049
ST	.054	.052	.051	.050	.050	.050	.048	.048	.047
KLD-U	.053	.054	.052	.051	.050	.051	.052	.052	.039
KLD-N	.058	.051	.045	.042	.039	.037	.035	.033	.026
Low Discrimination (a = 0.6)									
Z	.043	.045	.046	.046	.046	.048	.050	.054	.056
LR	.052	.053	.050	.051	.049	.050	.050	.050	.049
ST	.053	.054	.051	.052	.050	.051	.050	.049	.047
KLD-U	.041	.048	.050	.051	.049	.050	.048	.041	.024
KLD-N	.042	.041	.036	.034	.029	.029	.026	.021	.015
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.085	.089	.098	.100	.099	.097	.095	.093	.090
LR	.095	.093	.097	.098	.098	.098	.096	.093	.090
ST	.100	.097	.100	.098	.100	.101	.099	.097	.096
KLD-U	.081	.080	.080	.077	.073	.070	.072	.072	.064
KLD-N	.097	.088	.085	.079	.073	.066	.062	.056	.048
Moderate Discrimination (a = 1.0)									
Z	.070	.072	.074	.076	.075	.075	.072	.075	.075
LR	.078	.077	.076	.076	.076	.077	.074	.074	.072
ST	.080	.078	.078	.077	.077	.079	.076	.075	.074
KLD-U	.060	.066	.066	.066	.064	.064	.062	.061	.046
KLD-N	.067	.063	.058	.055	.049	.044	.039	.034	.025
Low Discrimination (a = 0.6)									
Z	.061	.060	.061	.059	.060	.062	.062	.067	.066
LR	.065	.063	.064	.062	.062	.063	.062	.063	.060
ST	.067	.065	.065	.062	.063	.065	.063	.063	.061
KLD-U	.042	.050	.055	.054	.053	.054	.050	.044	.027
KLD-N	.047	.045	.041	.035	.031	.028	.022	.018	.012
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.214	.235	.250	.263	.261	.257	.250	.239	.210
LR	.224	.237	.247	.258	.262	.260	.255	.242	.220
ST	.234	.242	.250	.260	.266	.266	.262	.252	.235
KLD-U	.187	.198	.202	.207	.202	.200	.199	.193	.156
KLD-N	.208	.213	.214	.214	.201	.188	.176	.158	.124

Continued on the next page

Table A47 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.142	.151	.157	.160	.159	.156	.154	.156	.144
LR	.148	.155	.157	.160	.161	.162	.158	.156	.143
ST	.153	.158	.159	.161	.164	.165	.162	.159	.152
KLD-U	.116	.130	.133	.133	.131	.129	.129	.124	.091
KLD-N	.123	.125	.121	.114	.105	.093	.085	.071	.052
Low Discrimination (a = 0.6)									
Z	.098	.100	.102	.101	.099	.098	.105	.106	.101
LR	.101	.102	.103	.103	.101	.101	.105	.102	.096
ST	.103	.104	.104	.105	.104	.104	.106	.104	.099
KLD-U	.062	.078	.084	.084	.083	.082	.080	.066	.042
KLD-N	.072	.070	.065	.058	.049	.042	.036	.026	.017
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.657	.706	.739	.757	.750	.745	.717	.672	.594
LR	.664	.703	.732	.754	.755	.753	.731	.694	.644
ST	.669	.703	.731	.756	.759	.758	.740	.708	.672
KLD-U	.619	.656	.678	.694	.687	.688	.672	.625	.497
KLD-N	.639	.673	.692	.699	.680	.664	.631	.565	.454
Moderate Discrimination (a = 1.0)									
Z	.440	.464	.477	.489	.481	.480	.469	.447	.400
LR	.446	.465	.479	.493	.492	.491	.478	.453	.422
ST	.450	.467	.482	.495	.496	.497	.484	.463	.443
KLD-U	.394	.420	.430	.439	.432	.435	.425	.378	.265
KLD-N	.394	.407	.406	.398	.371	.351	.315	.264	.187
Low Discrimination (a = 0.6)									
Z	.252	.262	.263	.261	.261	.258	.270	.264	.231
LR	.253	.263	.266	.266	.267	.262	.269	.259	.241
ST	.257	.266	.270	.271	.271	.266	.272	.264	.252
KLD-U	.188	.218	.226	.226	.226	.218	.212	.167	.103
KLD-N	.193	.193	.181	.164	.147	.125	.105	.078	.048
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.926	.951	.964	.970	.969	.959	.940	.910	.814
LR	.922	.945	.961	.969	.972	.968	.959	.942	.911
ST	.917	.941	.959	.969	.972	.969	.961	.948	.928
KLD-U	.917	.938	.952	.957	.959	.955	.943	.911	.750
KLD-N	.918	.940	.953	.957	.955	.946	.927	.886	.743
Moderate Discrimination (a = 1.0)									
Z	.778	.798	.812	.817	.815	.809	.785	.747	.622
LR	.778	.797	.812	.822	.824	.820	.798	.768	.718
ST	.777	.796	.812	.824	.826	.823	.802	.779	.744
KLD-U	.747	.765	.777	.784	.784	.784	.754	.675	.456
KLD-N	.737	.750	.753	.744	.726	.704	.644	.558	.385
Low Discrimination (a = 0.6)									
Z	.499	.504	.507	.508	.511	.516	.509	.480	.383
LR	.500	.506	.514	.517	.520	.522	.510	.489	.450
ST	.503	.509	.518	.522	.525	.525	.514	.498	.479
KLD-U	.432	.453	.462	.464	.465	.459	.413	.321	.190
KLD-N	.414	.406	.388	.362	.334	.297	.244	.181	.113

Table A48. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.041	.043	.046	.049	.048	.048	.050	.050	.053
LR	.055	.052	.052	.051	.048	.047	.048	.047	.048
ST	.058	.054	.052	.051	.048	.048	.048	.047	.046
KLD-U	.066	.058	.054	.049	.044	.043	.044	.047	.047
KLD-N	.085	.065	.054	.049	.043	.041	.040	.040	.038
Moderate Discrimination (a = 1.0)									
Z	.042	.044	.045	.046	.049	.050	.050	.053	.056
LR	.053	.051	.050	.048	.049	.049	.049	.048	.048
ST	.054	.052	.051	.049	.049	.050	.048	.047	.046
KLD-U	.052	.055	.053	.050	.050	.050	.051	.052	.040
KLD-N	.063	.056	.049	.045	.042	.041	.038	.036	.030
Low Discrimination (a = 0.6)									
Z	.045	.047	.047	.048	.048	.050	.052	.056	.058
LR	.052	.053	.051	.051	.049	.050	.051	.051	.051
ST	.053	.053	.051	.051	.050	.050	.050	.050	.048
KLD-U	.043	.052	.054	.055	.054	.054	.053	.044	.028
KLD-N	.044	.043	.038	.035	.031	.029	.027	.024	.018
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.089	.094	.102	.105	.104	.103	.102	.099	.093
LR	.095	.096	.098	.099	.099	.099	.098	.093	.093
ST	.101	.100	.101	.100	.101	.103	.102	.097	.099
KLD-U	.086	.086	.084	.080	.075	.072	.071	.071	.064
KLD-N	.097	.089	.085	.079	.074	.068	.064	.061	.052
Moderate Discrimination (a = 1.0)									
Z	.074	.076	.078	.079	.079	.078	.078	.080	.078
LR	.077	.076	.076	.076	.076	.076	.075	.073	.072
ST	.079	.077	.077	.076	.077	.077	.076	.075	.075
KLD-U	.062	.070	.069	.068	.067	.064	.063	.061	.046
KLD-N	.072	.069	.064	.061	.056	.050	.046	.043	.033
Low Discrimination (a = 0.6)									
Z	.063	.063	.063	.062	.062	.066	.066	.069	.067
LR	.064	.064	.064	.063	.062	.064	.063	.063	.061
ST	.065	.064	.064	.063	.063	.064	.064	.063	.062
KLD-U	.045	.056	.062	.060	.060	.062	.056	.047	.031
KLD-N	.048	.047	.043	.038	.034	.031	.027	.022	.016
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.222	.243	.259	.272	.272	.265	.259	.250	.210
LR	.223	.239	.249	.259	.264	.261	.255	.244	.222
ST	.234	.246	.253	.262	.269	.267	.263	.253	.239
KLD-U	.197	.208	.211	.211	.206	.200	.196	.196	.150
KLD-N	.207	.215	.216	.214	.205	.192	.179	.173	.131

Continued on the next page

Table A48 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.148	.159	.163	.164	.164	.162	.163	.164	.146
LR	.146	.154	.155	.157	.159	.160	.157	.154	.144
ST	.151	.156	.157	.159	.162	.164	.161	.159	.155
KLD-U	.121	.137	.139	.137	.135	.132	.129	.124	.092
KLD-N	.130	.134	.131	.124	.116	.107	.099	.091	.067
Low Discrimination (a = 0.6)									
Z	.103	.105	.106	.105	.103	.105	.111	.110	.100
LR	.102	.103	.104	.104	.102	.103	.106	.102	.098
ST	.102	.103	.104	.105	.104	.104	.106	.105	.103
KLD-U	.070	.088	.094	.095	.093	.092	.088	.072	.048
KLD-N	.074	.073	.069	.063	.054	.048	.043	.034	.024
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.660	.709	.742	.761	.752	.746	.723	.676	.575
LR	.660	.698	.726	.750	.748	.747	.725	.690	.630
ST	.666	.699	.725	.752	.754	.754	.734	.705	.661
KLD-U	.626	.660	.678	.690	.682	.678	.661	.618	.485
KLD-N	.635	.668	.687	.695	.678	.662	.635	.576	.475
Moderate Discrimination (a = 1.0)									
Z	.446	.470	.487	.499	.489	.488	.480	.454	.369
LR	.439	.459	.475	.489	.485	.485	.472	.450	.410
ST	.443	.460	.476	.492	.491	.491	.479	.461	.435
KLD-U	.401	.426	.439	.445	.436	.433	.418	.376	.262
KLD-N	.405	.419	.423	.419	.396	.374	.354	.302	.231
Low Discrimination (a = 0.6)									
Z	.260	.270	.270	.268	.268	.270	.279	.262	.211
LR	.255	.264	.267	.267	.268	.266	.270	.259	.237
ST	.256	.264	.268	.268	.270	.268	.274	.265	.255
KLD-U	.207	.238	.245	.245	.244	.234	.226	.182	.113
KLD-N	.197	.197	.189	.175	.159	.141	.128	.099	.066
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.923	.948	.964	.968	.967	.957	.939	.904	.777
LR	.917	.939	.958	.965	.968	.964	.954	.934	.900
ST	.911	.934	.955	.965	.969	.965	.957	.942	.918
KLD-U	.918	.936	.949	.953	.953	.948	.937	.897	.734
KLD-N	.914	.936	.950	.952	.951	.942	.925	.880	.760
Moderate Discrimination (a = 1.0)									
Z	.780	.799	.812	.819	.813	.811	.787	.727	.524
LR	.770	.788	.803	.814	.812	.810	.787	.757	.689
ST	.769	.787	.802	.815	.816	.814	.792	.771	.725
KLD-U	.748	.767	.777	.782	.777	.774	.744	.661	.443
KLD-N	.744	.757	.761	.758	.739	.723	.679	.596	.443
Low Discrimination (a = 0.6)									
Z	.508	.513	.513	.514	.519	.524	.508	.446	.311
LR	.500	.506	.511	.516	.520	.519	.506	.485	.416
ST	.500	.506	.513	.519	.522	.522	.511	.497	.473
KLD-U	.457	.479	.484	.484	.485	.472	.429	.339	.201
KLD-N	.420	.414	.399	.377	.355	.329	.282	.218	.142

Table A49. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.111	.054	.038	.036	.036	.037	.038	.047	.058
LR	.053	.055	.054	.051	.050	.050	.047	.044	.044
ST	.044	.049	.054	.054	.055	.056	.052	.047	.042
KLD-U	.040	.062	.051	.047	.047	.048	.049	.054	.062
KLD-N	.047	.061	.051	.046	.045	.045	.044	.047	.046
Moderate Discrimination (a = 1.0)									
Z	.084	.056	.041	.038	.038	.038	.042	.050	.057
LR	.056	.055	.053	.050	.050	.048	.047	.047	.045
ST	.050	.051	.052	.053	.053	.051	.049	.047	.044
KLD-U	.040	.056	.051	.049	.049	.049	.049	.056	.049
KLD-N	.044	.052	.047	.044	.043	.042	.040	.042	.036
Low Discrimination (a = 0.6)									
Z	.060	.060	.049	.041	.041	.041	.048	.053	.055
LR	.054	.055	.052	.050	.051	.048	.050	.048	.045
ST	.054	.054	.053	.051	.052	.049	.050	.048	.044
KLD-U	.028	.046	.054	.051	.053	.052	.055	.054	.033
KLD-N	.030	.040	.042	.041	.041	.039	.040	.036	.025
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.136	.119	.120	.129	.132	.136	.139	.147	.112
LR	.105	.140	.156	.166	.169	.163	.154	.135	.092
ST	.093	.135	.157	.171	.176	.168	.156	.131	.084
KLD-U	.081	.130	.127	.133	.136	.136	.136	.139	.091
KLD-N	.091	.133	.133	.135	.132	.126	.120	.117	.066
Moderate Discrimination (a = 1.0)									
Z	.101	.085	.082	.086	.090	.097	.100	.104	.100
LR	.086	.098	.106	.111	.112	.112	.104	.091	.078
ST	.079	.094	.106	.114	.114	.112	.102	.087	.072
KLD-U	.059	.089	.091	.093	.095	.098	.096	.091	.066
KLD-N	.068	.087	.088	.087	.083	.080	.073	.064	.046
Low Discrimination (a = 0.6)									
Z	.071	.071	.066	.064	.068	.070	.079	.080	.078
LR	.070	.076	.078	.079	.080	.079	.076	.068	.062
ST	.069	.075	.078	.080	.080	.077	.074	.065	.058
KLD-U	.035	.062	.072	.073	.073	.073	.073	.063	.037
KLD-N	.044	.059	.062	.060	.057	.052	.046	.038	.023
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.309	.373	.418	.450	.466	.465	.446	.405	.233
LR	.302	.421	.485	.512	.522	.507	.466	.378	.213
ST	.287	.419	.490	.521	.533	.516	.470	.368	.200
KLD-U	.283	.385	.427	.452	.468	.463	.443	.388	.195
KLD-N	.288	.402	.442	.456	.458	.441	.410	.346	.150

Continued on the next page

Table A49 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.197	.217	.242	.258	.274	.277	.269	.259	.202
LR	.198	.249	.287	.303	.312	.303	.273	.233	.165
ST	.190	.245	.289	.308	.317	.304	.269	.223	.155
KLD-U	.161	.228	.254	.266	.278	.277	.259	.230	.133
KLD-N	.171	.225	.251	.254	.253	.238	.211	.176	.094
Low Discrimination (a = 0.6)									
Z	.120	.129	.138	.144	.151	.158	.161	.159	.133
LR	.128	.147	.165	.170	.173	.169	.154	.137	.110
ST	.127	.147	.165	.171	.172	.166	.148	.131	.104
KLD-U	.074	.126	.150	.154	.158	.158	.146	.122	.064
KLD-N	.088	.119	.134	.133	.127	.117	.097	.074	.038
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.854	.898	.925	.961	.973	.967	.947	.866	.413
LR	.891	.953	.972	.979	.980	.972	.947	.842	.469
ST	.879	.942	.966	.979	.980	.973	.947	.830	.467
KLD-U	.876	.935	.960	.971	.973	.966	.946	.848	.406
KLD-N	.888	.944	.963	.970	.970	.960	.937	.811	.336
Moderate Discrimination (a = 1.0)									
Z	.632	.721	.769	.806	.811	.796	.754	.669	.427
LR	.662	.765	.810	.838	.835	.811	.749	.632	.410
ST	.656	.764	.812	.842	.838	.811	.744	.617	.407
KLD-U	.637	.736	.779	.811	.813	.795	.743	.621	.309
KLD-N	.637	.737	.773	.795	.786	.755	.684	.540	.249
Low Discrimination (a = 0.6)									
Z	.359	.421	.466	.493	.505	.494	.466	.410	.305
LR	.386	.461	.508	.533	.534	.505	.452	.375	.279
ST	.386	.462	.510	.535	.532	.499	.442	.362	.272
KLD-U	.325	.437	.484	.508	.512	.488	.435	.322	.150
KLD-N	.329	.413	.453	.463	.449	.405	.334	.226	.103
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.868	.849	.926	.997	1.000	1.000	.998	.970	.456
LR	.997	.999	1.000	1.000	1.000	1.000	.998	.967	.588
ST	.934	.907	.940	.992	1.000	1.000	.998	.961	.607
KLD-U	.993	.999	1.000	1.000	1.000	1.000	.998	.964	.505
KLD-N	.997	.999	1.000	1.000	1.000	1.000	.998	.948	.432
Moderate Discrimination (a = 1.0)									
Z	.927	.950	.978	.989	.988	.982	.963	.890	.516
LR	.964	.983	.989	.991	.990	.983	.959	.876	.591
ST	.957	.979	.989	.991	.990	.982	.957	.866	.628
KLD-U	.956	.979	.986	.989	.988	.981	.959	.852	.443
KLD-N	.958	.979	.985	.987	.984	.973	.940	.799	.386
Low Discrimination (a = 0.6)									
Z	.726	.795	.832	.843	.845	.817	.765	.665	.448
LR	.755	.823	.856	.863	.858	.818	.749	.638	.460
ST	.755	.824	.857	.863	.856	.813	.740	.624	.474
KLD-U	.728	.808	.841	.850	.848	.810	.728	.537	.248
KLD-N	.714	.788	.819	.817	.799	.738	.626	.430	.195

Table A50. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.105	.050	.040	.038	.039	.039	.041	.047	.059
LR	.048	.051	.053	.052	.052	.052	.051	.049	.049
ST	.045	.050	.054	.054	.054	.054	.052	.047	.043
KLD-U	.038	.062	.054	.051	.050	.050	.050	.055	.060
KLD-N	.041	.059	.056	.053	.051	.050	.049	.049	.047
Moderate Discrimination (a = 1.0)									
Z	.080	.053	.041	.040	.040	.040	.043	.050	.056
LR	.051	.052	.049	.051	.050	.050	.049	.050	.048
ST	.050	.051	.050	.053	.053	.052	.050	.048	.044
KLD-U	.038	.057	.053	.052	.051	.051	.052	.056	.047
KLD-N	.039	.051	.049	.049	.047	.046	.044	.044	.037
Low Discrimination (a = 0.6)									
Z	.059	.058	.048	.042	.042	.043	.049	.053	.055
LR	.051	.052	.051	.050	.051	.050	.051	.050	.047
ST	.053	.052	.052	.051	.052	.050	.050	.049	.045
KLD-U	.030	.049	.056	.056	.056	.056	.058	.056	.034
KLD-N	.029	.040	.044	.044	.044	.042	.041	.038	.026
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.135	.133	.138	.150	.152	.150	.148	.148	.113
LR	.100	.144	.164	.178	.181	.175	.166	.148	.102
ST	.095	.141	.167	.181	.184	.178	.166	.143	.093
KLD-U	.080	.149	.151	.157	.155	.151	.144	.137	.088
KLD-N	.084	.146	.157	.163	.159	.150	.137	.123	.069
Moderate Discrimination (a = 1.0)									
Z	.101	.090	.091	.096	.098	.101	.104	.101	.098
LR	.082	.096	.106	.115	.117	.116	.111	.097	.082
ST	.079	.096	.108	.117	.119	.117	.109	.094	.078
KLD-U	.060	.099	.102	.105	.104	.103	.099	.089	.064
KLD-N	.064	.091	.099	.102	.098	.093	.082	.069	.048
Low Discrimination (a = 0.6)									
Z	.073	.073	.069	.069	.071	.074	.081	.081	.077
LR	.066	.073	.078	.080	.083	.081	.080	.073	.065
ST	.068	.073	.078	.080	.083	.079	.077	.070	.062
KLD-U	.038	.068	.080	.080	.080	.078	.077	.066	.040
KLD-N	.041	.058	.066	.066	.064	.057	.052	.041	.025
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.328	.420	.470	.496	.501	.491	.462	.407	.236
LR	.312	.447	.511	.538	.544	.531	.491	.402	.231
ST	.302	.446	.517	.543	.551	.536	.492	.390	.218
KLD-U	.299	.441	.486	.503	.506	.493	.458	.385	.191
KLD-N	.296	.445	.499	.514	.512	.491	.446	.357	.156

Continued on the next page

Table A50 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.210	.237	.266	.282	.290	.288	.276	.256	.199
LR	.200	.255	.296	.316	.323	.315	.288	.245	.176
ST	.196	.255	.300	.321	.327	.317	.286	.239	.168
KLD-U	.177	.257	.284	.295	.299	.290	.264	.228	.131
KLD-N	.175	.244	.280	.289	.285	.267	.231	.185	.102
Low Discrimination (a = 0.6)									
Z	.129	.139	.151	.156	.161	.167	.162	.156	.131
LR	.127	.147	.168	.176	.180	.176	.160	.143	.114
ST	.127	.148	.168	.177	.179	.173	.156	.138	.109
KLD-U	.084	.143	.169	.173	.172	.167	.150	.124	.068
KLD-N	.089	.125	.145	.148	.141	.130	.106	.080	.044
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.135	.133	.138	.150	.152	.150	.148	.148	.113
LR	.100	.144	.164	.178	.181	.175	.166	.148	.102
ST	.095	.141	.167	.181	.184	.178	.166	.143	.093
KLD-U	.080	.149	.151	.157	.155	.151	.144	.137	.088
KLD-N	.084	.146	.157	.163	.159	.150	.137	.123	.069
Moderate Discrimination (a = 1.0)									
Z	.101	.090	.091	.096	.098	.101	.104	.101	.098
LR	.082	.096	.106	.115	.117	.116	.111	.097	.082
ST	.079	.096	.108	.117	.119	.117	.109	.094	.078
KLD-U	.060	.099	.102	.105	.104	.103	.099	.089	.064
KLD-N	.064	.091	.099	.102	.098	.093	.082	.069	.048
Low Discrimination (a = 0.6)									
Z	.073	.073	.069	.069	.071	.074	.081	.081	.077
LR	.066	.073	.078	.080	.083	.081	.080	.073	.065
ST	.068	.073	.078	.080	.083	.079	.077	.070	.062
KLD-U	.038	.068	.080	.080	.080	.078	.077	.066	.040
KLD-N	.041	.058	.066	.066	.064	.057	.052	.041	.025
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.997	.999	1.000	1.000	1.000	1.000	.999	.970	.464
LR	.997	.999	1.000	1.000	1.000	1.000	.998	.971	.614
ST	.997	.999	1.000	1.000	1.000	1.000	.998	.967	.641
KLD-U	.998	.999	1.000	1.000	1.000	1.000	.998	.963	.499
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.998	.950	.444
Moderate Discrimination (a = 1.0)									
Z	.964	.982	.988	.990	.988	.982	.963	.891	.524
LR	.968	.984	.990	.992	.990	.984	.963	.886	.617
ST	.967	.984	.990	.992	.991	.984	.962	.879	.658
KLD-U	.969	.984	.989	.991	.989	.982	.959	.851	.445
KLD-N	.966	.983	.989	.990	.987	.977	.944	.813	.409
Low Discrimination (a = 0.6)									
Z	.750	.812	.845	.853	.852	.822	.766	.668	.457
LR	.762	.828	.861	.868	.863	.826	.758	.652	.482
ST	.762	.828	.862	.868	.862	.823	.752	.642	.497
KLD-U	.761	.831	.860	.864	.857	.818	.735	.551	.264
KLD-N	.730	.804	.834	.836	.817	.756	.643	.457	.218

Table A51. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.104	.050	.040	.038	.039	.040	.042	.046	.057
LR	.048	.052	.052	.052	.052	.052	.051	.050	.050
ST	.046	.050	.053	.054	.054	.054	.052	.048	.045
KLD-U	.036	.061	.054	.052	.049	.051	.050	.055	.058
KLD-N	.040	.060	.057	.055	.052	.052	.050	.051	.047
Moderate Discrimination (a = 1.0)									
Z	.080	.053	.042	.041	.041	.042	.046	.051	.056
LR	.051	.052	.049	.051	.050	.050	.050	.050	.048
ST	.050	.051	.050	.053	.052	.051	.050	.049	.045
KLD-U	.037	.057	.054	.053	.051	.052	.053	.056	.046
KLD-N	.039	.052	.051	.052	.050	.048	.047	.046	.037
Low Discrimination (a = 0.6)									
Z	.060	.058	.048	.043	.044	.045	.049	.053	.054
LR	.050	.052	.051	.050	.050	.050	.051	.050	.048
ST	.051	.052	.051	.050	.051	.050	.050	.049	.045
KLD-U	.029	.048	.056	.056	.056	.056	.057	.055	.033
KLD-N	.030	.041	.045	.045	.045	.043	.042	.038	.026
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.136	.135	.140	.153	.154	.151	.152	.147	.111
LR	.102	.146	.165	.180	.181	.175	.169	.151	.104
ST	.096	.143	.167	.182	.184	.177	.168	.146	.096
KLD-U	.078	.152	.155	.162	.159	.153	.146	.135	.084
KLD-N	.083	.150	.163	.170	.166	.156	.144	.125	.068
Moderate Discrimination (a = 1.0)									
Z	.101	.093	.094	.100	.101	.104	.108	.101	.097
LR	.081	.098	.106	.115	.116	.116	.112	.098	.084
ST	.078	.096	.107	.116	.117	.116	.110	.095	.078
KLD-U	.059	.103	.105	.108	.106	.105	.100	.088	.062
KLD-N	.064	.095	.103	.106	.102	.097	.088	.072	.048
Low Discrimination (a = 0.6)									
Z	.074	.075	.071	.071	.074	.077	.081	.079	.076
LR	.067	.074	.078	.080	.082	.082	.080	.073	.066
ST	.068	.074	.077	.080	.081	.080	.077	.070	.063
KLD-U	.038	.069	.082	.082	.081	.078	.076	.064	.039
KLD-N	.043	.060	.067	.068	.066	.060	.054	.042	.026
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.327	.421	.472	.496	.504	.494	.464	.399	.233
LR	.315	.447	.510	.536	.544	.530	.492	.403	.235
ST	.304	.445	.514	.540	.549	.534	.492	.392	.223
KLD-U	.295	.448	.491	.508	.511	.496	.457	.376	.184
KLD-N	.295	.452	.506	.522	.522	.501	.454	.356	.155

Continued on the next page

Table A51 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.213	.242	.273	.288	.298	.296	.281	.254	.197
LR	.202	.256	.297	.317	.325	.316	.289	.248	.179
ST	.197	.255	.298	.318	.326	.316	.286	.241	.171
KLD-U	.178	.263	.291	.301	.305	.293	.266	.224	.128
KLD-N	.179	.252	.289	.299	.298	.278	.241	.190	.104
Low Discrimination (a = 0.6)									
Z	.130	.141	.156	.161	.167	.169	.163	.154	.129
LR	.126	.148	.169	.176	.181	.176	.162	.143	.116
ST	.126	.147	.168	.175	.179	.173	.158	.139	.111
KLD-U	.086	.145	.173	.175	.174	.166	.150	.121	.066
KLD-N	.091	.128	.150	.153	.148	.134	.111	.083	.046
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.901	.953	.970	.976	.975	.968	.947	.864	.416
LR	.904	.959	.975	.980	.980	.973	.953	.859	.499
ST	.901	.959	.975	.980	.981	.974	.952	.848	.506
KLD-U	.911	.959	.973	.977	.977	.968	.946	.840	.389
KLD-N	.908	.961	.975	.979	.978	.969	.943	.816	.344
Moderate Discrimination (a = 1.0)									
Z	.673	.760	.800	.823	.823	.805	.755	.664	.427
LR	.678	.777	.818	.842	.840	.819	.762	.652	.435
ST	.674	.777	.819	.844	.842	.819	.759	.642	.437
KLD-U	.685	.779	.812	.831	.828	.803	.742	.613	.303
KLD-N	.669	.774	.810	.828	.820	.788	.709	.559	.270
Low Discrimination (a = 0.6)									
Z	.385	.453	.495	.516	.523	.508	.469	.408	.299
LR	.391	.471	.518	.538	.543	.517	.464	.391	.293
ST	.390	.469	.517	.538	.541	.512	.458	.383	.291
KLD-U	.361	.479	.521	.534	.532	.501	.439	.327	.157
KLD-N	.343	.440	.486	.498	.484	.437	.360	.251	.123
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.997	.999	1.000	1.000	1.000	1.000	.998	.970	.460
LR	.997	.999	1.000	1.000	1.000	1.000	.998	.972	.619
ST	.996	.999	1.000	1.000	1.000	1.000	.998	.968	.649
KLD-U	.998	.999	1.000	1.000	1.000	1.000	.998	.961	.489
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.998	.950	.441
Moderate Discrimination (a = 1.0)									
Z	.964	.982	.987	.990	.988	.982	.962	.889	.516
LR	.966	.984	.989	.992	.990	.984	.962	.886	.618
ST	.965	.983	.989	.992	.990	.983	.961	.880	.660
KLD-U	.969	.984	.988	.990	.988	.981	.957	.847	.438
KLD-N	.966	.984	.988	.990	.987	.977	.945	.814	.412
Low Discrimination (a = 0.6)									
Z	.750	.813	.844	.854	.852	.822	.765	.662	.438
LR	.760	.825	.858	.866	.862	.825	.758	.652	.476
ST	.758	.825	.858	.866	.860	.822	.753	.644	.500
KLD-U	.760	.832	.858	.863	.855	.814	.730	.545	.258
KLD-N	.730	.806	.837	.839	.822	.761	.652	.465	.225

Table A52. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.054	.051	.049	.048	.047	.048	.048	.051	.054
LR	.055	.054	.052	.050	.047	.048	.047	.049	.047
ST	.049	.050	.050	.050	.049	.051	.049	.050	.050
KLD-U	.054	.053	.051	.049	.047	.050	.048	.051	.050
KLD-N	.065	.056	.050	.046	.043	.044	.042	.043	.042
Moderate Discrimination (a = 1.0)									
Z	.051	.049	.049	.048	.049	.047	.050	.052	.056
LR	.056	.053	.052	.050	.050	.048	.048	.048	.048
ST	.051	.050	.051	.050	.051	.050	.050	.049	.049
KLD-U	.049	.050	.051	.048	.048	.048	.049	.050	.046
KLD-N	.058	.050	.046	.042	.041	.039	.039	.038	.036
Low Discrimination (a = 0.6)									
Z	.048	.048	.048	.046	.048	.048	.050	.054	.056
LR	.053	.051	.051	.049	.049	.050	.049	.048	.048
ST	.051	.050	.052	.050	.051	.051	.050	.048	.050
KLD-U	.042	.047	.049	.048	.047	.049	.048	.045	.034
KLD-N	.049	.043	.040	.035	.034	.033	.032	.030	.028
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.130	.130	.130	.132	.134	.138	.141	.143	.152
LR	.141	.142	.139	.138	.139	.139	.138	.135	.134
ST	.134	.137	.136	.137	.140	.140	.137	.134	.133
KLD-U	.118	.118	.116	.115	.117	.119	.120	.118	.114
KLD-N	.139	.129	.120	.112	.108	.104	.100	.092	.085
Moderate Discrimination (a = 1.0)									
Z	.089	.090	.093	.092	.092	.095	.100	.107	.112
LR	.101	.100	.100	.098	.096	.096	.096	.095	.093
ST	.094	.095	.096	.096	.095	.095	.094	.092	.090
KLD-U	.078	.081	.082	.080	.078	.080	.083	.083	.073
KLD-N	.097	.087	.081	.073	.066	.062	.060	.056	.046
Low Discrimination (a = 0.6)									
Z	.067	.068	.069	.068	.068	.071	.075	.082	.083
LR	.074	.074	.074	.073	.071	.071	.072	.072	.069
ST	.071	.072	.072	.072	.071	.071	.070	.070	.067
KLD-U	.052	.060	.062	.061	.059	.060	.060	.057	.042
KLD-N	.068	.061	.054	.048	.042	.038	.034	.031	.023
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.384	.394	.407	.419	.421	.429	.431	.432	.422
LR	.401	.410	.419	.428	.428	.428	.425	.419	.397
ST	.397	.406	.417	.428	.431	.431	.426	.418	.396
KLD-U	.359	.367	.376	.383	.386	.393	.394	.385	.357
KLD-N	.391	.386	.383	.377	.368	.361	.349	.327	.290

Continued on the next page

Table A52 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.225	.234	.238	.243	.242	.244	.258	.260	.264
LR	.243	.248	.248	.252	.249	.245	.250	.242	.236
ST	.236	.242	.245	.250	.248	.243	.246	.236	.231
KLD-U	.201	.211	.214	.215	.213	.213	.223	.214	.188
KLD-N	.229	.222	.211	.200	.186	.175	.170	.154	.125
Low Discrimination (a = 0.6)									
Z	.130	.133	.133	.134	.135	.140	.145	.151	.156
LR	.142	.142	.141	.141	.139	.140	.138	.134	.134
ST	.138	.139	.140	.140	.138	.138	.135	.130	.130
KLD-U	.104	.116	.117	.117	.116	.118	.116	.105	.080
KLD-N	.128	.118	.106	.094	.083	.075	.066	.054	.041
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.917	.928	.939	.946	.950	.949	.946	.942	.931
LR	.929	.938	.945	.949	.950	.948	.945	.938	.926
ST	.918	.929	.939	.948	.951	.948	.945	.938	.926
KLD-U	.910	.920	.930	.935	.940	.940	.936	.928	.893
KLD-N	.924	.928	.931	.931	.932	.926	.915	.899	.851
Moderate Discrimination (a = 1.0)									
Z	.704	.715	.720	.731	.735	.739	.740	.737	.717
LR	.719	.726	.728	.738	.739	.738	.731	.718	.696
ST	.711	.721	.726	.737	.738	.736	.728	.714	.693
KLD-U	.673	.686	.688	.697	.701	.707	.704	.682	.580
KLD-N	.696	.692	.679	.672	.657	.648	.627	.583	.474
Low Discrimination (a = 0.6)									
Z	.400	.402	.409	.408	.412	.421	.429	.439	.425
LR	.415	.414	.420	.417	.418	.419	.415	.414	.402
ST	.412	.411	.419	.416	.416	.416	.409	.406	.397
KLD-U	.360	.370	.375	.372	.374	.378	.370	.340	.238
KLD-N	.378	.358	.342	.318	.297	.278	.250	.214	.150
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.990	.991	.995	.999	.999	1.000	.999	.999	.998
LR	.998	.999	.999	.999	.999	.999	.999	.999	.998
ST	.952	.957	.969	.990	.999	.999	.999	.999	.998
KLD-U	.996	.998	.999	.999	.999	.999	.999	.999	.985
KLD-N	.998	.998	.999	.999	.999	.999	.999	.997	.978
Moderate Discrimination (a = 1.0)									
Z	.959	.964	.969	.971	.973	.974	.971	.968	.960
LR	.965	.968	.971	.972	.973	.973	.969	.964	.956
ST	.958	.963	.969	.972	.973	.972	.968	.963	.955
KLD-U	.955	.958	.963	.964	.967	.968	.962	.948	.832
KLD-N	.959	.958	.958	.956	.954	.950	.937	.908	.772
Low Discrimination (a = 0.6)									
Z	.738	.738	.747	.747	.752	.759	.762	.759	.736
LR	.748	.746	.756	.753	.755	.754	.751	.742	.729
ST	.744	.743	.754	.752	.753	.751	.745	.737	.726
KLD-U	.707	.708	.716	.715	.720	.720	.705	.634	.443
KLD-N	.708	.687	.676	.651	.631	.604	.563	.484	.344

Table A53. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.048	.050	.050	.050	.049	.050	.049	.053	.055
LR	.050	.051	.051	.050	.049	.050	.049	.052	.051
ST	.050	.051	.051	.050	.049	.050	.049	.051	.051
KLD-U	.055	.054	.052	.049	.046	.049	.049	.052	.051
KLD-N	.061	.057	.053	.049	.046	.047	.046	.048	.047
Moderate Discrimination (a = 1.0)									
Z	.049	.049	.049	.049	.050	.047	.050	.052	.054
LR	.051	.051	.051	.051	.051	.049	.051	.050	.050
ST	.050	.051	.051	.050	.051	.049	.050	.049	.049
KLD-U	.051	.052	.052	.050	.050	.048	.051	.051	.048
KLD-N	.055	.051	.050	.046	.045	.042	.043	.042	.040
Low Discrimination (a = 0.6)									
Z	.048	.050	.050	.047	.048	.049	.050	.052	.055
LR	.050	.051	.051	.048	.050	.050	.050	.049	.049
ST	.051	.051	.051	.049	.051	.050	.050	.048	.049
KLD-U	.045	.051	.053	.050	.051	.052	.051	.048	.038
KLD-N	.047	.045	.043	.038	.037	.036	.035	.032	.030
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.140	.143	.145	.148	.148	.148	.150	.152	.156
LR	.144	.146	.146	.147	.149	.148	.148	.145	.146
ST	.144	.146	.145	.145	.148	.146	.145	.142	.143
KLD-U	.134	.133	.129	.127	.125	.125	.127	.124	.123
KLD-N	.143	.139	.134	.130	.124	.121	.118	.113	.109
Moderate Discrimination (a = 1.0)									
Z	.095	.097	.099	.099	.098	.098	.101	.107	.108
LR	.099	.100	.101	.100	.100	.100	.101	.101	.099
ST	.098	.099	.100	.098	.098	.099	.099	.098	.097
KLD-U	.088	.091	.092	.087	.085	.086	.087	.088	.081
KLD-N	.095	.092	.089	.082	.076	.074	.071	.069	.063
Low Discrimination (a = 0.6)									
Z	.071	.072	.072	.073	.071	.072	.076	.081	.080
LR	.072	.073	.073	.073	.072	.073	.074	.074	.072
ST	.072	.073	.072	.072	.072	.072	.073	.072	.070
KLD-U	.058	.067	.067	.067	.066	.066	.067	.063	.048
KLD-N	.065	.062	.056	.052	.047	.044	.042	.038	.031
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.417	.429	.437	.449	.445	.448	.450	.452	.434
LR	.424	.432	.438	.448	.448	.448	.447	.444	.423
ST	.425	.432	.436	.446	.446	.447	.444	.440	.420
KLD-U	.402	.408	.407	.409	.404	.409	.411	.407	.382
KLD-N	.417	.420	.416	.414	.403	.400	.393	.382	.356

Continued on the next page

Table A53 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.244	.253	.255	.257	.255	.252	.263	.264	.264
LR	.250	.258	.259	.260	.260	.256	.263	.255	.251
ST	.249	.256	.256	.257	.258	.254	.259	.250	.248
KLD-U	.229	.238	.236	.234	.230	.227	.235	.228	.210
KLD-N	.239	.239	.232	.223	.213	.201	.202	.190	.172
Low Discrimination (a = 0.6)									
Z	.142	.143	.143	.143	.141	.144	.148	.151	.154
LR	.144	.144	.144	.145	.144	.145	.145	.141	.143
ST	.143	.144	.143	.144	.143	.144	.142	.138	.140
KLD-U	.120	.133	.133	.131	.128	.131	.128	.118	.095
KLD-N	.127	.122	.114	.106	.097	.092	.084	.074	.062
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.938	.946	.951	.953	.954	.954	.950	.946	.938
LR	.939	.946	.950	.953	.954	.953	.950	.945	.936
ST	.938	.944	.948	.952	.954	.952	.949	.944	.936
KLD-U	.936	.940	.942	.943	.945	.944	.941	.935	.914
KLD-N	.938	.942	.944	.944	.944	.940	.934	.926	.906
Moderate Discrimination (a = 1.0)									
Z	.732	.739	.740	.746	.745	.747	.749	.743	.726
LR	.737	.742	.743	.751	.752	.751	.748	.737	.720
ST	.735	.741	.740	.748	.750	.749	.745	.734	.718
KLD-U	.718	.723	.721	.722	.721	.724	.722	.708	.632
KLD-N	.721	.721	.712	.707	.696	.690	.680	.656	.594
Low Discrimination (a = 0.6)									
Z	.425	.423	.427	.424	.425	.432	.435	.442	.425
LR	.427	.425	.430	.429	.430	.433	.429	.428	.418
ST	.427	.425	.429	.428	.429	.430	.426	.426	.417
KLD-U	.400	.406	.407	.402	.404	.406	.397	.371	.280
KLD-N	.392	.378	.366	.348	.334	.322	.301	.277	.220
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	.999	1.000	.999	1.000	.999	.999	.999
LR	.998	.999	.999	1.000	.999	1.000	.999	.999	.999
ST	.997	.998	.999	.999	.999	1.000	.999	.999	.999
KLD-U	.999	.999	.999	.999	.999	.999	.999	.999	.993
KLD-N	.998	.999	.999	.999	.999	.999	.999	.999	.995
Moderate Discrimination (a = 1.0)									
Z	.969	.970	.972	.972	.974	.974	.972	.969	.963
LR	.969	.970	.972	.973	.975	.974	.972	.968	.963
ST	.968	.969	.972	.973	.975	.974	.971	.968	.963
KLD-U	.966	.966	.968	.968	.970	.970	.967	.958	.886
KLD-N	.966	.965	.965	.964	.964	.962	.956	.946	.886
Low Discrimination (a = 0.6)									
Z	.755	.753	.761	.757	.761	.763	.768	.760	.743
LR	.756	.754	.764	.760	.764	.763	.762	.754	.747
ST	.756	.753	.763	.760	.764	.761	.760	.752	.749
KLD-U	.741	.738	.745	.740	.743	.741	.731	.675	.513
KLD-N	.722	.708	.703	.686	.674	.658	.634	.584	.472

