

J·T·L·A

The Journal of Technology, Learning, and Assessment

Volume 2, Number 5 · December 2003

Computerized Adaptive Testing: A Comparison of Three Content Balancing Methods

Chi-Keung Leung, Hua-Hua Chang,
and Kit-Tai Hau

www.jtla.org

A publication of the Technology and Assessment Study Collaborative
Caroline A. & Peter S. Lynch School of Education, Boston College



Volume 2, Number 5

Computerized Adaptive Testing: A Comparison of Three Content Balancing Methods

Chi-Keung Leung, Hua-Hua Chang, and Kit-Tai Hau

Editor: Michael Russell
russelmh@bc.edu
Technology and Assessment Study Collaborative
Lynch School of Education, Boston College
Chestnut Hill, MA 02467

Copy Editor: Kathleen O'Connor
Design and Layout: Thomas Hoffmann

JTLA is a free on-line journal, published by the Technology and Assessment Study Collaborative, Caroline A. & Peter S. Lynch School of Education, Boston College.

Copyright ©2003 by the Journal of Technology, Learning, and Assessment (ISSN 1540-2525).
Permission is hereby granted to copy any article provided that the Journal of Technology, Learning, and Assessment is credited and copies are not sold.

Preferred citation:

Leung, C.-K., Chang, H.-H., & Hau, K.-T. (2003). Computerized adaptive testing: A comparison of three content balancing methods. *Journal of Technology, Learning, and Assessment*, 2(5).
Available from <http://www.jtla.org>

Abstract:

Content balancing is often a practical consideration in the design of computerized adaptive testing (CAT). This study compared three content balancing methods, namely, the constrained CAT (CCAT), the modified constrained CAT (MCCAT), and the modified multinomial model (MMM), under various conditions of test length and target maximum exposure rate. Results of a series of simulation studies indicate that there is no systematic effect of content balancing method in measurement efficiency and pool utilization. However, among the three methods, the MMM appears to consistently over-expose fewer items.

Computerized Adaptive Testing: A Comparison of Three Content Balancing Methods

In the past two decades, the advancement in computer technology and psychometric theories has accelerated the change of test format from conventional paper-and-pencil tests to computerized adaptive testing (CAT)¹. In CAT, each examinee is presented with a tailor-made test in which one item at a time is adaptively selected on the basis of the currently available estimate of the examinee's ability (Lord, 1980; Weiss, 1982). One of the main advantages of this type of assessment over paper-and-pencil testing is that CAT enables more efficient and precise trait estimation (Owen, 1975; Wainer et al., 1990; Weiss, 1982). That is, it provides a means for more precisely estimating each student's ability in a particular domain without extending the length of the test.

Unfortunately, this assessment system is not without concerns. Perhaps the most salient issues raised in regards to the extended use of CAT are item overexposure and face validity (Leung, Chang, & Hau, 2000). Depending on the item selection algorithm used in CAT application programs, particular items in the item pool may be over-selected. That is, items that provide the most discriminating information to the CAT system about the examinee's ability may be administered to numerous participants and become familiar to test takers prior to testing, thus diminishing test security and reliability. In addition, if items are found to be over-selected and risk exposure, additional item development will be required, in effect increasing costs for CAT maintenance. It is inefficient to require additional development of items while a large proportion of the item pool remains unused. To limit overexposure and its effects, the item selection method needs to select discriminating items while considering pool utilization.

Item selection is also confounded by non-statistical issues such as content balancing. By nature of an adaptive test, examinees sitting to take the same test will be administered different items but each must receive the same distribution of items by content area. For example, for a 28 item mathematics test it would not be valid to administer 28 items on arithmetic to one student and 28 items on geometry to another. There must be a balance across content areas or domains measured.

As Davey and Parshall (1995) indicate, CAT may have conflicting goals of maximizing test efficiency and limiting the overexposure of individual items, while at the same time measuring the identical composite of traits across examinees through administration of items with the same content properties. Several methods have been developed in striving to achieve content balancing while maintaining test efficiency and controlling the exposure rate of items. To develop a better

understanding of how well different CAT program algorithms function, this simulation study evaluated the performance of three content balancing methods under different conditions of length and item exposure.