Table A54. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.048	.049	.049	.050	.049	.049	.051	.053	.054
LR	.050	.050	.050	.050	.049	.049	.050	.051	.051
ST	.050	.051	.050	.050	.050	.050	.050	.051	.051
KLD-U	.055	.054	.051	.049	.047	.048	.048	.050	.051
KLD-N	.061	.056	.052	.049	.046	.046	.047	.047	.047
Moderate Discrimination (a = 1.0)									
Z	.048	.049	.050	.049	.049	.047	.050	.053	.054
LR	.050	.051	.051	.050	.050	.048	.050	.050	.049
ST	.049	.050	.051	.050	.050	.048	.050	.049	.049
KLD-U	.050	.052	.053	.050	.050	.048	.051	.051	.048
KLD-N	.056	.054	.053	.048	.046	.044	.045	.045	.043
Low Discrimination (a = 0.6)									
Z	.049	.049	.049	.048	.048	.049	.052	.053	.054
LR	.051	.050	.050	.049	.050	.050	.051	.049	.049
ST	.051	.050	.050	.050	.050	.050	.050	.048	.049
KLD-U	.045	.051	.053	.052	.052	.052	.052	.048	.038
KLD-N	.049	.046	.044	.040	.038	.038	.036	.035	.031
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.141	.144	.147	.149	.148	.149	.151	.156	.154
LR	.143	.145	.146	.147	.147	.147	.147	.146	.146
ST	.144	.146	.146	.146	.147	.146	.146	.144	.144
KLD-U	.136	.133	.131	.128	.124	.125	.126	.126	.123
KLD-N	.143	.138	.135	.130	.123	.122	.120	.117	.113
Moderate Discrimination (a = 1.0)									
Z	.097	.098	.101	.100	.097	.100	.104	.108	.106
LR	.099	.100	.101	.099	.098	.100	.102	.100	.098
ST	.098	.099	.100	.098	.097	.099	.100	.097	.096
KLD-U	.090	.092	.093	.089	.085	.088	.089	.089	.081
KLD-N	.098	.096	.094	.087	.080	.080	.078	.076	.071
Low Discrimination (a = 0.6)									
Z	.072	.073	.073	.072	.071	.073	.079	.082	.081
LR	.073	.073	.073	.073	.072	.073	.075	.074	.072
ST	.072	.072	.072	.072	.071	.072	.074	.073	.071
KLD-U	.059	.069	.071	.068	.067	.068	.068	.065	.050
KLD-N	.066	.064	.059	.055	.050	.048	.046	.043	.036
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.419	.430	.439	.450	.445	.450	.452	.453	.431
LR	.422	.432	.438	.447	.446	.448	.446	.442	.422
ST	.425	.433	.438	.446	.445	.447	.443	.441	.420
KLD-U	.406	.409	.409	.410	.403	.409	.409	.407	.381
KLD-N	.416	.419	.416	.413	.402	.401	.396	.388	.364

Continued on the next page

Table A54 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.247	.253	.255	.260	.257	.256	.266	.266	.262
LR	.250	.255	.255	.260	.258	.256	.262	.254	.248
ST	.249	.254	.253	.257	.257	.254	.259	.250	.246
KLD-U	.233	.238	.237	.237	.232	.230	.237	.230	.209
KLD-N	.245	.244	.238	.232	.222	.214	.215	.205	.190
Low Discrimination (a = 0.6)									
Z	.143	.145	.144	.144	.142	.148	.153	.153	.154
LR	.144	.145	.143	.144	.143	.147	.146	.142	.142
ST	.143	.144	.142	.143	.142	.145	.144	.140	.140
KLD-U	.123	.136	.135	.135	.131	.133	.131	.120	.097
KLD-N	.130	.126	.117	.111	.101	.099	.092	.084	.073
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.936	.943	.948	.951	.952	.951	.949	.943	.934
LR	.936	.942	.946	.950	.952	.950	.947	.942	.933
ST	.935	.942	.945	.950	.951	.950	.947	.942	.933
KLD-U	.934	.938	.939	.940	.942	.942	.938	.932	.911
KLD-N	.935	.939	.941	.941	.941	.939	.933	.925	.908
Moderate Discrimination (a = 1.0)									
Z	.730	.736	.737	.744	.742	.745	.747	.738	.722
LR	.732	.737	.736	.744	.745	.746	.743	.731	.714
ST	.731	.735	.734	.742	.744	.744	.741	.730	.714
KLD-U	.718	.722	.717	.720	.719	.721	.720	.704	.632
KLD-N	.724	.725	.716	.713	.705	.701	.693	.674	.623
Low Discrimination (a = 0.6)									
Z	.425	.425	.428	.426	.426	.434	.439	.442	.425
LR	.426	.425	.428	.428	.429	.432	.428	.428	.418
ST	.424	.423	.427	.427	.428	.429	.425	.425	.417
KLD-U	.405	.412	.413	.409	.408	.410	.400	.375	.286
KLD-N	.395	.385	.375	.360	.348	.338	.320	.302	.250
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	.999	1.000	.999	.999	.999	.999	.998
LR	.998	.998	.999	.999	.999	.999	.999	.999	.998
ST	.997	.998	.999	.999	.999	.999	.999	.999	.998
KLD-U	.999	.999	.999	.999	.999	.999	.999	.999	.994
KLD-N	.998	.999	.999	.999	.999	.999	.999	.998	.995
Moderate Discrimination (a = 1.0)									
Z	.967	.969	.971	.971	.973	.972	.970	.966	.961
LR	.967	.968	.971	.971	.973	.972	.969	.965	.959
ST	.966	.967	.970	.971	.973	.972	.969	.965	.960
KLD-U	.965	.966	.967	.966	.969	.968	.965	.955	.890
KLD-N	.965	.966	.966	.964	.965	.962	.958	.948	.905
Low Discrimination (a = 0.6)									
Z	.752	.749	.758	.755	.759	.762	.765	.758	.740
LR	.752	.749	.759	.756	.761	.759	.757	.750	.744
ST	.750	.748	.757	.756	.760	.757	.755	.749	.746
KLD-U	.742	.738	.745	.740	.743	.740	.728	.678	.526
KLD-N	.723	.710	.708	.692	.685	.673	.654	.615	.521

Table A55. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.108	.052	.039	.037	.034	.036	.038	.046	.058
LR	.050	.052	.052	.052	.049	.050	.048	.044	.046
ST	.042	.045	.050	.055	.054	.055	.053	.046	.044
KLD-U	.037	.059	.050	.047	.044	.047	.050	.054	.063
KLD-N	.042	.058	.053	.050	.046	.048	.050	.051	.050
Moderate Discrimination (a = 1.0)									
Z	.085	.060	.045	.041	.037	.040	.044	.051	.062
LR	.052	.055	.055	.054	.050	.051	.049	.046	.048
ST	.046	.050	.054	.056	.054	.054	.051	.047	.048
KLD-U	.038	.058	.054	.050	.049	.050	.052	.054	.056
KLD-N	.041	.054	.052	.048	.046	.046	.047	.046	.046
Low Discrimination (a = 0.6)									
Z	.062	.059	.050	.043	.041	.043	.048	.055	.057
LR	.053	.053	.053	.052	.051	.050	.049	.049	.046
ST	.052	.052	.053	.053	.053	.052	.050	.049	.047
KLD-U	.032	.050	.056	.055	.054	.056	.055	.056	.041
KLD-N	.034	.044	.047	.046	.045	.046	.044	.043	.034
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.142	.162	.185	.199	.204	.206	.201	.186	.130
LR	.112	.181	.222	.242	.244	.239	.219	.171	.106
ST	.101	.173	.222	.247	.249	.242	.217	.162	.096
KLD-U	.093	.194	.203	.217	.220	.218	.208	.174	.101
KLD-N	.100	.194	.216	.227	.225	.218	.204	.162	.080
Moderate Discrimination (a = 1.0)									
Z	.112	.110	.118	.123	.130	.133	.136	.144	.120
LR	.098	.123	.142	.152	.155	.151	.140	.126	.092
ST	.090	.118	.140	.153	.156	.148	.135	.119	.085
KLD-U	.078	.128	.133	.135	.139	.137	.132	.126	.082
KLD-N	.085	.122	.133	.133	.134	.127	.117	.106	.064
Low Discrimination (a = 0.6)									
Z	.077	.079	.080	.079	.081	.089	.095	.099	.096
LR	.079	.084	.094	.097	.096	.097	.090	.082	.075
ST	.077	.082	.094	.096	.094	.093	.085	.078	.070
KLD-U	.051	.083	.094	.092	.092	.094	.089	.082	.056
KLD-N	.058	.075	.085	.082	.079	.077	.068	.057	.039
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.362	.558	.639	.677	.692	.685	.640	.523	.293
LR	.355	.585	.684	.722	.733	.718	.656	.497	.260
ST	.344	.582	.688	.728	.738	.722	.653	.482	.242
KLD-U	.362	.597	.663	.695	.709	.701	.652	.504	.233
KLD-N	.357	.604	.679	.706	.715	.701	.644	.480	.190

Continued on the next page

Table A55 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.259	.328	.375	.405	.422	.414	.398	.354	.270
LR	.261	.355	.415	.448	.461	.440	.402	.324	.223
ST	.252	.350	.416	.451	.462	.438	.394	.310	.209
KLD-U	.252	.360	.399	.422	.438	.422	.394	.324	.191
KLD-N	.248	.353	.400	.419	.428	.402	.364	.284	.153
Low Discrimination (a = 0.6)									
Z	.144	.173	.197	.211	.221	.224	.228	.213	.179
LR	.154	.192	.225	.240	.245	.236	.216	.186	.147
ST	.152	.190	.224	.239	.243	.230	.208	.176	.139
KLD-U	.121	.196	.224	.234	.240	.233	.215	.179	.105
KLD-N	.129	.178	.208	.215	.216	.200	.174	.133	.076
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.964	.993	.997	.999	.999	.998	.992	.937	.526
LR	.967	.995	.998	.999	.999	.998	.992	.924	.559
ST	.965	.995	.998	.999	.999	.998	.992	.916	.548
KLD-U	.973	.995	.998	.999	.999	.998	.992	.926	.476
KLD-N	.972	.996	.998	.999	.999	.998	.992	.908	.410
Moderate Discrimination (a = 1.0)									
Z	.811	.904	.939	.951	.950	.939	.905	.815	.601
LR	.821	.918	.950	.960	.957	.944	.903	.790	.560
ST	.817	.916	.951	.961	.958	.944	.899	.777	.538
KLD-U	.833	.915	.945	.955	.953	.942	.903	.786	.448
KLD-N	.821	.915	.945	.954	.950	.935	.887	.743	.397
Low Discrimination (a = 0.6)									
Z	.497	.603	.659	.697	.700	.681	.634	.558	.441
LR	.519	.633	.690	.727	.722	.688	.620	.523	.404
ST	.518	.634	.691	.727	.718	.682	.610	.509	.388
KLD-U	.513	.641	.690	.721	.720	.689	.620	.496	.267
KLD-N	.493	.617	.668	.698	.688	.641	.553	.416	.218
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.982	.978	.997	1.000	1.000	1.000	1.000	.992	.583
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.990	.695
ST	.995	.993	.998	1.000	1.000	1.000	1.000	.988	.703
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.989	.592
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.984	.525
Moderate Discrimination (a = 1.0)									
Z	.994	.999	1.000	1.000	1.000	.999	.995	.967	.746
LR	.996	.999	1.000	1.000	1.000	.999	.994	.959	.772
ST	.996	.999	1.000	1.000	1.000	.999	.994	.954	.771
KLD-U	.997	.999	1.000	1.000	1.000	.999	.994	.951	.610
KLD-N	.996	.999	1.000	1.000	1.000	.999	.992	.933	.573
Low Discrimination (a = 0.6)									
Z	.886	.932	.957	.962	.960	.946	.905	.831	.668
LR	.898	.942	.963	.968	.964	.946	.897	.810	.656
ST	.898	.941	.963	.968	.963	.944	.892	.799	.643
KLD-U	.903	.943	.963	.967	.964	.948	.896	.760	.425
KLD-N	.890	.936	.958	.961	.955	.930	.856	.695	.386

Table A56. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.101	.047	.039	.038	.036	.038	.042	.046	.058
LR	.046	.047	.051	.051	.050	.051	.052	.048	.049
ST	.044	.046	.053	.054	.053	.054	.054	.046	.044
KLD-U	.035	.057	.052	.049	.047	.048	.051	.053	.061
KLD-N	.036	.054	.055	.052	.050	.050	.052	.050	.049
Moderate Discrimination (a = 1.0)									
Z	.081	.054	.044	.042	.040	.042	.046	.051	.061
LR	.047	.051	.053	.052	.050	.051	.052	.049	.051
ST	.046	.049	.053	.054	.053	.053	.052	.048	.048
KLD-U	.036	.056	.054	.050	.048	.050	.052	.054	.054
KLD-N	.037	.052	.053	.051	.048	.049	.050	.048	.046
Low Discrimination (a = 0.6)									
Z	.060	.057	.049	.045	.043	.046	.050	.054	.057
LR	.050	.050	.052	.052	.052	.052	.052	.050	.048
ST	.051	.050	.052	.053	.053	.053	.051	.049	.047
KLD-U	.032	.050	.056	.055	.054	.056	.056	.054	.041
KLD-N	.033	.043	.047	.048	.048	.047	.046	.043	.034
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.139	.167	.196	.212	.220	.217	.211	.185	.129
LR	.108	.182	.230	.252	.259	.252	.232	.182	.114
ST	.104	.179	.230	.254	.261	.253	.229	.172	.103
KLD-U	.090	.198	.220	.233	.237	.230	.214	.172	.098
KLD-N	.092	.194	.231	.244	.246	.234	.213	.162	.079
Moderate Discrimination (a = 1.0)									
Z	.109	.114	.124	.133	.137	.140	.140	.142	.118
LR	.093	.122	.143	.156	.160	.156	.146	.133	.097
ST	.090	.120	.141	.155	.159	.153	.141	.125	.089
KLD-U	.075	.131	.138	.142	.144	.141	.131	.122	.080
KLD-N	.078	.124	.140	.146	.144	.137	.124	.108	.065
Low Discrimination (a = 0.6)									
Z	.079	.080	.086	.085	.088	.094	.096	.097	.094
LR	.076	.084	.097	.098	.100	.099	.092	.086	.077
ST	.075	.083	.096	.097	.097	.095	.088	.081	.072
KLD-U	.051	.086	.100	.096	.095	.095	.089	.081	.055
KLD-N	.055	.076	.090	.088	.084	.081	.071	.060	.041
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.367	.573	.661	.703	.709	.699	.648	.521	.292
LR	.362	.600	.699	.741	.745	.733	.672	.515	.275
ST	.354	.597	.702	.745	.748	.735	.670	.499	.256
KLD-U	.362	.613	.690	.725	.728	.715	.655	.501	.229
KLD-N	.354	.618	.702	.736	.736	.719	.654	.480	.190

Continued on the next page

Table A56 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.264	.340	.393	.422	.438	.425	.406	.352	.267
LR	.262	.362	.426	.460	.472	.451	.415	.337	.232
ST	.256	.358	.425	.460	.471	.449	.407	.323	.218
KLD-U	.253	.373	.417	.438	.451	.429	.395	.320	.186
KLD-N	.248	.367	.423	.443	.450	.421	.377	.289	.158
Low Discrimination (a = 0.6)									
Z	.149	.180	.206	.222	.231	.233	.229	.212	.177
LR	.154	.194	.227	.247	.251	.244	.223	.192	.154
ST	.154	.192	.224	.245	.248	.239	.215	.184	.145
KLD-U	.126	.202	.231	.243	.245	.237	.215	.178	.105
KLD-N	.127	.183	.215	.226	.225	.208	.179	.137	.079
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.969	.995	.998	.999	.999	.998	.992	.937	.528
LR	.971	.996	.998	1.000	.999	.999	.992	.930	.577
ST	.970	.996	.998	1.000	.999	.999	.992	.923	.569
KLD-U	.976	.996	.998	.999	.999	.999	.992	.924	.472
KLD-N	.974	.996	.999	.999	.999	.998	.992	.909	.411
Moderate Discrimination (a = 1.0)									
Z	.821	.914	.945	.955	.954	.940	.906	.813	.603
LR	.829	.923	.954	.962	.960	.947	.908	.798	.574
ST	.825	.922	.954	.962	.961	.946	.905	.787	.555
KLD-U	.842	.924	.951	.958	.957	.942	.902	.782	.445
KLD-N	.831	.925	.952	.959	.956	.939	.892	.748	.409
Low Discrimination (a = 0.6)									
Z	.511	.619	.674	.709	.712	.687	.636	.560	.442
LR	.528	.642	.700	.732	.732	.696	.630	.536	.416
ST	.526	.640	.699	.732	.729	.692	.620	.523	.402
KLD-U	.527	.656	.702	.729	.726	.690	.620	.498	.273
KLD-N	.502	.629	.684	.710	.702	.653	.561	.428	.235
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.992	.585
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.713
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.990	.724
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.989	.587
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.984	.525
Moderate Discrimination (a = 1.0)									
Z	.996	.999	1.000	1.000	1.000	.999	.995	.967	.748
LR	.997	.999	1.000	1.000	1.000	.999	.995	.962	.784
ST	.996	.999	1.000	1.000	1.000	.999	.995	.958	.783
KLD-U	.997	.999	1.000	1.000	1.000	.999	.994	.950	.609
KLD-N	.997	.999	1.000	1.000	1.000	.999	.993	.937	.588
Low Discrimination (a = 0.6)									
Z	.892	.937	.959	.964	.961	.947	.906	.832	.672
LR	.900	.944	.965	.969	.964	.949	.902	.816	.669
ST	.899	.944	.965	.969	.964	.948	.898	.808	.658
KLD-U	.908	.947	.965	.968	.964	.947	.895	.763	.432
KLD-N	.894	.940	.960	.964	.957	.932	.862	.710	.408

Table A57. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.100	.047	.038	.037	.037	.038	.043	.044	.056
LR	.046	.049	.051	.050	.051	.051	.053	.048	.050
ST	.044	.047	.052	.053	.054	.053	.054	.046	.045
KLD-U	.034	.057	.052	.049	.049	.049	.052	.053	.060
KLD-N	.036	.054	.056	.053	.051	.051	.054	.050	.049
Moderate Discrimination (a = 1.0)									
Z	.080	.055	.045	.042	.042	.042	.047	.051	.060
LR	.047	.051	.053	.051	.052	.051	.051	.049	.051
ST	.046	.050	.053	.054	.054	.052	.051	.048	.049
KLD-U	.035	.057	.054	.051	.050	.050	.052	.053	.054
KLD-N	.037	.053	.054	.052	.051	.050	.050	.048	.046
Low Discrimination (a = 0.6)									
Z	.060	.058	.050	.046	.043	.046	.049	.052	.056
LR	.050	.051	.053	.053	.051	.053	.051	.050	.049
ST	.051	.051	.054	.054	.052	.052	.051	.050	.048
KLD-U	.033	.052	.057	.057	.054	.056	.056	.055	.041
KLD-N	.033	.043	.048	.050	.046	.048	.046	.043	.034
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.138	.166	.198	.215	.220	.218	.213	.182	.127
LR	.109	.183	.231	.251	.258	.252	.236	.185	.117
ST	.105	.178	.230	.253	.259	.252	.231	.174	.105
KLD-U	.089	.199	.225	.237	.239	.234	.217	.171	.097
KLD-N	.092	.196	.236	.248	.248	.240	.218	.162	.079
Moderate Discrimination (a = 1.0)									
Z	.110	.117	.126	.135	.141	.143	.141	.140	.116
LR	.093	.124	.143	.154	.160	.157	.147	.133	.097
ST	.090	.121	.142	.153	.160	.154	.141	.125	.090
KLD-U	.074	.134	.141	.144	.148	.144	.133	.121	.078
KLD-N	.078	.127	.145	.149	.150	.143	.127	.109	.065
Low Discrimination (a = 0.6)									
Z	.079	.082	.086	.086	.091	.095	.096	.096	.091
LR	.076	.086	.098	.099	.101	.101	.096	.088	.078
ST	.076	.085	.097	.098	.099	.098	.092	.084	.074
KLD-U	.053	.088	.102	.100	.099	.097	.092	.082	.055
KLD-N	.057	.078	.091	.090	.087	.082	.073	.062	.043
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.367	.575	.661	.698	.706	.697	.646	.517	.288
LR	.365	.604	.700	.737	.744	.732	.673	.518	.279
ST	.356	.600	.702	.739	.745	.733	.669	.502	.259
KLD-U	.360	.619	.694	.725	.729	.717	.656	.498	.227
KLD-N	.353	.623	.708	.737	.738	.723	.657	.480	.190

Continued on the next page

Table A57 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.267	.344	.395	.425	.441	.428	.405	.349	.264
LR	.263	.362	.424	.458	.470	.451	.415	.338	.233
ST	.257	.359	.423	.459	.470	.448	.407	.325	.221
KLD-U	.254	.376	.420	.441	.453	.429	.393	.316	.184
KLD-N	.249	.371	.428	.450	.457	.427	.380	.290	.158
Low Discrimination (a = 0.6)									
Z	.150	.181	.210	.225	.237	.236	.228	.210	.176
LR	.155	.194	.229	.248	.257	.248	.226	.195	.155
ST	.155	.193	.228	.246	.254	.243	.219	.188	.148
KLD-U	.128	.206	.237	.247	.252	.242	.217	.179	.106
KLD-N	.129	.184	.219	.230	.232	.215	.182	.142	.082
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.969	.995	.998	.999	.999	.998	.992	.935	.522
LR	.971	.996	.998	.999	.999	.999	.993	.932	.582
ST	.970	.996	.998	.999	.999	.999	.992	.924	.574
KLD-U	.976	.996	.998	.999	.999	.998	.992	.923	.468
KLD-N	.975	.997	.999	.999	.999	.999	.992	.908	.410
Moderate Discrimination (a = 1.0)									
Z	.823	.913	.944	.953	.952	.940	.905	.811	.600
LR	.829	.922	.952	.959	.958	.946	.908	.798	.575
ST	.826	.921	.952	.960	.958	.946	.904	.788	.559
KLD-U	.842	.924	.950	.957	.955	.941	.900	.778	.440
KLD-N	.833	.926	.952	.958	.956	.940	.892	.749	.411
Low Discrimination (a = 0.6)									
Z	.514	.619	.674	.712	.713	.687	.636	.557	.440
LR	.530	.640	.699	.734	.732	.698	.634	.537	.420
ST	.528	.639	.698	.733	.730	.695	.626	.526	.407
KLD-U	.532	.658	.706	.734	.728	.693	.623	.499	.276
KLD-N	.504	.630	.686	.714	.704	.656	.569	.432	.241
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.578
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.717
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.990	.729
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.989	.583
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.984	.524
Moderate Discrimination (a = 1.0)									
Z	.996	.999	1.000	1.000	1.000	.999	.995	.966	.744
LR	.996	.999	1.000	1.000	1.000	1.000	.995	.962	.784
ST	.996	.999	1.000	1.000	1.000	.999	.995	.958	.786
KLD-U	.997	.999	1.000	1.000	1.000	.999	.994	.948	.604
KLD-N	.997	.999	1.000	1.000	1.000	.999	.993	.936	.589
Low Discrimination (a = 0.6)									
Z	.892	.938	.959	.963	.960	.946	.905	.831	.669
LR	.900	.945	.964	.968	.964	.949	.903	.820	.671
ST	.899	.944	.964	.968	.964	.948	.899	.811	.663
KLD-U	.909	.949	.966	.968	.964	.948	.896	.763	.437
KLD-N	.894	.941	.961	.964	.957	.934	.864	.714	.417

Table A58. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.056	.053	.050	.048	.048	.050	.049	.053	.055
LR	.057	.055	.052	.049	.049	.050	.048	.050	.048
ST	.051	.052	.051	.050	.051	.052	.051	.052	.050
KLD-U	.053	.053	.050	.051	.050	.051	.051	.054	.051
KLD-N	.063	.057	.051	.049	.048	.048	.048	.050	.048
Moderate Discrimination (a = 1.0)									
Z	.051	.047	.048	.048	.047	.048	.050	.053	.057
LR	.054	.051	.051	.050	.048	.049	.049	.049	.049
ST	.050	.050	.050	.051	.049	.051	.050	.051	.051
KLD-U	.052	.049	.049	.050	.049	.050	.050	.052	.050
KLD-N	.059	.051	.047	.045	.044	.044	.044	.046	.045
Low Discrimination (a = 0.6)									
Z	.050	.049	.049	.049	.048	.050	.051	.051	.056
LR	.056	.052	.052	.051	.050	.051	.050	.046	.048
ST	.053	.051	.052	.051	.051	.052	.051	.047	.049
KLD-U	.049	.051	.053	.052	.052	.053	.052	.048	.044
KLD-N	.056	.048	.045	.041	.040	.040	.040	.037	.038
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.179	.185	.183	.179	.182	.182	.186	.191	.206
LR	.188	.192	.188	.185	.182	.182	.181	.180	.185
ST	.179	.183	.183	.180	.179	.179	.179	.177	.181
KLD-U	.173	.177	.171	.169	.171	.171	.173	.174	.177
KLD-N	.193	.189	.177	.169	.166	.162	.160	.157	.158
Moderate Discrimination (a = 1.0)									
Z	.114	.114	.114	.113	.118	.119	.122	.131	.138
LR	.123	.123	.121	.118	.122	.121	.119	.120	.118
ST	.117	.118	.117	.116	.120	.119	.115	.116	.114
KLD-U	.111	.110	.108	.107	.110	.110	.112	.114	.108
KLD-N	.126	.117	.108	.102	.100	.097	.094	.094	.088
Low Discrimination (a = 0.6)									
Z	.076	.079	.080	.076	.075	.078	.086	.095	.097
LR	.086	.086	.086	.081	.080	.080	.084	.084	.081
ST	.081	.083	.083	.079	.079	.079	.081	.080	.077
KLD-U	.072	.078	.078	.073	.073	.075	.079	.078	.065
KLD-N	.085	.078	.072	.062	.057	.055	.057	.054	.047
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.561	.571	.568	.574	.565	.571	.578	.580	.583
LR	.566	.574	.572	.574	.563	.568	.572	.566	.563
ST	.562	.570	.568	.572	.562	.565	.569	.561	.558
KLD-U	.553	.558	.551	.558	.548	.554	.562	.559	.549
KLD-N	.574	.571	.558	.557	.541	.538	.542	.533	.518

Continued on the next page

Table A58 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.325	.330	.341	.333	.340	.338	.351	.359	.364
LR	.337	.339	.349	.341	.344	.340	.346	.341	.338
ST	.329	.333	.344	.337	.340	.337	.341	.334	.331
KLD-U	.318	.320	.329	.320	.325	.323	.333	.330	.316
KLD-N	.338	.329	.328	.310	.307	.295	.298	.291	.274
Low Discrimination (a = 0.6)									
Z	.175	.178	.179	.175	.183	.179	.191	.205	.208
LR	.186	.188	.188	.184	.188	.180	.184	.188	.184
ST	.181	.182	.183	.180	.185	.176	.178	.180	.178
KLD-U	.167	.174	.173	.169	.175	.170	.176	.176	.150
KLD-N	.182	.172	.160	.146	.144	.134	.134	.130	.113
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.990	.991	.991	.990	.992	.991	.992	.990	.989
LR	.989	.991	.991	.990	.991	.990	.991	.989	.988
ST	.989	.990	.991	.990	.991	.990	.991	.988	.987
KLD-U	.989	.990	.990	.989	.990	.990	.991	.989	.986
KLD-N	.990	.991	.990	.989	.990	.989	.989	.986	.983
Moderate Discrimination (a = 1.0)									
Z	.875	.879	.882	.880	.884	.878	.887	.882	.882
LR	.878	.881	.885	.882	.884	.878	.884	.873	.873
ST	.875	.879	.882	.881	.883	.876	.882	.870	.870
KLD-U	.870	.872	.874	.872	.876	.871	.880	.868	.840
KLD-N	.876	.874	.871	.864	.862	.852	.857	.841	.813
Low Discrimination (a = 0.6)									
Z	.556	.559	.566	.558	.558	.565	.574	.591	.580
LR	.567	.568	.576	.567	.565	.565	.565	.569	.563
ST	.561	.562	.571	.564	.561	.561	.558	.561	.557
KLD-U	.550	.552	.556	.548	.549	.553	.554	.545	.455
KLD-N	.553	.538	.529	.506	.491	.487	.479	.464	.399
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.996	.996	.997	.997	.997	.997	.997	.997	.997
LR	.996	.996	.997	.997	.997	.997	.997	.996	.996
ST	.996	.996	.997	.997	.997	.996	.997	.996	.996
KLD-U	.996	.996	.996	.997	.997	.997	.997	.996	.983
KLD-N	.996	.996	.996	.996	.996	.996	.995	.994	.983
Low Discrimination (a = 0.6)									
Z	.889	.894	.894	.894	.895	.894	.899	.897	.890
LR	.893	.897	.898	.897	.897	.892	.895	.890	.887
ST	.889	.894	.896	.896	.895	.890	.891	.886	.884
KLD-U	.885	.889	.889	.889	.891	.889	.888	.864	.739
KLD-N	.882	.879	.872	.864	.859	.851	.843	.817	.725

Table A59. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.050	.051	.052	.050	.049	.052	.051	.055	.054
LR	.051	.052	.052	.050	.049	.052	.051	.053	.051
ST	.051	.051	.052	.050	.050	.053	.051	.052	.050
KLD-U	.053	.052	.051	.049	.049	.051	.052	.054	.052
KLD-N	.057	.054	.052	.049	.048	.050	.050	.051	.051
Moderate Discrimination (a = 1.0)									
Z	.049	.049	.048	.049	.048	.049	.049	.052	.055
LR	.052	.051	.049	.051	.050	.051	.050	.050	.050
ST	.052	.051	.050	.051	.050	.051	.050	.050	.050
KLD-U	.052	.052	.047	.050	.050	.050	.051	.052	.050
KLD-N	.056	.052	.047	.048	.046	.046	.046	.048	.048
Low Discrimination (a = 0.6)									
Z	.051	.050	.050	.050	.049	.050	.051	.049	.056
LR	.052	.050	.052	.052	.051	.052	.051	.046	.051
ST	.052	.050	.052	.052	.051	.052	.051	.046	.050
KLD-U	.048	.050	.052	.052	.051	.053	.053	.048	.044
KLD-N	.053	.047	.046	.044	.042	.042	.041	.038	.040
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.186	.195	.194	.189	.191	.190	.193	.200	.209
LR	.191	.198	.196	.192	.192	.191	.193	.192	.200
ST	.188	.193	.192	.188	.187	.187	.189	.185	.191
KLD-U	.180	.184	.178	.173	.173	.173	.178	.177	.183
KLD-N	.191	.193	.185	.175	.171	.168	.170	.167	.173
Moderate Discrimination (a = 1.0)									
Z	.120	.121	.119	.119	.122	.123	.124	.133	.136
LR	.125	.125	.123	.122	.126	.126	.124	.127	.126
ST	.123	.123	.120	.118	.124	.123	.121	.122	.121
KLD-U	.116	.116	.112	.109	.113	.114	.114	.116	.113
KLD-N	.125	.121	.113	.107	.108	.104	.102	.102	.101
Low Discrimination (a = 0.6)									
Z	.082	.084	.084	.080	.079	.080	.087	.093	.093
LR	.085	.086	.085	.082	.082	.082	.086	.085	.082
ST	.082	.085	.083	.080	.080	.079	.083	.081	.078
KLD-U	.073	.081	.079	.075	.075	.075	.079	.078	.067
KLD-N	.083	.079	.074	.065	.061	.059	.062	.060	.054
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.579	.588	.584	.586	.579	.581	.587	.592	.593
LR	.583	.590	.588	.590	.581	.583	.588	.584	.586
ST	.580	.586	.582	.585	.576	.579	.582	.578	.579
KLD-U	.569	.574	.563	.565	.556	.558	.567	.569	.564
KLD-N	.584	.584	.571	.568	.553	.550	.556	.555	.546

Continued on the next page

Table A59 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.338	.344	.353	.345	.346	.346	.355	.362	.366
LR	.347	.350	.359	.352	.353	.352	.357	.354	.354
ST	.345	.347	.355	.349	.350	.349	.352	.348	.346
KLD-U	.334	.335	.338	.330	.332	.332	.340	.339	.328
KLD-N	.346	.342	.341	.325	.320	.314	.316	.314	.304
Low Discrimination (a = 0.6)									
Z	.184	.187	.188	.184	.187	.182	.194	.204	.207
LR	.190	.189	.192	.189	.192	.185	.190	.192	.191
ST	.188	.186	.188	.185	.189	.181	.185	.187	.186
KLD-U	.174	.180	.179	.176	.179	.173	.178	.177	.158
KLD-N	.183	.176	.168	.158	.154	.143	.146	.144	.132
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.991	.992	.992	.990	.992	.992	.992	.990	.990
LR	.991	.992	.992	.990	.992	.992	.992	.990	.990
ST	.990	.992	.992	.990	.992	.991	.992	.990	.989
KLD-U	.990	.991	.990	.989	.990	.990	.992	.989	.987
KLD-N	.991	.991	.991	.989	.990	.990	.991	.988	.985
Moderate Discrimination (a = 1.0)									
Z	.882	.886	.887	.885	.888	.881	.889	.883	.887
LR	.886	.889	.889	.889	.891	.885	.891	.881	.884
ST	.884	.886	.888	.888	.889	.883	.889	.879	.882
KLD-U	.881	.880	.880	.878	.881	.875	.884	.874	.854
KLD-N	.885	.882	.879	.874	.872	.864	.871	.859	.843
Low Discrimination (a = 0.6)									
Z	.571	.572	.577	.567	.565	.569	.578	.592	.584
LR	.576	.576	.583	.576	.573	.574	.575	.582	.576
ST	.574	.572	.579	.572	.571	.570	.570	.576	.571
KLD-U	.564	.563	.566	.558	.557	.559	.560	.555	.472
KLD-N	.562	.550	.542	.524	.512	.509	.504	.499	.448
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.997	.997	.997	.997	.997	.997	.997	.997	.997
LR	.997	.997	.997	.997	.997	.997	.997	.997	.997
ST	.997	.997	.997	.997	.997	.997	.997	.997	.996
KLD-U	.997	.997	.997	.997	.997	.997	.997	.996	.988
KLD-N	.997	.997	.996	.996	.997	.996	.996	.996	.990
Low Discrimination (a = 0.6)									
Z	.894	.898	.897	.896	.897	.896	.901	.899	.893
LR	.895	.900	.900	.900	.899	.897	.900	.895	.894
ST	.892	.898	.898	.899	.898	.896	.898	.893	.892
KLD-U	.889	.894	.892	.892	.893	.892	.893	.871	.765
KLD-N	.885	.885	.880	.873	.870	.864	.862	.845	.779

Table A60. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.050	.050	.051	.049	.049	.050	.050	.055	.053
LR	.051	.051	.052	.050	.050	.051	.050	.053	.051
ST	.051	.051	.052	.050	.051	.052	.051	.053	.051
KLD-U	.053	.051	.050	.049	.048	.050	.051	.054	.051
KLD-N	.056	.053	.052	.049	.048	.049	.049	.052	.050
Moderate Discrimination (a = 1.0)									
Z	.051	.049	.049	.051	.049	.049	.051	.054	.055
LR	.053	.050	.049	.052	.050	.050	.050	.050	.050
ST	.052	.051	.050	.052	.051	.051	.050	.051	.050
KLD-U	.053	.051	.048	.050	.050	.050	.052	.052	.051
KLD-N	.058	.054	.049	.049	.048	.047	.048	.048	.049
Low Discrimination (a = 0.6)									
Z	.050	.049	.050	.050	.048	.049	.052	.050	.054
LR	.052	.051	.051	.052	.050	.051	.051	.047	.050
ST	.052	.050	.051	.052	.050	.051	.051	.046	.049
KLD-U	.047	.050	.052	.053	.051	.053	.054	.048	.045
KLD-N	.053	.048	.047	.046	.042	.043	.043	.039	.042
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.186	.194	.193	.189	.190	.190	.194	.199	.208
LR	.191	.197	.196	.192	.192	.191	.193	.192	.199
ST	.190	.193	.192	.189	.187	.188	.189	.185	.191
KLD-U	.180	.182	.179	.172	.173	.174	.176	.178	.183
KLD-N	.191	.192	.184	.175	.172	.170	.170	.170	.174
Moderate Discrimination (a = 1.0)									
Z	.122	.124	.120	.120	.123	.127	.128	.134	.136
LR	.125	.126	.122	.121	.125	.127	.124	.125	.124
ST	.123	.123	.120	.119	.122	.125	.120	.121	.121
KLD-U	.118	.117	.112	.111	.114	.116	.114	.116	.114
KLD-N	.128	.123	.114	.110	.110	.109	.105	.106	.107
Low Discrimination (a = 0.6)									
Z	.081	.084	.084	.080	.080	.080	.089	.093	.094
LR	.084	.087	.085	.082	.082	.082	.087	.085	.083
ST	.082	.084	.082	.079	.081	.080	.084	.081	.080
KLD-U	.074	.082	.080	.076	.076	.076	.080	.078	.068
KLD-N	.082	.081	.075	.067	.065	.062	.065	.063	.059
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.574	.587	.584	.585	.576	.578	.588	.589	.591
LR	.580	.590	.588	.589	.579	.581	.587	.582	.584
ST	.577	.586	.583	.586	.574	.576	.583	.577	.576
KLD-U	.568	.572	.563	.565	.554	.557	.567	.567	.563
KLD-N	.581	.584	.571	.568	.552	.552	.559	.556	.551

Continued on the next page

Table A60 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length With 300-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.341	.347	.357	.349	.351	.351	.358	.364	.368
LR	.347	.350	.359	.353	.354	.351	.356	.354	.353
ST	.344	.346	.355	.349	.351	.348	.350	.348	.346
KLD-U	.336	.337	.341	.332	.336	.334	.339	.340	.329
KLD-N	.350	.346	.346	.331	.328	.322	.323	.324	.316
Low Discrimination (a = 0.6)									
Z	.186	.188	.189	.183	.187	.183	.196	.205	.206
LR	.191	.191	.192	.188	.192	.187	.192	.192	.192
ST	.187	.187	.187	.184	.188	.182	.187	.186	.185
KLD-U	.175	.183	.183	.175	.180	.176	.181	.179	.158
KLD-N	.186	.179	.172	.159	.159	.151	.154	.152	.140
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.990	.992	.991	.990	.991	.991	.992	.990	.989
LR	.990	.992	.991	.990	.991	.990	.991	.989	.989
ST	.990	.991	.991	.989	.991	.990	.991	.989	.988
KLD-U	.990	.990	.990	.988	.990	.990	.991	.988	.986
KLD-N	.990	.991	.990	.988	.989	.989	.990	.987	.985
Moderate Discrimination (a = 1.0)									
Z	.882	.886	.886	.884	.887	.881	.890	.884	.886
LR	.884	.887	.887	.886	.888	.883	.888	.881	.881
ST	.882	.885	.886	.884	.887	.882	.887	.879	.879
KLD-U	.879	.879	.878	.876	.880	.874	.883	.874	.853
KLD-N	.884	.883	.879	.874	.875	.868	.874	.864	.851
Low Discrimination (a = 0.6)									
Z	.570	.571	.574	.566	.564	.567	.578	.591	.580
LR	.575	.576	.580	.573	.572	.572	.575	.579	.574
ST	.571	.571	.575	.570	.569	.568	.569	.573	.568
KLD-U	.564	.565	.565	.558	.557	.560	.561	.556	.476
KLD-N	.563	.553	.546	.531	.520	.517	.515	.512	.468
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.997	.997	.996	.997	.997	.997	.997	.997	.996
LR	.997	.997	.996	.997	.997	.997	.997	.997	.996
ST	.996	.997	.996	.997	.997	.996	.997	.996	.996
KLD-U	.997	.997	.996	.996	.997	.996	.996	.996	.988
KLD-N	.997	.997	.996	.996	.996	.996	.996	.995	.991
Low Discrimination (a = 0.6)									
Z	.891	.897	.894	.894	.893	.894	.899	.896	.891
LR	.892	.899	.897	.898	.896	.895	.897	.893	.892
ST	.890	.896	.896	.897	.894	.892	.895	.891	.890
KLD-U	.889	.894	.891	.892	.891	.891	.892	.870	.771
KLD-N	.884	.886	.880	.876	.872	.868	.867	.853	.798

Table A61. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.074	.040	.035	.035	.038	.041	.044	.047	.055
LR	.085	.076	.071	.068	.066	.064	.064	.059	.057
ST	.079	.076	.076	.076	.075	.073	.073	.065	.060
KLD-U	.066	.057	.048	.044	.043	.043	.047	.050	.048
KLD-N	.083	.062	.047	.040	.036	.036	.035	.032	.031
Moderate Discrimination (a = 1.0)									
Z	.064	.043	.038	.040	.040	.043	.048	.052	.056
LR	.076	.069	.067	.065	.061	.061	.062	.059	.054
ST	.078	.073	.073	.074	.070	.069	.069	.062	.058
KLD-U	.045	.046	.044	.043	.041	.043	.048	.048	.032
KLD-N	.051	.045	.039	.036	.032	.031	.030	.027	.018
Low Discrimination (a = 0.6)									
Z	.056	.054	.047	.045	.045	.048	.051	.057	.053
LR	.068	.067	.064	.062	.060	.059	.058	.058	.049
ST	.074	.072	.070	.068	.066	.065	.062	.060	.059
KLD-U	.027	.036	.041	.044	.044	.045	.046	.038	.020
KLD-N	.024	.028	.030	.029	.027	.025	.022	.018	.010
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.098	.071	.073	.078	.086	.091	.093	.090	.083
LR	.120	.124	.127	.131	.133	.129	.125	.111	.097
ST	.112	.124	.133	.143	.149	.145	.139	.122	.107
KLD-U	.062	.061	.060	.063	.065	.068	.072	.074	.066
KLD-N	.084	.074	.063	.059	.055	.051	.047	.044	.041
Moderate Discrimination (a = 1.0)									
Z	.080	.064	.061	.065	.070	.071	.076	.080	.070
LR	.096	.096	.097	.099	.101	.096	.095	.088	.073
ST	.096	.100	.105	.110	.112	.107	.104	.094	.083
KLD-U	.039	.047	.048	.050	.051	.054	.059	.058	.037
KLD-N	.051	.051	.047	.043	.038	.034	.031	.027	.018
Low Discrimination (a = 0.6)									
Z	.067	.065	.062	.060	.062	.065	.067	.072	.057
LR	.080	.082	.081	.079	.078	.078	.075	.073	.054
ST	.085	.087	.088	.086	.086	.085	.080	.076	.069
KLD-U	.025	.036	.043	.045	.047	.050	.049	.041	.022
KLD-N	.026	.032	.033	.030	.028	.024	.020	.015	.009
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.187	.172	.196	.224	.250	.266	.251	.234	.156
LR	.246	.273	.304	.332	.340	.338	.308	.269	.192
ST	.235	.276	.313	.349	.365	.362	.330	.285	.215
KLD-U	.124	.138	.161	.183	.197	.210	.202	.197	.134
KLD-N	.163	.164	.170	.176	.170	.163	.141	.131	.089

Continued on the next page

Table A61 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.134	.125	.136	.151	.162	.171	.167	.158	.111
LR	.168	.183	.199	.210	.212	.211	.195	.171	.124
ST	.168	.191	.211	.226	.231	.228	.208	.180	.147
KLD-U	.067	.089	.103	.111	.118	.129	.129	.114	.065
KLD-N	.089	.100	.103	.098	.090	.085	.072	.058	.034
Low Discrimination (a = 0.6)									
Z	.100	.101	.104	.109	.114	.118	.119	.111	.074
LR	.116	.125	.133	.137	.139	.138	.129	.113	.074
ST	.122	.133	.143	.148	.150	.146	.135	.119	.102
KLD-U	.034	.054	.070	.077	.082	.086	.083	.062	.032
KLD-N	.039	.051	.058	.055	.049	.042	.034	.022	.012
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.483	.541	.585	.688	.782	.773	.725	.636	.262
LR	.678	.747	.801	.842	.852	.831	.781	.682	.404
ST	.619	.683	.737	.822	.861	.844	.796	.696	.469
KLD-U	.456	.526	.598	.672	.720	.720	.683	.605	.297
KLD-N	.541	.591	.625	.655	.668	.638	.577	.495	.226
Moderate Discrimination (a = 1.0)									
Z	.350	.393	.446	.513	.542	.538	.497	.416	.189
LR	.462	.532	.581	.612	.614	.594	.534	.446	.260
ST	.455	.530	.589	.631	.635	.613	.550	.461	.338
KLD-U	.264	.337	.388	.431	.460	.469	.433	.332	.152
KLD-N	.315	.366	.386	.391	.382	.353	.296	.213	.092
Low Discrimination (a = 0.6)									
Z	.238	.265	.296	.318	.322	.324	.304	.233	.115
LR	.283	.323	.353	.369	.364	.352	.318	.250	.132
ST	.292	.335	.369	.387	.381	.366	.327	.274	.213
KLD-U	.121	.183	.225	.247	.255	.257	.227	.150	.066
KLD-N	.136	.174	.191	.186	.166	.145	.112	.070	.030
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.684	.733	.802	.947	.979	.964	.935	.814	.282
LR	.924	.958	.979	.988	.989	.984	.965	.887	.517
ST	.680	.653	.691	.875	.984	.985	.968	.898	.613
KLD-U	.774	.833	.912	.957	.972	.966	.944	.837	.393
KLD-N	.862	.895	.927	.947	.954	.940	.902	.767	.315
Moderate Discrimination (a = 1.0)									
Z	.600	.657	.766	.864	.877	.854	.797	.592	.191
LR	.803	.860	.896	.910	.906	.880	.822	.682	.348
ST	.738	.795	.869	.914	.913	.888	.830	.712	.502
KLD-U	.609	.692	.766	.815	.832	.815	.747	.554	.224
KLD-N	.676	.728	.758	.770	.762	.711	.608	.416	.147
Low Discrimination (a = 0.6)									
Z	.465	.531	.592	.622	.622	.600	.525	.331	.126
LR	.556	.620	.660	.672	.661	.625	.552	.397	.177
ST	.562	.631	.674	.688	.674	.636	.564	.463	.336
KLD-U	.349	.448	.506	.540	.546	.521	.430	.261	.101
KLD-N	.362	.423	.445	.436	.401	.347	.256	.142	.050

Table A62. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.076	.044	.040	.041	.042	.044	.047	.049	.057
LR	.072	.065	.061	.062	.059	.058	.059	.056	.056
ST	.074	.071	.069	.069	.066	.064	.065	.060	.058
KLD-U	.072	.068	.060	.058	.054	.053	.057	.061	.057
KLD-N	.080	.067	.055	.051	.046	.045	.044	.041	.036
Moderate Discrimination (a = 1.0)									
Z	.067	.048	.043	.045	.044	.047	.050	.054	.057
LR	.067	.062	.060	.061	.057	.058	.057	.056	.053
ST	.070	.066	.066	.067	.063	.063	.062	.059	.055
KLD-U	.049	.055	.055	.055	.052	.053	.056	.057	.036
KLD-N	.051	.050	.046	.044	.040	.039	.037	.032	.021
Low Discrimination (a = 0.6)									
Z	.058	.055	.050	.048	.048	.049	.054	.059	.055
LR	.061	.061	.059	.057	.057	.055	.056	.056	.048
ST	.067	.065	.063	.062	.062	.059	.058	.057	.055
KLD-U	.030	.042	.050	.052	.051	.052	.053	.044	.023
KLD-N	.026	.031	.034	.034	.032	.030	.027	.021	.012
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.117	.098	.099	.104	.106	.102	.099	.095	.083
LR	.110	.115	.120	.127	.130	.126	.121	.112	.097
ST	.113	.126	.133	.139	.143	.139	.134	.124	.107
KLD-U	.087	.098	.098	.100	.099	.094	.092	.091	.075
KLD-N	.090	.093	.089	.088	.086	.078	.071	.062	.050
Moderate Discrimination (a = 1.0)									
Z	.094	.080	.077	.080	.080	.078	.079	.080	.071
LR	.088	.090	.091	.094	.096	.093	.091	.087	.075
ST	.092	.096	.099	.103	.105	.102	.100	.093	.084
KLD-U	.054	.070	.072	.074	.073	.070	.072	.068	.043
KLD-N	.055	.062	.061	.060	.057	.050	.044	.037	.023
Low Discrimination (a = 0.6)									
Z	.076	.073	.068	.067	.066	.068	.069	.073	.060
LR	.072	.075	.075	.075	.076	.075	.073	.072	.056
ST	.078	.080	.080	.081	.082	.081	.077	.075	.069
KLD-U	.030	.046	.057	.060	.059	.060	.057	.045	.025
KLD-N	.028	.036	.041	.039	.037	.032	.026	.019	.011
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.250	.255	.281	.299	.298	.293	.265	.239	.161
LR	.248	.284	.317	.337	.341	.336	.308	.271	.199
ST	.257	.303	.336	.359	.364	.359	.332	.291	.222
KLD-U	.207	.239	.264	.276	.276	.271	.250	.227	.151
KLD-N	.200	.228	.249	.257	.249	.234	.203	.170	.109

Continued on the next page

Table A62 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.171	.170	.179	.185	.184	.182	.172	.161	.116
LR	.163	.184	.200	.209	.212	.209	.194	.173	.130
ST	.168	.194	.214	.224	.228	.224	.208	.183	.151
KLD-U	.106	.144	.161	.164	.164	.161	.153	.130	.075
KLD-N	.104	.129	.141	.141	.132	.121	.101	.077	.044
Low Discrimination (a = 0.6)									
Z	.120	.120	.119	.122	.121	.123	.121	.111	.079
LR	.111	.121	.128	.134	.134	.135	.127	.112	.081
ST	.117	.126	.136	.143	.144	.143	.134	.119	.105
KLD-U	.047	.080	.098	.105	.104	.103	.093	.068	.036
KLD-N	.044	.064	.073	.073	.067	.058	.044	.028	.016
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.699	.757	.798	.822	.819	.791	.733	.641	.276
LR	.714	.780	.822	.849	.852	.831	.783	.688	.421
ST	.724	.790	.833	.860	.864	.846	.801	.706	.487
KLD-U	.673	.736	.777	.802	.806	.781	.732	.636	.328
KLD-N	.660	.725	.764	.785	.779	.739	.668	.550	.268
Moderate Discrimination (a = 1.0)									
Z	.474	.526	.561	.579	.575	.554	.507	.422	.202
LR	.481	.550	.593	.614	.616	.593	.538	.454	.280
ST	.491	.566	.611	.632	.635	.612	.557	.472	.356
KLD-U	.404	.486	.525	.543	.541	.522	.471	.362	.174
KLD-N	.386	.459	.492	.499	.479	.438	.362	.256	.120
Low Discrimination (a = 0.6)									
Z	.291	.314	.332	.345	.341	.333	.306	.239	.132
LR	.285	.322	.352	.368	.364	.352	.318	.259	.152
ST	.293	.334	.365	.383	.379	.365	.328	.279	.228
KLD-U	.176	.256	.293	.309	.302	.289	.246	.164	.074
KLD-N	.162	.213	.238	.238	.217	.184	.138	.084	.036
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.942	.967	.980	.985	.980	.964	.936	.820	.296
LR	.943	.967	.981	.988	.988	.983	.966	.893	.538
ST	.942	.966	.982	.989	.989	.985	.970	.906	.636
KLD-U	.940	.965	.978	.984	.983	.976	.954	.857	.428
KLD-N	.933	.960	.975	.981	.978	.966	.931	.804	.368
Moderate Discrimination (a = 1.0)									
Z	.817	.860	.883	.893	.884	.858	.800	.604	.208
LR	.825	.872	.897	.910	.903	.881	.825	.697	.380
ST	.829	.877	.902	.916	.911	.891	.835	.726	.536
KLD-U	.784	.840	.864	.877	.869	.843	.774	.586	.253
KLD-N	.765	.820	.845	.848	.826	.778	.671	.474	.192
Low Discrimination (a = 0.6)									
Z	.562	.612	.636	.646	.636	.606	.527	.347	.149
LR	.564	.626	.658	.671	.660	.625	.557	.419	.207
ST	.574	.638	.671	.684	.673	.638	.570	.476	.368
KLD-U	.465	.561	.595	.606	.593	.551	.451	.282	.114
KLD-N	.423	.498	.520	.515	.473	.401	.294	.166	.061

Table A63. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.075	.045	.040	.042	.044	.046	.048	.050	.057
LR	.069	.062	.059	.059	.056	.056	.057	.054	.055
ST	.071	.068	.065	.065	.062	.061	.062	.057	.055
KLD-U	.074	.077	.069	.065	.061	.060	.062	.065	.060
KLD-N	.082	.074	.062	.056	.052	.049	.048	.046	.039
Moderate Discrimination (a = 1.0)									
Z	.068	.050	.045	.046	.045	.048	.052	.056	.056
LR	.065	.060	.058	.058	.055	.055	.057	.055	.052
ST	.068	.064	.064	.064	.060	.060	.060	.056	.054
KLD-U	.052	.062	.062	.062	.057	.058	.062	.062	.040
KLD-N	.053	.054	.051	.048	.043	.042	.041	.036	.024
Low Discrimination (a = 0.6)									
Z	.060	.056	.051	.049	.050	.051	.055	.060	.054
LR	.059	.059	.056	.055	.054	.054	.054	.055	.046
ST	.064	.062	.060	.059	.058	.058	.056	.055	.055
KLD-U	.033	.045	.054	.057	.056	.057	.056	.047	.025
KLD-N	.029	.033	.036	.037	.035	.032	.030	.023	.013
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.118	.102	.104	.110	.112	.106	.103	.098	.083
LR	.107	.112	.118	.124	.127	.122	.118	.109	.096
ST	.109	.120	.128	.134	.139	.133	.129	.119	.106
KLD-U	.097	.118	.116	.117	.113	.106	.102	.099	.077
KLD-N	.096	.102	.100	.101	.097	.089	.080	.072	.052
Moderate Discrimination (a = 1.0)									
Z	.096	.082	.081	.083	.084	.081	.081	.083	.068
LR	.086	.086	.088	.091	.093	.091	.089	.086	.072
ST	.087	.091	.095	.099	.101	.099	.096	.091	.082
KLD-U	.059	.080	.084	.085	.083	.079	.078	.073	.045
KLD-N	.058	.066	.068	.068	.065	.057	.052	.042	.026
Low Discrimination (a = 0.6)									
Z	.077	.075	.071	.070	.069	.070	.071	.071	.054
LR	.069	.072	.073	.074	.073	.074	.071	.070	.052
ST	.073	.076	.077	.078	.078	.078	.074	.072	.069
KLD-U	.034	.053	.066	.069	.067	.065	.060	.048	.026
KLD-N	.030	.039	.045	.045	.040	.036	.029	.021	.012
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.258	.265	.288	.310	.309	.303	.273	.246	.158
LR	.245	.280	.308	.333	.337	.332	.302	.269	.196
ST	.250	.294	.325	.351	.356	.352	.322	.288	.220
KLD-U	.234	.274	.292	.304	.302	.296	.266	.242	.152
KLD-N	.215	.248	.269	.280	.273	.260	.226	.186	.111

Continued on the next page

Table A63 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.174	.175	.185	.194	.194	.188	.177	.163	.106
LR	.160	.179	.194	.207	.210	.205	.191	.172	.126
ST	.163	.187	.206	.218	.223	.218	.202	.181	.149
KLD-U	.121	.166	.180	.186	.184	.175	.163	.138	.077
KLD-N	.112	.139	.152	.158	.150	.135	.117	.086	.048
Low Discrimination (a = 0.6)									
Z	.122	.123	.124	.127	.126	.126	.123	.107	.069
LR	.108	.118	.125	.131	.131	.132	.126	.110	.073
ST	.112	.122	.130	.137	.138	.138	.130	.117	.105
KLD-U	.054	.090	.111	.118	.115	.111	.100	.072	.037
KLD-N	.048	.068	.078	.081	.074	.066	.052	.033	.017
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.704	.760	.800	.826	.819	.790	.737	.632	.266
LR	.706	.772	.813	.840	.841	.820	.774	.677	.417
ST	.713	.780	.821	.850	.853	.834	.790	.697	.484
KLD-U	.708	.764	.798	.820	.818	.791	.745	.636	.328
KLD-N	.677	.742	.779	.800	.794	.756	.693	.557	.269
Moderate Discrimination (a = 1.0)									
Z	.481	.533	.570	.586	.582	.557	.511	.406	.170
LR	.474	.542	.584	.605	.607	.583	.531	.445	.266
ST	.481	.554	.598	.620	.624	.601	.547	.463	.353
KLD-U	.440	.520	.555	.569	.565	.537	.487	.368	.175
KLD-N	.402	.476	.515	.523	.506	.464	.387	.269	.125
Low Discrimination (a = 0.6)									
Z	.297	.321	.342	.354	.348	.338	.302	.213	.100
LR	.278	.316	.346	.363	.360	.347	.314	.248	.130
ST	.282	.323	.356	.375	.372	.358	.323	.277	.226
KLD-U	.197	.279	.318	.332	.321	.302	.255	.170	.075
KLD-N	.170	.223	.254	.256	.234	.203	.152	.094	.037
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.942	.966	.979	.984	.980	.964	.932	.811	.285
LR	.939	.963	.978	.985	.986	.981	.961	.888	.531
ST	.936	.961	.977	.986	.987	.983	.966	.901	.633
KLD-U	.948	.969	.980	.984	.982	.976	.954	.854	.425
KLD-N	.937	.963	.976	.981	.978	.968	.934	.804	.367
Moderate Discrimination (a = 1.0)									
Z	.817	.860	.883	.891	.883	.857	.793	.563	.168
LR	.816	.864	.889	.901	.896	.872	.817	.684	.356
ST	.818	.868	.894	.906	.904	.881	.827	.719	.529
KLD-U	.805	.853	.876	.883	.875	.848	.779	.589	.249
KLD-N	.773	.830	.854	.856	.839	.794	.686	.484	.190
Low Discrimination (a = 0.6)									
Z	.566	.616	.641	.651	.639	.606	.497	.290	.106
LR	.554	.616	.648	.662	.650	.619	.547	.391	.170
ST	.560	.624	.657	.674	.662	.629	.562	.471	.360
KLD-U	.489	.584	.618	.625	.606	.563	.458	.283	.109
KLD-N	.433	.512	.538	.532	.495	.424	.315	.171	.058

Table A64. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.036	.037	.037	.038	.039	.041	.041	.044	.044
LR	.079	.076	.071	.069	.065	.064	.064	.065	.062
ST	.079	.079	.076	.077	.075	.074	.074	.075	.070
KLD-U	.064	.056	.048	.045	.042	.042	.044	.048	.044
KLD-N	.080	.061	.047	.040	.036	.034	.033	.032	.027
Moderate Discrimination (a = 1.0)									
Z	.033	.036	.038	.040	.040	.042	.043	.046	.048
LR	.070	.069	.067	.066	.064	.063	.062	.062	.061
ST	.075	.075	.074	.074	.072	.073	.071	.069	.067
KLD-U	.046	.046	.044	.044	.041	.042	.043	.044	.033
KLD-N	.053	.046	.038	.035	.030	.029	.027	.025	.019
Low Discrimination (a = 0.6)									
Z	.035	.038	.040	.041	.042	.043	.044	.047	.050
LR	.065	.065	.064	.063	.062	.061	.061	.060	.060
ST	.070	.071	.071	.071	.071	.069	.068	.066	.064
KLD-U	.036	.040	.041	.041	.040	.041	.040	.034	.021
KLD-N	.034	.032	.028	.027	.022	.021	.018	.016	.011
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.067	.069	.074	.077	.080	.087	.084	.084	.077
LR	.125	.122	.124	.124	.124	.127	.121	.119	.110
ST	.124	.126	.132	.135	.138	.142	.137	.135	.127
KLD-U	.066	.062	.062	.061	.061	.063	.065	.068	.059
KLD-N	.089	.071	.064	.056	.050	.046	.040	.035	.025
Moderate Discrimination (a = 1.0)									
Z	.053	.056	.059	.062	.064	.065	.067	.068	.067
LR	.097	.095	.095	.096	.093	.093	.092	.092	.087
ST	.102	.103	.104	.106	.105	.105	.103	.103	.096
KLD-U	.046	.046	.048	.048	.047	.048	.052	.051	.037
KLD-N	.059	.050	.045	.040	.034	.028	.026	.020	.013
Low Discrimination (a = 0.6)									
Z	.046	.051	.054	.052	.053	.053	.057	.060	.062
LR	.078	.080	.080	.076	.076	.073	.076	.076	.074
ST	.084	.087	.088	.085	.085	.082	.085	.083	.079
KLD-U	.032	.039	.041	.040	.040	.040	.042	.036	.023
KLD-N	.036	.036	.032	.025	.021	.017	.015	.011	.007
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.157	.170	.187	.207	.224	.235	.229	.216	.192
LR	.258	.269	.285	.302	.310	.309	.297	.282	.256
ST	.259	.274	.295	.317	.333	.333	.323	.308	.285
KLD-U	.134	.139	.150	.163	.172	.180	.184	.179	.142
KLD-N	.171	.162	.157	.154	.145	.134	.120	.099	.064

Continued on the next page

Table A64 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.107	.117	.127	.133	.140	.144	.147	.143	.135
LR	.174	.179	.188	.189	.190	.193	.191	.182	.170
ST	.183	.189	.200	.206	.208	.211	.208	.198	.185
KLD-U	.078	.088	.095	.098	.100	.107	.112	.105	.074
KLD-N	.101	.099	.093	.084	.072	.064	.054	.040	.026
Low Discrimination (a = 0.6)									
Z	.077	.083	.086	.088	.090	.092	.096	.096	.096
LR	.117	.123	.123	.123	.123	.121	.123	.118	.114
ST	.125	.132	.134	.135	.136	.133	.134	.127	.123
KLD-U	.044	.056	.059	.061	.063	.065	.067	.055	.037
KLD-N	.054	.053	.047	.040	.032	.026	.020	.015	.011
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.504	.549	.584	.654	.727	.734	.706	.660	.581
LR	.686	.724	.767	.799	.810	.804	.782	.743	.683
ST	.634	.671	.718	.787	.824	.821	.802	.768	.713
KLD-U	.471	.510	.564	.615	.656	.669	.662	.620	.464
KLD-N	.554	.566	.584	.594	.594	.570	.527	.444	.297
Moderate Discrimination (a = 1.0)									
Z	.345	.382	.415	.447	.463	.474	.463	.436	.379
LR	.476	.506	.533	.547	.549	.548	.534	.503	.452
ST	.474	.506	.544	.567	.573	.572	.557	.527	.477
KLD-U	.283	.321	.345	.366	.380	.399	.396	.349	.219
KLD-N	.335	.345	.339	.321	.296	.273	.232	.175	.112
Low Discrimination (a = 0.6)									
Z	.211	.227	.239	.245	.252	.257	.263	.258	.225
LR	.288	.302	.309	.312	.313	.310	.311	.298	.268
ST	.300	.316	.327	.332	.334	.330	.328	.314	.288
KLD-U	.139	.164	.175	.181	.189	.196	.193	.154	.091
KLD-N	.159	.157	.144	.122	.104	.087	.069	.053	.037
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.783	.782	.813	.935	.968	.961	.940	.908	.798
LR	.918	.947	.971	.980	.983	.980	.973	.959	.913
ST	.685	.680	.734	.889	.978	.983	.976	.965	.930
KLD-U	.782	.827	.891	.937	.954	.956	.947	.910	.704
KLD-N	.859	.879	.903	.922	.926	.919	.888	.809	.583
Moderate Discrimination (a = 1.0)									
Z	.643	.681	.738	.800	.819	.817	.793	.748	.579
LR	.799	.831	.852	.865	.867	.862	.841	.808	.725
ST	.731	.776	.831	.870	.878	.874	.854	.823	.763
KLD-U	.611	.660	.702	.737	.759	.770	.743	.645	.401
KLD-N	.675	.689	.686	.676	.658	.624	.547	.435	.280
Low Discrimination (a = 0.6)									
Z	.439	.467	.488	.507	.514	.522	.515	.483	.352
LR	.550	.569	.574	.584	.583	.584	.567	.542	.448
ST	.551	.580	.592	.605	.605	.604	.584	.559	.510
KLD-U	.348	.384	.400	.417	.431	.437	.403	.310	.168
KLD-N	.367	.363	.332	.300	.264	.230	.187	.144	.094

Table A65. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.035	.039	.040	.042	.042	.043	.043	.046	.046
LR	.067	.064	.063	.060	.059	.059	.059	.060	.057
ST	.076	.073	.070	.067	.065	.066	.066	.067	.063
KLD-U	.076	.067	.061	.057	.053	.053	.054	.061	.058
KLD-N	.085	.067	.055	.049	.045	.043	.042	.043	.038
Moderate Discrimination (a = 1.0)									
Z	.036	.039	.040	.043	.044	.044	.046	.048	.050
LR	.063	.062	.060	.059	.058	.058	.058	.058	.058
ST	.070	.068	.066	.066	.065	.064	.064	.063	.061
KLD-U	.055	.057	.054	.053	.051	.051	.052	.054	.043
KLD-N	.058	.051	.045	.042	.038	.036	.035	.032	.025
Low Discrimination (a = 0.6)									
Z	.039	.041	.042	.044	.045	.045	.046	.048	.051
LR	.059	.058	.057	.058	.058	.056	.057	.057	.057
ST	.066	.064	.063	.065	.065	.062	.062	.061	.060
KLD-U	.040	.047	.048	.050	.050	.048	.047	.042	.026
KLD-N	.037	.036	.034	.031	.028	.025	.023	.019	.013
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.083	.089	.096	.098	.095	.097	.090	.088	.079
LR	.115	.114	.118	.120	.118	.121	.115	.114	.107
ST	.131	.129	.131	.131	.130	.134	.129	.130	.125
KLD-U	.100	.097	.096	.093	.087	.089	.084	.089	.080
KLD-N	.100	.091	.086	.082	.074	.072	.063	.062	.052
Moderate Discrimination (a = 1.0)									
Z	.067	.071	.073	.074	.072	.070	.070	.070	.066
LR	.089	.089	.089	.090	.089	.088	.087	.088	.084
ST	.100	.100	.098	.099	.099	.098	.097	.098	.094
KLD-U	.064	.070	.070	.069	.065	.063	.065	.065	.049
KLD-N	.065	.062	.059	.056	.048	.043	.040	.035	.026
Low Discrimination (a = 0.6)									
Z	.055	.058	.061	.058	.058	.056	.059	.061	.060
LR	.072	.073	.074	.071	.071	.069	.071	.072	.070
ST	.080	.081	.081	.078	.079	.077	.079	.079	.076
KLD-U	.040	.050	.056	.053	.053	.050	.051	.046	.029
KLD-N	.038	.039	.038	.033	.029	.025	.022	.018	.011
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.218	.238	.258	.268	.266	.261	.248	.231	.197
LR	.264	.277	.294	.304	.307	.305	.297	.283	.257
ST	.290	.300	.315	.322	.330	.328	.322	.310	.291
KLD-U	.224	.233	.242	.244	.242	.240	.234	.231	.193
KLD-N	.217	.222	.227	.224	.215	.204	.186	.174	.138

Continued on the next page

Table A65 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.144	.152	.163	.161	.159	.157	.156	.148	.135
LR	.174	.179	.187	.186	.187	.188	.186	.178	.170
ST	.190	.193	.201	.201	.203	.204	.203	.196	.191
KLD-U	.126	.139	.146	.142	.139	.138	.139	.132	.101
KLD-N	.123	.125	.127	.118	.108	.100	.090	.075	.058
Low Discrimination (a = 0.6)									
Z	.096	.101	.100	.102	.099	.098	.100	.097	.095
LR	.114	.118	.116	.118	.118	.117	.119	.114	.113
ST	.126	.128	.126	.129	.128	.128	.129	.125	.124
KLD-U	.063	.080	.085	.086	.085	.082	.083	.068	.047
KLD-N	.061	.064	.061	.056	.049	.042	.036	.027	.019
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.678	.721	.757	.770	.769	.757	.725	.675	.589
LR	.723	.754	.785	.803	.808	.802	.783	.748	.699
ST	.740	.768	.797	.816	.824	.819	.805	.775	.739
KLD-U	.678	.706	.733	.745	.750	.743	.727	.691	.572
KLD-N	.667	.694	.717	.722	.716	.694	.664	.607	.493
Moderate Discrimination (a = 1.0)									
Z	.455	.484	.503	.506	.499	.495	.482	.449	.391
LR	.496	.521	.540	.547	.546	.545	.534	.510	.469
ST	.517	.540	.559	.567	.568	.569	.559	.540	.504
KLD-U	.423	.453	.467	.468	.462	.462	.456	.414	.289
KLD-N	.406	.427	.432	.419	.396	.374	.344	.288	.204
Low Discrimination (a = 0.6)									
Z	.262	.274	.278	.277	.273	.271	.271	.260	.232
LR	.291	.301	.307	.310	.308	.307	.308	.298	.280
ST	.308	.319	.323	.327	.327	.326	.326	.317	.303
KLD-U	.201	.233	.241	.242	.238	.237	.229	.188	.117
KLD-N	.186	.195	.188	.173	.154	.137	.116	.089	.059
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.933	.956	.970	.974	.971	.959	.939	.910	.804
LR	.939	.958	.973	.979	.981	.980	.972	.962	.930
ST	.938	.957	.974	.981	.983	.982	.977	.970	.948
KLD-U	.936	.953	.965	.971	.973	.970	.962	.940	.801
KLD-N	.927	.947	.960	.965	.964	.957	.942	.910	.769
Moderate Discrimination (a = 1.0)									
Z	.798	.820	.833	.837	.834	.827	.802	.756	.627
LR	.820	.840	.854	.862	.863	.859	.844	.818	.771
ST	.828	.849	.863	.872	.875	.872	.858	.838	.803
KLD-U	.775	.797	.808	.813	.812	.812	.790	.718	.501
KLD-N	.757	.775	.779	.772	.754	.735	.683	.602	.425
Low Discrimination (a = 0.6)									
Z	.525	.539	.537	.541	.534	.536	.524	.492	.394
LR	.556	.571	.573	.581	.578	.580	.567	.550	.508
ST	.574	.589	.591	.601	.598	.600	.588	.573	.547
KLD-U	.462	.490	.493	.496	.492	.490	.454	.364	.215
KLD-N	.423	.430	.412	.388	.360	.328	.280	.216	.138

Table A66. Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.038	.040	.042	.044	.044	.046	.044	.049	.047
LR	.063	.062	.060	.059	.056	.056	.055	.059	.054
ST	.071	.068	.066	.065	.062	.063	.061	.064	.058
KLD-U	.085	.077	.071	.064	.059	.058	.059	.067	.064
KLD-N	.092	.073	.062	.055	.049	.048	.046	.050	.044
Moderate Discrimination (a = 1.0)									
Z	.037	.041	.042	.044	.045	.047	.048	.048	.052
LR	.060	.059	.057	.057	.056	.056	.056	.054	.056
ST	.066	.065	.063	.062	.062	.062	.061	.059	.058
KLD-U	.059	.063	.061	.060	.057	.057	.057	.058	.048
KLD-N	.063	.056	.050	.046	.042	.040	.037	.036	.031
Low Discrimination (a = 0.6)									
Z	.041	.042	.044	.045	.047	.048	.049	.051	.052
LR	.057	.056	.055	.056	.056	.056	.055	.055	.055
ST	.063	.061	.061	.061	.061	.060	.059	.057	.056
KLD-U	.044	.050	.053	.055	.055	.054	.052	.046	.029
KLD-N	.041	.039	.036	.034	.031	.028	.026	.023	.016
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.087	.093	.100	.104	.100	.101	.094	.091	.084
LR	.112	.112	.115	.117	.115	.118	.111	.109	.105
ST	.126	.124	.126	.126	.125	.129	.124	.122	.120
KLD-U	.117	.115	.113	.109	.101	.101	.095	.097	.092
KLD-N	.112	.101	.098	.094	.085	.084	.074	.074	.066
Moderate Discrimination (a = 1.0)									
Z	.070	.075	.076	.078	.076	.074	.073	.072	.072
LR	.085	.088	.087	.088	.087	.085	.084	.083	.083
ST	.095	.096	.094	.095	.094	.094	.093	.092	.090
KLD-U	.072	.082	.081	.079	.076	.073	.071	.071	.058
KLD-N	.070	.068	.065	.061	.056	.050	.046	.043	.035
Low Discrimination (a = 0.6)									
Z	.058	.061	.064	.061	.062	.059	.062	.063	.063
LR	.069	.071	.072	.068	.070	.066	.070	.069	.070
ST	.076	.078	.078	.074	.076	.072	.076	.075	.074
KLD-U	.044	.057	.062	.060	.060	.056	.057	.050	.033
KLD-N	.042	.043	.042	.036	.033	.028	.026	.022	.016
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.226	.246	.268	.278	.277	.271	.255	.242	.202
LR	.258	.273	.290	.299	.304	.300	.289	.277	.251
ST	.281	.291	.306	.315	.322	.319	.312	.301	.283
KLD-U	.255	.263	.271	.271	.266	.262	.252	.249	.212
KLD-N	.235	.240	.248	.246	.238	.227	.209	.205	.164

Continued on the next page

Table A66 (continued). Mean Type I Error Rates and Power for 15-Item Fixed-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.150	.158	.170	.169	.167	.165	.162	.156	.143
LR	.168	.174	.182	.183	.184	.184	.181	.175	.168
ST	.183	.186	.194	.194	.197	.198	.197	.190	.186
KLD-U	.144	.159	.166	.162	.157	.156	.151	.144	.114
KLD-N	.132	.136	.139	.132	.123	.114	.106	.097	.075
Low Discrimination (a = 0.6)									
Z	.101	.105	.107	.106	.105	.103	.104	.100	.097
LR	.111	.115	.115	.116	.116	.116	.116	.111	.112
ST	.119	.123	.122	.124	.125	.124	.124	.119	.120
KLD-U	.072	.092	.098	.098	.097	.093	.090	.076	.054
KLD-N	.066	.069	.068	.063	.056	.049	.045	.035	.026
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.681	.724	.761	.774	.771	.758	.727	.678	.579
LR	.712	.744	.776	.793	.798	.791	.770	.735	.676
ST	.726	.756	.786	.805	.812	.807	.791	.761	.715
KLD-U	.710	.733	.756	.764	.764	.757	.738	.705	.585
KLD-N	.683	.711	.734	.739	.734	.716	.689	.644	.542
Moderate Discrimination (a = 1.0)									
Z	.460	.494	.512	.516	.506	.504	.488	.461	.372
LR	.485	.514	.531	.539	.536	.536	.524	.501	.457
ST	.502	.530	.546	.555	.555	.557	.546	.524	.490
KLD-U	.452	.484	.497	.498	.488	.486	.472	.435	.307
KLD-N	.420	.445	.453	.445	.422	.402	.382	.331	.251
Low Discrimination (a = 0.6)									
Z	.271	.282	.286	.287	.282	.278	.277	.263	.214
LR	.286	.296	.301	.304	.302	.301	.300	.294	.273
ST	.300	.309	.314	.320	.318	.317	.317	.310	.296
KLD-U	.223	.256	.265	.266	.260	.254	.243	.203	.130
KLD-N	.197	.203	.201	.190	.170	.156	.144	.114	.082
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.930	.955	.968	.973	.971	.960	.940	.904	.776
LR	.933	.952	.969	.976	.978	.975	.968	.953	.918
ST	.931	.950	.969	.977	.980	.978	.973	.962	.936
KLD-U	.944	.959	.968	.972	.973	.970	.962	.939	.813
KLD-N	.931	.950	.962	.966	.966	.960	.948	.916	.810
Moderate Discrimination (a = 1.0)									
Z	.798	.822	.833	.836	.834	.825	.804	.743	.545
LR	.810	.831	.844	.851	.853	.849	.832	.805	.745
ST	.817	.838	.851	.860	.864	.860	.846	.824	.782
KLD-U	.795	.814	.822	.823	.823	.817	.796	.726	.518
KLD-N	.766	.786	.790	.783	.770	.752	.717	.638	.488
Low Discrimination (a = 0.6)									
Z	.532	.548	.544	.548	.543	.541	.520	.460	.319
LR	.547	.562	.563	.572	.570	.570	.557	.542	.471
ST	.560	.575	.577	.588	.587	.588	.575	.562	.531
KLD-U	.488	.519	.518	.521	.516	.507	.469	.382	.226
KLD-N	.434	.443	.428	.412	.386	.362	.318	.260	.170

Table A67. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.080	.048	.046	.045	.045	.046	.046	.050	.064
LR	.061	.057	.059	.058	.056	.058	.055	.054	.055
ST	.060	.059	.063	.063	.061	.062	.059	.056	.052
KLD-U	.053	.053	.048	.046	.044	.045	.045	.049	.058
KLD-N	.050	.050	.046	.043	.041	.041	.040	.041	.042
Moderate Discrimination (a = 1.0)									
Z	.067	.050	.047	.048	.046	.047	.049	.051	.061
LR	.060	.056	.058	.059	.057	.055	.056	.054	.055
ST	.060	.059	.062	.063	.061	.060	.060	.056	.054
KLD-U	.043	.049	.046	.046	.045	.045	.048	.051	.047
KLD-N	.039	.043	.041	.041	.040	.038	.039	.038	.033
Low Discrimination (a = 0.6)									
Z	.062	.054	.051	.049	.048	.049	.051	.053	.062
LR	.060	.058	.058	.057	.056	.055	.056	.054	.057
ST	.063	.060	.061	.060	.059	.058	.058	.054	.055
KLD-U	.028	.044	.049	.047	.046	.047	.049	.047	.033
KLD-N	.024	.035	.038	.038	.036	.035	.034	.030	.022
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.142	.145	.152	.163	.170	.169	.165	.158	.152
LR	.131	.165	.182	.192	.197	.194	.186	.164	.132
ST	.130	.171	.192	.202	.208	.204	.194	.167	.124
KLD-U	.110	.125	.127	.136	.140	.139	.137	.137	.108
KLD-N	.102	.124	.127	.132	.132	.126	.120	.116	.081
Moderate Discrimination (a = 1.0)									
Z	.106	.100	.104	.113	.113	.112	.109	.107	.111
LR	.101	.112	.121	.131	.131	.128	.120	.111	.100
ST	.102	.116	.128	.138	.138	.135	.126	.113	.097
KLD-U	.068	.084	.086	.091	.092	.091	.090	.091	.072
KLD-N	.065	.079	.082	.085	.082	.077	.070	.067	.049
Low Discrimination (a = 0.6)									
Z	.081	.079	.078	.080	.080	.080	.085	.085	.088
LR	.080	.084	.086	.090	.089	.089	.090	.084	.080
ST	.083	.087	.091	.094	.095	.092	.093	.085	.077
KLD-U	.035	.059	.066	.067	.067	.067	.071	.064	.040
KLD-N	.035	.050	.055	.055	.053	.048	.046	.037	.023
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.388	.442	.488	.530	.542	.532	.503	.457	.346
LR	.391	.484	.538	.573	.583	.568	.531	.458	.314
ST	.392	.494	.550	.586	.597	.582	.542	.460	.298
KLD-U	.352	.395	.435	.474	.486	.480	.456	.416	.255
KLD-N	.338	.401	.438	.468	.469	.453	.421	.374	.201

Continued on the next page

Table A67 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.239	.266	.294	.317	.325	.324	.304	.275	.241
LR	.240	.290	.327	.350	.357	.352	.321	.277	.220
ST	.243	.299	.338	.362	.369	.363	.330	.279	.213
KLD-U	.185	.229	.252	.270	.279	.279	.266	.239	.156
KLD-N	.179	.222	.246	.257	.256	.246	.225	.189	.116
Low Discrimination (a = 0.6)									
Z	.145	.159	.173	.183	.188	.185	.181	.170	.156
LR	.146	.170	.188	.201	.206	.197	.186	.166	.143
ST	.150	.175	.195	.209	.214	.203	.190	.166	.139
KLD-U	.080	.126	.147	.155	.159	.155	.151	.128	.073
KLD-N	.080	.111	.128	.133	.130	.119	.105	.079	.043
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.899	.927	.950	.976	.984	.980	.967	.920	.648
LR	.935	.969	.981	.986	.987	.983	.970	.912	.660
ST	.927	.960	.976	.985	.988	.984	.971	.909	.649
KLD-U	.899	.938	.960	.973	.977	.973	.959	.897	.537
KLD-N	.908	.946	.962	.971	.973	.967	.950	.868	.459
Moderate Discrimination (a = 1.0)									
Z	.713	.787	.836	.859	.859	.844	.801	.721	.556
LR	.736	.818	.860	.878	.877	.859	.810	.712	.540
ST	.740	.822	.865	.884	.883	.864	.814	.710	.532
KLD-U	.669	.747	.798	.823	.825	.812	.769	.664	.378
KLD-N	.665	.747	.793	.810	.803	.778	.719	.593	.316
Low Discrimination (a = 0.6)									
Z	.441	.506	.551	.574	.574	.557	.515	.460	.368
LR	.454	.528	.576	.599	.596	.571	.518	.448	.349
ST	.461	.536	.585	.609	.604	.577	.521	.445	.344
KLD-U	.356	.456	.503	.526	.526	.510	.460	.354	.175
KLD-N	.340	.426	.469	.482	.468	.433	.365	.256	.122
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.910	.932	.974	.999	1.000	1.000	1.000	.992	.714
LR	.997	.999	1.000	1.000	1.000	1.000	.999	.991	.800
ST	.917	.902	.936	.991	1.000	1.000	.999	.989	.810
KLD-U	.990	.996	.999	1.000	1.000	1.000	.999	.986	.661
KLD-N	.995	.998	.999	1.000	1.000	1.000	.999	.978	.586
Moderate Discrimination (a = 1.0)									
Z	.944	.966	.988	.994	.994	.990	.978	.936	.697
LR	.979	.990	.994	.995	.994	.991	.978	.929	.746
ST	.973	.986	.994	.995	.995	.991	.978	.926	.761
KLD-U	.959	.979	.989	.991	.991	.987	.972	.896	.537
KLD-N	.963	.980	.988	.989	.988	.982	.959	.856	.486
Low Discrimination (a = 0.6)									
Z	.803	.862	.888	.896	.890	.868	.819	.733	.548
LR	.816	.875	.899	.906	.898	.874	.816	.719	.560
ST	.820	.879	.903	.910	.901	.876	.816	.714	.571
KLD-U	.757	.830	.861	.871	.865	.839	.769	.589	.281
KLD-N	.738	.812	.840	.843	.825	.780	.677	.482	.224

Table A68. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.078	.049	.046	.046	.046	.046	.047	.050	.065
LR	.058	.054	.055	.054	.054	.053	.053	.053	.055
ST	.057	.056	.059	.058	.058	.057	.056	.054	.052
KLD-U	.055	.061	.055	.052	.051	.050	.051	.054	.061
KLD-N	.050	.054	.052	.049	.047	.046	.045	.045	.044
Moderate Discrimination (a = 1.0)									
Z	.066	.051	.048	.048	.047	.047	.050	.052	.061
LR	.057	.053	.054	.055	.053	.053	.054	.053	.055
ST	.057	.055	.057	.059	.057	.056	.057	.054	.053
KLD-U	.046	.055	.053	.052	.050	.050	.053	.055	.051
KLD-N	.041	.046	.047	.046	.044	.043	.044	.041	.035
Low Discrimination (a = 0.6)									
Z	.060	.055	.051	.049	.048	.049	.052	.054	.062
LR	.056	.056	.055	.054	.053	.053	.054	.053	.056
ST	.059	.058	.058	.056	.056	.055	.055	.053	.054
KLD-U	.031	.048	.054	.052	.050	.050	.053	.051	.035
KLD-N	.026	.036	.041	.041	.040	.038	.037	.032	.022
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.147	.158	.168	.179	.180	.176	.170	.160	.152
LR	.128	.164	.183	.194	.196	.192	.184	.164	.132
ST	.126	.168	.190	.203	.206	.201	.191	.168	.125
KLD-U	.122	.156	.162	.169	.168	.163	.156	.145	.111
KLD-N	.107	.145	.155	.163	.160	.152	.142	.126	.085
Moderate Discrimination (a = 1.0)									
Z	.109	.108	.110	.120	.118	.115	.111	.109	.110
LR	.097	.110	.118	.129	.129	.125	.119	.111	.100
ST	.098	.113	.123	.135	.135	.131	.124	.113	.097
KLD-U	.077	.105	.105	.110	.108	.104	.100	.096	.075
KLD-N	.069	.090	.095	.102	.098	.091	.083	.074	.053
Low Discrimination (a = 0.6)									
Z	.083	.083	.081	.083	.084	.083	.086	.086	.088
LR	.076	.082	.085	.088	.089	.088	.088	.084	.079
ST	.079	.085	.088	.092	.093	.090	.090	.084	.077
KLD-U	.041	.070	.077	.077	.076	.074	.075	.068	.044
KLD-N	.038	.055	.061	.063	.061	.055	.051	.042	.027
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.413	.482	.526	.559	.564	.549	.516	.461	.348
LR	.396	.496	.546	.580	.588	.572	.535	.461	.317
ST	.397	.506	.558	.592	.600	.584	.546	.464	.302
KLD-U	.403	.472	.511	.540	.543	.526	.492	.430	.263
KLD-N	.368	.458	.501	.530	.530	.507	.466	.392	.211