Content Balancing Methods

The three content balancing methods examined in this study include the constrained CAT (CCAT), the modified multinomial model (MMM), and the modified constrained CAT (MCCAT). Kingsbury and Zara (1989) proposed the popular constrained CAT (CCAT) method. This content-balancing algorithm selects the most optimal item from the content area with the current exposure rate farthest below its target administration percentage. Chen, Ankenmann, and Spray (1999) argued that the CCAT may yield undesirable order effects as the sequence of content areas is highly predictable. Instead, they developed a modified multinomial model (MMM) to meet the balanced content requirement. Subsequently, Leung, Chang, and Hau (2000) proposed a modified constrained CAT (MCCAT) to eliminate the predictability of the sequence of content areas of CCAT and to satisfy the practical constraint of content balancing as well. The degree to which each method is beneficial continues to be studied in light of item selection and item exposure control.

Efficiency and Exposure Control

To attain high efficiency in CAT, many item selection algorithms adopt an approach in which an item is selected if it has the maximum Fisher information at the current ability estimate based on the responses to previously administered items. Item information increases as the item difficulty approaches the examinee's ability, the discrimination parameter increases, and the probability of guessing correctly is close to zero (Hambleton & Swaminathan, 1985, pp. 104–105). Unfortunately, it has been noted that this information criterion would cause unbalanced item exposure distributions (Davey & Parshall, 1995; McBride & Martin, 1983; Simpson & Hetter, 1985; van der Linden, 1998). By selecting the item that has the maximum Fisher information, the program is selecting the item that will provide the most information of value at the given ability level. Thus, highly discriminating items may be overly exposed while some less discriminating items may never be used. As described previously, overexposing items introduces the possibility of damaged test security and of increased cost in developing and maintaining item pools.

To directly control the exposure rate of frequently selected items in maximum information item selection, Simpson and Hetter's (1985) probabilistic method, known as the SH method, is utilized. While the SH method has proven to be capable of controlling overexposure, it should be noted that this method cannot directly increase the usage of those items that are rarely selected. Acknowledging these limitations, the SH method provides a usable construct for reducing item overexposure at target rates through an exposure control algorithm.

Prior Research on Item Selection and Content Balancing

Prior research has been conducted comparing the three content-balancing methods using three different item selection designs: multi-stage a -stratified design (ASTR), a -stratified with b -blocking design (BASTR), and content-stratified, a -stratified with b -blocking design (CBASTR). In essence the three designs build on each other such that the ASTR, which was proposed by Chang and Ying (1999), partitions items into several strata in an ascending order of the item discrimination (a) parameter. Each test administered consisted of matching numbers of stages and strata, with items of the first stage being selected from the first stratum that mainly contains less discriminating items, and so on. One major rationale for such a design is that in early stages, the gain in information by using the most informative item may not be realized because the ability estimation is still relatively inaccurate. Thus, items with high discrimination values are saved for later stages. Through simulation studies, the ASTR has been shown to be effective in both reducing item-overexposure rate and enhancing pool utilization. BASTR extends ASTR by first dividing pools of items into several strata based on b parameter (item difficulty) and then stratifying by the a parameter (Chang, Qian, & Ying, 2001). Yi and Chang (in press) extended BASTR by dividing items into strata based on three factors, namely content, b parameter, and then a parameter. The research on content balancing under these different stratification designs indicates that the CCAT, MMM, and MCCAT have similar effects on measurement efficiency but the CCAT is consistently less effective than the other two methods in terms of pool utilization and control of item overexposure rate (Leung, Chang, & Hau, 2003). The current research study examined the same three content balancing methods using the maximum information item selection method. In the discussion, the results of this current simulation study and the prior study, comparing content balancing methods using various stratification designs, will be compared and contrasted, elaborating on the effectiveness of each method in addressing the issues of item exposure and face validity.

Method

Content Balancing Methods

The three content balancing methods studied are as follows:

- (1) The Constrained CAT (CCAT): The selection of an optimal item is restricted to the content area with current exposure rate farthest below its target percentage for the test.
- (2) The Modified Multinomial Model (MMM): A cumulative distribution is first formed based on the target percentages of the content areas that sum to 1.0. Then, a random number from the uniform distribution $U(0,1)$ is used to determine the corresponding content area in the cumulative distribution where the next optimal item will be selected. When a content area has reached its target percentage, a new multinomial distribution is formed by adjusting the unfulfilled percentages of the remaining content areas. As random mechanism is incorporated in this method, the sequence of content areas varies.
- (3) The Modified Constrained CAT (MCCAT): Instead of being restricted to the content area that has current exposure rate farthest below its target percentage, an optimal item can be chosen from all the content areas that still have quota not fully used up. As a result, the undesirable order effect of CCAT is eliminated.

Exposure Control

The foundation of the SH control algorithm rests on the concept of conditional probability: $P(A) = P(A|S) * P(S)$, where $P(S)$ is the probability that the item selected is the best next item for a randomly sampled examinee from a typical population, and $P(A|S)$ is the probability that the item is administered when selected. The procedure attempts to control $P(A)$, the overall probability that an item is administered, by assigning an exposure control parameter $P(A|S)$ to the item. The exposure control parameters for all items are determined through a series of prior adjustment simulations so that the probability of administration for each item is restricted to the pre-specified maximum exposure rate (Simpson & Hetter, 1985).

Simulation Design

Item pool: A pool of 700 calibrated mathematics items from four major content areas was used. The content areas included Numbers, Measurement, Data Handling, and Geometry which contained 234, 166, 150, and 150 items, respectively.

Test length: To investigate the effect of the content balancing methods on short, moderate, and long tests, three test lengths of respectively 16, 28, and 40 items were studied.

Content specifications: For each 16-item test, the numbers of items from the four content areas (Numbers, Measurement, Data Handling, and Geometry) were 6, 4, 3, and 3. The numbers of items from Numbers, Measurement, Data Handling, and Geometry were 11, 7, 5, and 5, respectively, for the test of 28 items, and 14, 10, 8, and 8, respectively, for the test of 40 items.

Exposure rate: Two target maximum exposure rates of .1 and .2, respectively representing stringent and less stringent exposure control conditions, were studied.

Ability traits: A sample of 5000 simulees with abilities (θ s) randomly generated from $N(0,1)$ was drawn. Each simulee received an adaptive test from each of the 18 combinations (3 methods $\times$ 3 test lengths $\times$ 2 exposure rates) of conditions.

Ability estimation: An interim Bayes estimate of θ was used during testing and then a final estimate was obtained by maximum likelihood estimation.

Evaluative Criteria

The performances of the content balancing methods were evaluated in terms of (i) correlation of true and estimated theta, (ii) average bias, (iii) mean square error, (iv) scaled chi-square statistic (Chang & Ying, 1999), (v) number of under-utilized items, and (vi) number of over-exposed items.

Correlation of True and Estimated Theta

Irrespective of the item selection design, CAT should always provide highly correlated estimates for individual examinee abilities, otherwise the test results could not be reliably used for inference or decision-making. In this study, Pearson correlation coefficient was used for comparison.

Average Bias and Mean Square Error

Bias was estimated using Formula 1. For Equation 1, we let θ_i , $i = 1, \dots, N$ be the true abilities of N examinees and $\hat{\theta}_i$ be the respective estimators from the CAT. Then the estimated bias is computed as

$$Bias = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_i) \quad (1)$$

Mean square error (MSE) was calculated using Equation 2:

$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2 \quad (2)$$

A smaller bias and MSE indicates a better item selection method.

Scaled Chi-Square Statistics

Chang and Ying (1999) have proposed that a uniform exposure rate distribution should be the most desirable in order to have a maximum item pool utilization. They introduced a scaled chi-square to measure the overall item pool usage efficiency. Equation 1 reflects the discrepancy between the observed and the ideal exposure rates.

$$\chi^2 = \sum_{j=1}^N \frac{\left(er_j - \frac{L}{N}\right)^2}{\frac{L}{N}} \quad (3)$$

If the entire pool size is N and the test length is L , then the optimum uniform exposure rate is $\frac{L}{N}$. er_j represents the observed exposure rate for the j^{th} item. As χ^2 decreases in value, pool utilization improves and the utility of the item selection method increases.