Continued on the next page

Table A68 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.254	.288	.314	.336	.336	.333	.310	.276	.242
LR	.241	.295	.330	.353	.356	.351	.323	.277	.222
ST	.242	.302	.339	.362	.367	.361	.332	.280	.215
KLD-U	.220	.278	.298	.314	.314	.309	.286	.248	.163
KLD-N	.196	.255	.282	.298	.293	.281	.250	.202	.126
Low Discrimination (a = 0.6)									
Z	.152	.169	.180	.191	.194	.189	.182	.171	.156
LR	.144	.170	.186	.201	.205	.197	.186	.166	.143
ST	.147	.174	.191	.207	.212	.202	.189	.166	.140
KLD-U	.094	.151	.168	.177	.178	.168	.159	.134	.078
KLD-N	.085	.124	.142	.151	.149	.134	.116	.088	.050
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.944	.971	.981	.986	.986	.981	.968	.922	.654
LR	.944	.973	.983	.987	.988	.983	.970	.914	.665
ST	.945	.973	.983	.988	.988	.984	.972	.912	.654
KLD-U	.945	.970	.980	.984	.984	.979	.964	.902	.549
KLD-N	.938	.968	.979	.983	.983	.976	.958	.877	.477
Moderate Discrimination (a = 1.0)									
Z	.749	.816	.854	.870	.865	.849	.804	.724	.562
LR	.748	.824	.863	.880	.876	.859	.811	.715	.546
ST	.751	.829	.868	.885	.881	.865	.815	.714	.540
KLD-U	.738	.804	.841	.856	.850	.832	.783	.676	.397
KLD-N	.710	.790	.830	.844	.835	.808	.743	.612	.342
Low Discrimination (a = 0.6)									
Z	.459	.526	.564	.584	.583	.563	.519	.462	.372
LR	.456	.533	.576	.598	.595	.572	.519	.450	.354
ST	.461	.539	.583	.606	.604	.577	.522	.447	.349
KLD-U	.403	.506	.542	.560	.554	.529	.473	.368	.188
KLD-N	.365	.458	.501	.518	.501	.462	.387	.279	.142
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	1.000	1.000	1.000	1.000	.999	.993	.721
LR	.998	.999	1.000	1.000	1.000	1.000	1.000	.991	.805
ST	.998	.999	1.000	1.000	1.000	1.000	.999	.990	.815
KLD-U	.999	1.000	1.000	1.000	1.000	1.000	.999	.987	.672
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.999	.980	.604
Moderate Discrimination (a = 1.0)									
Z	.980	.991	.994	.994	.994	.990	.978	.938	.706
LR	.980	.991	.994	.995	.994	.991	.979	.932	.756
ST	.980	.991	.995	.995	.995	.991	.979	.929	.770
KLD-U	.979	.990	.993	.993	.993	.988	.974	.903	.558
KLD-N	.975	.988	.992	.992	.991	.984	.963	.870	.518
Low Discrimination (a = 0.6)									
Z	.818	.870	.894	.899	.893	.869	.822	.736	.558
LR	.818	.875	.899	.904	.898	.872	.818	.724	.573
ST	.822	.878	.902	.907	.901	.874	.818	.720	.581
KLD-U	.799	.859	.883	.887	.877	.848	.780	.608	.303
KLD-N	.763	.835	.861	.863	.845	.798	.699	.516	.255

Table A69. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.078	.051	.047	.046	.047	.047	.048	.051	.065
LR	.058	.054	.054	.052	.053	.053	.052	.052	.055
ST	.056	.055	.057	.055	.056	.055	.055	.053	.051
KLD-U	.055	.065	.059	.055	.054	.053	.054	.056	.062
KLD-N	.051	.057	.055	.051	.050	.049	.048	.048	.045
Moderate Discrimination (a = 1.0)									
Z	.066	.051	.050	.050	.047	.048	.051	.052	.062
LR	.056	.052	.054	.054	.052	.052	.054	.052	.054
ST	.056	.054	.056	.057	.054	.054	.056	.053	.052
KLD-U	.046	.058	.057	.056	.052	.053	.056	.057	.052
KLD-N	.042	.048	.050	.049	.046	.046	.047	.042	.036
Low Discrimination (a = 0.6)									
Z	.061	.055	.052	.049	.049	.050	.052	.055	.062
LR	.055	.054	.054	.052	.052	.052	.053	.053	.055
ST	.056	.055	.056	.054	.054	.054	.054	.053	.052
KLD-U	.032	.050	.056	.055	.052	.053	.055	.053	.036
KLD-N	.028	.038	.042	.043	.041	.040	.038	.033	.024
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.147	.160	.171	.183	.185	.182	.174	.162	.152
LR	.128	.161	.180	.192	.194	.192	.183	.164	.132
ST	.126	.164	.187	.198	.202	.199	.190	.166	.125
KLD-U	.126	.166	.172	.179	.177	.173	.166	.149	.112
KLD-N	.109	.151	.164	.172	.169	.163	.153	.131	.086
Moderate Discrimination (a = 1.0)									
Z	.110	.110	.113	.122	.121	.117	.113	.110	.110
LR	.096	.108	.117	.128	.127	.124	.118	.110	.099
ST	.096	.110	.121	.132	.132	.128	.122	.112	.097
KLD-U	.082	.113	.112	.118	.115	.110	.105	.099	.076
KLD-N	.071	.094	.100	.108	.104	.097	.088	.077	.054
Low Discrimination (a = 0.6)									
Z	.083	.083	.084	.085	.086	.083	.088	.087	.088
LR	.075	.081	.085	.087	.088	.085	.088	.083	.078
ST	.076	.082	.087	.090	.091	.088	.089	.083	.076
KLD-U	.044	.075	.083	.083	.082	.077	.079	.070	.045
KLD-N	.040	.057	.064	.067	.065	.059	.056	.044	.028
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.414	.488	.528	.565	.568	.553	.518	.462	.348
LR	.394	.494	.541	.577	.582	.567	.531	.460	.316
ST	.393	.500	.550	.586	.591	.578	.541	.462	.301
KLD-U	.411	.488	.525	.556	.555	.539	.503	.436	.263
KLD-N	.372	.471	.514	.545	.542	.522	.479	.400	.211

Continued on the next page

Table A69 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.257	.293	.319	.340	.341	.338	.313	.279	.242
LR	.240	.293	.326	.349	.352	.348	.320	.277	.221
ST	.240	.298	.334	.357	.361	.356	.327	.279	.215
KLD-U	.230	.294	.312	.327	.326	.321	.294	.253	.164
KLD-N	.201	.265	.293	.310	.306	.294	.261	.209	.128
Low Discrimination (a = 0.6)									
Z	.153	.171	.184	.194	.199	.192	.185	.172	.156
LR	.142	.168	.185	.198	.203	.195	.185	.165	.142
ST	.144	.171	.189	.203	.208	.199	.188	.165	.138
KLD-U	.101	.160	.178	.186	.187	.175	.164	.137	.079
KLD-N	.089	.129	.148	.158	.157	.141	.122	.093	.053
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.944	.970	.981	.985	.985	.981	.967	.922	.654
LR	.941	.971	.982	.986	.986	.982	.969	.914	.665
ST	.940	.971	.982	.986	.986	.983	.970	.911	.654
KLD-U	.947	.971	.981	.984	.984	.979	.964	.902	.549
KLD-N	.939	.968	.980	.983	.983	.977	.958	.878	.478
Moderate Discrimination (a = 1.0)									
Z	.749	.815	.853	.871	.864	.848	.804	.723	.560
LR	.742	.819	.858	.876	.870	.855	.807	.714	.544
ST	.743	.822	.861	.880	.875	.860	.811	.712	.538
KLD-U	.748	.811	.847	.863	.855	.837	.786	.678	.398
KLD-N	.714	.794	.835	.851	.840	.815	.750	.616	.346
Low Discrimination (a = 0.6)									
Z	.462	.528	.567	.586	.585	.567	.519	.463	.367
LR	.451	.527	.570	.592	.592	.569	.515	.448	.351
ST	.454	.532	.576	.599	.598	.574	.517	.445	.348
KLD-U	.415	.518	.556	.570	.565	.538	.479	.372	.193
KLD-N	.371	.467	.511	.527	.514	.473	.397	.288	.150
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	1.000	1.000	1.000	1.000	.999	.992	.721
LR	.998	.999	1.000	1.000	1.000	1.000	.999	.991	.805
ST	.997	.999	1.000	1.000	1.000	1.000	.999	.989	.815
KLD-U	.999	1.000	1.000	1.000	1.000	1.000	.999	.987	.672
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.999	.980	.605
Moderate Discrimination (a = 1.0)									
Z	.979	.990	.994	.994	.993	.990	.978	.937	.699
LR	.978	.990	.994	.994	.994	.990	.978	.931	.752
ST	.978	.990	.994	.995	.994	.990	.978	.928	.768
KLD-U	.979	.990	.993	.993	.992	.988	.974	.903	.559
KLD-N	.975	.989	.992	.992	.991	.985	.963	.872	.520
Low Discrimination (a = 0.6)									
Z	.816	.871	.892	.899	.891	.869	.820	.734	.540
LR	.812	.871	.894	.902	.893	.869	.815	.722	.565
ST	.814	.873	.896	.904	.895	.871	.815	.718	.580
KLD-U	.804	.864	.886	.890	.879	.850	.781	.611	.307
KLD-N	.765	.837	.864	.866	.850	.803	.708	.527	.266

Table A70. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.044	.044	.047	.046	.046	.046	.048	.047	.046
LR	.059	.059	.060	.057	.057	.057	.059	.058	.057
ST	.063	.062	.064	.062	.061	.062	.064	.063	.061
KLD-U	.050	.049	.048	.044	.044	.044	.046	.046	.047
KLD-N	.052	.048	.046	.042	.040	.040	.041	.039	.038
Moderate Discrimination (a = 1.0)									
Z	.046	.046	.046	.046	.046	.046	.046	.046	.047
LR	.059	.058	.057	.057	.056	.056	.056	.056	.057
ST	.063	.062	.061	.062	.061	.060	.061	.061	.061
KLD-U	.047	.047	.045	.044	.043	.043	.044	.045	.045
KLD-N	.047	.044	.041	.040	.037	.036	.035	.034	.033
Low Discrimination (a = 0.6)									
Z	.047	.046	.047	.046	.046	.046	.046	.047	.047
LR	.057	.056	.057	.056	.056	.055	.056	.056	.055
ST	.060	.060	.061	.060	.060	.060	.060	.060	.059
KLD-U	.040	.043	.045	.043	.043	.043	.044	.043	.034
KLD-N	.038	.036	.036	.033	.032	.030	.028	.027	.022
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.134	.138	.144	.146	.150	.150	.148	.148	.140
LR	.163	.165	.170	.171	.173	.174	.172	.172	.165
ST	.172	.174	.178	.179	.182	.184	.182	.182	.175
KLD-U	.116	.118	.120	.119	.121	.121	.121	.125	.118
KLD-N	.123	.122	.120	.115	.113	.108	.103	.101	.087
Moderate Discrimination (a = 1.0)									
Z	.093	.093	.096	.099	.099	.097	.096	.098	.096
LR	.112	.112	.113	.116	.116	.114	.114	.114	.112
ST	.118	.118	.120	.123	.123	.122	.120	.121	.119
KLD-U	.075	.077	.079	.080	.078	.076	.078	.081	.076
KLD-N	.080	.076	.076	.073	.067	.062	.059	.056	.048
Low Discrimination (a = 0.6)									
Z	.069	.072	.070	.070	.070	.070	.072	.071	.072
LR	.082	.086	.083	.084	.083	.084	.085	.084	.084
ST	.087	.090	.088	.089	.089	.089	.090	.089	.088
KLD-U	.048	.057	.056	.056	.055	.056	.059	.057	.046
KLD-N	.052	.053	.048	.044	.040	.037	.035	.030	.025
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.418	.431	.453	.466	.471	.470	.466	.457	.433
LR	.468	.478	.496	.507	.512	.510	.505	.497	.474
ST	.480	.490	.508	.519	.526	.524	.520	.512	.491
KLD-U	.376	.384	.398	.408	.413	.416	.415	.410	.386
KLD-N	.394	.393	.400	.400	.394	.386	.375	.355	.316

Continued on the next page

Table A70 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.249	.255	.266	.269	.268	.269	.267	.259	.256
LR	.284	.288	.297	.300	.300	.302	.299	.290	.286
ST	.296	.299	.307	.311	.312	.314	.312	.303	.298
KLD-U	.211	.215	.225	.225	.223	.226	.230	.223	.209
KLD-N	.220	.217	.219	.211	.199	.191	.184	.166	.146
Low Discrimination (a = 0.6)									
Z	.144	.147	.144	.148	.146	.148	.147	.149	.145
LR	.165	.168	.164	.170	.168	.170	.168	.168	.164
ST	.172	.176	.172	.178	.177	.179	.177	.177	.172
KLD-U	.106	.117	.115	.118	.116	.118	.120	.118	.092
KLD-N	.112	.111	.102	.098	.087	.081	.074	.067	.049
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.926	.940	.952	.961	.968	.966	.964	.957	.947
LR	.952	.959	.968	.971	.975	.972	.971	.965	.957
ST	.946	.955	.965	.972	.976	.974	.972	.968	.960
KLD-U	.916	.929	.940	.948	.955	.953	.951	.944	.922
KLD-N	.928	.935	.942	.944	.948	.942	.936	.920	.886
Moderate Discrimination (a = 1.0)									
Z	.744	.760	.774	.779	.780	.782	.775	.767	.742
LR	.778	.790	.800	.806	.808	.809	.801	.793	.772
ST	.784	.796	.807	.813	.816	.818	.809	.802	.782
KLD-U	.695	.711	.725	.730	.732	.739	.735	.725	.636
KLD-N	.709	.715	.717	.708	.696	.688	.667	.637	.541
Low Discrimination (a = 0.6)									
Z	.449	.452	.458	.455	.460	.467	.467	.457	.450
LR	.484	.486	.493	.492	.498	.502	.502	.488	.480
ST	.494	.497	.505	.505	.511	.516	.514	.501	.492
KLD-U	.384	.394	.398	.397	.401	.414	.414	.385	.285
KLD-N	.385	.377	.366	.346	.330	.321	.299	.262	.192
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.980	.986	.991	.999	1.000	1.000	1.000	.999	.999
LR	.998	.999	1.000	1.000	1.000	1.000	1.000	1.000	.999
ST	.943	.947	.964	.990	1.000	1.000	1.000	1.000	.999
KLD-U	.993	.996	.998	.999	1.000	1.000	1.000	.999	.992
KLD-N	.997	.998	.999	.999	.999	.999	.999	.999	.986
Moderate Discrimination (a = 1.0)									
Z	.968	.974	.980	.984	.983	.983	.980	.978	.972
LR	.979	.982	.985	.987	.986	.986	.984	.982	.977
ST	.974	.979	.985	.987	.987	.987	.985	.983	.978
KLD-U	.960	.967	.971	.975	.976	.977	.974	.967	.871
KLD-N	.965	.969	.969	.969	.967	.964	.956	.941	.822
Low Discrimination (a = 0.6)									
Z	.790	.799	.799	.804	.805	.808	.803	.799	.780
LR	.815	.822	.823	.828	.828	.831	.826	.820	.801
ST	.819	.828	.830	.835	.836	.838	.833	.827	.809
KLD-U	.741	.750	.751	.756	.761	.767	.760	.709	.502
KLD-N	.737	.732	.718	.704	.687	.672	.640	.576	.405

Table A71. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.045	.045	.046	.047	.047	.047	.048	.047	.047
LR	.056	.055	.055	.054	.055	.054	.055	.054	.054
ST	.059	.059	.059	.058	.058	.057	.059	.058	.058
KLD-U	.060	.057	.054	.051	.050	.048	.052	.052	.053
KLD-N	.059	.055	.051	.048	.046	.044	.046	.045	.044
Moderate Discrimination (a = 1.0)									
Z	.047	.047	.046	.048	.046	.047	.047	.048	.047
LR	.056	.056	.054	.055	.054	.054	.054	.054	.054
ST	.060	.059	.057	.058	.057	.058	.057	.058	.057
KLD-U	.055	.054	.051	.050	.048	.048	.049	.052	.050
KLD-N	.054	.050	.046	.045	.042	.040	.040	.040	.038
Low Discrimination (a = 0.6)									
Z	.047	.046	.047	.047	.047	.046	.047	.047	.047
LR	.055	.053	.054	.054	.053	.053	.054	.054	.053
ST	.058	.056	.057	.057	.056	.056	.057	.056	.056
KLD-U	.044	.048	.049	.048	.047	.047	.048	.048	.038
KLD-N	.041	.040	.039	.036	.035	.033	.032	.031	.026
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.146	.150	.155	.156	.158	.158	.155	.154	.146
LR	.163	.166	.170	.169	.173	.172	.171	.171	.162
ST	.173	.175	.177	.177	.180	.180	.180	.180	.172
KLD-U	.148	.149	.147	.143	.142	.143	.141	.145	.139
KLD-N	.143	.143	.141	.137	.134	.131	.126	.128	.119
Moderate Discrimination (a = 1.0)									
Z	.098	.100	.101	.104	.103	.101	.099	.100	.098
LR	.110	.110	.112	.115	.114	.113	.111	.112	.110
ST	.117	.117	.118	.120	.120	.119	.117	.118	.117
KLD-U	.095	.095	.095	.095	.092	.090	.089	.093	.089
KLD-N	.090	.089	.087	.086	.081	.076	.072	.072	.067
Low Discrimination (a = 0.6)									
Z	.073	.074	.073	.073	.072	.072	.073	.072	.071
LR	.080	.083	.080	.081	.080	.081	.081	.081	.080
ST	.085	.087	.085	.085	.085	.085	.085	.085	.085
KLD-U	.058	.068	.066	.065	.063	.064	.065	.064	.053
KLD-N	.056	.058	.054	.051	.047	.045	.043	.039	.034
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.456	.465	.485	.487	.492	.487	.481	.472	.448
LR	.482	.490	.507	.510	.515	.511	.507	.499	.479
ST	.496	.503	.518	.522	.527	.524	.520	.513	.494
KLD-U	.453	.455	.465	.461	.463	.458	.458	.454	.434
KLD-N	.445	.447	.457	.450	.447	.438	.431	.421	.396

Continued on the next page

Table A71 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.267	.274	.281	.283	.280	.278	.277	.269	.263
LR	.287	.293	.298	.300	.300	.299	.299	.290	.285
ST	.298	.304	.308	.310	.310	.311	.310	.302	.298
KLD-U	.256	.260	.263	.261	.255	.253	.257	.251	.239
KLD-N	.245	.248	.248	.243	.232	.224	.221	.210	.198
Low Discrimination (a = 0.6)									
Z	.153	.155	.153	.156	.153	.153	.151	.151	.148
LR	.164	.167	.165	.168	.167	.168	.166	.166	.163
ST	.171	.174	.171	.176	.174	.175	.174	.173	.170
KLD-U	.128	.139	.136	.138	.134	.136	.134	.132	.109
KLD-N	.121	.122	.115	.114	.105	.101	.094	.089	.074
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.954	.961	.967	.969	.971	.969	.967	.960	.951
LR	.959	.964	.970	.972	.974	.973	.971	.966	.958
ST	.960	.965	.971	.973	.976	.974	.973	.968	.962
KLD-U	.955	.960	.963	.964	.967	.964	.963	.958	.943
KLD-N	.952	.958	.961	.961	.963	.959	.956	.949	.934
Moderate Discrimination (a = 1.0)									
Z	.772	.782	.790	.792	.790	.790	.786	.777	.758
LR	.788	.796	.804	.807	.806	.808	.802	.795	.780
ST	.795	.803	.810	.814	.814	.816	.811	.804	.790
KLD-U	.761	.767	.771	.770	.767	.769	.768	.760	.699
KLD-N	.750	.754	.756	.751	.741	.736	.727	.711	.660
Low Discrimination (a = 0.6)									
Z	.469	.470	.474	.470	.475	.477	.477	.466	.460
LR	.488	.489	.494	.492	.497	.501	.501	.491	.483
ST	.498	.499	.504	.503	.508	.512	.513	.502	.496
KLD-U	.435	.440	.442	.436	.440	.443	.446	.419	.333
KLD-N	.412	.407	.401	.385	.377	.368	.356	.328	.279
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.999	.999	1.000	1.000	1.000	1.000	1.000	.999	.999
LR	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	.998	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000	.997
KLD-N	.999	.999	1.000	1.000	1.000	1.000	1.000	.999	.998
Moderate Discrimination (a = 1.0)									
Z	.979	.982	.983	.984	.984	.984	.982	.980	.974
LR	.980	.983	.984	.986	.986	.986	.984	.983	.978
ST	.981	.983	.985	.987	.987	.986	.985	.984	.980
KLD-U	.977	.979	.980	.981	.981	.981	.980	.975	.924
KLD-N	.975	.977	.978	.978	.976	.975	.972	.966	.925
Low Discrimination (a = 0.6)									
Z	.805	.812	.810	.812	.811	.816	.810	.806	.792
LR	.817	.825	.823	.826	.827	.831	.825	.822	.809
ST	.822	.830	.829	.832	.834	.838	.832	.830	.817
KLD-U	.784	.790	.786	.786	.787	.793	.786	.747	.583
KLD-N	.762	.762	.752	.742	.731	.725	.708	.674	.556

Table A72. Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.046	.046	.048	.047	.048	.047	.049	.048	.048
LR	.054	.054	.054	.053	.054	.053	.054	.053	.053
ST	.057	.057	.057	.056	.056	.055	.056	.056	.056
KLD-U	.064	.062	.058	.054	.053	.052	.054	.056	.057
KLD-N	.063	.059	.055	.050	.049	.047	.048	.048	.048
Moderate Discrimination (a = 1.0)									
Z	.049	.048	.047	.048	.048	.048	.048	.048	.048
LR	.056	.054	.053	.053	.052	.052	.053	.053	.052
ST	.059	.057	.056	.056	.055	.055	.055	.056	.055
KLD-U	.059	.057	.055	.053	.051	.051	.052	.054	.053
KLD-N	.057	.052	.049	.047	.044	.043	.043	.043	.041
Low Discrimination (a = 0.6)									
Z	.048	.048	.049	.048	.047	.048	.048	.049	.048
LR	.053	.052	.053	.053	.052	.053	.053	.053	.052
ST	.055	.055	.056	.055	.054	.056	.055	.056	.054
KLD-U	.046	.052	.053	.051	.049	.050	.051	.051	.040
KLD-N	.044	.042	.042	.039	.036	.036	.035	.034	.028
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.149	.152	.159	.159	.159	.159	.158	.158	.150
LR	.160	.162	.168	.167	.168	.169	.169	.168	.162
ST	.168	.170	.174	.173	.174	.176	.176	.176	.170
KLD-U	.159	.158	.158	.152	.150	.149	.150	.154	.148
KLD-N	.150	.149	.150	.145	.142	.139	.137	.139	.132
Moderate Discrimination (a = 1.0)									
Z	.102	.102	.104	.106	.106	.102	.101	.104	.100
LR	.109	.110	.110	.112	.112	.111	.108	.111	.108
ST	.115	.115	.115	.117	.116	.115	.114	.116	.114
KLD-U	.102	.103	.102	.101	.099	.095	.095	.100	.094
KLD-N	.094	.094	.092	.090	.087	.082	.080	.081	.077
Low Discrimination (a = 0.6)									
Z	.074	.076	.075	.075	.073	.074	.074	.074	.073
LR	.079	.081	.080	.080	.078	.080	.079	.079	.079
ST	.082	.084	.082	.083	.082	.083	.083	.083	.083
KLD-U	.062	.072	.071	.071	.068	.068	.068	.068	.057
KLD-N	.059	.060	.057	.055	.050	.048	.048	.044	.040
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.459	.470	.489	.493	.494	.488	.484	.476	.451
LR	.476	.486	.503	.506	.508	.504	.501	.493	.472
ST	.488	.497	.513	.515	.518	.514	.511	.505	.485
KLD-U	.470	.472	.480	.476	.476	.469	.470	.467	.446
KLD-N	.455	.460	.469	.464	.461	.449	.447	.441	.417

Continued on the next page

Table A72 (continued). Mean Type I Error Rates and Power for 30-Item Fixed-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (*a* = 1.0)									
Z	.270	.278	.284	.286	.285	.283	.281	.275	.268
LR	.283	.290	.293	.296	.297	.296	.295	.289	.283
ST	.292	.298	.301	.304	.305	.306	.303	.298	.294
KLD-U	.268	.273	.273	.271	.267	.266	.268	.264	.250
KLD-N	.252	.256	.256	.252	.244	.238	.236	.228	.217
Low Discrimination (*a* = 0.6)									
Z	.154	.158	.157	.158	.157	.158	.156	.155	.152
LR	.162	.165	.163	.165	.165	.167	.166	.164	.161
ST	.168	.170	.168	.170	.170	.173	.172	.171	.167
KLD-U	.134	.147	.146	.146	.143	.144	.142	.140	.117
KLD-N	.124	.126	.121	.119	.112	.108	.103	.099	.087
Power: Medium Change, Δθ = 1.0									
High Discrimination (*a* = 1.5)									
Z	.953	.960	.966	.968	.969	.968	.966	.959	.950
LR	.956	.962	.967	.970	.972	.970	.969	.963	.955
ST	.958	.962	.968	.971	.973	.972	.970	.965	.958
KLD-U	.957	.961	.964	.965	.966	.965	.964	.958	.945
KLD-N	.953	.958	.962	.962	.963	.961	.959	.951	.939
Moderate Discrimination (*a* = 1.0)									
Z	.772	.782	.791	.792	.791	.790	.787	.778	.758
LR	.782	.791	.798	.800	.800	.801	.798	.790	.772
ST	.788	.796	.802	.806	.806	.808	.805	.797	.782
KLD-U	.768	.775	.779	.777	.774	.774	.776	.767	.707
KLD-N	.754	.760	.763	.758	.749	.746	.742	.728	.685
Low Discrimination (*a* = 0.6)									
Z	.472	.474	.479	.476	.477	.480	.480	.469	.463
LR	.484	.484	.489	.489	.492	.495	.496	.484	.478
ST	.492	.491	.497	.496	.500	.504	.505	.494	.488
KLD-U	.447	.454	.456	.452	.451	.456	.456	.431	.349
KLD-N	.418	.415	.412	.400	.391	.385	.378	.354	.312
Power: Large Change, Δθ = 1.5									
High Discrimination (*a* = 1.5)									
Z	.999	.999	.999	1.000	1.000	1.000	1.000	1.000	.999
LR	.998	.999	.999	1.000	1.000	1.000	1.000	1.000	.999
ST	.998	.999	.999	1.000	1.000	1.000	1.000	1.000	.999
KLD-U	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000	.998
KLD-N	.999	.999	.999	1.000	1.000	1.000	1.000	.999	.998
Moderate Discrimination (*a* = 1.0)									
Z	.978	.980	.982	.984	.982	.983	.980	.979	.973
LR	.979	.981	.983	.985	.984	.985	.982	.980	.976
ST	.979	.982	.983	.985	.984	.986	.983	.982	.978
KLD-U	.977	.980	.980	.982	.980	.981	.979	.975	.932
KLD-N	.974	.977	.977	.978	.976	.976	.973	.968	.937
Low Discrimination (*a* = 0.6)									
Z	.805	.812	.810	.812	.811	.816	.810	.806	.791
LR	.811	.819	.817	.821	.821	.825	.821	.817	.803
ST	.815	.822	.822	.826	.826	.830	.827	.823	.810
KLD-U	.790	.798	.793	.794	.792	.799	.791	.755	.606
KLD-N	.765	.766	.757	.751	.742	.740	.726	.700	.604

Table A73. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.077	.054	.048	.048	.047	.049	.048	.052	.064
LR	.054	.055	.054	.054	.053	.055	.053	.053	.055
ST	.052	.055	.056	.057	.056	.058	.056	.054	.051
KLD-U	.046	.058	.047	.045	.045	.048	.047	.052	.062
KLD-N	.039	.052	.046	.044	.044	.045	.044	.047	.048
Moderate Discrimination (a = 1.0)									
Z	.064	.053	.048	.048	.047	.048	.048	.052	.058
LR	.054	.054	.053	.054	.053	.054	.052	.053	.053
ST	.054	.055	.055	.056	.056	.056	.054	.054	.051
KLD-U	.043	.053	.047	.046	.046	.046	.047	.052	.054
KLD-N	.037	.045	.043	.043	.042	.042	.041	.043	.041
Low Discrimination (a = 0.6)									
Z	.059	.052	.050	.050	.050	.051	.051	.054	.058
LR	.056	.053	.053	.054	.054	.054	.053	.053	.054
ST	.058	.055	.055	.056	.056	.056	.054	.053	.053
KLD-U	.033	.048	.049	.049	.048	.049	.050	.053	.043
KLD-N	.028	.036	.040	.042	.041	.041	.039	.038	.030
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.165	.218	.243	.260	.266	.267	.243	.226	.182
LR	.146	.223	.260	.279	.284	.285	.256	.226	.161
ST	.144	.226	.267	.286	.292	.292	.263	.226	.153
KLD-U	.140	.203	.217	.230	.236	.238	.218	.204	.140
KLD-N	.122	.196	.216	.228	.230	.228	.205	.188	.113
Moderate Discrimination (a = 1.0)									
Z	.123	.136	.150	.156	.164	.158	.154	.142	.130
LR	.114	.140	.160	.169	.176	.169	.161	.143	.119
ST	.114	.142	.165	.174	.181	.174	.165	.144	.116
KLD-U	.096	.125	.132	.138	.143	.138	.134	.125	.099
KLD-N	.084	.115	.128	.132	.134	.126	.119	.105	.076
Low Discrimination (a = 0.6)									
Z	.087	.092	.098	.102	.104	.104	.101	.096	.094
LR	.084	.094	.103	.108	.111	.110	.104	.095	.088
ST	.086	.096	.105	.111	.114	.112	.106	.096	.087
KLD-U	.052	.082	.089	.090	.092	.091	.089	.082	.062
KLD-N	.048	.068	.078	.080	.080	.077	.070	.060	.043
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.507	.669	.731	.764	.772	.760	.718	.626	.446
LR	.494	.681	.749	.780	.788	.774	.729	.619	.408
ST	.492	.685	.756	.786	.795	.780	.734	.617	.391
KLD-U	.496	.641	.699	.730	.741	.729	.688	.593	.357
KLD-N	.463	.638	.698	.726	.734	.716	.671	.567	.301

Continued on the next page

Table A73 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.328	.409	.460	.488	.498	.488	.452	.397	.316
LR	.320	.419	.478	.508	.517	.504	.462	.394	.295
ST	.321	.424	.486	.517	.525	.511	.467	.394	.287
KLD-U	.301	.382	.427	.453	.461	.451	.418	.363	.243
KLD-N	.274	.370	.420	.443	.445	.429	.388	.324	.199
Low Discrimination (a = 0.6)									
Z	.184	.223	.250	.268	.271	.270	.253	.228	.198
LR	.183	.228	.259	.279	.282	.279	.257	.225	.187
ST	.186	.232	.264	.284	.288	.284	.259	.225	.184
KLD-U	.137	.204	.230	.244	.244	.243	.228	.196	.129
KLD-N	.126	.180	.209	.225	.220	.214	.190	.152	.095
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.989	.996	.999	.999	1.000	.999	.998	.981	.814
LR	.991	.998	.999	1.000	1.000	.999	.998	.979	.798
ST	.990	.998	.999	1.000	1.000	.999	.998	.978	.784
KLD-U	.988	.997	.999	.999	.999	.999	.997	.976	.706
KLD-N	.988	.997	.999	.999	.999	.999	.997	.968	.646
Moderate Discrimination (a = 1.0)									
Z	.895	.949	.968	.973	.972	.964	.941	.879	.736
LR	.897	.952	.971	.976	.975	.966	.943	.873	.708
ST	.897	.954	.972	.977	.975	.967	.943	.871	.695
KLD-U	.884	.938	.959	.966	.965	.957	.931	.854	.589
KLD-N	.873	.936	.958	.964	.961	.950	.918	.822	.537
Low Discrimination (a = 0.6)									
Z	.605	.696	.752	.776	.777	.750	.704	.620	.508
LR	.609	.705	.762	.787	.786	.756	.704	.613	.488
ST	.613	.709	.766	.791	.789	.759	.706	.611	.481
KLD-U	.567	.673	.727	.751	.750	.723	.672	.561	.328
KLD-N	.535	.647	.707	.729	.722	.684	.616	.486	.278
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.984	.995	1.000	1.000	1.000	1.000	1.000	.999	.890
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.921
ST	.992	.992	.998	1.000	1.000	1.000	1.000	.999	.922
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.831
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.786
Moderate Discrimination (a = 1.0)									
Z	.998	.999	1.000	1.000	1.000	1.000	.999	.989	.906
LR	.999	1.000	1.000	1.000	1.000	1.000	.999	.987	.906
ST	.999	1.000	1.000	1.000	1.000	1.000	.998	.986	.901
KLD-U	.998	1.000	1.000	1.000	1.000	1.000	.998	.982	.771
KLD-N	.998	1.000	1.000	1.000	1.000	1.000	.997	.974	.744
Low Discrimination (a = 0.6)									
Z	.939	.969	.980	.981	.980	.971	.944	.885	.776
LR	.940	.970	.982	.982	.981	.972	.943	.878	.763
ST	.941	.971	.982	.983	.982	.972	.943	.876	.757
KLD-U	.929	.963	.975	.977	.974	.964	.931	.825	.521
KLD-N	.918	.958	.971	.972	.968	.953	.904	.776	.488

Table A74. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.076	.055	.048	.048	.046	.050	.050	.054	.064
LR	.053	.054	.052	.052	.051	.054	.053	.054	.055
ST	.051	.055	.054	.055	.053	.056	.054	.054	.052
KLD-U	.048	.060	.052	.050	.048	.051	.052	.056	.064
KLD-N	.040	.054	.050	.049	.046	.049	.048	.049	.049
Moderate Discrimination (a = 1.0)									
Z	.063	.053	.048	.047	.048	.048	.048	.053	.058
LR	.053	.053	.052	.051	.052	.051	.050	.053	.052
ST	.053	.053	.053	.053	.054	.053	.051	.053	.050
KLD-U	.045	.056	.051	.049	.049	.049	.049	.054	.055
KLD-N	.038	.047	.047	.046	.045	.045	.043	.045	.041
Low Discrimination (a = 0.6)									
Z	.059	.052	.050	.051	.050	.051	.050	.053	.058
LR	.054	.052	.052	.053	.053	.054	.051	.052	.054
ST	.056	.053	.053	.054	.055	.055	.052	.052	.052
KLD-U	.036	.050	.052	.051	.051	.052	.051	.054	.044
KLD-N	.030	.038	.042	.044	.044	.043	.041	.039	.031
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.166	.224	.253	.270	.274	.272	.250	.228	.182
LR	.146	.224	.262	.282	.285	.282	.260	.226	.162
ST	.143	.226	.269	.288	.292	.289	.265	.227	.154
KLD-U	.144	.222	.243	.258	.260	.256	.235	.209	.142
KLD-N	.123	.210	.238	.253	.254	.248	.224	.194	.115
Moderate Discrimination (a = 1.0)									
Z	.124	.139	.154	.162	.167	.161	.157	.143	.129
LR	.113	.138	.159	.169	.174	.168	.162	.143	.119
ST	.112	.140	.163	.173	.178	.172	.165	.144	.116
KLD-U	.102	.139	.147	.153	.155	.148	.144	.129	.102
KLD-N	.087	.123	.139	.146	.147	.138	.130	.111	.079
Low Discrimination (a = 0.6)									
Z	.088	.094	.100	.104	.105	.105	.102	.097	.093
LR	.082	.092	.102	.108	.109	.109	.104	.095	.087
ST	.084	.094	.104	.110	.111	.111	.105	.095	.086
KLD-U	.056	.089	.096	.099	.097	.096	.092	.084	.064
KLD-N	.050	.071	.082	.088	.085	.083	.074	.062	.045
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.514	.684	.746	.775	.781	.766	.723	.628	.447
LR	.496	.688	.756	.785	.791	.776	.731	.621	.409
ST	.494	.691	.761	.790	.797	.781	.735	.620	.393
KLD-U	.510	.676	.735	.762	.767	.750	.705	.600	.361
KLD-N	.473	.668	.731	.757	.761	.741	.692	.575	.307

Continued on the next page

Table A74 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.334	.420	.471	.499	.506	.494	.456	.399	.317
LR	.321	.422	.481	.511	.519	.506	.463	.396	.296
ST	.321	.426	.487	.518	.525	.512	.468	.396	.288
KLD-U	.319	.414	.458	.482	.487	.471	.432	.371	.249
KLD-N	.286	.393	.446	.471	.473	.452	.406	.334	.207
Low Discrimination (a = 0.6)									
Z	.189	.227	.254	.272	.277	.274	.255	.229	.198
LR	.183	.227	.258	.279	.284	.279	.256	.224	.187
ST	.186	.230	.261	.283	.288	.283	.258	.224	.184
KLD-U	.150	.219	.244	.258	.259	.254	.233	.200	.133
KLD-N	.133	.190	.221	.238	.236	.226	.199	.157	.102
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.992	.999	.999	1.000	1.000	.999	.998	.982	.815
LR	.992	.999	1.000	1.000	1.000	.999	.998	.979	.800
ST	.992	.999	1.000	1.000	1.000	.999	.998	.978	.786
KLD-U	.992	.999	.999	1.000	1.000	.999	.997	.976	.710
KLD-N	.991	.998	.999	1.000	.999	.999	.997	.970	.652
Moderate Discrimination (a = 1.0)									
Z	.901	.952	.969	.975	.974	.965	.942	.880	.737
LR	.899	.954	.971	.976	.975	.967	.943	.874	.710
ST	.899	.955	.972	.977	.976	.967	.944	.873	.698
KLD-U	.899	.949	.966	.972	.970	.960	.934	.859	.598
KLD-N	.885	.946	.964	.970	.967	.956	.924	.829	.551
Low Discrimination (a = 0.6)									
Z	.613	.704	.758	.781	.780	.752	.707	.621	.510
LR	.610	.706	.764	.787	.786	.755	.706	.613	.491
ST	.613	.710	.767	.790	.789	.758	.707	.611	.485
KLD-U	.591	.695	.746	.766	.763	.731	.679	.567	.339
KLD-N	.548	.663	.724	.745	.738	.695	.627	.498	.297
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.891
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.922
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.923
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.834
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.792
Moderate Discrimination (a = 1.0)									
Z	.999	1.000	1.000	1.000	1.000	1.000	.999	.989	.908
LR	.999	1.000	1.000	1.000	1.000	1.000	.999	.988	.908
ST	.999	1.000	1.000	1.000	1.000	1.000	.999	.987	.903
KLD-U	.999	1.000	1.000	1.000	1.000	1.000	.998	.982	.780
KLD-N	.999	1.000	1.000	1.000	1.000	1.000	.998	.976	.759
Low Discrimination (a = 0.6)									
Z	.942	.970	.980	.982	.980	.970	.944	.886	.778
LR	.942	.970	.981	.982	.980	.971	.943	.879	.766
ST	.943	.971	.981	.983	.981	.971	.943	.877	.760
KLD-U	.939	.967	.978	.979	.976	.966	.933	.832	.539
KLD-N	.925	.962	.974	.975	.971	.955	.909	.789	.517

Table A75. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.077	.055	.050	.048	.046	.050	.050	.054	.064
LR	.053	.053	.052	.051	.049	.053	.051	.054	.054
ST	.051	.054	.054	.053	.051	.054	.053	.054	.051
KLD-U	.048	.061	.054	.052	.050	.052	.052	.057	.063
KLD-N	.040	.054	.052	.050	.048	.050	.049	.051	.049
Moderate Discrimination (a = 1.0)									
Z	.063	.054	.048	.048	.049	.049	.049	.052	.058
LR	.053	.052	.050	.051	.051	.051	.050	.052	.052
ST	.052	.053	.052	.052	.053	.052	.051	.052	.050
KLD-U	.045	.058	.052	.052	.051	.051	.051	.055	.055
KLD-N	.038	.048	.048	.048	.048	.047	.045	.046	.041
Low Discrimination (a = 0.6)									
Z	.060	.053	.052	.052	.050	.051	.051	.053	.057
LR	.054	.052	.052	.053	.052	.052	.051	.052	.052
ST	.055	.052	.053	.054	.053	.053	.051	.051	.051
KLD-U	.037	.052	.054	.054	.052	.053	.053	.055	.044
KLD-N	.031	.040	.044	.046	.045	.044	.042	.040	.032
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.166	.225	.257	.275	.277	.275	.253	.228	.182
LR	.146	.222	.262	.281	.284	.282	.258	.226	.162
ST	.144	.223	.267	.286	.289	.287	.263	.226	.154
KLD-U	.144	.226	.252	.266	.267	.263	.240	.211	.143
KLD-N	.123	.214	.246	.261	.261	.256	.231	.196	.115
Moderate Discrimination (a = 1.0)									
Z	.124	.140	.156	.163	.169	.164	.157	.144	.130
LR	.112	.137	.158	.166	.174	.167	.160	.142	.119
ST	.111	.138	.161	.170	.177	.170	.162	.142	.115
KLD-U	.103	.142	.153	.157	.161	.154	.146	.131	.102
KLD-N	.088	.126	.144	.150	.153	.144	.133	.114	.080
Low Discrimination (a = 0.6)									
Z	.089	.095	.101	.106	.105	.107	.103	.097	.093
LR	.083	.093	.100	.107	.107	.108	.103	.095	.086
ST	.084	.094	.102	.109	.109	.110	.104	.094	.084
KLD-U	.059	.092	.100	.102	.100	.100	.094	.086	.064
KLD-N	.052	.074	.085	.090	.088	.086	.077	.064	.047
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.515	.685	.748	.776	.782	.768	.722	.628	.447
LR	.497	.685	.754	.782	.788	.774	.727	.620	.409
ST	.495	.687	.758	.786	.792	.778	.731	.618	.393
KLD-U	.512	.682	.742	.768	.772	.756	.708	.601	.362
KLD-N	.475	.672	.737	.763	.766	.748	.696	.577	.308

Continued on the next page

Table A75 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.334	.420	.475	.502	.507	.496	.458	.399	.318
LR	.320	.418	.479	.508	.515	.502	.462	.394	.297
ST	.319	.421	.484	.514	.520	.506	.466	.394	.289
KLD-U	.323	.419	.467	.490	.493	.478	.438	.371	.250
KLD-N	.288	.397	.452	.478	.479	.460	.413	.336	.209
Low Discrimination (a = 0.6)									
Z	.190	.229	.256	.276	.278	.276	.257	.228	.198
LR	.182	.226	.256	.278	.281	.278	.256	.222	.187
ST	.183	.228	.259	.281	.284	.281	.257	.222	.183
KLD-U	.154	.226	.250	.266	.264	.259	.236	.201	.135
KLD-N	.134	.194	.225	.243	.241	.231	.202	.159	.105
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.992	.998	.999	1.000	1.000	.999	.998	.982	.815
LR	.992	.998	.999	1.000	1.000	.999	.998	.979	.800
ST	.991	.998	.999	1.000	1.000	.999	.998	.978	.786
KLD-U	.992	.998	.999	1.000	1.000	.999	.997	.976	.710
KLD-N	.991	.998	.999	1.000	.999	.999	.997	.970	.652
Moderate Discrimination (a = 1.0)									
Z	.901	.952	.968	.974	.973	.964	.941	.880	.737
LR	.898	.952	.969	.975	.974	.965	.941	.874	.710
ST	.898	.953	.969	.976	.974	.966	.942	.872	.698
KLD-U	.902	.951	.966	.972	.970	.960	.934	.859	.600
KLD-N	.888	.946	.963	.970	.968	.956	.924	.830	.554
Low Discrimination (a = 0.6)									
Z	.614	.704	.759	.781	.781	.751	.707	.622	.511
LR	.607	.703	.760	.783	.783	.752	.704	.612	.490
ST	.608	.705	.763	.786	.786	.754	.705	.609	.482
KLD-U	.597	.700	.751	.769	.767	.733	.682	.570	.342
KLD-N	.552	.666	.728	.748	.742	.699	.631	.502	.303
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.891
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.922
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.923
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.834
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.792
Moderate Discrimination (a = 1.0)									
Z	.999	1.000	1.000	1.000	1.000	1.000	.999	.989	.907
LR	.999	1.000	1.000	1.000	1.000	1.000	.998	.987	.908
ST	.999	1.000	1.000	1.000	1.000	1.000	.998	.987	.903
KLD-U	.999	1.000	1.000	1.000	1.000	1.000	.998	.983	.782
KLD-N	.999	1.000	1.000	1.000	1.000	1.000	.998	.976	.761
Low Discrimination (a = 0.6)									
Z	.941	.968	.979	.981	.979	.970	.945	.886	.778
LR	.939	.969	.980	.981	.979	.970	.942	.879	.766
ST	.940	.969	.980	.982	.980	.971	.942	.877	.760
KLD-U	.940	.967	.978	.979	.976	.966	.934	.833	.543
KLD-N	.926	.961	.974	.976	.971	.956	.910	.792	.527