Under-Utilized and Over-Exposed Items

The exposure rate of an item is defined as the ratio of the number of times the item is administered to examinees over the total number of examinees taking the test. If there are too many items with low exposure rates, then the item pool is not well utilized, which challenges directly the cost effectiveness of the item pool and the appropriateness of the item selection method. In this study, an item was considered as under-utilized if its exposure rate was below .02. The smaller the number of under-utilized items, the better the item selection method is.

If an item has a high exposure rate, then the item has a greater risk of being known to prospective examinees, which in turn would cause test security and validity problems. Here an item was considered as overly exposed if its exposure rate was greater than the corresponding target maximum exposure rate (.1 or .2, depending upon the trial). The smaller the number of over-exposed items, the better the item selection method performs.

Results

The results of the study are summarized in Table 1. As seen in Table 1, all three content balancing methods yielded estimated thetas that are highly correlated with examinees' "true" ability. Across all three methods, the correlations are nearly identical for a given test. As expected, the correlations increase as the test length increases. Similarly, the correlations are slightly higher when the target exposure rate is set as .2 as compared to .1.

As seen in Table 1, the estimated bias for all three content balancing methods are close to zero, indicating that all three methods appear virtually unbiased. In addition, the MSE are relatively small for all three methods. Again, not surprising, the MSEs decrease as the test lengths increase and as the target maximum exposure rate increases from .1 to .2.

Table 1 Summary Statistics for Three Content Balancing Methods

	CCAT	MMM	MCCAT
16-item test	$r = .1$ ($r = .2$)	$r = .1$ ($r = .2$)	$r = .1$ ($r = .2$)
Correlation	.954 (.961)	.954 (.960)	.955 (.960)
Bias	-.006 (.009)	.005 (.008)	.007 (.005)
MSE	.102 (.089)	.101 (.089)	.101 (.090)
Scaled χ^2	48.9 (98.3)	47.6 (95.2)	47.7 (95.5)
N(exp<.02)	521 (583)	520 (576)	520 (578)
N(exp> r)	60 (29)	52 (21)	55 (23)
28-item test			
Correlation	.971 (.975)	.970 (.973)	.970 (.975)
Bias	-.000 (.003)	-.001 (.004)	-.001 (.001)
MSE	.064 (.053)	.065 (.054)	.064 (.053)
Scaled χ^2	37.7 (90.8)	37.1 (88.7)	37.0 (88.3)
N(exp<.02)	393 (501)	380 (499)	386 (498)
N(exp> r)	119 (44)	108 (37)	109 (35)
40-item test			
Correlation	.976 (.981)	.975 (.980)	.976 (.981)
Bias	.003 (.002)	.005 (-.002)	-.004 (.000)
MSE	.054 (.040)	.054 (.040)	.054 (.040)
Scaled χ^2	27.7 (80.7)	26.1 (80.5)	26.2 (80.2)
N(exp<.02)	271 (432)	258 (432)	260 (430)
N(exp> r)	178 (65)	157 (55)	171 (64)

With respect to pool utilization, the CCAT yielded slightly higher values in scaled χ^2 and larger numbers of under-utilized items than the MMM and the MCCAT. Nevertheless, the difference appears relatively minor as reflected in Figures 1 and 2. As evidenced by larger χ^2 values and larger numbers of under-utilized items, the item exposure distribution becomes more skewed when the target maximum exposure rate increases from .1 to .2. To the contrary, when the test length increases, the item exposure distribution becomes more even.

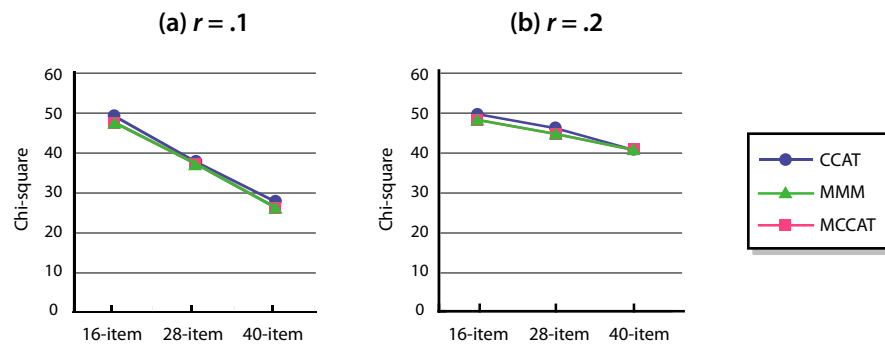
Figure 1

Figure 1: Chi-square statistics across content balancing method and test length for target maximum exposure rates of .1 and .2.

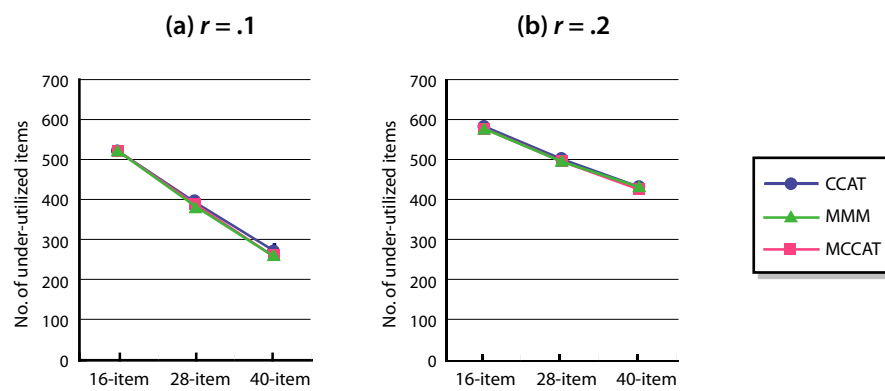
Figure 2

Figure 2: Number of under-utilized items across content balancing method and test length for target maximum exposure rates of .1 and .2.

As seen in Figure 3, the MMM method yielded fewer over-exposed items while the CCAT generally produced the most over-exposed items. As test length increased, the discrepancy between the numbers of over-exposed items also increased among the three methods. It should also be noted that the number of over-exposed items was noticeably smaller for all three methods when the target exposure rate was set at .2 as compared to .1.

Figure 3

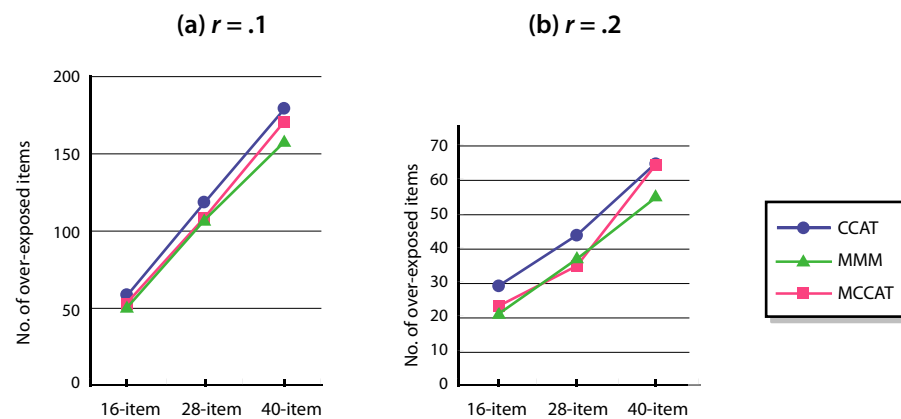


Figure 3: Number of over-exposed items across content balancing method and test length for target maximum exposure rates of .1 and .2.

Note. Y-axis scales differ in *a* and *b*.