Table A76. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.047	.047	.047	.047	.048	.048	.047	.048	.049
LR	.054	.053	.053	.053	.053	.053	.053	.054	.055
ST	.056	.056	.056	.055	.056	.056	.056	.056	.058
KLD-U	.048	.047	.046	.044	.045	.045	.045	.046	.049
KLD-N	.048	.047	.045	.043	.043	.042	.042	.042	.043
Moderate Discrimination (a = 1.0)									
Z	.047	.050	.048	.047	.047	.048	.048	.048	.048
LR	.054	.056	.054	.053	.053	.054	.054	.054	.053
ST	.056	.059	.056	.056	.056	.056	.056	.057	.056
KLD-U	.047	.048	.046	.046	.044	.045	.046	.046	.046
KLD-N	.045	.046	.043	.042	.040	.040	.040	.040	.038
Low Discrimination (a = 0.6)									
Z	.047	.048	.048	.047	.048	.048	.049	.047	.048
LR	.053	.053	.053	.053	.053	.054	.056	.052	.054
ST	.055	.055	.056	.056	.056	.056	.058	.054	.056
KLD-U	.043	.046	.045	.045	.045	.045	.047	.045	.042
KLD-N	.040	.040	.039	.037	.037	.036	.036	.033	.031
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.211	.215	.217	.216	.220	.221	.220	.211	.214
LR	.228	.231	.232	.232	.236	.236	.236	.227	.231
ST	.234	.238	.238	.237	.242	.242	.242	.234	.239
KLD-U	.189	.191	.191	.189	.192	.192	.194	.188	.193
KLD-N	.192	.193	.190	.186	.186	.183	.180	.171	.173
Moderate Discrimination (a = 1.0)									
Z	.129	.132	.132	.135	.135	.135	.130	.135	.131
LR	.141	.144	.143	.147	.147	.146	.142	.146	.142
ST	.146	.149	.147	.152	.151	.152	.147	.152	.147
KLD-U	.112	.115	.114	.116	.116	.116	.113	.117	.114
KLD-N	.114	.114	.110	.110	.107	.104	.097	.098	.094
Low Discrimination (a = 0.6)									
Z	.087	.085	.089	.088	.087	.087	.086	.087	.088
LR	.095	.093	.097	.095	.095	.096	.095	.095	.096
ST	.098	.096	.100	.099	.099	.099	.098	.098	.100
KLD-U	.071	.073	.076	.074	.074	.074	.074	.075	.071
KLD-N	.070	.068	.068	.064	.061	.058	.055	.053	.049
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.643	.653	.655	.668	.669	.670	.669	.659	.650
LR	.663	.672	.674	.686	.687	.690	.688	.679	.671
ST	.670	.679	.681	.692	.693	.696	.695	.686	.678
KLD-U	.610	.619	.618	.629	.630	.632	.635	.627	.620
KLD-N	.616	.622	.617	.624	.619	.616	.613	.601	.589

Continued on the next page

Table A76 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.387	.391	.397	.404	.397	.406	.398	.399	.391
LR	.408	.412	.417	.424	.417	.426	.419	.420	.411
ST	.416	.419	.424	.431	.425	.434	.427	.428	.419
KLD-U	.353	.358	.361	.366	.359	.369	.364	.367	.356
KLD-N	.356	.357	.354	.354	.341	.343	.332	.328	.309
Low Discrimination (a = 0.6)									
Z	.208	.207	.211	.212	.210	.208	.210	.209	.207
LR	.222	.222	.225	.227	.225	.225	.225	.223	.221
ST	.227	.227	.231	.233	.231	.231	.231	.229	.227
KLD-U	.179	.182	.185	.185	.183	.181	.185	.185	.167
KLD-N	.176	.172	.172	.165	.157	.150	.146	.140	.127
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.996	.996	.997	.998	.998	.998	.997	.997	.996
LR	.997	.997	.997	.998	.998	.998	.998	.997	.997
ST	.996	.997	.997	.998	.998	.998	.998	.998	.997
KLD-U	.994	.995	.996	.996	.996	.997	.996	.996	.995
KLD-N	.995	.995	.996	.996	.996	.996	.996	.995	.993
Moderate Discrimination (a = 1.0)									
Z	.923	.926	.928	.929	.930	.933	.929	.926	.920
LR	.930	.932	.934	.936	.937	.939	.936	.933	.928
ST	.932	.934	.936	.938	.939	.941	.938	.935	.930
KLD-U	.907	.910	.912	.913	.915	.919	.916	.912	.892
KLD-N	.908	.909	.909	.907	.905	.906	.899	.889	.867
Low Discrimination (a = 0.6)									
Z	.641	.642	.646	.646	.641	.646	.647	.641	.637
LR	.658	.659	.663	.664	.661	.665	.665	.659	.655
ST	.664	.665	.669	.670	.668	.672	.671	.666	.662
KLD-U	.604	.605	.606	.606	.602	.608	.611	.601	.524
KLD-N	.594	.588	.581	.572	.556	.550	.542	.521	.467
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	.998	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.999	.999	.999	.999	.999	.999	.999	.999	.999
LR	.999	.999	.999	.999	.999	.999	.999	.999	.999
ST	.999	.999	.999	.999	.999	.999	.999	.999	.999
KLD-U	.998	.998	.999	.999	.999	.999	.999	.998	.992
KLD-N	.998	.998	.999	.999	.998	.998	.998	.998	.992
Low Discrimination (a = 0.6)									
Z	.935	.938	.937	.939	.938	.940	.937	.936	.932
LR	.940	.943	.942	.945	.944	.945	.943	.942	.938
ST	.942	.944	.944	.947	.945	.947	.945	.944	.939
KLD-U	.920	.924	.923	.924	.923	.926	.925	.913	.808
KLD-N	.916	.917	.912	.908	.902	.900	.894	.878	.796

Table A77. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.048	.048	.047	.048	.048	.047	.047	.048	.050
LR	.053	.052	.051	.052	.051	.051	.051	.052	.054
ST	.054	.054	.053	.054	.053	.053	.053	.054	.056
KLD-U	.054	.053	.050	.049	.048	.047	.048	.050	.053
KLD-N	.053	.052	.048	.047	.046	.045	.044	.046	.048
Moderate Discrimination (a = 1.0)									
Z	.048	.050	.048	.048	.047	.048	.048	.048	.048
LR	.052	.054	.052	.052	.051	.052	.052	.052	.052
ST	.054	.056	.054	.053	.054	.054	.054	.055	.054
KLD-U	.051	.052	.049	.049	.048	.048	.048	.050	.049
KLD-N	.048	.049	.046	.045	.043	.043	.043	.043	.042
Low Discrimination (a = 0.6)									
Z	.048	.048	.048	.048	.047	.048	.049	.047	.048
LR	.052	.052	.052	.051	.051	.052	.053	.052	.051
ST	.054	.053	.054	.054	.053	.054	.056	.054	.053
KLD-U	.047	.049	.048	.048	.047	.048	.049	.048	.045
KLD-N	.043	.042	.042	.040	.038	.038	.038	.036	.034
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.220	.223	.223	.222	.225	.225	.224	.216	.219
LR	.230	.234	.233	.232	.234	.235	.235	.227	.230
ST	.237	.240	.239	.236	.240	.240	.240	.233	.236
KLD-U	.213	.214	.211	.206	.208	.206	.207	.203	.209
KLD-N	.210	.211	.208	.201	.202	.197	.196	.191	.195
Moderate Discrimination (a = 1.0)									
Z	.133	.136	.135	.138	.138	.138	.133	.136	.135
LR	.140	.144	.142	.145	.145	.146	.141	.145	.143
ST	.145	.148	.146	.149	.149	.150	.145	.149	.147
KLD-U	.126	.128	.125	.127	.126	.125	.121	.126	.125
KLD-N	.121	.123	.120	.120	.117	.114	.109	.111	.108
Low Discrimination (a = 0.6)									
Z	.088	.086	.090	.088	.089	.087	.087	.088	.088
LR	.094	.091	.095	.094	.094	.093	.094	.093	.094
ST	.096	.094	.098	.096	.097	.096	.097	.096	.097
KLD-U	.079	.080	.083	.080	.080	.079	.079	.080	.076
KLD-N	.074	.072	.073	.069	.066	.064	.061	.060	.058
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.658	.668	.668	.678	.676	.676	.677	.667	.661
LR	.671	.680	.680	.689	.687	.688	.690	.680	.675
ST	.678	.686	.685	.694	.692	.694	.696	.687	.682
KLD-U	.647	.654	.651	.656	.652	.653	.656	.650	.648
KLD-N	.644	.651	.646	.649	.642	.641	.640	.632	.629

Continued on the next page

Table A77 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.398	.402	.406	.411	.404	.414	.403	.407	.399
LR	.412	.415	.418	.423	.417	.428	.417	.421	.413
ST	.420	.422	.424	.430	.424	.434	.424	.428	.420
KLD-U	.384	.386	.387	.389	.380	.389	.381	.387	.379
KLD-N	.376	.377	.376	.375	.363	.366	.356	.355	.347
Low Discrimination (a = 0.6)									
Z	.214	.213	.215	.217	.212	.213	.214	.213	.210
LR	.223	.221	.224	.226	.222	.224	.224	.223	.221
ST	.228	.226	.229	.231	.227	.229	.229	.228	.227
KLD-U	.196	.198	.200	.198	.193	.194	.197	.197	.180
KLD-N	.185	.182	.181	.177	.168	.164	.163	.159	.147
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.997	.997	.997	.998	.998	.998	.997	.997	.996
LR	.997	.997	.997	.998	.998	.998	.998	.997	.997
ST	.997	.997	.997	.998	.998	.998	.998	.998	.997
KLD-U	.996	.997	.997	.997	.997	.997	.997	.997	.996
KLD-N	.996	.997	.997	.997	.997	.997	.997	.996	.995
Moderate Discrimination (a = 1.0)									
Z	.928	.930	.931	.931	.932	.934	.932	.929	.925
LR	.932	.934	.934	.935	.936	.938	.936	.933	.930
ST	.934	.936	.936	.937	.938	.940	.938	.935	.932
KLD-U	.922	.923	.923	.923	.923	.926	.924	.922	.908
KLD-N	.919	.920	.918	.917	.915	.917	.912	.908	.897
Low Discrimination (a = 0.6)									
Z	.651	.649	.652	.651	.647	.651	.651	.646	.645
LR	.662	.659	.663	.663	.660	.664	.664	.660	.658
ST	.668	.665	.669	.669	.666	.671	.669	.665	.664
KLD-U	.632	.628	.628	.626	.621	.626	.628	.619	.556
KLD-N	.612	.605	.602	.593	.581	.579	.574	.561	.524
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.999	.999	.999	.999	.999	.999	.999	.999	.999
LR	.999	.999	.999	.999	.999	.999	.999	.999	.999
ST	.999	.999	.999	.999	.999	.999	.999	.999	.999
KLD-U	.999	.999	.999	.999	.999	.999	.999	.999	.996
KLD-N	.999	.999	.999	.999	.999	.999	.998	.998	.996
Low Discrimination (a = 0.6)									
Z	.938	.940	.938	.941	.938	.940	.939	.938	.935
LR	.941	.943	.942	.944	.942	.944	.943	.943	.939
ST	.942	.944	.943	.946	.944	.946	.944	.944	.941
KLD-U	.931	.933	.930	.932	.930	.932	.931	.923	.840
KLD-N	.923	.924	.920	.918	.914	.913	.908	.901	.848

Table A78. Mean Type I Error Rates and Power for 50-Item Fixed-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.048	.049	.047	.049	.049	.049	.048	.049	.050
LR	.051	.053	.051	.052	.051	.052	.051	.052	.052
ST	.053	.054	.052	.052	.053	.053	.053	.053	.054
KLD-U	.055	.055	.052	.051	.050	.050	.050	.052	.054
KLD-N	.054	.054	.050	.048	.048	.047	.047	.047	.049
Moderate Discrimination (a = 1.0)									
Z	.048	.050	.048	.049	.048	.049	.049	.049	.048
LR	.051	.053	.051	.052	.051	.052	.051	.052	.051
ST	.052	.054	.052	.053	.052	.053	.053	.054	.052
KLD-U	.053	.054	.052	.052	.049	.050	.050	.052	.051
KLD-N	.051	.051	.048	.047	.045	.045	.045	.045	.043
Low Discrimination (a = 0.6)									
Z	.049	.048	.049	.049	.047	.049	.050	.049	.048
LR	.052	.051	.051	.051	.050	.052	.052	.051	.051
ST	.053	.052	.053	.053	.051	.053	.054	.053	.052
KLD-U	.049	.050	.051	.050	.048	.050	.051	.050	.046
KLD-N	.045	.044	.043	.042	.039	.040	.040	.038	.035
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.222	.225	.226	.224	.227	.226	.227	.218	.220
LR	.228	.232	.232	.230	.233	.232	.234	.225	.228
ST	.233	.237	.235	.233	.236	.237	.238	.230	.232
KLD-U	.220	.220	.217	.212	.214	.212	.214	.209	.214
KLD-N	.215	.216	.212	.207	.207	.203	.204	.197	.202
Moderate Discrimination (a = 1.0)									
Z	.135	.138	.137	.139	.139	.140	.134	.138	.135
LR	.140	.142	.140	.143	.144	.145	.139	.144	.140
ST	.143	.146	.143	.146	.147	.148	.142	.147	.144
KLD-U	.131	.133	.130	.130	.130	.129	.125	.130	.129
KLD-N	.125	.126	.123	.122	.121	.118	.113	.116	.114
Low Discrimination (a = 0.6)									
Z	.090	.088	.091	.090	.090	.088	.089	.089	.090
LR	.093	.090	.094	.093	.093	.092	.093	.093	.094
ST	.095	.092	.096	.094	.095	.094	.095	.095	.096
KLD-U	.082	.083	.086	.083	.082	.082	.083	.083	.079
KLD-N	.076	.074	.074	.071	.069	.067	.065	.065	.062
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.658	.668	.669	.678	.675	.678	.676	.667	.660
LR	.666	.675	.675	.684	.682	.685	.684	.675	.669
ST	.672	.680	.679	.688	.686	.690	.689	.680	.674
KLD-U	.653	.660	.656	.660	.657	.660	.661	.655	.653
KLD-N	.648	.655	.651	.654	.648	.649	.647	.640	.638

Continued on the next page

Table A78 (continued). Mean Type I Error Rates and Power for 50-Item Fixed-Length With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.401	.406	.410	.414	.407	.414	.407	.407	.401
LR	.410	.414	.417	.421	.414	.423	.415	.416	.409
ST	.415	.419	.421	.426	.419	.428	.420	.422	.415
KLD-U	.392	.395	.396	.397	.388	.395	.390	.393	.386
KLD-N	.382	.384	.384	.383	.370	.374	.366	.366	.359
Low Discrimination (a = 0.6)									
Z	.216	.215	.218	.219	.215	.217	.217	.215	.212
LR	.221	.220	.222	.224	.221	.223	.224	.221	.219
ST	.224	.223	.225	.228	.225	.228	.228	.225	.223
KLD-U	.202	.204	.205	.205	.200	.201	.203	.202	.185
KLD-N	.189	.186	.185	.182	.174	.172	.170	.166	.158
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.996	.997	.997	.997	.998	.998	.997	.997	.996
LR	.996	.997	.997	.998	.998	.998	.997	.997	.996
ST	.996	.997	.997	.998	.998	.998	.997	.997	.997
KLD-U	.996	.997	.997	.997	.997	.997	.997	.997	.996
KLD-N	.996	.996	.996	.997	.997	.997	.996	.996	.995
Moderate Discrimination (a = 1.0)									
Z	.927	.930	.930	.930	.930	.933	.931	.928	.923
LR	.930	.932	.932	.932	.933	.935	.934	.930	.926
ST	.931	.933	.933	.934	.934	.937	.935	.932	.928
KLD-U	.924	.925	.924	.924	.923	.926	.925	.923	.909
KLD-N	.920	.921	.919	.919	.916	.918	.914	.910	.900
Low Discrimination (a = 0.6)									
Z	.653	.651	.653	.652	.648	.654	.651	.647	.645
LR	.659	.657	.658	.659	.655	.662	.660	.655	.653
ST	.662	.661	.662	.663	.659	.666	.664	.660	.659
KLD-U	.638	.635	.634	.632	.627	.634	.633	.625	.564
KLD-N	.615	.609	.605	.599	.587	.589	.583	.574	.543
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.999	.999	.999	.999	.999	.999	.999	.999	.999
LR	.999	.999	.999	.999	.999	.999	.999	.999	.999
ST	.999	.999	.999	.999	.999	.999	.999	.999	.999
KLD-U	.999	.999	.999	.999	.999	.999	.999	.999	.996
KLD-N	.999	.999	.999	.999	.999	.999	.998	.998	.997
Low Discrimination (a = 0.6)									
Z	.937	.939	.938	.940	.938	.939	.938	.938	.934
LR	.938	.941	.941	.942	.941	.942	.940	.941	.936
ST	.940	.942	.942	.943	.942	.943	.942	.942	.938
KLD-U	.932	.934	.932	.933	.930	.933	.932	.924	.849
KLD-N	.923	.924	.920	.920	.914	.915	.911	.906	.862

Table A79. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.084	.045	.039	.038	.040	.044	.048	.050	.056
LR	.066	.058	.053	.049	.046	.046	.046	.041	.039
ST	.055	.052	.053	.053	.052	.052	.051	.044	.039
KLD-U	.064	.056	.048	.044	.042	.043	.047	.050	.048
KLD-N	.084	.064	.049	.040	.036	.036	.036	.033	.032
Moderate Discrimination (a = 1.0)									
Z	.070	.047	.041	.041	.042	.044	.051	.055	.057
LR	.063	.057	.054	.051	.048	.047	.048	.044	.040
ST	.056	.053	.054	.054	.051	.050	.050	.044	.040
KLD-U	.046	.048	.046	.045	.043	.045	.050	.050	.032
KLD-N	.056	.049	.043	.038	.034	.034	.032	.030	.020
Low Discrimination (a = 0.6)									
Z	.057	.056	.048	.045	.043	.047	.051	.056	.051
LR	.058	.058	.054	.052	.049	.049	.048	.046	.038
ST	.056	.056	.053	.053	.049	.049	.047	.044	.042
KLD-U	.028	.038	.044	.048	.047	.049	.049	.040	.022
KLD-N	.031	.034	.036	.036	.033	.031	.028	.024	.015
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.106	.076	.075	.082	.089	.096	.099	.099	.085
LR	.097	.098	.099	.103	.103	.101	.097	.085	.070
ST	.081	.090	.097	.106	.112	.110	.104	.091	.076
KLD-U	.061	.061	.060	.063	.065	.069	.072	.075	.065
KLD-N	.088	.077	.065	.060	.056	.053	.047	.044	.041
Moderate Discrimination (a = 1.0)									
Z	.085	.067	.062	.066	.072	.073	.080	.087	.073
LR	.082	.080	.080	.081	.082	.078	.076	.070	.055
ST	.072	.075	.079	.083	.085	.080	.077	.069	.060
KLD-U	.042	.050	.050	.052	.054	.058	.062	.060	.038
KLD-N	.059	.058	.051	.046	.042	.037	.033	.029	.019
Low Discrimination (a = 0.6)									
Z	.068	.066	.060	.057	.060	.066	.068	.074	.055
LR	.070	.071	.069	.068	.066	.068	.063	.059	.042
ST	.066	.069	.069	.068	.066	.067	.061	.057	.051
KLD-U	.028	.040	.047	.049	.052	.055	.053	.044	.023
KLD-N	.037	.043	.043	.038	.034	.030	.024	.019	.011
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.193	.178	.199	.227	.255	.274	.262	.247	.159
LR	.204	.228	.254	.279	.283	.282	.256	.221	.148
ST	.180	.218	.250	.284	.298	.298	.268	.231	.163
KLD-U	.124	.139	.160	.183	.196	.211	.205	.198	.134
KLD-N	.171	.171	.175	.180	.173	.166	.142	.132	.091

Continued on the next page

Table A79 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.134	.125	.136	.151	.162	.173	.176	.170	.114
LR	.142	.156	.170	.179	.180	.179	.167	.143	.097
ST	.130	.150	.169	.182	.185	.183	.167	.141	.111
KLD-U	.070	.092	.108	.115	.122	.133	.136	.117	.065
KLD-N	.101	.111	.112	.105	.095	.088	.076	.059	.034
Low Discrimination (a = 0.6)									
Z	.099	.098	.100	.106	.112	.119	.122	.113	.071
LR	.102	.110	.116	.120	.121	.120	.112	.095	.059
ST	.098	.107	.115	.119	.120	.118	.108	.093	.077
KLD-U	.038	.059	.075	.084	.089	.093	.088	.065	.033
KLD-N	.055	.068	.071	.066	.058	.050	.039	.026	.014
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.490	.544	.584	.692	.790	.780	.737	.645	.266
LR	.624	.701	.757	.804	.814	.788	.734	.623	.334
ST	.546	.620	.673	.763	.818	.799	.745	.634	.385
KLD-U	.455	.527	.597	.675	.724	.723	.687	.605	.298
KLD-N	.553	.602	.631	.662	.675	.643	.583	.498	.232
Moderate Discrimination (a = 1.0)									
Z	.349	.388	.445	.514	.547	.544	.512	.427	.192
LR	.421	.486	.540	.568	.570	.548	.491	.396	.217
ST	.389	.461	.524	.569	.576	.552	.488	.396	.275
KLD-U	.271	.343	.399	.443	.473	.480	.444	.336	.154
KLD-N	.338	.383	.406	.408	.396	.362	.302	.215	.094
Low Discrimination (a = 0.6)									
Z	.232	.257	.285	.308	.316	.325	.312	.234	.113
LR	.259	.296	.324	.336	.330	.322	.289	.218	.110
ST	.250	.291	.323	.337	.328	.316	.282	.228	.170
KLD-U	.130	.196	.237	.260	.265	.269	.236	.155	.068
KLD-N	.170	.207	.220	.210	.183	.160	.124	.077	.032
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.688	.735	.801	.950	.980	.966	.939	.818	.283
LR	.911	.950	.973	.984	.984	.978	.954	.851	.440
ST	.624	.606	.630	.815	.969	.978	.956	.862	.521
KLD-U	.779	.837	.918	.962	.973	.968	.945	.838	.393
KLD-N	.873	.903	.933	.953	.957	.944	.905	.772	.322
Moderate Discrimination (a = 1.0)									
Z	.596	.653	.761	.867	.879	.858	.806	.598	.194
LR	.774	.836	.875	.892	.886	.857	.788	.634	.300
ST	.680	.740	.825	.886	.887	.858	.785	.653	.428
KLD-U	.620	.706	.781	.828	.841	.824	.751	.559	.228
KLD-N	.697	.748	.778	.787	.773	.722	.612	.422	.151
Low Discrimination (a = 0.6)									
Z	.451	.519	.579	.615	.622	.604	.529	.328	.122
LR	.524	.590	.627	.640	.630	.594	.516	.354	.147
ST	.507	.581	.625	.640	.626	.584	.511	.402	.274
KLD-U	.364	.467	.524	.558	.564	.535	.440	.266	.102
KLD-N	.410	.468	.484	.471	.430	.372	.272	.151	.052

Table A80. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.080	.046	.041	.042	.043	.046	.048	.050	.057
LR	.061	.054	.051	.051	.049	.048	.050	.046	.045
ST	.058	.054	.052	.052	.050	.050	.050	.045	.041
KLD-U	.061	.058	.050	.048	.046	.046	.048	.050	.047
KLD-N	.077	.065	.052	.048	.045	.043	.043	.039	.034
Moderate Discrimination (a = 1.0)									
Z	.070	.048	.043	.044	.043	.045	.048	.054	.055
LR	.057	.053	.052	.052	.048	.049	.048	.047	.043
ST	.055	.053	.053	.053	.050	.050	.048	.044	.042
KLD-U	.046	.053	.053	.053	.050	.051	.054	.054	.034
KLD-N	.052	.052	.048	.046	.042	.040	.039	.034	.023
Low Discrimination (a = 0.6)									
Z	.056	.054	.047	.045	.045	.046	.050	.056	.050
LR	.054	.054	.053	.052	.051	.050	.050	.049	.041
ST	.055	.054	.053	.053	.052	.049	.048	.047	.044
KLD-U	.031	.043	.052	.056	.054	.055	.055	.046	.024
KLD-N	.029	.036	.039	.040	.038	.034	.031	.026	.015
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.121	.099	.099	.106	.108	.104	.101	.101	.082
LR	.095	.100	.104	.111	.113	.109	.104	.098	.082
ST	.090	.100	.107	.114	.117	.113	.108	.100	.086
KLD-U	.072	.082	.081	.084	.083	.078	.078	.076	.060
KLD-N	.088	.090	.086	.086	.082	.074	.067	.058	.047
Moderate Discrimination (a = 1.0)									
Z	.096	.080	.077	.078	.079	.076	.079	.083	.070
LR	.078	.078	.080	.083	.084	.081	.080	.076	.062
ST	.075	.079	.082	.085	.087	.084	.081	.076	.067
KLD-U	.051	.067	.070	.071	.070	.067	.068	.064	.040
KLD-N	.058	.066	.065	.064	.059	.052	.045	.039	.024
Low Discrimination (a = 0.6)									
Z	.074	.072	.065	.062	.062	.066	.068	.070	.056
LR	.066	.069	.068	.069	.069	.070	.067	.064	.048
ST	.065	.068	.068	.069	.069	.070	.065	.064	.058
KLD-U	.032	.049	.060	.062	.062	.064	.059	.047	.026
KLD-N	.034	.044	.048	.046	.042	.038	.030	.022	.013
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.255	.260	.284	.302	.301	.296	.272	.248	.160
LR	.225	.259	.289	.309	.311	.307	.280	.246	.173
ST	.217	.262	.293	.314	.318	.314	.288	.252	.184
KLD-U	.178	.211	.233	.246	.244	.237	.221	.198	.127
KLD-N	.197	.225	.244	.253	.243	.226	.196	.162	.105

Continued on the next page

Table A80 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.167	.165	.177	.183	.180	.178	.175	.165	.113
LR	.144	.163	.182	.188	.190	.187	.177	.155	.112
ST	.141	.164	.186	.192	.196	.192	.180	.157	.127
KLD-U	.101	.138	.156	.160	.157	.155	.148	.125	.070
KLD-N	.110	.135	.148	.147	.135	.123	.104	.078	.044
Low Discrimination (a = 0.6)									
Z	.115	.115	.113	.116	.114	.119	.120	.107	.073
LR	.102	.111	.117	.125	.124	.126	.118	.101	.070
ST	.100	.109	.117	.125	.124	.125	.116	.102	.088
KLD-U	.049	.083	.102	.109	.107	.109	.097	.069	.036
KLD-N	.056	.075	.084	.082	.074	.065	.049	.030	.017
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.700	.758	.795	.821	.821	.793	.740	.641	.274
LR	.687	.757	.798	.827	.832	.808	.759	.656	.385
ST	.683	.757	.800	.830	.837	.814	.766	.663	.429
KLD-U	.636	.706	.745	.776	.777	.750	.701	.596	.290
KLD-N	.654	.720	.757	.780	.772	.732	.660	.538	.263
Moderate Discrimination (a = 1.0)									
Z	.466	.516	.556	.572	.570	.552	.511	.420	.198
LR	.449	.518	.566	.584	.587	.564	.509	.422	.250
ST	.446	.520	.570	.590	.595	.570	.512	.428	.314
KLD-U	.392	.475	.521	.535	.532	.515	.462	.350	.166
KLD-N	.396	.466	.505	.508	.487	.445	.367	.257	.122
Low Discrimination (a = 0.6)									
Z	.284	.304	.320	.331	.327	.328	.304	.229	.123
LR	.269	.307	.333	.349	.344	.336	.301	.239	.136
ST	.266	.305	.335	.351	.345	.336	.300	.252	.202
KLD-U	.185	.265	.302	.315	.307	.298	.253	.167	.076
KLD-N	.189	.238	.259	.255	.230	.196	.147	.089	.038
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.943	.966	.979	.986	.982	.966	.937	.820	.292
LR	.938	.963	.977	.986	.986	.980	.960	.877	.499
ST	.933	.960	.976	.986	.986	.981	.961	.883	.572
KLD-U	.932	.959	.972	.981	.979	.970	.945	.832	.384
KLD-N	.933	.959	.972	.980	.977	.964	.928	.799	.363
Moderate Discrimination (a = 1.0)									
Z	.810	.854	.880	.891	.882	.856	.799	.598	.203
LR	.804	.855	.883	.898	.890	.865	.804	.667	.345
ST	.800	.854	.884	.900	.892	.868	.805	.688	.485
KLD-U	.776	.834	.862	.873	.865	.838	.765	.576	.246
KLD-N	.771	.825	.851	.854	.832	.782	.674	.478	.197
Low Discrimination (a = 0.6)									
Z	.549	.601	.621	.632	.626	.603	.520	.333	.139
LR	.544	.610	.638	.652	.642	.608	.537	.392	.186
ST	.540	.609	.640	.655	.643	.606	.538	.441	.330
KLD-U	.475	.574	.604	.614	.604	.562	.460	.285	.115
KLD-N	.454	.528	.542	.534	.493	.417	.306	.172	.064

Table A81. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.079	.047	.042	.044	.045	.046	.048	.051	.055
LR	.060	.055	.051	.050	.049	.049	.049	.047	.046
ST	.058	.055	.053	.052	.050	.050	.048	.045	.042
KLD-U	.059	.060	.053	.048	.046	.045	.047	.050	.044
KLD-N	.074	.066	.055	.048	.045	.044	.042	.040	.034
Moderate Discrimination (a = 1.0)									
Z	.069	.049	.043	.044	.044	.045	.050	.053	.051
LR	.057	.053	.051	.050	.048	.047	.049	.048	.044
ST	.055	.053	.052	.052	.049	.048	.048	.046	.043
KLD-U	.045	.054	.054	.054	.050	.050	.054	.053	.033
KLD-N	.054	.056	.052	.050	.045	.043	.042	.037	.025
Low Discrimination (a = 0.6)									
Z	.057	.054	.048	.046	.046	.047	.052	.054	.047
LR	.053	.053	.051	.052	.050	.050	.051	.049	.041
ST	.054	.052	.051	.052	.050	.049	.048	.046	.045
KLD-U	.034	.045	.055	.060	.059	.059	.058	.048	.026
KLD-N	.031	.037	.041	.042	.039	.037	.033	.027	.016
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.121	.102	.104	.111	.113	.108	.107	.101	.079
LR	.095	.101	.105	.111	.115	.111	.107	.099	.083
ST	.090	.101	.108	.114	.119	.116	.110	.102	.089
KLD-U	.075	.089	.088	.089	.088	.082	.078	.077	.054
KLD-N	.087	.092	.090	.090	.088	.079	.070	.062	.044
Moderate Discrimination (a = 1.0)									
Z	.096	.080	.079	.080	.081	.080	.082	.081	.062
LR	.076	.076	.080	.082	.085	.082	.081	.078	.061
ST	.073	.076	.081	.083	.086	.084	.081	.078	.068
KLD-U	.052	.069	.074	.074	.072	.068	.069	.063	.037
KLD-N	.061	.069	.072	.070	.066	.059	.054	.044	.027
Low Discrimination (a = 0.6)									
Z	.073	.072	.067	.065	.064	.068	.068	.066	.048
LR	.064	.068	.068	.069	.069	.070	.066	.064	.046
ST	.062	.066	.067	.068	.068	.068	.065	.063	.060
KLD-U	.035	.055	.068	.071	.070	.068	.062	.050	.026
KLD-N	.035	.046	.052	.050	.044	.041	.033	.024	.013
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.260	.265	.289	.310	.309	.303	.280	.247	.152
LR	.226	.260	.288	.311	.314	.308	.282	.249	.175
ST	.219	.261	.292	.317	.322	.316	.289	.257	.190
KLD-U	.190	.225	.245	.253	.251	.242	.222	.194	.114
KLD-N	.200	.231	.251	.259	.252	.235	.206	.164	.096

Continued on the next page

Table A81 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.169	.168	.180	.188	.188	.184	.180	.160	.099
LR	.144	.163	.180	.190	.192	.188	.177	.157	.110
ST	.139	.162	.182	.192	.196	.192	.178	.160	.128
KLD-U	.106	.146	.162	.168	.165	.156	.148	.121	.064
KLD-N	.117	.144	.159	.162	.153	.137	.119	.088	.049
Low Discrimination (a = 0.6)									
Z	.117	.117	.116	.119	.118	.122	.119	.098	.062
LR	.101	.110	.117	.124	.124	.126	.118	.100	.065
ST	.098	.108	.115	.122	.124	.124	.117	.104	.092
KLD-U	.057	.093	.114	.122	.118	.116	.104	.074	.038
KLD-N	.059	.079	.088	.089	.080	.072	.056	.036	.018
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.701	.759	.796	.822	.819	.792	.742	.625	.261
LR	.683	.754	.793	.824	.825	.801	.756	.651	.388
ST	.678	.753	.793	.825	.830	.807	.764	.661	.439
KLD-U	.651	.716	.751	.776	.775	.745	.698	.571	.265
KLD-N	.655	.724	.759	.780	.773	.733	.668	.525	.242
Moderate Discrimination (a = 1.0)									
Z	.471	.520	.564	.576	.576	.555	.510	.392	.162
LR	.449	.514	.565	.580	.585	.561	.509	.416	.240
ST	.444	.514	.567	.584	.590	.566	.512	.428	.316
KLD-U	.407	.486	.531	.540	.536	.510	.458	.333	.151
KLD-N	.410	.482	.526	.528	.511	.468	.392	.271	.127
Low Discrimination (a = 0.6)									
Z	.287	.309	.327	.337	.333	.332	.292	.197	.090
LR	.267	.305	.332	.348	.343	.335	.301	.233	.118
ST	.260	.300	.329	.348	.342	.333	.301	.257	.202
KLD-U	.204	.286	.324	.337	.324	.310	.262	.174	.077
KLD-N	.193	.246	.271	.269	.243	.214	.160	.098	.039
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.941	.965	.977	.984	.980	.964	.931	.807	.278
LR	.933	.960	.974	.983	.983	.977	.955	.875	.500
ST	.927	.955	.972	.983	.984	.979	.958	.884	.581
KLD-U	.935	.960	.972	.978	.974	.967	.936	.811	.354
KLD-N	.932	.959	.971	.978	.974	.963	.921	.781	.336
Moderate Discrimination (a = 1.0)									
Z	.808	.854	.878	.887	.880	.854	.785	.546	.158
LR	.797	.849	.878	.890	.885	.860	.799	.659	.329
ST	.791	.846	.877	.891	.887	.862	.802	.690	.484
KLD-U	.781	.836	.862	.868	.860	.832	.752	.553	.220
KLD-N	.776	.833	.857	.859	.842	.797	.688	.490	.196
Low Discrimination (a = 0.6)									
Z	.551	.600	.622	.635	.628	.597	.476	.269	.094
LR	.539	.603	.632	.649	.638	.606	.530	.369	.153
ST	.531	.597	.630	.648	.637	.604	.536	.443	.326
KLD-U	.496	.591	.622	.632	.614	.572	.466	.286	.111
KLD-N	.461	.537	.553	.548	.511	.440	.325	.177	.061

Table A82. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.044	.043	.042	.043	.041	.044	.044	.048	.049
LR	.061	.058	.054	.052	.047	.046	.046	.047	.044
ST	.053	.053	.052	.053	.050	.049	.051	.050	.046
KLD-U	.064	.057	.050	.047	.042	.043	.045	.050	.045
KLD-N	.082	.063	.049	.042	.035	.035	.034	.034	.028
Moderate Discrimination (a = 1.0)									
Z	.037	.040	.041	.045	.043	.045	.047	.051	.054
LR	.058	.054	.052	.051	.047	.048	.048	.047	.045
ST	.052	.051	.050	.052	.049	.050	.049	.047	.044
KLD-U	.047	.047	.045	.046	.042	.044	.044	.045	.034
KLD-N	.059	.050	.041	.038	.031	.031	.029	.027	.022
Low Discrimination (a = 0.6)									
Z	.038	.040	.044	.045	.044	.046	.048	.052	.056
LR	.054	.052	.052	.051	.050	.048	.048	.048	.046
ST	.052	.051	.052	.053	.051	.049	.048	.047	.044
KLD-U	.038	.042	.043	.045	.043	.043	.042	.036	.022
KLD-N	.044	.039	.034	.032	.027	.025	.022	.020	.016
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.076	.075	.078	.080	.084	.092	.090	.093	.091
LR	.101	.098	.097	.097	.094	.097	.091	.089	.083
ST	.089	.090	.093	.097	.098	.103	.097	.095	.089
KLD-U	.068	.063	.064	.063	.062	.065	.067	.070	.059
KLD-N	.096	.076	.066	.058	.050	.047	.041	.036	.026
Moderate Discrimination (a = 1.0)									
Z	.056	.059	.064	.066	.067	.070	.074	.079	.078
LR	.079	.076	.077	.076	.073	.074	.073	.073	.065
ST	.072	.072	.075	.076	.074	.076	.074	.073	.066
KLD-U	.047	.048	.050	.050	.048	.051	.054	.054	.037
KLD-N	.067	.056	.050	.043	.035	.030	.027	.022	.014
Low Discrimination (a = 0.6)									
Z	.050	.055	.057	.055	.057	.058	.064	.069	.071
LR	.066	.068	.066	.062	.062	.060	.062	.061	.058
ST	.062	.066	.066	.062	.063	.061	.062	.060	.057
KLD-U	.034	.042	.044	.042	.044	.044	.045	.038	.024
KLD-N	.050	.047	.040	.031	.025	.021	.018	.014	.009
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.165	.174	.192	.212	.232	.247	.241	.235	.214
LR	.211	.221	.236	.250	.255	.256	.244	.231	.206
ST	.197	.208	.226	.248	.262	.264	.255	.240	.220
KLD-U	.135	.141	.153	.166	.175	.186	.187	.182	.144
KLD-N	.179	.167	.162	.157	.147	.136	.119	.101	.064

Continued on the next page

Table A82 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.111	.121	.133	.141	.146	.153	.158	.160	.152
LR	.144	.151	.158	.158	.159	.158	.156	.148	.137
ST	.134	.143	.153	.158	.160	.161	.157	.148	.138
KLD-U	.080	.092	.100	.102	.103	.110	.115	.106	.076
KLD-N	.115	.108	.101	.090	.075	.065	.056	.041	.027
Low Discrimination (a = 0.6)									
Z	.080	.086	.091	.093	.096	.098	.107	.110	.110
LR	.101	.103	.104	.103	.103	.101	.103	.098	.093
ST	.096	.101	.104	.105	.104	.101	.102	.096	.092
KLD-U	.049	.060	.065	.066	.067	.068	.070	.057	.038
KLD-N	.076	.069	.060	.049	.038	.030	.024	.018	.012
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.511	.553	.590	.661	.740	.746	.719	.682	.596
LR	.632	.674	.724	.756	.768	.759	.732	.690	.619
ST	.556	.593	.646	.715	.765	.765	.743	.703	.638
KLD-U	.477	.515	.576	.627	.667	.678	.669	.625	.464
KLD-N	.565	.578	.598	.605	.604	.576	.530	.448	.297
Moderate Discrimination (a = 1.0)									
Z	.353	.391	.425	.458	.477	.488	.485	.466	.394
LR	.430	.458	.484	.496	.498	.494	.480	.448	.392
ST	.399	.430	.465	.492	.502	.496	.480	.446	.399
KLD-U	.292	.332	.355	.376	.392	.407	.403	.351	.219
KLD-N	.358	.366	.356	.334	.306	.276	.235	.178	.112
Low Discrimination (a = 0.6)									
Z	.218	.234	.248	.255	.264	.273	.284	.284	.241
LR	.259	.268	.274	.276	.278	.276	.273	.260	.227
ST	.250	.263	.273	.278	.280	.275	.270	.257	.233
KLD-U	.151	.176	.186	.191	.199	.206	.197	.156	.091
KLD-N	.198	.189	.168	.140	.116	.095	.074	.056	.039
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.787	.784	.814	.939	.970	.963	.944	.914	.802
LR	.904	.937	.963	.974	.975	.973	.964	.943	.884
ST	.620	.626	.674	.824	.951	.973	.965	.947	.899
KLD-U	.789	.839	.902	.943	.957	.959	.949	.910	.702
KLD-N	.869	.889	.911	.929	.932	.922	.890	.810	.580
Moderate Discrimination (a = 1.0)									
Z	.652	.687	.747	.810	.826	.826	.808	.766	.581
LR	.769	.801	.822	.836	.836	.831	.808	.769	.668
ST	.668	.712	.774	.826	.837	.830	.806	.768	.695
KLD-U	.627	.674	.715	.751	.766	.776	.746	.647	.401
KLD-N	.699	.709	.702	.693	.666	.629	.550	.438	.281
Low Discrimination (a = 0.6)									
Z	.446	.478	.500	.520	.529	.542	.542	.506	.362
LR	.510	.529	.535	.543	.542	.540	.525	.492	.393
ST	.486	.518	.532	.544	.542	.537	.519	.490	.437
KLD-U	.366	.402	.419	.435	.446	.447	.408	.311	.166
KLD-N	.416	.404	.370	.330	.286	.243	.194	.148	.093