Discussion

Content balancing is a common requirement of many large-scale educational tests. Prior research has examined the performance of three content balancing methods (CCAT, MMM, and MCCAT) under three different stratification conditions (ASTR, BASTR, and CBASTR). This prior research found that all three content balancing methods yielded similar correlations between estimated and “true” ability and produced comparable bias and mean squared error estimates across all three stratification conditions. Differences, however, were found with respect to item pool utilization, such that the MMM method generally performed best across all three stratification designs. This prior research focused on a test containing 35 items and a targeted maximum item exposure of .2 (Leung, Chang, & Hau, 2003).

The study presented in this article compared the performance of CCAT, MMM, and MCCAT under a single condition of item selection design (maximum information selection SH method) but under three different test lengths (16, 28, and 40 items) and two target exposure control levels (.1 and .2). Results of the present study indicate that the three content balancing methods, when used in conjunction with the traditional maximum information selection approach, offered comparable estimation accuracy and precision in terms of MSE, bias, and correlation coefficient. This finding is consistent with prior research that found that all three methods resulted in similar measurement efficiency when used in conjunction with three different stratification designs (Leung, Chang, & Hau, 2003). In addition, the present study also found that the test length and target maximum exposure rate are two significant factors that affect measurement performance. Specifically, the present study found that accuracy and precision increased as test length increased from 16 to 28 to 40 items. Similarly, the accuracy and precision increased as the target maximum exposure rate increased from .1 to .2.

Similar to prior research (Leung, Chang, & Hau, 2003), the three content balancing methods examined in this study resulted in different numbers of over-exposed items. As found previously, CCAT generally resulted in higher numbers of over-exposed items. In contrast, the MMM method tended to over-expose a smaller number of items and thus appears to be favorable with respect to item security control.

It should also be noted that the content area of items presented via the CCAT method was highly predictable. In general, the first few items in each test presented via CCAT came from the content area with the largest pre-specified percentage. The content of subsequent items then tended to cycle in a predictable manner. Again, this finding is consistent with the prior research.

The results of the present study and the prior research differ with respect to pool utilization. When employing a maximum information item selection approach, the three content balancing methods showed no difference in terms of the scaled chi-square statistics, an indicator of pool utilization. However, when used in conjunction with three different stratification designs (ASTR, BASTR, and CBASTR), the

content balancing method appeared to have a systematic effect on pool utilization: The CCAT generally performed worse while MMM performed best in conjunction with these three stratification designs. One possible explanation for the difference is that the maximum information item selection ignores content sequence and instead focuses on selecting items that are highly discriminating within a given content area, whether selected at random via the MMM method or in sequence via the CCAT method. In contrast, the stratified designs utilize highly discriminating items in a progressive way by stratifying items into several groups and dividing a test into matching number of stages. Since the MMM ignores content sequence at the start of the test and modifies the probability with which a content area is selected as the test proceeds, the MMM involves more items within each stratum. On the contrary, by imposing a regular sequence to the content of the items, the CCAT restricts the number of items eligible for selection within each stratum each time an item is selected. By decreasing the pool of eligible items in a systematic manner, the CCAT appears to reduce overall pool utilization when used in conjunction with the three stratification designs.

The current findings, together with those of prior research, suggest that the MMM reduces the predictability of item content sequence and the number of over-exposed items, regardless of the item selection approach, test length, or target maximum item exposure rate. However, since the present study involves only one item pool, the advantages of the MMM over the other two methods need to be cross-examined using different item pools and under additional testing conditions.

Endnote

- ¹ CAT was first developed under the item response theory models (Lord, 1970).

References

- Chang, H.-H., Qian, J., & Ying, Z. (2001). A-stratified multistage computerized adaptive testing with *b*-blocking. *Applied Psychological Measurement*, 25, 333–341.
- Chang, H. H., & Ying, Z. (1999). A-stratified multistage computerized adaptive testing. *Applied Psychological Measurement*, 20, 213–229.
- Chen, S., Ankenmann, R. D., & Spray, J. A. (1999, April). *Exploring the relationship between item exposure rate and test overlap rate in computerized adaptive testing*. Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, Canada.
- Davey, T., & Parshall, C. G. (1995, April). *New algorithms for item selection and exposure control with computerized adaptive testing*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer-Nijhoff.
- Kingsbury, G. G., & Zara, A. R. (1989). Procedures for selecting items for computerized adaptive tests. *Applied Measurement in Education*, 2, 359–375.
- Leung, C.-K., Chang, H.-H., & Hau, K.-T. (2000, April). *Content balancing in stratified computerized adaptive testing designs*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Leung, C.-K., Chang, H.-H., & Hau, K.-T. (2003). Incorporation of content balancing requirements in stratification designs for computerized adaptive testing. *Educational and Psychological Measurement*, 63, 257–270.
- Lord, M. F. (1970). Some test theory for tailored testing. In W. H. Holzman (Ed.), *Computer assisted instruction, testing, and guidance*. New York: Harper and Row.
- Lord, M. F. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- McBride, J. R., & Martin, J. T. (1983). Reliability and validity of adaptive ability tests in a military setting. In D. J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing*. New York: Academic Press.
- Owen, R. J. (1975). A Bayesian sequential procedure for quantal response in the context of adaptive mental testing. *Journal of the American Statistical Association*, 70, 351–356.

- Sympson, J. B., & Hetter, R. D. (1985). Controlling item-exposure rates in computerized adaptive testing. *Proceedings of the 27th Annual Meeting of the Military Testing Association* (pp. 973–977). San Diego, CA: Navy Personnel Research and Development Center.
- van der Linden, W. J. (1998). Bayesian item selection criteria for adaptive testing. *Psychometrika*, 63, 201–216.
- Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., Mislevy, R. J., Steinberg, L., & Thissen, D. (1990). *Computerized adaptive testing: A primer*. Hillsdale, NJ: Erlbaum.
- Weiss, D. J. (1982). Improving measurement quality and efficiency with adaptive testing. *Applied Psychological Measurement*, 6, 473–492.
- Yi, Q., & Chang, H.-H. (in press). Multiple stratification CAT designs with content control. *British Journal of Mathematical and Statistical Psychology*.

Acknowledgements

The research was supported in part by the research grant (RG13/ 2002–03) of the Hong Kong Institute of Education. The authors appreciate the comments provided on earlier drafts by Michael Russell and Kathleen O'Connor, the editors, and the three anonymous reviewers.

Authors' Note

Parts of this research were presented at the NCME Annual Meeting in Chicago, April 2003. Please send all correspondence to Dr. Chi-Keung Leung at the Department of Mathematics, The Hong Kong Institute of Education, Tai Po, Hong Kong (Email: ckleung@ied.edu.hk).

Author Biographies

Chi-Keung Leung is lecturer in the Department of Mathematics at The Hong Kong Institute of Education. His research interest includes computerized adaptive testing, mathematics assessment and problem solving.

Hua-Hua Chang is an associate professor in the Department of Educational Psychology, University of Texas at Austin. His specialization includes computerized adaptive testing, large-scale assessment, and item response theory.

Kit-Tai Hau is a professor and the Chair of the Department of Educational Psychology at the Chinese University of Hong Kong. His research interest includes structural equation modeling, academic achievement motivation, policy on educational assessment, and adolescent suicide.



The Journal of Technology, Learning, and Assessment

Editorial Board

Michael Russell, Editor
Boston College

Allan Collins
Northwestern University

Cathleen Norris
University of North Texas

Edys S. Quellmalz
SRI International

Elliot Soloway
University of Michigan

George Madaus
Boston College

Gerald A. Tindal
University of Oregon

James Pellegrino
University of Illinois at Chicago

Katerine Bielaczyc
Harvard University

Larry Cuban
Stanford University

Lawrence M. Rudner
University of Maryland

Mark R. Wilson
UC Berkeley

Marshall S. Smith
Stanford University

Paul Holland
ETS

Randy Elliot Bennett
ETS

Robert J. Mislevy
University of Maryland

Ronald H. Stevens
UCLA

Seymour A. Papert
MIT

Terry P. Vendlinski
UCLA

Walt Haney
Boston College

Walter F. Heinecke
University of Virginia

www.jtla.org