Table A83. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.038	.042	.045	.046	.046	.046	.045	.051	.050
LR	.056	.054	.053	.051	.048	.048	.048	.052	.047
ST	.057	.055	.054	.051	.049	.048	.048	.051	.045
KLD-U	.064	.056	.052	.048	.044	.044	.045	.052	.047
KLD-N	.082	.065	.055	.048	.043	.042	.040	.042	.037
Moderate Discrimination (a = 1.0)									
Z	.039	.042	.044	.047	.046	.047	.048	.051	.054
LR	.053	.052	.050	.050	.049	.049	.049	.049	.048
ST	.054	.053	.051	.050	.049	.049	.048	.047	.046
KLD-U	.052	.054	.051	.051	.048	.048	.050	.051	.040
KLD-N	.060	.052	.047	.043	.039	.037	.036	.034	.027
Low Discrimination (a = 0.6)									
Z	.041	.043	.045	.047	.048	.047	.049	.053	.056
LR	.052	.051	.050	.051	.050	.049	.049	.050	.048
ST	.053	.051	.050	.051	.051	.049	.049	.048	.046
KLD-U	.042	.048	.050	.052	.051	.049	.048	.044	.027
KLD-N	.042	.040	.038	.036	.031	.028	.026	.024	.017
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.088	.095	.101	.104	.100	.104	.097	.097	.091
LR	.100	.100	.102	.104	.102	.106	.100	.099	.094
ST	.105	.103	.103	.105	.103	.108	.103	.102	.101
KLD-U	.084	.080	.081	.079	.073	.074	.071	.076	.066
KLD-N	.099	.088	.085	.080	.072	.069	.061	.060	.049
Moderate Discrimination (a = 1.0)									
Z	.069	.075	.077	.078	.076	.074	.075	.080	.074
LR	.077	.078	.078	.078	.077	.075	.076	.079	.072
ST	.080	.080	.079	.079	.077	.077	.077	.080	.075
KLD-U	.061	.067	.068	.065	.061	.060	.061	.064	.045
KLD-N	.069	.065	.063	.057	.050	.045	.040	.037	.026
Low Discrimination (a = 0.6)									
Z	.059	.062	.064	.061	.061	.060	.064	.068	.068
LR	.065	.065	.066	.062	.063	.061	.064	.065	.062
ST	.066	.066	.067	.062	.064	.063	.064	.064	.062
KLD-U	.042	.052	.058	.055	.055	.053	.054	.047	.030
KLD-N	.047	.047	.045	.038	.034	.028	.025	.020	.013
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.223	.244	.266	.277	.276	.272	.259	.250	.214
LR	.235	.249	.267	.276	.279	.278	.269	.258	.231
ST	.243	.254	.269	.276	.282	.283	.276	.265	.247
KLD-U	.191	.202	.212	.217	.213	.212	.205	.203	.164
KLD-N	.211	.216	.223	.222	.210	.199	.180	.168	.132

Continued on the next page

Table A83 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.147	.158	.171	.169	.166	.163	.165	.162	.149
LR	.154	.159	.168	.167	.166	.167	.168	.161	.151
ST	.158	.162	.169	.168	.169	.170	.171	.164	.161
KLD-U	.119	.134	.142	.137	.132	.131	.135	.125	.095
KLD-N	.128	.131	.134	.123	.110	.101	.092	.076	.058
Low Discrimination (a = 0.6)									
Z	.099	.105	.105	.108	.105	.103	.109	.108	.107
LR	.102	.106	.105	.108	.107	.106	.109	.104	.102
ST	.105	.108	.106	.109	.108	.107	.110	.105	.106
KLD-U	.066	.083	.089	.090	.088	.085	.086	.070	.048
KLD-N	.074	.076	.070	.063	.054	.045	.041	.028	.020
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.683	.726	.765	.777	.778	.765	.736	.692	.602
LR	.696	.728	.762	.779	.787	.779	.757	.721	.668
ST	.698	.726	.760	.780	.790	.783	.765	.735	.695
KLD-U	.643	.671	.706	.715	.720	.711	.696	.656	.524
KLD-N	.663	.688	.718	.719	.713	.687	.657	.599	.482
Moderate Discrimination (a = 1.0)									
Z	.460	.492	.512	.517	.510	.504	.499	.471	.403
LR	.465	.492	.510	.519	.519	.513	.506	.479	.434
ST	.468	.493	.509	.521	.523	.518	.510	.488	.458
KLD-U	.413	.446	.460	.459	.453	.451	.446	.399	.274
KLD-N	.416	.438	.442	.426	.401	.376	.346	.289	.203
Low Discrimination (a = 0.6)									
Z	.271	.280	.286	.287	.283	.283	.288	.282	.248
LR	.272	.278	.285	.290	.289	.287	.288	.278	.257
ST	.275	.281	.288	.293	.293	.290	.290	.283	.271
KLD-U	.210	.239	.249	.251	.245	.242	.232	.189	.117
KLD-N	.213	.214	.205	.186	.164	.145	.121	.092	.060
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.934	.957	.971	.975	.973	.963	.944	.914	.807
LR	.931	.952	.968	.976	.977	.975	.967	.954	.917
ST	.926	.948	.967	.975	.977	.976	.969	.959	.934
KLD-U	.926	.946	.960	.965	.966	.963	.953	.925	.765
KLD-N	.926	.946	.961	.964	.963	.955	.939	.905	.760
Moderate Discrimination (a = 1.0)									
Z	.802	.823	.837	.842	.838	.832	.813	.769	.628
LR	.802	.819	.837	.844	.844	.842	.825	.797	.739
ST	.799	.817	.836	.844	.845	.844	.828	.808	.767
KLD-U	.771	.790	.802	.806	.804	.804	.781	.706	.484
KLD-N	.766	.781	.785	.776	.755	.737	.684	.602	.426
Low Discrimination (a = 0.6)									
Z	.530	.549	.547	.552	.548	.552	.544	.511	.404
LR	.530	.548	.550	.558	.557	.558	.545	.521	.474
ST	.531	.549	.552	.562	.560	.561	.549	.532	.506
KLD-U	.471	.502	.505	.506	.500	.497	.458	.363	.213
KLD-N	.450	.455	.436	.405	.375	.338	.285	.218	.136

Table A84. Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.041	.044	.046	.048	.047	.049	.047	.054	.050
LR	.054	.053	.053	.052	.047	.049	.047	.051	.047
ST	.057	.054	.053	.052	.048	.048	.047	.050	.045
KLD-U	.066	.059	.054	.049	.043	.043	.043	.052	.046
KLD-N	.083	.065	.055	.048	.042	.041	.039	.044	.038
Moderate Discrimination (a = 1.0)									
Z	.040	.044	.045	.048	.048	.050	.050	.053	.056
LR	.051	.052	.050	.050	.049	.050	.048	.048	.049
ST	.053	.053	.050	.051	.050	.050	.048	.047	.047
KLD-U	.050	.055	.052	.051	.048	.048	.048	.049	.039
KLD-N	.063	.058	.051	.047	.043	.041	.038	.038	.032
Low Discrimination (a = 0.6)									
Z	.043	.044	.046	.048	.050	.050	.052	.056	.056
LR	.051	.050	.049	.050	.050	.049	.049	.049	.048
ST	.052	.051	.050	.051	.051	.050	.049	.048	.046
KLD-U	.045	.051	.054	.057	.057	.055	.053	.047	.030
KLD-N	.046	.043	.039	.037	.034	.030	.029	.026	.018
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.092	.099	.106	.110	.105	.108	.101	.102	.090
LR	.100	.101	.103	.105	.101	.106	.100	.097	.095
ST	.105	.103	.105	.105	.102	.107	.103	.100	.102
KLD-U	.090	.090	.086	.084	.075	.076	.070	.074	.065
KLD-N	.100	.092	.085	.082	.073	.072	.063	.064	.054
Moderate Discrimination (a = 1.0)									
Z	.072	.078	.080	.083	.080	.077	.080	.084	.078
LR	.076	.078	.078	.078	.077	.075	.076	.078	.073
ST	.079	.081	.080	.079	.078	.078	.077	.079	.077
KLD-U	.062	.070	.071	.068	.063	.061	.060	.062	.047
KLD-N	.073	.071	.068	.064	.057	.051	.046	.046	.035
Low Discrimination (a = 0.6)									
Z	.061	.065	.067	.063	.065	.064	.068	.071	.069
LR	.062	.064	.066	.062	.064	.061	.064	.063	.063
ST	.064	.065	.066	.062	.064	.061	.065	.063	.064
KLD-U	.046	.059	.065	.062	.062	.058	.059	.050	.034
KLD-N	.049	.049	.047	.040	.036	.031	.029	.024	.018
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.230	.252	.278	.288	.286	.280	.269	.262	.207
LR	.234	.249	.269	.277	.281	.277	.268	.258	.229
ST	.243	.253	.271	.277	.284	.281	.274	.264	.248
KLD-U	.203	.214	.223	.222	.218	.210	.202	.208	.160
KLD-N	.213	.219	.227	.224	.216	.201	.186	.184	.140

Continued on the next page

Table A84 (continued). Mean Type I Error Rates and Power for 15-Item Variable-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.151	.164	.177	.177	.173	.172	.174	.172	.149
LR	.151	.159	.169	.169	.168	.169	.167	.162	.152
ST	.156	.162	.170	.170	.170	.172	.170	.167	.164
KLD-U	.123	.141	.148	.143	.137	.136	.132	.127	.095
KLD-N	.135	.141	.144	.136	.125	.114	.106	.097	.075
Low Discrimination (a = 0.6)									
Z	.104	.110	.112	.112	.111	.110	.115	.112	.104
LR	.102	.105	.106	.107	.107	.107	.109	.102	.102
ST	.103	.107	.107	.109	.109	.108	.110	.106	.108
KLD-U	.074	.094	.102	.102	.101	.096	.093	.077	.054
KLD-N	.076	.079	.075	.069	.060	.052	.049	.038	.027
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.687	.729	.769	.781	.777	.766	.740	.690	.581
LR	.688	.721	.758	.773	.777	.770	.749	.713	.649
ST	.690	.719	.754	.773	.779	.774	.756	.727	.679
KLD-U	.653	.676	.708	.712	.713	.703	.687	.647	.514
KLD-N	.660	.684	.714	.715	.708	.686	.662	.606	.505
Moderate Discrimination (a = 1.0)									
Z	.466	.502	.521	.526	.517	.516	.510	.477	.372
LR	.460	.491	.507	.516	.513	.513	.500	.477	.422
ST	.465	.492	.507	.518	.518	.518	.506	.488	.450
KLD-U	.418	.454	.465	.465	.455	.448	.439	.396	.269
KLD-N	.428	.454	.459	.451	.426	.402	.384	.330	.250
Low Discrimination (a = 0.6)									
Z	.279	.290	.294	.296	.292	.293	.297	.281	.221
LR	.270	.278	.284	.288	.286	.288	.285	.277	.251
ST	.273	.279	.285	.290	.289	.290	.289	.285	.270
KLD-U	.230	.261	.271	.273	.264	.258	.245	.204	.129
KLD-N	.217	.219	.213	.200	.178	.164	.148	.116	.082
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.932	.956	.970	.974	.971	.962	.942	.904	.779
LR	.926	.947	.964	.972	.974	.971	.962	.944	.908
ST	.920	.941	.962	.971	.974	.972	.965	.952	.925
KLD-U	.928	.946	.957	.961	.961	.957	.947	.914	.752
KLD-N	.924	.944	.957	.960	.959	.952	.938	.899	.776
Moderate Discrimination (a = 1.0)									
Z	.805	.825	.838	.840	.838	.832	.815	.750	.541
LR	.796	.813	.829	.834	.837	.833	.816	.786	.715
ST	.794	.811	.828	.835	.840	.837	.821	.799	.752
KLD-U	.775	.792	.801	.801	.799	.792	.771	.691	.471
KLD-N	.774	.789	.794	.786	.770	.752	.717	.640	.488
Low Discrimination (a = 0.6)									
Z	.538	.555	.554	.559	.556	.560	.543	.473	.320
LR	.524	.543	.544	.552	.552	.553	.539	.516	.437
ST	.524	.544	.545	.554	.555	.557	.546	.530	.496
KLD-U	.493	.526	.528	.529	.522	.511	.472	.381	.223
KLD-N	.454	.462	.444	.423	.397	.371	.323	.261	.168

Table A85. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.092	.050	.041	.041	.040	.042	.043	.051	.062
LR	.057	.055	.053	.052	.050	.050	.048	.047	.045
ST	.046	.049	.053	.055	.054	.054	.052	.050	.044
KLD-U	.054	.056	.050	.048	.046	.047	.047	.052	.058
KLD-N	.061	.058	.050	.047	.044	.045	.043	.046	.046
Moderate Discrimination (a = 1.0)									
Z	.077	.051	.044	.041	.040	.041	.045	.052	.061
LR	.056	.053	.053	.052	.049	.047	.048	.046	.045
ST	.049	.049	.053	.054	.052	.050	.051	.046	.044
KLD-U	.045	.053	.050	.049	.049	.048	.051	.053	.049
KLD-N	.050	.051	.046	.045	.043	.042	.043	.042	.038
Low Discrimination (a = 0.6)									
Z	.061	.057	.048	.043	.041	.043	.048	.054	.061
LR	.055	.055	.052	.051	.049	.049	.048	.046	.048
ST	.054	.054	.053	.052	.051	.050	.049	.046	.047
KLD-U	.030	.048	.052	.051	.052	.052	.054	.051	.036
KLD-N	.031	.040	.041	.041	.039	.038	.037	.033	.026
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.136	.129	.132	.143	.150	.155	.157	.164	.146
LR	.123	.151	.163	.172	.176	.173	.164	.144	.108
ST	.109	.146	.164	.176	.182	.176	.165	.139	.095
KLD-U	.112	.128	.128	.137	.142	.143	.141	.140	.104
KLD-N	.120	.138	.136	.139	.138	.133	.124	.119	.080
Moderate Discrimination (a = 1.0)									
Z	.103	.089	.087	.094	.099	.103	.107	.118	.112
LR	.094	.102	.109	.116	.116	.114	.106	.096	.081
ST	.086	.099	.108	.118	.118	.114	.104	.092	.074
KLD-U	.072	.090	.089	.095	.097	.098	.096	.095	.071
KLD-N	.083	.092	.090	.091	.087	.082	.073	.069	.050
Low Discrimination (a = 0.6)									
Z	.074	.072	.067	.068	.070	.074	.086	.090	.087
LR	.076	.079	.080	.082	.081	.079	.081	.073	.066
ST	.073	.078	.080	.083	.081	.078	.077	.070	.061
KLD-U	.040	.065	.071	.074	.073	.073	.077	.068	.041
KLD-N	.046	.059	.061	.060	.056	.051	.047	.040	.025
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.364	.408	.452	.491	.510	.505	.487	.456	.322
LR	.363	.455	.505	.536	.546	.529	.493	.418	.265
ST	.350	.452	.509	.542	.556	.535	.495	.408	.240
KLD-U	.356	.399	.439	.475	.492	.484	.461	.418	.243
KLD-N	.364	.423	.456	.480	.484	.463	.429	.380	.196

Continued on the next page

Table A85 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.217	.235	.259	.284	.294	.300	.298	.284	.236
LR	.221	.268	.297	.321	.324	.318	.294	.247	.185
ST	.212	.265	.300	.325	.326	.320	.288	.237	.171
KLD-U	.192	.237	.257	.279	.287	.288	.276	.243	.152
KLD-N	.204	.243	.258	.270	.265	.253	.232	.191	.114
Low Discrimination (a = 0.6)									
Z	.127	.135	.148	.159	.168	.172	.178	.176	.151
LR	.135	.156	.174	.184	.188	.180	.166	.148	.120
ST	.134	.155	.175	.184	.186	.176	.160	.140	.112
KLD-U	.086	.133	.155	.163	.169	.165	.157	.131	.072
KLD-N	.095	.123	.138	.140	.135	.123	.107	.080	.042
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.883	.907	.929	.967	.982	.976	.964	.914	.604
LR	.927	.964	.978	.983	.984	.979	.963	.889	.591
ST	.910	.950	.970	.982	.985	.979	.962	.879	.564
KLD-U	.905	.944	.965	.976	.979	.974	.960	.892	.516
KLD-N	.922	.955	.969	.976	.977	.969	.952	.867	.447
Moderate Discrimination (a = 1.0)									
Z	.678	.752	.806	.836	.838	.826	.791	.715	.528
LR	.708	.795	.838	.858	.855	.834	.781	.669	.480
ST	.700	.793	.838	.861	.857	.834	.775	.654	.458
KLD-U	.680	.756	.806	.831	.832	.818	.773	.660	.366
KLD-N	.688	.765	.805	.820	.810	.782	.723	.588	.304
Low Discrimination (a = 0.6)									
Z	.398	.459	.510	.537	.543	.536	.510	.451	.348
LR	.426	.496	.549	.571	.566	.541	.488	.406	.308
ST	.424	.498	.551	.572	.564	.534	.476	.390	.295
KLD-U	.368	.466	.519	.542	.542	.523	.471	.352	.174
KLD-N	.363	.441	.485	.496	.478	.440	.370	.253	.121
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.861	.868	.931	.998	1.000	1.000	.999	.991	.667
LR	.997	.999	1.000	1.000	1.000	1.000	.999	.985	.740
ST	.902	.880	.914	.984	1.000	1.000	.999	.982	.733
KLD-U	.994	.999	1.000	1.000	1.000	1.000	.999	.984	.638
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.999	.976	.571
Moderate Discrimination (a = 1.0)									
Z	.925	.948	.980	.992	.992	.988	.976	.928	.657
LR	.974	.988	.993	.994	.993	.989	.972	.908	.687
ST	.965	.981	.992	.994	.993	.988	.970	.899	.694
KLD-U	.965	.983	.990	.992	.992	.988	.972	.890	.524
KLD-N	.969	.984	.990	.991	.989	.982	.959	.850	.473
Low Discrimination (a = 0.6)									
Z	.766	.836	.868	.880	.874	.856	.807	.716	.510
LR	.794	.860	.886	.894	.883	.853	.788	.681	.507
ST	.792	.860	.886	.894	.881	.849	.779	.664	.508
KLD-U	.767	.841	.870	.880	.872	.844	.769	.585	.278
KLD-N	.752	.823	.849	.852	.830	.784	.677	.478	.220

Table A86. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ								
	−2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.082	.047	.042	.041	.042	.043	.044	.052	.061
LR	.052	.050	.051	.050	.050	.050	.050	.050	.049
ST	.048	.050	.053	.052	.053	.052	.052	.049	.044
KLD-U	.051	.057	.051	.049	.047	.048	.048	.052	.055
KLD-N	.054	.057	.054	.050	.049	.049	.048	.049	.046
Moderate Discrimination (a = 1.0)									
Z	.072	.049	.044	.042	.042	.043	.048	.051	.060
LR	.051	.049	.051	.049	.049	.048	.051	.048	.048
ST	.049	.049	.052	.051	.051	.050	.051	.047	.045
KLD-U	.043	.053	.051	.050	.049	.049	.052	.052	.047
KLD-N	.045	.050	.050	.049	.048	.046	.047	.044	.039
Low Discrimination (a = 0.6)									
Z	.060	.056	.048	.045	.043	.045	.050	.055	.060
LR	.053	.053	.052	.053	.051	.051	.052	.050	.050
ST	.054	.054	.053	.054	.052	.051	.051	.049	.048
KLD-U	.034	.052	.056	.057	.055	.055	.057	.054	.037
KLD-N	.032	.041	.044	.045	.043	.042	.041	.036	.027
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.139	.145	.154	.165	.167	.166	.165	.163	.144
LR	.118	.154	.172	.183	.185	.182	.173	.155	.118
ST	.112	.154	.175	.186	.189	.185	.174	.150	.104
KLD-U	.115	.149	.152	.159	.158	.156	.148	.137	.099
KLD-N	.114	.153	.161	.168	.164	.156	.144	.126	.082
Moderate Discrimination (a = 1.0)									
Z	.106	.098	.099	.108	.110	.109	.112	.116	.109
LR	.090	.102	.110	.120	.121	.118	.112	.104	.087
ST	.087	.101	.111	.121	.123	.118	.109	.099	.080
KLD-U	.076	.102	.101	.105	.105	.102	.098	.093	.067
KLD-N	.079	.099	.103	.106	.103	.096	.086	.076	.053
Low Discrimination (a = 0.6)									
Z	.077	.076	.073	.074	.076	.080	.089	.088	.086
LR	.072	.078	.082	.085	.086	.085	.086	.078	.072
ST	.072	.078	.082	.085	.086	.083	.083	.075	.068
KLD-U	.044	.074	.081	.083	.082	.080	.082	.071	.045
KLD-N	.046	.063	.068	.068	.065	.059	.054	.044	.029
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.392	.458	.500	.534	.543	.527	.503	.456	.322
LR	.377	.477	.527	.559	.569	.549	.516	.441	.283
ST	.370	.480	.533	.565	.576	.554	.518	.430	.260
KLD-U	.390	.458	.495	.522	.529	.510	.478	.414	.235
KLD-N	.381	.470	.510	.535	.537	.510	.469	.391	.201

Continued on the next page

Table A86 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.236	.267	.289	.313	.314	.314	.308	.281	.234
LR	.224	.278	.308	.335	.334	.330	.307	.261	.197
ST	.219	.278	.311	.339	.336	.331	.304	.253	.185
KLD-U	.214	.271	.289	.307	.304	.300	.280	.239	.150
KLD-N	.211	.269	.291	.308	.298	.285	.255	.204	.124
Low Discrimination (a = 0.6)									
Z	.139	.150	.162	.173	.181	.183	.181	.174	.149
LR	.136	.160	.177	.191	.195	.190	.175	.156	.128
ST	.135	.160	.178	.192	.195	.187	.170	.151	.122
KLD-U	.100	.156	.176	.184	.187	.178	.164	.137	.078
KLD-N	.100	.134	.152	.158	.155	.139	.117	.090	.050
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.937	.966	.978	.982	.984	.977	.965	.914	.610
LR	.937	.969	.980	.984	.986	.980	.966	.900	.616
ST	.936	.969	.980	.985	.986	.981	.966	.892	.596
KLD-U	.941	.966	.978	.982	.983	.976	.960	.889	.508
KLD-N	.940	.968	.979	.983	.983	.975	.958	.872	.458
Moderate Discrimination (a = 1.0)									
Z	.726	.795	.835	.854	.850	.833	.795	.714	.534
LR	.726	.807	.847	.866	.860	.842	.793	.687	.504
ST	.724	.808	.848	.868	.862	.842	.790	.677	.488
KLD-U	.730	.797	.832	.850	.844	.822	.774	.656	.366
KLD-N	.722	.797	.835	.849	.838	.808	.746	.608	.330
Low Discrimination (a = 0.6)									
Z	.427	.491	.532	.556	.558	.549	.512	.450	.351
LR	.436	.512	.559	.582	.579	.556	.502	.422	.326
ST	.435	.513	.561	.584	.578	.551	.495	.412	.316
KLD-U	.411	.512	.551	.571	.565	.540	.481	.366	.187
KLD-N	.384	.472	.516	.529	.509	.467	.391	.276	.141
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	1.000	1.000	1.000	1.000	.999	.991	.675
LR	.998	.999	1.000	1.000	1.000	1.000	.999	.988	.763
ST	.997	.999	1.000	1.000	1.000	1.000	.999	.985	.764
KLD-U	.998	1.000	1.000	1.000	1.000	1.000	.999	.984	.630
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.999	.978	.584
Moderate Discrimination (a = 1.0)									
Z	.977	.988	.993	.993	.993	.989	.976	.930	.668
LR	.977	.989	.994	.994	.993	.989	.974	.918	.713
ST	.976	.989	.994	.994	.993	.989	.974	.911	.724
KLD-U	.978	.989	.993	.993	.992	.987	.972	.890	.526
KLD-N	.977	.989	.993	.993	.991	.985	.963	.864	.504
Low Discrimination (a = 0.6)									
Z	.794	.852	.878	.886	.880	.860	.810	.721	.520
LR	.803	.864	.891	.897	.888	.861	.800	.698	.536
ST	.802	.865	.891	.898	.887	.859	.794	.688	.539
KLD-U	.803	.864	.888	.893	.881	.853	.779	.604	.299
KLD-N	.774	.842	.867	.869	.848	.800	.698	.512	.252

Table A87. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.082	.047	.044	.042	.043	.043	.046	.051	.059
LR	.051	.050	.052	.049	.050	.050	.050	.050	.049
ST	.048	.051	.053	.052	.052	.052	.052	.049	.045
KLD-U	.049	.058	.053	.049	.049	.048	.049	.052	.054
KLD-N	.052	.058	.056	.051	.051	.050	.049	.050	.046
Moderate Discrimination (a = 1.0)									
Z	.070	.049	.046	.043	.042	.044	.049	.051	.058
LR	.050	.050	.052	.051	.050	.049	.052	.049	.048
ST	.049	.050	.053	.053	.051	.050	.052	.048	.046
KLD-U	.042	.054	.053	.051	.049	.049	.053	.052	.046
KLD-N	.046	.052	.053	.052	.050	.048	.050	.045	.040
Low Discrimination (a = 0.6)									
Z	.059	.056	.048	.045	.042	.046	.050	.054	.057
LR	.051	.052	.052	.051	.049	.050	.050	.050	.050
ST	.052	.052	.052	.052	.050	.051	.050	.049	.048
KLD-U	.032	.051	.056	.056	.053	.055	.056	.053	.036
KLD-N	.033	.043	.046	.047	.043	.044	.042	.038	.029
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.141	.146	.158	.170	.171	.174	.170	.162	.142
LR	.119	.153	.172	.183	.185	.185	.174	.155	.119
ST	.113	.152	.174	.186	.188	.186	.175	.151	.107
KLD-U	.114	.154	.158	.165	.161	.162	.152	.136	.096
KLD-N	.114	.155	.166	.172	.168	.164	.151	.128	.081
Moderate Discrimination (a = 1.0)									
Z	.106	.099	.101	.110	.112	.111	.114	.112	.106
LR	.091	.102	.111	.121	.122	.119	.114	.105	.088
ST	.088	.101	.111	.121	.123	.118	.111	.100	.083
KLD-U	.076	.105	.104	.109	.108	.103	.098	.091	.065
KLD-N	.081	.103	.107	.112	.110	.100	.091	.079	.054
Low Discrimination (a = 0.6)									
Z	.078	.076	.075	.076	.078	.082	.089	.085	.083
LR	.071	.078	.082	.085	.086	.085	.086	.077	.071
ST	.070	.077	.082	.084	.085	.082	.082	.075	.067
KLD-U	.045	.076	.084	.084	.083	.080	.080	.069	.043
KLD-N	.049	.066	.072	.072	.070	.064	.059	.046	.030
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.395	.464	.504	.539	.549	.533	.508	.455	.320
LR	.378	.477	.525	.559	.568	.549	.514	.442	.286
ST	.370	.479	.529	.563	.573	.553	.516	.434	.265
KLD-U	.390	.467	.501	.530	.535	.514	.480	.411	.229
KLD-N	.381	.476	.516	.543	.544	.519	.476	.394	.197

Continued on the next page

Table A87 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.237	.270	.293	.317	.317	.320	.309	.278	.228
LR	.226	.281	.312	.337	.336	.332	.309	.264	.199
ST	.221	.280	.314	.338	.337	.332	.306	.258	.187
KLD-U	.216	.281	.295	.313	.309	.304	.282	.235	.144
KLD-N	.215	.279	.302	.319	.311	.299	.266	.210	.125
Low Discrimination (a = 0.6)									
Z	.140	.152	.164	.175	.184	.186	.180	.170	.146
LR	.135	.160	.177	.190	.196	.190	.176	.155	.127
ST	.134	.160	.177	.190	.195	.186	.172	.150	.122
KLD-U	.102	.160	.178	.186	.188	.179	.162	.134	.074
KLD-N	.104	.142	.158	.166	.163	.148	.124	.094	.053
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.936	.966	.977	.982	.983	.977	.964	.913	.608
LR	.934	.967	.979	.983	.984	.979	.964	.901	.620
ST	.932	.967	.979	.983	.984	.979	.964	.893	.602
KLD-U	.941	.967	.978	.981	.982	.976	.958	.885	.499
KLD-N	.939	.968	.979	.982	.982	.976	.956	.869	.452
Moderate Discrimination (a = 1.0)									
Z	.724	.792	.833	.854	.846	.832	.792	.708	.525
LR	.724	.804	.845	.865	.858	.841	.792	.689	.507
ST	.721	.804	.846	.867	.859	.840	.789	.680	.494
KLD-U	.733	.798	.834	.852	.842	.822	.769	.649	.358
KLD-N	.725	.801	.838	.855	.842	.814	.750	.611	.334
Low Discrimination (a = 0.6)									
Z	.428	.492	.534	.558	.561	.547	.508	.447	.341
LR	.432	.508	.556	.577	.577	.554	.500	.423	.323
ST	.431	.508	.556	.578	.576	.550	.494	.414	.317
KLD-U	.413	.514	.555	.571	.565	.537	.476	.360	.184
KLD-N	.391	.481	.524	.538	.522	.478	.402	.286	.148
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	1.000	1.000	1.000	1.000	.999	.990	.674
LR	.997	.999	1.000	1.000	1.000	1.000	.999	.988	.766
ST	.997	.999	1.000	1.000	1.000	1.000	.999	.985	.770
KLD-U	.998	1.000	1.000	1.000	1.000	1.000	.999	.982	.620
KLD-N	.998	.999	1.000	1.000	1.000	1.000	.999	.977	.576
Moderate Discrimination (a = 1.0)									
Z	.976	.988	.992	.993	.992	.987	.974	.927	.656
LR	.976	.989	.993	.993	.993	.988	.974	.919	.714
ST	.975	.988	.993	.994	.993	.988	.974	.912	.728
KLD-U	.978	.989	.992	.993	.991	.986	.969	.886	.517
KLD-N	.977	.989	.993	.993	.991	.985	.963	.867	.507
Low Discrimination (a = 0.6)									
Z	.791	.848	.876	.885	.878	.857	.807	.712	.497
LR	.797	.859	.887	.894	.885	.858	.798	.697	.528
ST	.795	.860	.887	.894	.884	.856	.793	.688	.540
KLD-U	.801	.862	.886	.891	.878	.848	.773	.597	.292
KLD-N	.775	.843	.870	.872	.852	.805	.707	.524	.261

Table A88. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.054	.050	.052	.049	.048	.048	.050	.050	.054
LR	.057	.053	.055	.050	.049	.048	.050	.047	.047
ST	.050	.049	.051	.049	.051	.050	.051	.050	.048
KLD-U	.055	.052	.052	.048	.047	.047	.050	.049	.049
KLD-N	.066	.055	.052	.045	.044	.042	.044	.042	.041
Moderate Discrimination (a = 1.0)									
Z	.050	.049	.048	.049	.046	.047	.050	.051	.056
LR	.055	.053	.050	.051	.048	.048	.049	.048	.048
ST	.051	.052	.050	.051	.049	.050	.050	.050	.049
KLD-U	.051	.052	.049	.050	.048	.048	.048	.050	.047
KLD-N	.059	.051	.045	.044	.040	.039	.039	.039	.038
Low Discrimination (a = 0.6)									
Z	.049	.048	.050	.050	.050	.048	.050	.055	.057
LR	.055	.052	.053	.052	.052	.049	.049	.050	.048
ST	.052	.050	.052	.052	.053	.050	.050	.051	.049
KLD-U	.044	.047	.049	.048	.049	.046	.049	.047	.036
KLD-N	.048	.042	.039	.036	.034	.031	.031	.031	.027
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.136	.136	.142	.142	.148	.151	.151	.159	.160
LR	.150	.150	.152	.150	.152	.152	.146	.151	.143
ST	.140	.143	.146	.146	.149	.150	.144	.148	.139
KLD-U	.122	.123	.126	.123	.126	.127	.126	.130	.120
KLD-N	.144	.136	.131	.121	.118	.112	.105	.104	.090
Moderate Discrimination (a = 1.0)									
Z	.091	.091	.095	.098	.099	.100	.103	.111	.115
LR	.104	.102	.102	.104	.102	.100	.100	.101	.095
ST	.097	.098	.100	.102	.101	.099	.097	.098	.093
KLD-U	.081	.082	.085	.086	.084	.083	.084	.087	.077
KLD-N	.099	.089	.085	.080	.072	.065	.061	.059	.049
Low Discrimination (a = 0.6)									
Z	.070	.073	.072	.072	.072	.076	.080	.084	.091
LR	.078	.079	.078	.076	.075	.076	.076	.074	.075
ST	.074	.077	.076	.075	.074	.075	.074	.072	.072
KLD-U	.054	.063	.063	.062	.061	.063	.063	.059	.049
KLD-N	.066	.061	.054	.047	.042	.039	.036	.031	.028
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.414	.422	.442	.456	.464	.466	.470	.465	.456
LR	.435	.441	.455	.466	.468	.466	.464	.452	.432
ST	.425	.434	.450	.464	.467	.464	.461	.450	.430
KLD-U	.386	.392	.407	.417	.422	.423	.426	.416	.388
KLD-N	.421	.413	.415	.413	.405	.392	.383	.358	.320

Continued on the next page

Table A88 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.243	.246	.259	.264	.264	.269	.275	.276	.283
LR	.261	.262	.269	.273	.270	.270	.268	.257	.254
ST	.255	.257	.266	.270	.269	.268	.265	.252	.250
KLD-U	.222	.223	.236	.237	.234	.236	.240	.227	.210
KLD-N	.247	.235	.233	.223	.207	.196	.188	.166	.147
Low Discrimination (a = 0.6)									
Z	.142	.147	.142	.151	.146	.151	.158	.169	.168
LR	.154	.157	.149	.156	.151	.152	.151	.152	.144
ST	.149	.154	.146	.154	.148	.150	.147	.148	.140
KLD-U	.116	.129	.123	.128	.123	.124	.126	.121	.091
KLD-N	.132	.126	.108	.103	.090	.082	.074	.067	.050
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.926	.938	.950	.959	.968	.965	.964	.959	.949
LR	.946	.953	.961	.963	.968	.964	.962	.956	.945
ST	.930	.941	.954	.961	.967	.964	.962	.956	.945
KLD-U	.925	.937	.948	.952	.959	.956	.954	.947	.921
KLD-N	.940	.944	.949	.949	.953	.946	.939	.923	.888
Moderate Discrimination (a = 1.0)									
Z	.731	.752	.765	.773	.776	.780	.779	.779	.759
LR	.750	.765	.772	.778	.779	.780	.773	.765	.739
ST	.743	.761	.769	.777	.778	.777	.770	.761	.736
KLD-U	.706	.726	.738	.743	.744	.749	.745	.732	.637
KLD-N	.728	.734	.730	.721	.706	.694	.673	.641	.540
Low Discrimination (a = 0.6)									
Z	.441	.446	.453	.452	.458	.470	.481	.481	.474
LR	.456	.458	.463	.462	.465	.468	.468	.454	.447
ST	.450	.455	.461	.460	.463	.463	.462	.448	.442
KLD-U	.400	.412	.414	.412	.414	.422	.420	.386	.286
KLD-N	.409	.395	.379	.354	.334	.320	.298	.261	.192
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.980	.984	.990	.999	1.000	1.000	1.000	1.000	.999
LR	.998	.999	1.000	1.000	1.000	1.000	1.000	.999	.999
ST	.926	.930	.950	.982	.999	1.000	1.000	.999	.999
KLD-U	.997	.998	.999	1.000	1.000	1.000	1.000	.999	.991
KLD-N	.998	.999	.999	.999	1.000	1.000	.999	.999	.986
Moderate Discrimination (a = 1.0)									
Z	.966	.973	.980	.983	.982	.983	.981	.979	.972
LR	.974	.978	.982	.983	.982	.982	.980	.977	.969
ST	.965	.972	.980	.983	.982	.982	.979	.977	.969
KLD-U	.966	.972	.976	.978	.978	.978	.976	.968	.870
KLD-N	.970	.973	.973	.972	.969	.966	.958	.943	.820
Low Discrimination (a = 0.6)									
Z	.784	.796	.795	.801	.804	.809	.812	.810	.783
LR	.793	.802	.801	.806	.806	.806	.803	.794	.772
ST	.788	.799	.800	.804	.803	.802	.798	.789	.769
KLD-U	.755	.766	.765	.767	.770	.773	.763	.707	.502
KLD-N	.752	.746	.729	.710	.689	.672	.640	.574	.405

Table A89. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.048	.048	.050	.049	.050	.049	.051	.051	.053
LR	.050	.050	.051	.049	.050	.049	.051	.050	.049
ST	.050	.050	.051	.049	.050	.049	.050	.049	.048
KLD-U	.055	.052	.052	.048	.047	.046	.048	.048	.049
KLD-N	.061	.055	.052	.048	.047	.045	.046	.045	.045
Moderate Discrimination (a = 1.0)									
Z	.049	.049	.049	.051	.048	.048	.050	.053	.054
LR	.051	.052	.050	.052	.049	.050	.051	.051	.050
ST	.052	.052	.050	.052	.049	.050	.050	.050	.049
KLD-U	.053	.053	.050	.051	.048	.047	.050	.051	.048
KLD-N	.058	.054	.049	.049	.044	.043	.044	.044	.041
Low Discrimination (a = 0.6)									
Z	.050	.048	.050	.050	.049	.049	.050	.053	.053
LR	.052	.049	.051	.051	.051	.049	.050	.050	.049
ST	.052	.050	.052	.051	.051	.050	.050	.050	.048
KLD-U	.047	.051	.053	.052	.051	.050	.051	.051	.039
KLD-N	.048	.044	.043	.039	.038	.035	.035	.035	.031
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.146	.152	.155	.157	.160	.161	.157	.168	.163
LR	.151	.154	.158	.157	.160	.162	.156	.161	.154
ST	.152	.154	.156	.155	.157	.159	.153	.157	.151
KLD-U	.140	.140	.138	.135	.134	.134	.130	.137	.130
KLD-N	.149	.146	.143	.138	.133	.129	.121	.126	.117
Moderate Discrimination (a = 1.0)									
Z	.099	.099	.103	.107	.105	.104	.104	.113	.113
LR	.103	.102	.104	.106	.105	.105	.103	.106	.102
ST	.102	.101	.103	.105	.104	.104	.101	.104	.100
KLD-U	.091	.092	.093	.094	.090	.089	.087	.091	.083
KLD-N	.100	.095	.093	.092	.084	.079	.074	.074	.068
Low Discrimination (a = 0.6)									
Z	.076	.076	.076	.075	.074	.076	.078	.082	.084
LR	.078	.078	.077	.076	.075	.077	.076	.076	.076
ST	.077	.078	.077	.076	.075	.075	.074	.075	.074
KLD-U	.063	.073	.071	.070	.068	.068	.068	.065	.055
KLD-N	.068	.066	.060	.055	.050	.047	.044	.040	.037
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.452	.462	.481	.485	.491	.485	.487	.483	.467
LR	.459	.466	.482	.485	.490	.485	.484	.474	.458
ST	.460	.466	.480	.482	.488	.483	.482	.472	.456
KLD-U	.437	.437	.449	.445	.446	.440	.443	.437	.416
KLD-N	.450	.450	.459	.449	.445	.431	.427	.414	.392

Continued on the next page

Table A89 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.263	.269	.278	.284	.280	.280	.284	.285	.282
LR	.269	.275	.279	.285	.283	.282	.282	.274	.268
ST	.270	.274	.277	.282	.280	.279	.278	.270	.266
KLD-U	.251	.253	.257	.258	.250	.249	.252	.243	.228
KLD-N	.261	.258	.258	.252	.238	.228	.224	.210	.199
Low Discrimination (a = 0.6)									
Z	.152	.158	.153	.159	.154	.156	.160	.165	.162
LR	.154	.159	.154	.161	.157	.158	.158	.156	.151
ST	.154	.158	.153	.161	.156	.156	.154	.154	.150
KLD-U	.134	.148	.142	.147	.140	.141	.139	.134	.109
KLD-N	.137	.136	.123	.121	.108	.103	.095	.090	.073
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.952	.961	.966	.967	.971	.967	.967	.961	.952
LR	.952	.959	.965	.967	.970	.967	.967	.960	.952
ST	.952	.958	.964	.966	.970	.967	.967	.960	.952
KLD-U	.951	.956	.960	.960	.963	.960	.960	.953	.937
KLD-N	.952	.958	.961	.960	.962	.957	.955	.947	.932
Moderate Discrimination (a = 1.0)									
Z	.766	.778	.785	.790	.787	.788	.790	.786	.767
LR	.770	.781	.785	.791	.790	.789	.788	.781	.761
ST	.768	.780	.783	.789	.788	.788	.786	.779	.761
KLD-U	.753	.761	.765	.766	.762	.762	.763	.753	.684
KLD-N	.758	.765	.763	.758	.746	.738	.731	.714	.661
Low Discrimination (a = 0.6)									
Z	.465	.467	.473	.468	.471	.475	.486	.480	.474
LR	.468	.469	.476	.474	.476	.478	.482	.468	.464
ST	.468	.468	.474	.472	.475	.476	.479	.466	.464
KLD-U	.445	.452	.454	.447	.448	.448	.451	.418	.335
KLD-N	.434	.423	.414	.394	.382	.368	.356	.325	.280
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	1.000	1.000	1.000	1.000	1.000	1.000	.999
LR	.998	.999	1.000	1.000	1.000	1.000	1.000	1.000	.999
ST	.998	.998	.999	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	.999	.999	1.000	1.000	1.000	1.000	1.000	.999	.997
KLD-N	.999	.999	1.000	1.000	1.000	1.000	1.000	.999	.998
Moderate Discrimination (a = 1.0)									
Z	.978	.980	.982	.984	.983	.983	.982	.980	.974
LR	.978	.980	.982	.984	.983	.983	.982	.980	.974
ST	.977	.980	.982	.983	.983	.982	.982	.980	.974
KLD-U	.976	.978	.980	.980	.979	.979	.979	.973	.917
KLD-N	.976	.978	.979	.978	.977	.975	.972	.967	.924
Low Discrimination (a = 0.6)									
Z	.799	.809	.807	.807	.807	.813	.814	.811	.788
LR	.800	.811	.809	.811	.810	.814	.811	.806	.790
ST	.799	.809	.808	.810	.810	.812	.810	.805	.791
KLD-U	.789	.799	.794	.792	.792	.795	.787	.746	.583
KLD-N	.773	.774	.761	.746	.734	.725	.707	.672	.556

Table A90. Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.048	.047	.051	.048	.049	.048	.051	.053	.052
LR	.050	.050	.051	.049	.049	.048	.050	.050	.049
ST	.049	.050	.051	.049	.049	.048	.050	.050	.048
KLD-U	.056	.054	.052	.047	.046	.045	.048	.049	.050
KLD-N	.063	.057	.054	.049	.047	.046	.048	.048	.047
Moderate Discrimination (a = 1.0)									
Z	.050	.048	.049	.050	.047	.049	.051	.054	.053
LR	.052	.052	.050	.051	.049	.050	.050	.052	.050
ST	.052	.051	.050	.051	.049	.050	.050	.051	.049
KLD-U	.054	.054	.052	.051	.048	.048	.049	.052	.048
KLD-N	.061	.056	.053	.050	.046	.046	.046	.047	.045
Low Discrimination (a = 0.6)									
Z	.048	.048	.049	.050	.049	.049	.050	.054	.052
LR	.051	.050	.051	.052	.051	.051	.050	.052	.049
ST	.051	.050	.051	.052	.051	.051	.050	.051	.049
KLD-U	.046	.052	.053	.053	.052	.051	.052	.052	.040
KLD-N	.050	.047	.045	.042	.039	.038	.038	.038	.033
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.149	.151	.158	.158	.159	.160	.160	.171	.163
LR	.150	.152	.158	.156	.157	.158	.156	.160	.155
ST	.151	.152	.157	.155	.155	.156	.153	.158	.152
KLD-U	.144	.141	.142	.136	.135	.134	.132	.139	.132
KLD-N	.153	.149	.149	.141	.137	.134	.129	.133	.126
Moderate Discrimination (a = 1.0)									
Z	.101	.101	.103	.107	.107	.105	.107	.116	.112
LR	.104	.104	.104	.107	.107	.105	.104	.108	.103
ST	.104	.104	.103	.105	.104	.103	.101	.105	.101
KLD-U	.096	.096	.095	.096	.093	.091	.090	.094	.085
KLD-N	.105	.100	.098	.096	.090	.085	.081	.084	.077
Low Discrimination (a = 0.6)									
Z	.075	.077	.077	.076	.074	.077	.080	.083	.084
LR	.077	.078	.078	.076	.075	.077	.077	.077	.076
ST	.076	.077	.077	.075	.075	.076	.076	.076	.076
KLD-U	.063	.073	.074	.072	.069	.069	.069	.067	.056
KLD-N	.071	.068	.064	.058	.053	.051	.050	.046	.043
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.452	.462	.483	.486	.491	.483	.488	.485	.463
LR	.455	.464	.481	.484	.488	.480	.482	.473	.455
ST	.457	.463	.478	.482	.486	.478	.480	.471	.451
KLD-U	.441	.441	.452	.449	.448	.439	.445	.439	.416
KLD-N	.456	.456	.465	.458	.453	.437	.439	.428	.406

Continued on the next page

Table A90 (continued). Mean Type I Error Rates and Power for 30-Item Variable-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.266	.272	.279	.284	.283	.282	.289	.286	.282
LR	.271	.276	.280	.284	.283	.282	.284	.277	.270
ST	.270	.274	.277	.280	.280	.280	.279	.272	.268
KLD-U	.255	.258	.261	.261	.256	.254	.256	.248	.232
KLD-N	.267	.266	.265	.260	.250	.242	.239	.227	.216
Low Discrimination (a = 0.6)									
Z	.152	.158	.155	.160	.155	.160	.166	.168	.163
LR	.155	.160	.156	.162	.158	.161	.160	.158	.152
ST	.154	.159	.154	.160	.156	.159	.158	.157	.152
KLD-U	.134	.150	.147	.150	.143	.144	.142	.136	.111
KLD-N	.141	.141	.130	.128	.116	.110	.104	.100	.086
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.951	.959	.964	.966	.968	.966	.966	.958	.950
LR	.950	.957	.962	.965	.968	.966	.964	.957	.950
ST	.949	.956	.961	.964	.967	.965	.964	.957	.950
KLD-U	.950	.955	.958	.958	.961	.958	.958	.950	.934
KLD-N	.952	.957	.960	.960	.961	.958	.957	.948	.934
Moderate Discrimination (a = 1.0)									
Z	.762	.775	.784	.787	.786	.786	.790	.783	.762
LR	.766	.777	.784	.788	.787	.788	.786	.778	.757
ST	.764	.776	.781	.786	.785	.785	.784	.776	.758
KLD-U	.754	.763	.768	.766	.762	.762	.764	.751	.685
KLD-N	.760	.768	.769	.764	.754	.748	.746	.729	.686
Low Discrimination (a = 0.6)									
Z	.465	.469	.473	.472	.472	.478	.488	.478	.472
LR	.469	.471	.476	.474	.476	.479	.482	.469	.465
ST	.468	.468	.473	.473	.474	.476	.479	.467	.465
KLD-U	.448	.456	.458	.453	.450	.451	.450	.421	.341
KLD-N	.440	.432	.425	.409	.396	.386	.378	.352	.312
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.998	.999	.999	1.000	1.000	1.000	1.000	1.000	.999
LR	.998	.999	.999	1.000	1.000	1.000	1.000	1.000	.999
ST	.998	.998	.999	1.000	1.000	1.000	1.000	1.000	.999
KLD-U	.999	.999	.999	1.000	1.000	1.000	1.000	.999	.997
KLD-N	.999	.999	.999	1.000	1.000	1.000	1.000	.999	.998
Moderate Discrimination (a = 1.0)									
Z	.976	.979	.982	.983	.981	.982	.981	.978	.972
LR	.976	.979	.981	.982	.981	.982	.980	.978	.972
ST	.975	.978	.980	.982	.981	.981	.980	.978	.972
KLD-U	.975	.978	.979	.979	.977	.979	.977	.972	.922
KLD-N	.976	.978	.979	.979	.976	.976	.973	.969	.936
Low Discrimination (a = 0.6)									
Z	.797	.806	.804	.807	.807	.813	.812	.807	.785
LR	.798	.808	.806	.809	.810	.813	.809	.803	.787
ST	.797	.805	.804	.808	.808	.812	.808	.803	.788
KLD-U	.790	.798	.793	.793	.791	.795	.786	.745	.596
KLD-N	.776	.776	.764	.756	.745	.740	.726	.698	.604

Table A91. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.084	.053	.043	.040	.039	.040	.042	.050	.061
LR	.052	.054	.054	.051	.049	.050	.048	.047	.045
ST	.041	.049	.053	.053	.054	.055	.052	.050	.045
KLD-U	.044	.056	.048	.044	.045	.048	.047	.051	.059
KLD-N	.048	.058	.051	.046	.046	.048	.046	.049	.051
Moderate Discrimination (a = 1.0)									
Z	.077	.054	.043	.041	.039	.042	.044	.054	.059
LR	.053	.053	.053	.052	.050	.051	.048	.049	.045
ST	.045	.049	.051	.054	.053	.054	.050	.051	.047
KLD-U	.042	.053	.049	.049	.047	.049	.049	.054	.053
KLD-N	.047	.052	.049	.048	.046	.047	.046	.048	.046
Low Discrimination (a = 0.6)									
Z	.060	.057	.047	.042	.041	.042	.046	.055	.057
LR	.054	.052	.052	.052	.051	.050	.048	.049	.046
ST	.052	.050	.051	.053	.052	.051	.049	.049	.047
KLD-U	.035	.048	.053	.053	.053	.053	.053	.056	.043
KLD-N	.038	.043	.046	.047	.046	.045	.044	.044	.037
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.149	.185	.206	.221	.225	.234	.219	.222	.177
LR	.138	.204	.239	.254	.254	.259	.232	.203	.136
ST	.125	.197	.239	.256	.258	.260	.231	.194	.122
KLD-U	.132	.195	.211	.222	.227	.233	.214	.204	.132
KLD-N	.135	.204	.224	.231	.230	.231	.206	.192	.111
Moderate Discrimination (a = 1.0)									
Z	.114	.116	.125	.133	.139	.139	.143	.150	.139
LR	.111	.132	.149	.157	.163	.154	.147	.132	.104
ST	.101	.126	.147	.156	.162	.151	.141	.124	.096
KLD-U	.096	.125	.132	.138	.144	.139	.135	.128	.096
KLD-N	.104	.129	.138	.141	.141	.131	.123	.111	.078
Low Discrimination (a = 0.6)									
Z	.075	.079	.080	.082	.086	.091	.098	.103	.098
LR	.080	.088	.097	.101	.104	.102	.096	.086	.076
ST	.077	.086	.095	.099	.101	.098	.091	.081	.072
KLD-U	.054	.083	.092	.094	.096	.096	.094	.084	.062
KLD-N	.064	.079	.086	.086	.085	.081	.074	.062	.047
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.464	.621	.685	.718	.731	.721	.686	.609	.426
LR	.460	.646	.716	.748	.756	.742	.696	.580	.360
ST	.449	.645	.719	.752	.760	.744	.695	.567	.334
KLD-U	.478	.630	.688	.719	.731	.718	.682	.585	.339
KLD-N	.473	.645	.702	.728	.734	.716	.673	.567	.294

Continued on the next page

Table A91 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Peaked Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.292	.364	.411	.442	.456	.452	.430	.393	.317
LR	.301	.394	.448	.477	.489	.473	.432	.362	.260
ST	.291	.389	.446	.478	.488	.470	.424	.347	.242
KLD-U	.297	.380	.420	.450	.463	.450	.418	.358	.232
KLD-N	.299	.389	.430	.453	.457	.437	.395	.325	.198
Low Discrimination (a = 0.6)									
Z	.155	.188	.208	.228	.232	.242	.244	.228	.195
LR	.172	.215	.240	.260	.259	.258	.237	.202	.163
ST	.169	.212	.238	.257	.255	.250	.227	.191	.152
KLD-U	.140	.209	.232	.248	.246	.247	.232	.194	.126
KLD-N	.149	.200	.220	.234	.225	.217	.194	.153	.096
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.986	.994	.998	.999	1.000	.999	.997	.979	.788
LR	.988	.998	.999	.999	1.000	.999	.997	.973	.756
ST	.987	.998	.999	.999	1.000	.999	.997	.970	.729
KLD-U	.988	.998	.999	.999	1.000	.999	.997	.974	.685
KLD-N	.989	.998	.999	.999	1.000	.999	.997	.969	.635
Moderate Discrimination (a = 1.0)									
Z	.873	.936	.958	.966	.964	.956	.932	.871	.724
LR	.882	.945	.965	.970	.969	.959	.931	.852	.669
ST	.878	.944	.965	.970	.968	.958	.928	.842	.644
KLD-U	.882	.938	.960	.966	.964	.956	.929	.849	.575
KLD-N	.883	.942	.962	.966	.962	.952	.919	.824	.535
Low Discrimination (a = 0.6)									
Z	.552	.648	.704	.735	.741	.720	.682	.609	.486
LR	.584	.679	.735	.762	.763	.732	.675	.580	.448
ST	.580	.677	.734	.760	.759	.726	.664	.564	.430
KLD-U	.570	.677	.728	.752	.755	.728	.673	.559	.323
KLD-N	.560	.664	.716	.736	.729	.690	.619	.490	.279
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	.960	.971	.997	1.000	1.000	1.000	1.000	.999	.868
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.898
ST	.989	.989	.996	1.000	1.000	1.000	1.000	.998	.890
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.814
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.776
Moderate Discrimination (a = 1.0)									
Z	.996	.998	1.000	1.000	1.000	1.000	.998	.987	.892
LR	.999	1.000	1.000	1.000	1.000	1.000	.998	.983	.883
ST	.998	1.000	1.000	1.000	1.000	1.000	.998	.981	.870
KLD-U	.998	1.000	1.000	1.000	1.000	1.000	.998	.980	.756
KLD-N	.998	1.000	1.000	1.000	1.000	1.000	.997	.974	.740
Low Discrimination (a = 0.6)									
Z	.921	.958	.972	.974	.972	.963	.935	.873	.746
LR	.932	.965	.977	.978	.975	.965	.932	.856	.729
ST	.931	.964	.976	.978	.974	.963	.927	.847	.714
KLD-U	.930	.964	.976	.977	.974	.965	.930	.822	.517
KLD-N	.926	.961	.973	.973	.968	.954	.905	.778	.492

Table A92. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.077	.048	.042	.042	.040	.043	.044	.050	.059
LR	.046	.051	.051	.052	.049	.053	.052	.051	.050
ST	.042	.050	.052	.054	.051	.054	.053	.050	.046
KLD-U	.042	.055	.048	.048	.045	.049	.050	.052	.058
KLD-N	.041	.056	.052	.051	.048	.052	.050	.051	.050
Moderate Discrimination (a = 1.0)									
Z	.072	.052	.044	.043	.042	.044	.047	.054	.058
LR	.049	.051	.050	.051	.050	.051	.051	.052	.048
ST	.047	.050	.051	.052	.052	.052	.050	.051	.046
KLD-U	.040	.053	.050	.048	.048	.049	.049	.053	.051
KLD-N	.042	.052	.051	.050	.049	.049	.048	.050	.045
Low Discrimination (a = 0.6)									
Z	.061	.055	.047	.045	.044	.046	.048	.053	.057
LR	.051	.050	.051	.052	.052	.052	.049	.050	.048
ST	.051	.049	.052	.053	.053	.052	.050	.049	.047
KLD-U	.035	.049	.054	.054	.053	.054	.053	.055	.044
KLD-N	.036	.043	.047	.048	.048	.047	.045	.044	.037
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.147	.195	.222	.237	.240	.245	.229	.221	.173
LR	.135	.210	.250	.267	.267	.270	.246	.216	.146
ST	.127	.206	.249	.267	.267	.268	.243	.206	.131
KLD-U	.130	.207	.230	.242	.243	.243	.223	.202	.129
KLD-N	.128	.214	.242	.255	.251	.249	.223	.195	.112
Moderate Discrimination (a = 1.0)									
Z	.115	.122	.135	.145	.152	.148	.152	.149	.137
LR	.107	.132	.153	.163	.168	.160	.156	.139	.111
ST	.103	.129	.151	.162	.166	.158	.150	.131	.102
KLD-U	.097	.132	.141	.148	.152	.144	.140	.127	.094
KLD-N	.099	.133	.148	.154	.154	.142	.134	.116	.080
Low Discrimination (a = 0.6)									
Z	.078	.082	.084	.090	.093	.098	.102	.100	.096
LR	.078	.087	.097	.102	.105	.106	.099	.089	.080
ST	.078	.086	.096	.101	.104	.102	.095	.084	.075
KLD-U	.055	.087	.095	.099	.099	.099	.093	.082	.062
KLD-N	.061	.080	.090	.093	.091	.087	.078	.064	.048
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.473	.643	.706	.738	.747	.735	.695	.607	.423
LR	.473	.666	.734	.763	.772	.757	.713	.600	.379
ST	.463	.664	.736	.764	.772	.757	.710	.586	.351
KLD-U	.484	.657	.716	.744	.750	.735	.691	.583	.336
KLD-N	.476	.670	.730	.754	.758	.740	.691	.572	.298

Continued on the next page

Table A92 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Peaked Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.304	.385	.433	.465	.477	.468	.440	.393	.315
LR	.305	.406	.460	.492	.502	.487	.447	.378	.275
ST	.300	.402	.458	.491	.500	.483	.439	.365	.257
KLD-U	.306	.401	.442	.470	.479	.461	.423	.358	.232
KLD-N	.302	.406	.453	.479	.483	.458	.412	.336	.206
Low Discrimination (a = 0.6)									
Z	.164	.199	.222	.240	.246	.251	.247	.227	.195
LR	.173	.217	.244	.265	.267	.263	.244	.209	.169
ST	.171	.215	.243	.263	.265	.257	.236	.200	.160
KLD-U	.148	.219	.242	.256	.257	.250	.232	.194	.126
KLD-N	.151	.207	.232	.246	.240	.228	.203	.159	.103
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.989	.998	.999	.999	.999	.999	.997	.979	.788
LR	.990	.998	.999	1.000	1.000	.999	.997	.976	.774
ST	.989	.998	.999	1.000	1.000	.999	.997	.973	.747
KLD-U	.991	.998	.999	.999	.999	.999	.997	.974	.682
KLD-N	.991	.999	.999	1.000	.999	.999	.997	.969	.641
Moderate Discrimination (a = 1.0)									
Z	.884	.943	.962	.968	.967	.959	.935	.874	.725
LR	.889	.949	.967	.972	.971	.963	.936	.862	.684
ST	.886	.948	.966	.972	.971	.962	.933	.853	.663
KLD-U	.892	.947	.964	.969	.968	.958	.930	.849	.576
KLD-N	.891	.949	.966	.971	.968	.957	.925	.831	.550
Low Discrimination (a = 0.6)									
Z	.567	.663	.720	.746	.751	.727	.688	.610	.487
LR	.586	.686	.743	.768	.770	.738	.685	.588	.459
ST	.585	.686	.743	.767	.768	.733	.678	.578	.445
KLD-U	.584	.690	.741	.762	.761	.729	.675	.558	.328
KLD-N	.567	.676	.730	.750	.744	.700	.630	.502	.297
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.870
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.908
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.901
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.812
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.780
Moderate Discrimination (a = 1.0)									
Z	.998	1.000	1.000	1.000	1.000	1.000	.998	.988	.894
LR	.999	1.000	1.000	1.000	1.000	1.000	.998	.985	.892
ST	.998	1.000	1.000	1.000	1.000	1.000	.998	.983	.882
KLD-U	.999	1.000	1.000	1.000	1.000	1.000	.998	.980	.759
KLD-N	.999	1.000	1.000	1.000	1.000	1.000	.998	.976	.755
Low Discrimination (a = 0.6)									
Z	.927	.961	.974	.976	.973	.964	.936	.875	.750
LR	.934	.966	.977	.979	.976	.966	.935	.863	.740
ST	.934	.966	.977	.979	.976	.966	.932	.856	.728
KLD-U	.938	.966	.977	.979	.975	.965	.930	.825	.527
KLD-N	.930	.964	.975	.976	.971	.956	.910	.790	.520

Table A93. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.077	.049	.044	.041	.040	.045	.044	.051	.058
LR	.046	.051	.052	.051	.048	.053	.050	.051	.050
ST	.043	.051	.053	.053	.050	.054	.052	.051	.046
KLD-U	.040	.054	.050	.048	.046	.050	.048	.052	.056
KLD-N	.040	.055	.054	.051	.048	.052	.049	.052	.050
Moderate Discrimination (a = 1.0)									
Z	.070	.050	.044	.043	.043	.045	.048	.053	.057
LR	.048	.050	.050	.050	.051	.051	.051	.052	.049
ST	.046	.049	.051	.052	.053	.052	.051	.050	.047
KLD-U	.040	.052	.049	.049	.049	.049	.050	.053	.050
KLD-N	.042	.052	.051	.051	.051	.051	.049	.050	.046
Low Discrimination (a = 0.6)									
Z	.060	.055	.048	.046	.044	.046	.048	.052	.055
LR	.050	.050	.052	.053	.051	.052	.050	.049	.048
ST	.051	.050	.053	.054	.052	.052	.050	.048	.047
KLD-U	.035	.049	.054	.055	.053	.054	.052	.053	.042
KLD-N	.037	.044	.049	.050	.048	.048	.046	.044	.038
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.148	.196	.225	.242	.245	.250	.234	.220	.172
LR	.135	.210	.250	.269	.267	.270	.247	.216	.147
ST	.129	.206	.249	.269	.267	.268	.244	.207	.133
KLD-U	.127	.208	.235	.248	.246	.248	.225	.200	.127
KLD-N	.125	.214	.247	.260	.254	.253	.227	.194	.110
Moderate Discrimination (a = 1.0)									
Z	.114	.122	.137	.146	.152	.151	.152	.148	.133
LR	.106	.131	.151	.162	.167	.162	.154	.140	.110
ST	.103	.128	.150	.161	.166	.159	.149	.133	.102
KLD-U	.095	.132	.144	.150	.153	.147	.138	.126	.092
KLD-N	.099	.134	.151	.158	.158	.148	.135	.118	.081
Low Discrimination (a = 0.6)									
Z	.079	.082	.086	.090	.093	.100	.102	.100	.094
LR	.078	.088	.096	.103	.103	.106	.100	.089	.079
ST	.078	.086	.095	.102	.101	.103	.096	.084	.075
KLD-U	.056	.088	.096	.100	.098	.099	.094	.082	.060
KLD-N	.063	.082	.091	.095	.092	.089	.081	.066	.050
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.475	.646	.710	.738	.749	.737	.696	.607	.422
LR	.474	.665	.734	.762	.770	.756	.710	.600	.381
ST	.464	.663	.736	.763	.770	.756	.708	.589	.355
KLD-U	.482	.660	.721	.745	.752	.737	.690	.580	.332
KLD-N	.473	.670	.733	.756	.760	.743	.692	.570	.294

Continued on the next page

Table A93 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Peaked Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.304	.383	.433	.465	.475	.467	.440	.389	.312
LR	.305	.402	.458	.490	.498	.484	.446	.376	.276
ST	.301	.399	.458	.490	.497	.481	.439	.365	.261
KLD-U	.306	.401	.446	.473	.478	.462	.424	.353	.229
KLD-N	.304	.409	.458	.485	.487	.464	.418	.336	.206
Low Discrimination (a = 0.6)									
Z	.165	.201	.226	.244	.248	.253	.246	.223	.192
LR	.172	.216	.245	.266	.267	.264	.244	.208	.170
ST	.172	.215	.244	.264	.264	.258	.237	.200	.162
KLD-U	.150	.221	.246	.260	.257	.252	.232	.192	.125
KLD-N	.152	.209	.237	.252	.245	.233	.206	.160	.105
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.989	.998	.999	.999	.999	.999	.997	.979	.788
LR	.990	.998	.999	1.000	1.000	.999	.997	.976	.776
ST	.989	.998	.999	1.000	1.000	.999	.997	.974	.750
KLD-U	.991	.998	.999	1.000	.999	.999	.997	.973	.677
KLD-N	.991	.998	.999	1.000	1.000	.999	.997	.969	.635
Moderate Discrimination (a = 1.0)									
Z	.883	.942	.960	.968	.966	.958	.932	.871	.721
LR	.888	.947	.965	.972	.970	.961	.934	.861	.685
ST	.886	.946	.965	.971	.970	.960	.933	.854	.667
KLD-U	.893	.946	.963	.969	.967	.957	.928	.846	.572
KLD-N	.893	.949	.965	.971	.968	.957	.925	.831	.552
Low Discrimination (a = 0.6)									
Z	.570	.664	.722	.747	.752	.728	.687	.607	.486
LR	.587	.685	.742	.766	.768	.737	.684	.589	.461
ST	.585	.684	.742	.765	.766	.733	.678	.579	.448
KLD-U	.587	.692	.743	.762	.760	.728	.672	.556	.326
KLD-N	.572	.679	.735	.753	.746	.704	.634	.505	.304
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.870
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.908
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.903
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.808
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.776
Moderate Discrimination (a = 1.0)									
Z	.998	1.000	1.000	1.000	1.000	1.000	.998	.987	.891
LR	.999	1.000	1.000	1.000	1.000	1.000	.998	.985	.892
ST	.998	1.000	1.000	1.000	1.000	1.000	.998	.984	.884
KLD-U	.999	1.000	1.000	1.000	1.000	1.000	.998	.980	.755
KLD-N	.999	1.000	1.000	1.000	1.000	1.000	.998	.976	.756
Low Discrimination (a = 0.6)									
Z	.927	.961	.973	.976	.973	.964	.936	.874	.748
LR	.933	.964	.976	.978	.975	.966	.934	.864	.741
ST	.932	.964	.976	.978	.975	.965	.932	.858	.732
KLD-U	.937	.965	.976	.978	.974	.964	.929	.823	.526
KLD-N	.930	.963	.974	.976	.971	.957	.912	.792	.529

Table A94. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.056	.050	.048	.047	.048	.048	.048	.049	.055
LR	.056	.052	.050	.049	.048	.049	.047	.047	.049
ST	.050	.049	.049	.049	.050	.051	.050	.049	.052
KLD-U	.052	.051	.050	.047	.048	.049	.048	.048	.051
KLD-N	.063	.055	.050	.046	.046	.046	.045	.045	.049
Moderate Discrimination (a = 1.0)									
Z	.052	.052	.049	.048	.048	.049	.050	.053	.057
LR	.055	.055	.052	.050	.049	.049	.049	.049	.048
ST	.051	.052	.051	.051	.050	.051	.050	.050	.050
KLD-U	.052	.053	.052	.050	.049	.050	.050	.051	.050
KLD-N	.060	.054	.049	.046	.045	.044	.044	.045	.044
Low Discrimination (a = 0.6)									
Z	.048	.050	.048	.048	.048	.049	.052	.054	.058
LR	.054	.053	.052	.050	.050	.050	.052	.049	.048
ST	.051	.052	.052	.051	.051	.051	.052	.049	.049
KLD-U	.049	.052	.052	.049	.050	.050	.052	.049	.045
KLD-N	.056	.050	.045	.041	.040	.040	.041	.040	.039
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.201	.203	.204	.202	.208	.212	.216	.212	.227
LR	.210	.211	.211	.207	.210	.212	.212	.203	.208
ST	.201	.204	.206	.204	.208	.209	.208	.198	.202
KLD-U	.193	.197	.194	.190	.196	.196	.199	.192	.196
KLD-N	.216	.210	.200	.190	.192	.187	.185	.174	.176
Moderate Discrimination (a = 1.0)									
Z	.123	.127	.126	.129	.130	.136	.131	.144	.148
LR	.132	.135	.132	.133	.132	.134	.126	.131	.128
ST	.126	.128	.128	.130	.130	.131	.121	.126	.123
KLD-U	.119	.123	.120	.122	.123	.125	.117	.122	.117
KLD-N	.137	.130	.121	.117	.113	.110	.099	.100	.097
Low Discrimination (a = 0.6)									
Z	.082	.083	.086	.086	.085	.088	.091	.100	.106
LR	.091	.089	.093	.090	.089	.089	.086	.088	.087
ST	.086	.086	.089	.088	.087	.086	.084	.084	.083
KLD-U	.078	.081	.084	.081	.080	.080	.079	.079	.072
KLD-N	.091	.081	.078	.070	.065	.061	.059	.057	.052
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.626	.630	.632	.644	.651	.651	.659	.649	.649
LR	.632	.636	.635	.646	.650	.648	.652	.638	.631
ST	.627	.631	.632	.644	.648	.647	.649	.634	.626
KLD-U	.616	.621	.617	.628	.634	.630	.639	.626	.618
KLD-N	.638	.634	.624	.628	.628	.617	.619	.601	.590

Continued on the next page

Table A94 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Flat Bank Using Fisher Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.368	.376	.381	.387	.384	.399	.394	.409	.406
LR	.378	.384	.385	.390	.385	.395	.385	.390	.377
ST	.371	.378	.380	.387	.381	.390	.379	.383	.370
KLD-U	.360	.368	.369	.373	.369	.379	.372	.377	.358
KLD-N	.382	.377	.367	.363	.349	.351	.336	.333	.312
Low Discrimination (a = 0.6)									
Z	.199	.198	.202	.205	.203	.205	.214	.225	.229
LR	.210	.207	.210	.212	.208	.206	.208	.206	.203
ST	.203	.201	.205	.207	.204	.201	.201	.198	.195
KLD-U	.190	.193	.196	.196	.190	.189	.193	.190	.169
KLD-N	.206	.192	.184	.175	.161	.154	.150	.143	.130
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.996	.996	.996	.997	.997	.997	.997	.997	.996
LR	.996	.996	.996	.997	.997	.997	.997	.996	.996
ST	.995	.996	.996	.997	.997	.997	.997	.996	.995
KLD-U	.995	.996	.996	.997	.997	.997	.997	.996	.995
KLD-N	.996	.996	.996	.996	.996	.996	.996	.995	.993
Moderate Discrimination (a = 1.0)									
Z	.915	.918	.922	.922	.926	.929	.926	.926	.922
LR	.917	.918	.923	.922	.924	.927	.923	.919	.914
ST	.915	.916	.921	.921	.923	.925	.921	.917	.912
KLD-U	.912	.913	.917	.917	.919	.922	.919	.914	.892
KLD-N	.917	.914	.915	.911	.910	.908	.902	.893	.870
Low Discrimination (a = 0.6)									
Z	.626	.627	.630	.629	.626	.636	.649	.649	.641
LR	.634	.634	.636	.634	.630	.636	.640	.629	.622
ST	.629	.629	.632	.630	.627	.630	.632	.622	.615
KLD-U	.618	.619	.619	.615	.610	.617	.622	.605	.524
KLD-N	.622	.608	.596	.579	.560	.553	.550	.525	.468
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	.997	.998	.999	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.999	.999	.999	.999	.999	.999	.999	.999	.998
LR	.999	.999	.999	.999	.999	.999	.999	.999	.998
ST	.998	.998	.999	.999	.999	.999	.999	.999	.998
KLD-U	.999	.998	.999	.999	.999	.999	.999	.998	.992
KLD-N	.999	.998	.999	.999	.998	.998	.998	.998	.992
Low Discrimination (a = 0.6)									
Z	.929	.932	.930	.933	.932	.937	.937	.938	.928
LR	.931	.933	.931	.934	.932	.936	.933	.932	.924
ST	.929	.932	.930	.933	.931	.934	.931	.930	.922
KLD-U	.926	.928	.925	.928	.926	.931	.928	.914	.805
KLD-N	.924	.922	.914	.911	.904	.904	.897	.880	.793

Table A95. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.050	.049	.048	.048	.048	.049	.049	.053	.056
LR	.050	.050	.048	.049	.048	.050	.048	.050	.052
ST	.050	.050	.049	.049	.048	.049	.048	.050	.051
KLD-U	.052	.052	.049	.047	.047	.048	.047	.050	.052
KLD-N	.056	.054	.050	.047	.046	.046	.046	.048	.051
Moderate Discrimination (a = 1.0)									
Z	.051	.052	.050	.049	.049	.051	.052	.055	.056
LR	.052	.053	.051	.050	.050	.051	.051	.052	.051
ST	.052	.052	.051	.050	.051	.051	.050	.051	.050
KLD-U	.052	.053	.051	.049	.048	.050	.050	.051	.050
KLD-N	.057	.055	.051	.048	.046	.047	.047	.048	.048
Low Discrimination (a = 0.6)									
Z	.048	.049	.049	.048	.047	.049	.051	.052	.054
LR	.051	.051	.051	.050	.050	.052	.053	.050	.050
ST	.051	.051	.051	.050	.050	.052	.052	.050	.049
KLD-U	.049	.052	.052	.051	.050	.052	.053	.051	.046
KLD-N	.052	.050	.047	.043	.041	.042	.042	.041	.041
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.212	.215	.216	.213	.220	.221	.223	.223	.232
LR	.216	.219	.218	.215	.221	.221	.220	.214	.220
ST	.213	.216	.214	.212	.217	.216	.215	.209	.212
KLD-U	.206	.208	.206	.198	.204	.201	.204	.200	.204
KLD-N	.218	.217	.211	.200	.203	.196	.195	.189	.193
Moderate Discrimination (a = 1.0)									
Z	.131	.134	.134	.137	.137	.140	.136	.148	.149
LR	.134	.136	.135	.138	.138	.140	.133	.138	.136
ST	.132	.134	.132	.135	.135	.136	.128	.133	.131
KLD-U	.126	.129	.126	.127	.126	.128	.121	.126	.122
KLD-N	.137	.135	.128	.126	.122	.119	.110	.112	.111
Low Discrimination (a = 0.6)									
Z	.085	.086	.090	.088	.088	.088	.092	.097	.100
LR	.090	.089	.093	.091	.092	.090	.091	.090	.090
ST	.088	.087	.090	.088	.090	.088	.088	.087	.087
KLD-U	.081	.085	.089	.084	.085	.083	.083	.082	.076
KLD-N	.090	.084	.082	.075	.071	.067	.065	.063	.061
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.645	.652	.652	.659	.663	.663	.668	.660	.660
LR	.650	.654	.654	.661	.663	.663	.666	.653	.652
ST	.647	.651	.651	.658	.659	.659	.662	.648	.646
KLD-U	.641	.644	.640	.641	.645	.640	.650	.638	.637
KLD-N	.652	.652	.644	.643	.642	.633	.639	.623	.622

Continued on the next page

Table A95 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Flat Bank Using KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.386	.394	.398	.401	.397	.409	.404	.417	.411
LR	.391	.397	.401	.403	.398	.409	.401	.406	.396
ST	.389	.394	.397	.398	.394	.404	.394	.399	.388
KLD-U	.379	.384	.385	.385	.380	.389	.380	.385	.372
KLD-N	.394	.394	.389	.382	.371	.374	.360	.361	.349
Low Discrimination (a = 0.6)									
Z	.206	.206	.210	.211	.205	.208	.218	.226	.225
LR	.214	.211	.215	.217	.212	.215	.218	.216	.212
ST	.211	.207	.210	.213	.207	.210	.212	.211	.206
KLD-U	.202	.203	.206	.204	.196	.199	.202	.200	.180
KLD-N	.209	.198	.193	.186	.172	.168	.166	.162	.150
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.996	.997	.997	.997	.997	.997	.997	.997	.996
LR	.996	.996	.997	.997	.997	.997	.997	.997	.996
ST	.996	.996	.996	.997	.997	.997	.997	.996	.996
KLD-U	.996	.996	.996	.997	.997	.997	.997	.996	.995
KLD-N	.996	.997	.996	.997	.997	.997	.997	.996	.995
Moderate Discrimination (a = 1.0)									
Z	.922	.925	.926	.926	.928	.931	.930	.928	.924
LR	.922	.925	.927	.926	.929	.931	.929	.926	.921
ST	.922	.924	.926	.925	.928	.930	.928	.924	.920
KLD-U	.920	.921	.921	.920	.922	.925	.923	.920	.904
KLD-N	.923	.923	.922	.919	.918	.919	.914	.910	.899
Low Discrimination (a = 0.6)									
Z	.637	.636	.639	.637	.630	.639	.650	.647	.642
LR	.644	.643	.646	.645	.640	.645	.649	.641	.638
ST	.642	.640	.642	.642	.637	.640	.644	.635	.634
KLD-U	.638	.636	.634	.630	.624	.628	.633	.618	.550
KLD-N	.633	.622	.614	.600	.584	.581	.581	.565	.524
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.999	.999	.999	.999	.999	.999	.999	.999	.999
LR	.999	.999	.999	.999	.999	.999	.999	.999	.999
ST	.998	.999	.999	.999	.999	.999	.999	.999	.999
KLD-U	.999	.999	.999	.999	.999	.999	.999	.999	.995
KLD-N	.999	.999	.999	.999	.999	.999	.999	.998	.996
Low Discrimination (a = 0.6)									
Z	.931	.934	.932	.934	.933	.936	.936	.938	.928
LR	.934	.935	.935	.937	.935	.938	.936	.937	.930
ST	.933	.934	.933	.936	.934	.937	.934	.936	.929
KLD-U	.932	.933	.931	.932	.930	.934	.930	.923	.834
KLD-N	.928	.927	.921	.919	.914	.915	.910	.902	.845

Table A96. Mean Type I Error Rates and Power for 50-Item Variable-Length AMC With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Type I Error									
High Discrimination (a = 1.5)									
Z	.048	.050	.048	.048	.047	.049	.049	.052	.054
LR	.049	.051	.049	.049	.048	.050	.049	.050	.051
ST	.049	.051	.049	.049	.048	.050	.049	.049	.051
KLD-U	.052	.053	.049	.048	.047	.049	.048	.050	.051
KLD-N	.055	.055	.050	.047	.046	.047	.046	.048	.050
Moderate Discrimination (a = 1.0)									
Z	.050	.051	.050	.050	.048	.050	.051	.054	.055
LR	.052	.053	.051	.052	.050	.050	.051	.051	.051
ST	.052	.053	.051	.051	.050	.051	.050	.051	.050
KLD-U	.052	.053	.051	.051	.049	.050	.050	.051	.051
KLD-N	.058	.055	.052	.050	.048	.047	.048	.049	.050
Low Discrimination (a = 0.6)									
Z	.049	.048	.049	.048	.046	.050	.052	.053	.053
LR	.051	.050	.050	.050	.048	.051	.052	.050	.049
ST	.051	.050	.050	.050	.049	.052	.051	.050	.049
KLD-U	.049	.052	.052	.052	.049	.052	.052	.051	.045
KLD-N	.054	.050	.048	.045	.042	.044	.043	.043	.042
Power: Very Small Change, $\Delta\theta$ = 0.25									
High Discrimination (a = 1.5)									
Z	.210	.215	.216	.212	.219	.221	.225	.223	.229
LR	.215	.219	.219	.215	.220	.222	.223	.214	.220
ST	.212	.216	.214	.212	.215	.217	.218	.209	.213
KLD-U	.209	.211	.206	.200	.204	.204	.206	.202	.205
KLD-N	.218	.217	.211	.201	.202	.199	.198	.192	.196
Moderate Discrimination (a = 1.0)									
Z	.130	.134	.133	.136	.136	.140	.136	.147	.147
LR	.134	.138	.136	.138	.138	.140	.133	.138	.136
ST	.131	.135	.132	.134	.134	.136	.128	.133	.131
KLD-U	.128	.130	.127	.127	.126	.128	.121	.126	.123
KLD-N	.138	.136	.130	.127	.124	.122	.113	.117	.116
Low Discrimination (a = 0.6)									
Z	.086	.087	.090	.087	.088	.088	.093	.098	.100
LR	.089	.088	.091	.089	.090	.089	.090	.089	.089
ST	.087	.086	.089	.087	.088	.087	.088	.087	.087
KLD-U	.082	.086	.087	.084	.084	.082	.084	.082	.076
KLD-N	.091	.085	.082	.076	.073	.069	.068	.068	.066
Power: Small Change, $\Delta\theta$ = 0.5									
High Discrimination (a = 1.5)									
Z	.641	.647	.650	.655	.660	.661	.668	.657	.655
LR	.648	.652	.654	.659	.662	.662	.666	.652	.650
ST	.645	.649	.650	.656	.658	.658	.662	.647	.644
KLD-U	.641	.644	.642	.641	.645	.642	.650	.638	.636
KLD-N	.650	.650	.646	.642	.643	.635	.641	.626	.625

Continued on the next page

Table A96 (continued). Mean Type I Error Rates and Power for 50-Item Variable-Length With 500-Item Flat Bank Using Modified KL Information Item Selection

Discrimination and Test	θ −2.0	−1.5	−1.0	−0.5	0.0	0.5	1.0	1.5	2.0
Moderate Discrimination (a = 1.0)									
Z	.385	.392	.397	.402	.394	.408	.406	.412	.408
LR	.392	.397	.401	.404	.397	.409	.403	.404	.396
ST	.389	.394	.397	.400	.393	.404	.396	.398	.389
KLD-U	.381	.386	.385	.387	.379	.390	.383	.385	.374
KLD-N	.396	.396	.391	.388	.374	.379	.369	.368	.360
Low Discrimination (a = 0.6)									
Z	.206	.206	.210	.213	.205	.209	.220	.224	.223
LR	.212	.209	.212	.215	.209	.213	.215	.212	.209
ST	.209	.206	.208	.212	.206	.209	.211	.207	.204
KLD-U	.204	.205	.206	.205	.197	.200	.202	.199	.179
KLD-N	.211	.201	.196	.190	.177	.175	.173	.169	.160
Power: Medium Change, $\Delta\theta$ = 1.0									
High Discrimination (a = 1.5)									
Z	.996	.996	.996	.997	.997	.997	.997	.996	.996
LR	.996	.996	.996	.997	.997	.997	.997	.997	.996
ST	.996	.996	.996	.996	.997	.997	.996	.996	.996
KLD-U	.996	.996	.996	.996	.997	.997	.996	.996	.995
KLD-N	.996	.996	.996	.996	.996	.996	.996	.996	.995
Moderate Discrimination (a = 1.0)									
Z	.920	.923	.924	.925	.925	.928	.928	.926	.921
LR	.922	.924	.925	.926	.926	.929	.927	.925	.919
ST	.921	.923	.923	.924	.925	.927	.926	.923	.917
KLD-U	.919	.920	.920	.921	.920	.923	.921	.918	.903
KLD-N	.923	.922	.921	.920	.918	.918	.915	.912	.901
Low Discrimination (a = 0.6)									
Z	.634	.637	.636	.635	.629	.638	.647	.644	.642
LR	.640	.640	.640	.639	.635	.641	.644	.635	.635
ST	.637	.637	.636	.637	.633	.638	.640	.631	.631
KLD-U	.637	.636	.632	.628	.622	.628	.631	.615	.552
KLD-N	.635	.626	.615	.604	.590	.590	.590	.576	.543
Power: Large Change, $\Delta\theta$ = 1.5									
High Discrimination (a = 1.5)									
Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
LR	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ST	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-U	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
KLD-N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Moderate Discrimination (a = 1.0)									
Z	.998	.999	.999	.999	.999	.999	.999	.999	.999
LR	.999	.999	.999	.999	.999	.999	.999	.999	.999
ST	.998	.999	.999	.999	.999	.999	.999	.999	.999
KLD-U	.998	.999	.999	.999	.999	.999	.999	.999	.995
KLD-N	.999	.999	.999	.998	.999	.999	.999	.998	.997
Low Discrimination (a = 0.6)									
Z	.930	.932	.929	.932	.930	.934	.934	.936	.926
LR	.931	.933	.931	.934	.932	.934	.933	.934	.927
ST	.930	.932	.930	.933	.931	.934	.932	.933	.927
KLD-U	.931	.932	.928	.930	.928	.931	.929	.921	.838
KLD-N	.928	.926	.921	.921	.914	.916	.912	.908	.859

Table A97. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 300-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Fisher Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.049	.049	.052	.057	.054	.055	.058	.055	.054
LR	.067	.063	.060	.057	.056	.057	.053	.054	.054
ST	.071	.068	.066	.059	.058	.058	.054	.054	.055
KLD-U	.050	.043	.038	.049	.046	.043	.050	.049	.046
KLD-N	.044	.033	.023	.043	.038	.031	.044	.042	.036
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.085	.071	.064	.152	.106	.081	.215	.138	.095
LR	.117	.091	.073	.161	.111	.083	.215	.139	.095
ST	.125	.098	.080	.165	.115	.085	.216	.141	.096
KLD-U	.065	.049	.039	.122	.081	.058	.185	.117	.076
KLD-N	.056	.037	.023	.110	.068	.042	.172	.105	.062
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.209	.143	.103	.441	.272	.162	.615	.394	.221
LR	.274	.178	.116	.460	.284	.167	.615	.396	.222
ST	.287	.191	.126	.466	.290	.170	.615	.398	.224
KLD-U	.165	.100	.062	.388	.223	.121	.574	.355	.188
KLD-N	.145	.077	.038	.366	.197	.094	.554	.330	.160
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.589	.413	.255	.888	.741	.473	.944	.886	.644
LR	.705	.488	.285	.905	.753	.480	.947	.883	.643
ST	.698	.509	.307	.904	.756	.485	.946	.882	.644
KLD-U	.563	.346	.180	.875	.685	.397	.931	.854	.590
KLD-N	.533	.294	.122	.860	.651	.339	.918	.835	.545
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.771	.665	.466	.924	.923	.778	.959	.972	.910
LR	.884	.761	.514	.957	.937	.784	.972	.974	.909
ST	.802	.774	.548	.942	.941	.788	.973	.974	.908
KLD-U	.819	.649	.386	.940	.902	.704	.956	.951	.862
KLD-N	.808	.594	.291	.930	.886	.649	.945	.944	.838

Table A98. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 300-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.051	.052	.054	.058	.054	.055	.058	.055	.054
LR	.060	.058	.057	.054	.054	.055	.052	.052	.053
ST	.065	.063	.061	.056	.055	.056	.053	.053	.054
KLD-U	.059	.051	.044	.054	.050	.047	.052	.051	.048
KLD-N	.050	.038	.026	.047	.041	.034	.046	.044	.038
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.099	.079	.069	.159	.109	.082	.218	.141	.095
LR	.113	.087	.071	.161	.110	.081	.215	.139	.094
ST	.123	.094	.076	.165	.112	.083	.216	.140	.095
KLD-U	.089	.064	.048	.140	.092	.064	.197	.125	.080
KLD-N	.074	.047	.028	.126	.077	.047	.184	.111	.065
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.249	.162	.111	.459	.282	.166	.622	.399	.223
LR	.278	.177	.114	.465	.285	.165	.617	.397	.221
ST	.296	.190	.123	.471	.290	.168	.616	.399	.223
KLD-U	.226	.132	.077	.429	.249	.134	.593	.371	.196
KLD-N	.197	.103	.048	.405	.220	.104	.572	.345	.167
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.676	.463	.276	.901	.753	.480	.945	.889	.647
LR	.718	.495	.288	.908	.756	.481	.947	.884	.643
ST	.736	.520	.308	.909	.760	.485	.946	.883	.644
KLD-U	.667	.423	.217	.891	.715	.422	.933	.862	.602
KLD-N	.628	.363	.151	.874	.681	.363	.920	.843	.558
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.850	.727	.495	.943	.929	.783	.960	.972	.912
LR	.890	.769	.519	.958	.938	.786	.972	.975	.909
ST	.903	.798	.554	.962	.943	.789	.974	.975	.908
KLD-U	.872	.716	.436	.943	.909	.723	.956	.953	.868
KLD-N	.849	.660	.338	.932	.895	.671	.945	.946	.846

Table A99. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 300-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Modified KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.053	.053	.055	.058	.055	.056	.058	.055	.054
LR	.058	.056	.055	.054	.053	.054	.052	.052	.052
ST	.062	.060	.058	.054	.054	.054	.052	.052	.053
KLD-U	.064	.056	.047	.056	.053	.049	.053	.052	.049
KLD-N	.054	.042	.028	.049	.043	.035	.047	.045	.038
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.103	.082	.070	.162	.111	.083	.220	.142	.096
LR	.110	.084	.068	.160	.109	.080	.214	.139	.093
ST	.119	.091	.073	.162	.111	.082	.215	.139	.094
KLD-U	.099	.071	.052	.146	.096	.067	.200	.128	.083
KLD-N	.083	.053	.031	.132	.081	.049	.187	.114	.067
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.255	.166	.112	.461	.285	.167	.622	.400	.225
LR	.273	.173	.111	.462	.283	.164	.615	.395	.221
ST	.289	.185	.120	.466	.286	.166	.614	.396	.221
KLD-U	.245	.146	.085	.438	.258	.139	.596	.375	.200
KLD-N	.214	.114	.053	.413	.229	.108	.575	.349	.171
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.676	.464	.274	.901	.753	.481	.945	.888	.647
LR	.710	.487	.281	.906	.753	.478	.947	.883	.641
ST	.727	.510	.301	.908	.755	.480	.946	.881	.641
KLD-U	.682	.442	.231	.891	.721	.429	.933	.862	.605
KLD-N	.642	.380	.162	.875	.687	.371	.920	.844	.562
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.847	.719	.484	.943	.928	.780	.960	.972	.911
LR	.887	.759	.507	.958	.937	.781	.972	.974	.907
ST	.899	.789	.545	.962	.942	.785	.974	.975	.906
KLD-U	.873	.723	.449	.943	.910	.727	.956	.953	.869
KLD-N	.850	.668	.351	.932	.895	.677	.945	.946	.848

Table A100. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 300-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Fisher Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.040	.041	.043	.047	.047	.047	.049	.048	.048
LR	.068	.065	.062	.058	.057	.056	.054	.054	.053
ST	.074	.071	.069	.062	.061	.060	.056	.056	.056
KLD-U	.048	.042	.037	.047	.045	.041	.048	.047	.045
KLD-N	.043	.033	.022	.043	.038	.030	.046	.042	.036
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.077	.062	.054	.137	.093	.068	.192	.122	.082
LR	.118	.091	.075	.158	.108	.080	.203	.132	.089
ST	.128	.100	.082	.166	.115	.085	.207	.136	.092
KLD-U	.061	.046	.036	.112	.074	.052	.169	.106	.069
KLD-N	.051	.034	.021	.103	.063	.038	.161	.096	.055
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.196	.127	.087	.421	.243	.135	.589	.353	.188
LR	.272	.174	.116	.458	.271	.154	.604	.371	.201
ST	.289	.188	.125	.470	.281	.161	.609	.377	.206
KLD-U	.153	.090	.054	.370	.201	.104	.553	.319	.161
KLD-N	.126	.066	.031	.348	.176	.076	.540	.299	.135
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.621	.402	.226	.941	.727	.415	.992	.890	.580
LR	.735	.491	.280	.952	.756	.447	.992	.898	.598
ST	.730	.508	.296	.953	.764	.458	.992	.901	.604
KLD-U	.558	.318	.150	.921	.669	.343	.989	.868	.531
KLD-N	.498	.249	.091	.910	.625	.275	.988	.853	.482
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.877	.715	.446	.997	.968	.750	1.000	.997	.901
LR	.953	.805	.520	.999	.974	.775	1.000	.997	.908
ST	.874	.803	.541	.989	.974	.782	1.000	.998	.910
KLD-U	.866	.640	.334	.996	.943	.665	1.000	.995	.864
KLD-N	.835	.552	.222	.996	.925	.583	1.000	.994	.837

Table A101. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 300-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.042	.043	.045	.048	.047	.047	.050	.049	.048
LR	.061	.059	.058	.055	.054	.054	.053	.053	.052
ST	.067	.064	.063	.058	.057	.057	.054	.054	.054
KLD-U	.059	.051	.043	.053	.050	.046	.051	.050	.048
KLD-N	.051	.039	.026	.048	.042	.034	.048	.044	.038
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.088	.069	.058	.144	.096	.070	.195	.124	.083
LR	.112	.086	.070	.158	.106	.077	.202	.131	.088
ST	.125	.095	.077	.165	.112	.082	.206	.134	.090
KLD-U	.087	.062	.046	.133	.087	.060	.180	.114	.074
KLD-N	.072	.046	.026	.122	.074	.043	.172	.104	.060
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.233	.145	.095	.439	.252	.140	.595	.359	.191
LR	.275	.172	.112	.462	.270	.153	.606	.371	.200
ST	.297	.187	.122	.474	.280	.159	.609	.376	.204
KLD-U	.219	.125	.071	.416	.230	.119	.571	.338	.172
KLD-N	.187	.095	.042	.396	.205	.091	.559	.318	.146
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.696	.449	.247	.949	.741	.426	.992	.893	.586
LR	.745	.497	.280	.954	.759	.447	.992	.898	.598
ST	.765	.520	.297	.956	.767	.457	.992	.900	.603
KLD-U	.678	.404	.193	.942	.712	.379	.990	.880	.550
KLD-N	.632	.336	.125	.935	.680	.317	.989	.868	.507
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.932	.769	.479	.999	.971	.759	1.000	.997	.903
LR	.958	.812	.527	.999	.974	.775	1.000	.997	.908
ST	.962	.828	.547	.999	.975	.782	1.000	.998	.909
KLD-U	.932	.727	.401	.999	.959	.703	1.000	.996	.876
KLD-N	.914	.662	.289	.999	.952	.642	1.000	.995	.855

Table A102. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 300-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Modified KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.044	.045	.047	.048	.048	.048	.050	.049	.049
LR	.058	.056	.056	.054	.053	.052	.052	.052	.051
ST	.064	.061	.060	.056	.055	.055	.053	.053	.053
KLD-U	.066	.057	.047	.056	.053	.048	.052	.051	.049
KLD-N	.056	.043	.029	.051	.045	.036	.049	.046	.039
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.093	.073	.060	.146	.098	.072	.196	.125	.084
LR	.109	.084	.068	.155	.105	.076	.201	.129	.087
ST	.120	.091	.073	.161	.109	.080	.203	.132	.089
KLD-U	.100	.071	.051	.141	.093	.064	.184	.118	.077
KLD-N	.083	.053	.030	.130	.080	.047	.177	.108	.063
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.241	.151	.099	.443	.256	.143	.596	.361	.193
LR	.269	.168	.109	.458	.267	.151	.602	.368	.198
ST	.289	.181	.117	.467	.275	.156	.605	.372	.201
KLD-U	.242	.140	.080	.430	.241	.126	.577	.344	.177
KLD-N	.209	.108	.048	.411	.216	.097	.565	.325	.152
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.698	.454	.252	.947	.742	.430	.992	.892	.587
LR	.733	.486	.274	.951	.753	.443	.992	.896	.594
ST	.751	.506	.288	.952	.759	.450	.992	.897	.598
KLD-U	.697	.428	.211	.943	.721	.392	.990	.882	.557
KLD-N	.658	.364	.140	.937	.692	.333	.989	.871	.516
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.926	.757	.473	.999	.970	.758	1.000	.997	.902
LR	.952	.800	.514	.999	.972	.768	1.000	.997	.905
ST	.956	.814	.535	.999	.973	.774	1.000	.997	.906
KLD-U	.935	.740	.421	.999	.960	.711	1.000	.996	.878
KLD-N	.922	.685	.314	.999	.955	.659	1.000	.995	.860

Table A103. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 300-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Fisher Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.051	.051	.051	.051	.050	.050	.050	.052	.051
LR	.050	.051	.051	.050	.050	.050	.049	.051	.051
ST	.050	.050	.050	.050	.050	.050	.049	.051	.051
KLD-U	.051	.045	.041	.051	.050	.047	.050	.051	.050
KLD-N	.045	.037	.029	.048	.043	.037	.050	.047	.043
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.088	.072	.063	.130	.094	.072	.179	.125	.086
LR	.093	.075	.063	.142	.100	.074	.193	.131	.088
ST	.092	.075	.062	.141	.098	.073	.190	.127	.086
KLD-U	.067	.051	.042	.123	.086	.062	.181	.121	.081
KLD-N	.058	.041	.030	.117	.075	.049	.180	.113	.069
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.211	.143	.100	.396	.244	.144	.563	.358	.199
LR	.230	.153	.102	.423	.258	.150	.579	.370	.204
ST	.231	.153	.102	.423	.255	.148	.576	.365	.200
KLD-U	.168	.104	.066	.389	.232	.128	.568	.356	.194
KLD-N	.149	.084	.048	.377	.208	.103	.564	.339	.170
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.589	.411	.249	.867	.709	.435	.934	.869	.608
LR	.660	.448	.260	.890	.723	.448	.937	.867	.614
ST	.637	.450	.264	.885	.721	.444	.935	.862	.608
KLD-U	.570	.355	.189	.876	.694	.407	.928	.853	.595
KLD-N	.541	.308	.142	.864	.662	.353	.919	.838	.555
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.768	.661	.457	.896	.909	.748	.948	.967	.894
LR	.864	.731	.484	.950	.925	.758	.965	.969	.894
ST	.756	.727	.497	.926	.927	.756	.964	.968	.890
KLD-U	.824	.658	.398	.940	.904	.711	.953	.950	.863
KLD-N	.813	.608	.320	.930	.888	.658	.945	.944	.841

Table A104. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 300-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.051	.050	.050	.051	.049	.050	.049	.051	.051
LR	.050	.050	.050	.051	.050	.050	.049	.051	.051
ST	.050	.051	.050	.050	.050	.051	.050	.050	.051
KLD-U	.050	.050	.045	.052	.051	.050	.050	.051	.050
KLD-N	.050	.040	.031	.051	.045	.039	.050	.048	.043
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.097	.077	.064	.141	.098	.074	.186	.128	.089
LR	.098	.077	.063	.151	.102	.076	.201	.134	.090
ST	.100	.078	.064	.150	.102	.075	.199	.130	.087
KLD-U	.075	.063	.050	.134	.092	.068	.188	.123	.083
KLD-N	.074	.050	.033	.132	.083	.052	.188	.118	.072
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.244	.158	.104	.423	.256	.150	.575	.367	.204
LR	.251	.160	.104	.445	.268	.155	.594	.380	.210
ST	.256	.164	.106	.444	.268	.153	.589	.374	.205
KLD-U	.200	.130	.080	.418	.247	.139	.580	.362	.198
KLD-N	.198	.108	.055	.413	.229	.112	.577	.353	.175
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.670	.454	.263	.888	.725	.449	.935	.872	.617
LR	.691	.469	.269	.898	.737	.458	.940	.873	.623
ST	.697	.482	.277	.897	.737	.456	.938	.869	.618
KLD-U	.637	.419	.223	.885	.712	.429	.928	.856	.602
KLD-N	.628	.371	.166	.875	.688	.374	.920	.846	.567
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.847	.719	.479	.937	.919	.758	.953	.967	.897
LR	.878	.749	.497	.953	.930	.767	.967	.971	.898
ST	.886	.770	.518	.956	.934	.766	.968	.970	.895
KLD-U	.859	.712	.443	.940	.907	.727	.953	.950	.866
KLD-N	.848	.665	.356	.932	.895	.677	.945	.946	.848

Table A105. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 300-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Modified KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.050	.050	.050	.051	.050	.050	.049	.051	.051
LR	.050	.051	.049	.051	.050	.050	.050	.051	.051
ST	.049	.051	.050	.051	.050	.050	.050	.051	.052
KLD-U	.050	.050	.048	.052	.051	.050	.051	.051	.051
KLD-N	.050	.043	.032	.052	.047	.039	.050	.049	.044
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.097	.078	.064	.142	.100	.075	.186	.130	.089
LR	.098	.078	.063	.152	.103	.076	.202	.134	.091
ST	.100	.079	.064	.151	.102	.074	.199	.131	.089
KLD-U	.077	.065	.054	.136	.093	.068	.190	.124	.085
KLD-N	.076	.056	.036	.136	.086	.054	.191	.121	.074
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.244	.159	.104	.423	.260	.152	.573	.369	.206
LR	.251	.162	.103	.446	.270	.155	.595	.380	.212
ST	.254	.166	.106	.444	.268	.153	.590	.374	.208
KLD-U	.203	.134	.087	.418	.250	.140	.581	.363	.202
KLD-N	.201	.117	.059	.418	.237	.116	.579	.357	.179
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.663	.451	.258	.886	.725	.451	.934	.871	.617
LR	.686	.468	.265	.898	.736	.458	.941	.872	.625
ST	.692	.480	.276	.897	.735	.455	.939	.868	.620
KLD-U	.634	.421	.235	.882	.711	.428	.928	.854	.606
KLD-N	.625	.386	.174	.875	.692	.380	.920	.846	.571
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.841	.707	.464	.936	.918	.755	.952	.966	.896
LR	.875	.744	.488	.954	.930	.765	.968	.971	.898
ST	.883	.766	.514	.957	.934	.765	.969	.970	.896
KLD-U	.851	.706	.454	.938	.905	.724	.952	.949	.867
KLD-N	.842	.672	.365	.932	.896	.682	.945	.946	.849

Table A106. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 300-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Fisher Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.045	.045	.047	.050	.050	.050	.052	.050	.050
LR	.049	.049	.050	.050	.050	.050	.051	.050	.051
ST	.050	.049	.050	.050	.050	.050	.051	.050	.051
KLD-U	.049	.044	.039	.050	.049	.045	.052	.050	.050
KLD-N	.044	.036	.028	.048	.043	.036	.051	.047	.043
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.083	.067	.058	.137	.097	.072	.186	.120	.083
LR	.089	.072	.062	.138	.097	.072	.185	.121	.083
ST	.091	.072	.061	.136	.094	.071	.180	.117	.080
KLD-U	.063	.048	.039	.117	.080	.057	.173	.110	.074
KLD-N	.053	.037	.027	.110	.070	.044	.170	.103	.063
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.206	.135	.094	.415	.245	.140	.572	.342	.186
LR	.221	.144	.098	.417	.246	.139	.569	.342	.186
ST	.224	.144	.097	.417	.242	.136	.565	.336	.180
KLD-U	.156	.094	.057	.378	.210	.110	.555	.324	.170
KLD-N	.130	.071	.039	.359	.186	.085	.548	.308	.146
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.633	.415	.239	.939	.726	.416	.991	.881	.567
LR	.679	.439	.247	.941	.726	.415	.990	.880	.567
ST	.659	.435	.246	.938	.723	.411	.990	.878	.562
KLD-U	.565	.326	.158	.926	.680	.353	.989	.869	.540
KLD-N	.505	.260	.107	.915	.636	.287	.989	.857	.494
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.882	.725	.460	.997	.968	.749	1.000	.997	.894
LR	.938	.767	.479	.999	.968	.748	1.000	.997	.894
ST	.836	.748	.479	.985	.966	.745	1.000	.997	.891
KLD-U	.872	.649	.346	.997	.946	.674	1.000	.995	.869
KLD-N	.841	.564	.246	.996	.928	.594	1.000	.994	.844

Table A107. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 300-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.046	.047	.048	.051	.050	.050	.052	.050	.051
LR	.049	.050	.050	.050	.051	.050	.051	.051	.051
ST	.050	.050	.051	.050	.050	.050	.051	.051	.051
KLD-U	.050	.050	.045	.051	.051	.049	.051	.051	.050
KLD-N	.051	.041	.030	.051	.046	.038	.051	.049	.044
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.094	.074	.062	.148	.100	.074	.194	.124	.085
LR	.095	.075	.063	.147	.100	.073	.194	.125	.084
ST	.099	.077	.064	.145	.098	.072	.189	.122	.081
KLD-U	.074	.062	.048	.127	.087	.063	.178	.114	.076
KLD-N	.073	.048	.031	.126	.079	.049	.177	.109	.066
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.242	.153	.101	.440	.256	.146	.586	.350	.191
LR	.245	.155	.102	.439	.257	.144	.586	.353	.190
ST	.252	.159	.104	.437	.254	.143	.581	.349	.186
KLD-U	.194	.124	.073	.404	.230	.124	.565	.334	.175
KLD-N	.188	.099	.048	.400	.212	.097	.563	.325	.156
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.704	.461	.258	.948	.741	.429	.991	.885	.575
LR	.715	.469	.261	.947	.742	.428	.991	.887	.577
ST	.722	.475	.266	.947	.740	.426	.991	.885	.573
KLD-U	.646	.402	.198	.938	.710	.386	.990	.876	.550
KLD-N	.633	.344	.137	.935	.686	.327	.990	.870	.517
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.934	.776	.491	.999	.971	.758	1.000	.997	.897
LR	.950	.793	.503	.999	.971	.758	1.000	.997	.898
ST	.952	.798	.510	.999	.970	.757	1.000	.997	.896
KLD-U	.920	.725	.406	.998	.958	.708	1.000	.996	.876
KLD-N	.914	.667	.304	.999	.953	.649	1.000	.996	.860

Table A108. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 300-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Modified KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.047	.048	.050	.050	.050	.050	.051	.051	.050
LR	.050	.049	.051	.050	.050	.050	.051	.051	.051
ST	.051	.049	.051	.050	.050	.050	.051	.051	.050
KLD-U	.050	.050	.049	.050	.050	.049	.051	.051	.050
KLD-N	.051	.045	.032	.050	.048	.040	.051	.050	.045
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.099	.078	.064	.149	.101	.075	.194	.126	.085
LR	.097	.075	.063	.146	.100	.073	.193	.124	.084
ST	.100	.077	.064	.145	.098	.072	.189	.122	.082
KLD-U	.077	.063	.053	.128	.088	.065	.177	.114	.077
KLD-N	.074	.055	.034	.127	.084	.052	.178	.113	.069
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.250	.159	.106	.441	.258	.147	.584	.354	.191
LR	.246	.154	.103	.438	.255	.144	.585	.353	.191
ST	.254	.158	.104	.438	.253	.143	.580	.349	.186
KLD-U	.197	.127	.082	.405	.231	.127	.564	.336	.177
KLD-N	.192	.111	.053	.402	.223	.104	.564	.332	.161
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.705	.465	.262	.945	.738	.430	.991	.885	.574
LR	.709	.462	.261	.944	.736	.427	.990	.885	.575
ST	.717	.470	.265	.944	.735	.425	.990	.883	.570
KLD-U	.642	.404	.215	.935	.708	.391	.989	.875	.551
KLD-N	.634	.369	.150	.934	.697	.342	.989	.873	.525
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.927	.764	.484	.999	.969	.755	1.000	.997	.895
LR	.944	.781	.498	.999	.968	.754	1.000	.997	.896
ST	.946	.788	.507	.999	.968	.753	1.000	.997	.893
KLD-U	.914	.719	.425	.998	.957	.709	1.000	.995	.876
KLD-N	.912	.689	.326	.998	.956	.665	1.000	.996	.865

Table A109. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 500-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Fisher Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.045	.047	.051	.052	.052	.053	.054	.052	.053
LR	.068	.064	.060	.057	.057	.056	.054	.053	.054
ST	.073	.070	.066	.059	.059	.059	.055	.054	.055
KLD-U	.050	.043	.038	.049	.047	.043	.050	.048	.047
KLD-N	.045	.034	.024	.044	.039	.032	.045	.042	.037
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.085	.071	.064	.157	.108	.082	.230	.146	.098
LR	.122	.094	.076	.171	.117	.086	.236	.150	.100
ST	.130	.101	.082	.177	.121	.089	.239	.153	.101
KLD-U	.066	.049	.040	.129	.085	.060	.203	.126	.081
KLD-N	.057	.038	.024	.118	.073	.045	.192	.113	.067
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.215	.146	.106	.470	.287	.171	.666	.426	.238
LR	.289	.186	.123	.493	.304	.178	.669	.433	.242
ST	.303	.199	.133	.500	.311	.182	.671	.437	.245
KLD-U	.172	.103	.065	.417	.239	.130	.630	.389	.206
KLD-N	.152	.081	.040	.396	.215	.103	.613	.366	.179
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.608	.432	.268	.917	.775	.505	.975	.920	.688
LR	.735	.515	.305	.932	.788	.515	.974	.918	.690
ST	.725	.534	.327	.928	.791	.520	.972	.917	.692
KLD-U	.586	.363	.192	.901	.721	.430	.962	.894	.639
KLD-N	.557	.311	.134	.889	.691	.373	.955	.880	.600
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.793	.688	.490	.947	.943	.812	.985	.988	.936
LR	.910	.790	.547	.976	.955	.818	.991	.988	.935
ST	.816	.796	.581	.949	.955	.821	.989	.987	.934
KLD-U	.843	.673	.411	.959	.922	.740	.981	.972	.896
KLD-N	.834	.620	.318	.950	.910	.691	.976	.968	.879

Table A110. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 500-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.049	.050	.053	.053	.052	.053	.054	.052	.053
LR	.061	.059	.057	.054	.054	.054	.053	.052	.053
ST	.066	.064	.061	.056	.056	.056	.054	.053	.054
KLD-U	.060	.052	.044	.054	.052	.047	.053	.051	.049
KLD-N	.052	.040	.027	.048	.043	.035	.048	.044	.039
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.100	.080	.069	.166	.112	.084	.235	.148	.099
LR	.117	.089	.072	.171	.115	.084	.237	.149	.099
ST	.129	.097	.078	.175	.119	.087	.239	.152	.100
KLD-U	.093	.066	.049	.150	.098	.067	.219	.135	.086
KLD-N	.078	.050	.030	.137	.084	.050	.207	.122	.071
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.260	.169	.115	.491	.299	.176	.674	.433	.242
LR	.294	.186	.120	.499	.305	.177	.673	.435	.242
ST	.314	.199	.130	.505	.311	.181	.674	.438	.244
KLD-U	.240	.140	.081	.464	.270	.145	.652	.409	.217
KLD-N	.211	.110	.052	.440	.242	.116	.634	.385	.189
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.704	.489	.292	.933	.788	.515	.976	.922	.692
LR	.749	.524	.308	.934	.791	.517	.974	.919	.691
ST	.768	.548	.328	.933	.794	.521	.972	.918	.692
KLD-U	.697	.448	.234	.918	.753	.458	.964	.901	.653
KLD-N	.660	.388	.168	.904	.724	.401	.956	.888	.615
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.874	.756	.525	.968	.952	.818	.988	.988	.937
LR	.916	.799	.554	.977	.957	.820	.991	.988	.935
ST	.930	.825	.588	.978	.958	.823	.991	.987	.934
KLD-U	.896	.744	.469	.962	.930	.760	.981	.973	.901
KLD-N	.877	.691	.372	.953	.919	.715	.977	.970	.886

Table A111. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 500-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Modified KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.050	.052	.054	.053	.053	.054	.055	.052	.053
LR	.059	.057	.055	.054	.053	.053	.052	.051	.052
ST	.063	.061	.059	.055	.055	.054	.053	.052	.053
KLD-U	.066	.057	.048	.057	.054	.049	.054	.052	.050
KLD-N	.056	.044	.030	.051	.045	.036	.049	.045	.040
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.104	.082	.070	.168	.114	.085	.238	.149	.100
LR	.115	.087	.070	.170	.114	.083	.236	.148	.098
ST	.124	.093	.075	.173	.117	.085	.238	.150	.099
KLD-U	.105	.074	.054	.158	.103	.071	.224	.139	.089
KLD-N	.088	.056	.033	.144	.088	.053	.211	.126	.074
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.268	.173	.116	.494	.302	.178	.675	.434	.243
LR	.289	.183	.117	.496	.303	.176	.671	.433	.241
ST	.306	.194	.126	.500	.308	.178	.671	.435	.242
KLD-U	.262	.155	.090	.475	.280	.152	.656	.414	.221
KLD-N	.230	.122	.057	.451	.252	.121	.638	.390	.193
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.704	.488	.290	.932	.788	.516	.976	.921	.692
LR	.740	.515	.300	.933	.787	.513	.974	.917	.688
ST	.758	.538	.321	.932	.789	.516	.972	.916	.689
KLD-U	.712	.468	.250	.918	.758	.467	.964	.902	.657
KLD-N	.674	.407	.180	.905	.729	.411	.956	.889	.619
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.871	.746	.512	.968	.950	.815	.988	.988	.936
LR	.912	.789	.540	.977	.956	.816	.991	.988	.934
ST	.926	.816	.578	.978	.957	.818	.991	.987	.933
KLD-U	.897	.751	.482	.962	.930	.764	.981	.973	.902
KLD-N	.879	.701	.386	.953	.920	.721	.977	.970	.888

Table A112. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 500-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Fisher Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.040	.041	.042	.046	.046	.046	.048	.048	.048
LR	.069	.065	.062	.058	.057	.056	.053	.054	.053
ST	.075	.072	.069	.062	.061	.060	.056	.056	.056
KLD-U	.048	.043	.037	.046	.045	.042	.046	.046	.045
KLD-N	.043	.034	.023	.043	.038	.031	.044	.042	.037
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.078	.062	.054	.144	.096	.071	.216	.133	.087
LR	.122	.093	.076	.169	.114	.084	.232	.144	.095
ST	.133	.103	.084	.179	.121	.089	.238	.149	.098
KLD-U	.063	.047	.037	.120	.078	.054	.191	.115	.074
KLD-N	.053	.035	.022	.110	.067	.040	.184	.106	.061
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.202	.133	.089	.452	.262	.146	.660	.397	.209
LR	.285	.184	.120	.494	.294	.167	.679	.417	.224
ST	.303	.199	.131	.508	.306	.176	.686	.425	.230
KLD-U	.160	.095	.056	.401	.221	.114	.624	.361	.181
KLD-N	.134	.071	.033	.379	.195	.087	.613	.341	.156
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.633	.423	.242	.953	.767	.457	.997	.927	.643
LR	.755	.516	.301	.966	.795	.492	.997	.934	.661
ST	.749	.533	.319	.965	.803	.504	.997	.936	.667
KLD-U	.581	.340	.165	.940	.714	.386	.996	.911	.596
KLD-N	.526	.270	.104	.931	.675	.320	.995	.900	.552
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.876	.735	.476	.995	.978	.798	1.000	.999	.937
LR	.958	.828	.556	.999	.983	.822	1.000	.999	.943
ST	.869	.822	.577	.982	.983	.828	1.000	.999	.944
KLD-U	.879	.670	.366	.997	.960	.722	1.000	.998	.910
KLD-N	.854	.586	.253	.997	.947	.652	1.000	.998	.892

Table A113. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 500-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.042	.043	.045	.047	.047	.047	.048	.048	.048
LR	.061	.059	.058	.055	.055	.054	.052	.052	.052
ST	.068	.065	.063	.058	.058	.057	.054	.054	.054
KLD-U	.060	.052	.044	.053	.051	.046	.050	.049	.048
KLD-N	.052	.040	.027	.049	.044	.035	.048	.045	.039
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.091	.070	.058	.153	.100	.072	.222	.136	.088
LR	.116	.088	.072	.169	.112	.081	.232	.143	.094
ST	.130	.098	.079	.177	.118	.085	.238	.148	.096
KLD-U	.091	.065	.047	.144	.093	.063	.209	.125	.079
KLD-N	.076	.048	.028	.134	.080	.048	.201	.116	.066
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.243	.153	.099	.475	.275	.153	.670	.405	.213
LR	.288	.182	.117	.500	.294	.166	.682	.418	.223
ST	.312	.198	.127	.513	.306	.173	.688	.425	.228
KLD-U	.232	.134	.075	.456	.255	.132	.652	.385	.195
KLD-N	.201	.103	.046	.437	.230	.104	.642	.366	.169
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.716	.474	.267	.963	.782	.471	.997	.930	.649
LR	.767	.523	.301	.968	.798	.493	.997	.934	.661
ST	.787	.547	.320	.969	.806	.504	.997	.936	.667
KLD-U	.705	.433	.214	.960	.759	.426	.997	.922	.618
KLD-N	.664	.366	.144	.955	.732	.368	.996	.914	.581
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.935	.793	.514	.999	.981	.808	1.000	.999	.939
LR	.964	.837	.563	.999	.983	.823	1.000	.999	.942
ST	.968	.851	.585	.999	.984	.829	1.000	.999	.944
KLD-U	.941	.758	.440	.999	.973	.760	1.000	.999	.920
KLD-N	.927	.698	.330	.999	.969	.713	1.000	.998	.908

Table A114. Average Across θ Levels of Type I Error Rate and Power for Fixed-Length AMC of n = 15, 30, and 50 Items With 500-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Modified KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.044	.045	.047	.048	.048	.048	.049	.049	.049
LR	.058	.057	.056	.054	.053	.053	.052	.052	.051
ST	.064	.062	.060	.056	.056	.055	.053	.053	.053
KLD-U	.067	.058	.049	.057	.054	.049	.052	.051	.050
KLD-N	.058	.045	.030	.052	.047	.037	.049	.047	.041
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.095	.074	.061	.156	.103	.074	.224	.137	.089
LR	.113	.086	.069	.166	.110	.079	.230	.142	.093
ST	.125	.094	.075	.173	.115	.083	.235	.145	.095
KLD-U	.104	.074	.053	.153	.099	.067	.215	.130	.083
KLD-N	.087	.055	.032	.142	.086	.051	.207	.120	.069
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.252	.160	.103	.478	.279	.156	.670	.407	.216
LR	.282	.178	.114	.494	.291	.164	.677	.415	.222
ST	.303	.191	.122	.505	.300	.170	.682	.421	.225
KLD-U	.256	.150	.086	.469	.267	.140	.657	.392	.201
KLD-N	.224	.117	.053	.451	.242	.111	.648	.374	.176
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.717	.479	.271	.962	.782	.475	.997	.929	.650
LR	.755	.514	.295	.965	.792	.488	.997	.932	.658
ST	.773	.534	.311	.966	.799	.496	.997	.933	.662
KLD-U	.724	.457	.233	.960	.766	.439	.997	.922	.625
KLD-N	.688	.395	.162	.956	.743	.385	.996	.915	.589
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.931	.782	.506	.999	.980	.808	1.000	.999	.938
LR	.958	.824	.550	.999	.982	.817	1.000	.999	.940
ST	.962	.838	.572	.999	.982	.822	1.000	.999	.941
KLD-U	.944	.770	.461	.999	.974	.769	1.000	.999	.922
KLD-N	.934	.721	.357	.999	.971	.728	1.000	.998	.911

Table A115. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 500-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Fisher Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.049	.050	.051	.051	.050	.051	.050	.050	.050
LR	.049	.050	.050	.051	.050	.051	.050	.050	.051
ST	.050	.050	.050	.051	.050	.051	.050	.050	.051
KLD-U	.049	.045	.040	.051	.050	.047	.049	.049	.050
KLD-N	.046	.037	.030	.049	.044	.036	.049	.048	.043
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.090	.074	.064	.146	.101	.077	.204	.133	.088
LR	.095	.076	.064	.153	.104	.078	.213	.139	.092
ST	.096	.075	.064	.150	.101	.076	.209	.134	.089
KLD-U	.066	.052	.043	.131	.089	.065	.197	.126	.084
KLD-N	.059	.042	.031	.125	.080	.050	.196	.122	.074
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.222	.149	.104	.444	.267	.157	.629	.395	.213
LR	.239	.157	.106	.457	.275	.161	.634	.404	.223
ST	.243	.157	.106	.454	.272	.158	.629	.397	.217
KLD-U	.172	.107	.069	.419	.246	.137	.619	.385	.208
KLD-N	.156	.087	.050	.408	.225	.109	.615	.376	.188
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.614	.435	.265	.903	.752	.477	.971	.909	.653
LR	.687	.471	.276	.918	.758	.483	.968	.905	.662
ST	.665	.470	.280	.909	.752	.478	.964	.900	.655
KLD-U	.588	.372	.202	.901	.725	.440	.960	.891	.641
KLD-N	.564	.323	.154	.892	.699	.383	.954	.883	.609
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.796	.690	.485	.924	.932	.790	.977	.986	.924
LR	.892	.760	.514	.969	.944	.794	.988	.985	.923
ST	.773	.749	.528	.933	.942	.790	.985	.983	.920
KLD-U	.846	.682	.425	.957	.922	.745	.979	.970	.895
KLD-N	.840	.632	.346	.949	.910	.696	.975	.968	.881

Table A116. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 500-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.050	.050	.050	.050	.050	.051	.049	.051	.051
LR	.051	.050	.050	.050	.049	.052	.050	.050	.051
ST	.050	.050	.051	.050	.049	.052	.050	.050	.051
KLD-U	.050	.050	.046	.051	.049	.051	.050	.049	.050
KLD-N	.050	.042	.032	.051	.046	.039	.050	.048	.044
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.102	.080	.066	.156	.107	.080	.212	.140	.091
LR	.102	.078	.065	.160	.107	.080	.223	.143	.094
ST	.104	.079	.066	.159	.106	.079	.218	.139	.091
KLD-U	.077	.063	.051	.142	.094	.071	.205	.131	.086
KLD-N	.075	.052	.035	.141	.089	.055	.208	.129	.077
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.264	.167	.110	.470	.284	.166	.641	.409	.221
LR	.267	.167	.110	.478	.286	.168	.651	.417	.228
ST	.271	.170	.112	.476	.284	.165	.645	.411	.223
KLD-U	.211	.135	.085	.448	.262	.151	.633	.397	.214
KLD-N	.205	.114	.059	.445	.249	.122	.632	.393	.197
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.705	.485	.283	.924	.771	.492	.972	.913	.662
LR	.723	.494	.291	.924	.770	.497	.970	.910	.669
ST	.731	.505	.299	.921	.767	.494	.967	.906	.665
KLD-U	.664	.439	.241	.909	.741	.465	.960	.895	.648
KLD-N	.653	.395	.182	.902	.726	.409	.955	.890	.622
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.875	.753	.514	.962	.945	.800	.985	.986	.926
LR	.907	.779	.534	.972	.949	.804	.990	.986	.926
ST	.915	.797	.556	.972	.949	.803	.989	.985	.924
KLD-U	.884	.737	.477	.957	.924	.763	.979	.971	.898
KLD-N	.875	.696	.390	.951	.918	.718	.975	.970	.888

Table A117. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 500-Item Peaked Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Modified KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.051	.050	.050	.051	.050	.051	.050	.050	.050
LR	.051	.050	.050	.050	.050	.051	.050	.050	.050
ST	.050	.050	.050	.051	.050	.051	.050	.050	.051
KLD-U	.050	.050	.049	.051	.050	.050	.049	.049	.050
KLD-N	.050	.045	.034	.051	.048	.040	.050	.049	.045
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.105	.080	.066	.159	.108	.080	.215	.139	.092
LR	.103	.078	.065	.160	.108	.080	.224	.143	.094
ST	.105	.079	.065	.159	.106	.078	.219	.139	.091
KLD-U	.080	.064	.056	.144	.095	.072	.207	.131	.086
KLD-N	.078	.058	.038	.144	.093	.059	.209	.131	.079
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.268	.168	.110	.474	.285	.166	.642	.408	.222
LR	.268	.167	.110	.478	.288	.167	.650	.415	.228
ST	.274	.170	.112	.476	.286	.165	.645	.410	.224
KLD-U	.215	.137	.093	.451	.264	.151	.633	.397	.215
KLD-N	.210	.125	.064	.449	.258	.128	.632	.396	.200
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.702	.480	.278	.923	.767	.491	.972	.911	.663
LR	.719	.491	.287	.923	.769	.495	.970	.909	.669
ST	.728	.502	.297	.921	.767	.492	.967	.906	.665
KLD-U	.661	.439	.255	.907	.740	.464	.959	.893	.648
KLD-N	.651	.413	.193	.901	.730	.419	.954	.890	.626
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.870	.739	.497	.962	.943	.795	.985	.986	.926
LR	.904	.772	.524	.972	.949	.800	.990	.986	.926
ST	.913	.792	.550	.972	.949	.800	.989	.985	.924
KLD-U	.876	.729	.488	.955	.922	.759	.979	.970	.897
KLD-N	.868	.704	.401	.950	.918	.723	.975	.970	.889

Table A118. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 500-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Fisher Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.044	.045	.046	.050	.049	.051	.050	.051	.051
LR	.051	.050	.050	.051	.050	.051	.050	.050	.051
ST	.051	.049	.050	.050	.050	.051	.050	.051	.051
KLD-U	.049	.044	.039	.050	.049	.046	.049	.051	.050
KLD-N	.045	.036	.029	.048	.044	.035	.049	.048	.043
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.084	.068	.060	.147	.100	.077	.209	.133	.090
LR	.094	.074	.063	.149	.101	.076	.209	.131	.089
ST	.094	.073	.062	.145	.098	.074	.205	.127	.086
KLD-U	.065	.049	.040	.125	.083	.060	.195	.121	.079
KLD-N	.055	.038	.028	.118	.073	.045	.192	.114	.068
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.212	.142	.097	.451	.264	.153	.644	.389	.209
LR	.234	.152	.101	.453	.265	.152	.641	.385	.208
ST	.236	.150	.100	.449	.261	.149	.638	.380	.202
KLD-U	.163	.098	.060	.409	.229	.120	.625	.370	.190
KLD-N	.137	.075	.042	.391	.205	.093	.620	.352	.166
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.644	.438	.256	.953	.766	.462	.997	.923	.635
LR	.706	.464	.266	.958	.767	.460	.996	.921	.633
ST	.680	.457	.264	.953	.764	.456	.996	.919	.627
KLD-U	.589	.347	.173	.945	.724	.396	.996	.914	.605
KLD-N	.533	.280	.119	.937	.685	.327	.996	.904	.562
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.880	.745	.492	.994	.978	.799	1.000	.999	.933
LR	.946	.793	.512	.999	.979	.798	1.000	.999	.932
ST	.831	.768	.512	.976	.976	.795	.999	.999	.930
KLD-U	.883	.678	.378	.998	.962	.730	1.000	.998	.912
KLD-N	.859	.596	.276	.998	.949	.657	1.000	.998	.894

Table A119. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 500-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.046	.047	.048	.050	.050	.050	.050	.052	.050
LR	.051	.050	.050	.050	.050	.050	.049	.051	.051
ST	.051	.050	.050	.050	.050	.050	.049	.051	.051
KLD-U	.050	.049	.046	.049	.050	.050	.049	.050	.051
KLD-N	.050	.042	.031	.049	.047	.039	.049	.050	.044
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.097	.075	.063	.158	.105	.077	.219	.138	.090
LR	.101	.077	.064	.157	.104	.077	.218	.136	.091
ST	.104	.078	.064	.155	.103	.076	.214	.133	.088
KLD-U	.076	.061	.050	.135	.090	.067	.203	.126	.083
KLD-N	.074	.050	.033	.134	.084	.052	.202	.122	.073
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.253	.161	.106	.477	.278	.158	.658	.402	.213
LR	.260	.162	.105	.476	.277	.157	.657	.400	.214
ST	.266	.166	.107	.474	.275	.155	.653	.395	.209
KLD-U	.202	.128	.078	.439	.249	.137	.642	.382	.199
KLD-N	.196	.106	.052	.435	.236	.110	.639	.375	.178
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.725	.485	.279	.963	.782	.473	.997	.927	.640
LR	.742	.493	.281	.962	.782	.473	.997	.926	.644
ST	.748	.499	.285	.962	.780	.471	.997	.925	.640
KLD-U	.671	.422	.219	.956	.752	.433	.996	.920	.621
KLD-N	.658	.371	.156	.954	.737	.375	.996	.916	.589
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.938	.798	.526	.999	.981	.806	1.000	.999	.934
LR	.958	.816	.538	.999	.981	.807	1.000	.999	.935
ST	.959	.821	.545	.999	.980	.806	1.000	.999	.934
KLD-U	.930	.750	.446	.999	.971	.764	1.000	.998	.920
KLD-N	.924	.701	.344	.999	.970	.716	1.000	.998	.909

Table A120. Average Across θ Levels of Type I Error Rate and Power for Variable-Length AMC of n = 15, 30, and 50 Items With 500-Item Flat Bank of High (HD), Medium (MD) and Low Discrimination (LD) Using Modified KL Information Item Selection

	n = 15			n = 30			n = 50		
Test	HD	MD	LD	HD	MD	LD	HD	MD	LD
Type I Error									
Z	.047	.048	.050	.050	.050	.050	.049	.051	.050
LR	.050	.050	.050	.049	.051	.051	.049	.051	.050
ST	.050	.050	.050	.049	.050	.051	.049	.051	.050
KLD-U	.051	.049	.050	.050	.051	.050	.050	.051	.050
KLD-N	.050	.046	.034	.051	.050	.041	.049	.051	.046
Power: Very Small Change, $\Delta\theta$ = 0.25									
Z	.102	.079	.066	.159	.107	.078	.219	.138	.091
LR	.101	.077	.063	.156	.105	.077	.219	.137	.090
ST	.103	.078	.064	.154	.103	.076	.214	.133	.087
KLD-U	.079	.063	.055	.137	.093	.068	.205	.126	.083
KLD-N	.076	.057	.036	.139	.091	.056	.204	.125	.075
Power: Small Change, $\Delta\theta$ = 0.5									
Z	.261	.168	.110	.477	.280	.160	.655	.400	.213
LR	.260	.163	.105	.473	.279	.158	.656	.400	.212
ST	.266	.167	.107	.472	.276	.156	.652	.395	.208
KLD-U	.207	.131	.088	.441	.253	.140	.642	.383	.200
KLD-N	.201	.119	.058	.444	.248	.117	.640	.380	.184
Power: Medium Change, $\Delta\theta$ = 1.0									
Z	.724	.490	.282	.961	.779	.474	.996	.924	.638
LR	.733	.489	.279	.960	.779	.473	.996	.925	.639
ST	.739	.496	.283	.959	.777	.471	.996	.923	.636
KLD-U	.668	.423	.237	.954	.753	.436	.996	.918	.620
KLD-N	.660	.398	.171	.954	.747	.392	.996	.917	.596
Power: Large Change, $\Delta\theta$ = 1.5									
Z	.932	.787	.518	.999	.979	.804	1.000	.999	.931
LR	.952	.807	.529	.999	.979	.805	1.000	.999	.932
ST	.953	.813	.539	.999	.979	.804	1.000	.999	.931
KLD-U	.925	.743	.465	.999	.971	.765	1.000	.998	.919
KLD-N	.923	.723	.367	.999	.971	.732	1.000	.998	.912