

## **Redeveloping the exposure control parameters of CAT items when a pool is modified**

Shun-Wen Chang  
National Taiwan Normal University

Deborah J. Harris  
ACT, Inc.

Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, April 2002.

## **Abstract**

Both pool management and exposure control method implementation have been the focus of easing test security concerns in CAT. Maintaining pools of high quality is an ongoing process and sometimes only a small modification may be needed. This study explored how the deletion of a single item and the unused items might alter the exposure control parameters of the remaining items derived by the SLC algorithm.

Simulations were carried out for this study. The findings indicated that the original exposure control parameters were no longer appropriate for a modified item pool and the derivation for the exposure control parameters ought to be repeated. Removing the unused items seemed to cause no effects on the performance of the SLC procedure in the observed exposure rates and measurement precision. The characteristics of the unused items were examined and compared with those of the high, moderate and low exposure groups, respectively. Results from this study have offered a better understanding of the effects of item interactions on the values of the exposure control parameters. Also, this study should provide psychometric researchers and test practitioners valuable insights for a better pool construction design.

**Keywords:** computerized adaptive testing, item exposure rate, exposure control parameters, item pool

Test security is an issue of great concern for large scale, high-stakes tests in the continuous testing context of computerized adaptive tests (CAT). Efforts have been made to prevent some popular items from being overly exposed to examinees. This issue of controlling exposure rates of CAT items has been studied extensively through the management of item pools, as well as the incorporation of exposure control methods into the item selection process.

For pool management, the focus is on the CAT pool design, creation, rotation and maintenance over time (Guo, Way, & Reshetar, 2000; Mills & Steffen, 2000; Stocking & Lewis, 2000; van der Linden, 2000; Veldkamp & van der Linden, 2000; Way, Swanson, Steffen, & Stocking, 2001). Exposure rates of items can be limited through rotation and regulation of multiple parallel pools. It was found that through pool rotation, the impact of collusion among test-takers was reduced dramatically (Mills & Steffen, 2000). Since adaptive tests are administered on a continuous basis, a good pool management system is essential for secure and effective operational CAT administrations.

In addition to well-managed item pools, methods have been developed by embedding statistical mechanisms in the item selection procedure to control the exposure rates of items to a desired maximum value,  $r$ , that is specified in advance of testing (Davey & Parshall, 1995; Hetter & Simpson, 1997; Stocking & Lewis, 1998, 2000; Simpson & Hetter, 1985). Most methods control the exposure rate of an item through the *exposure control parameter*, which dictates the probability of administering the item, given it is selected. Provided that an item has been selected, whether to administer this item to the examinee depends upon the exposure control parameter of the item. For very popular items, the exposure control parameters could be as low as the pre-specified desired exposure rate, indicating that these items cannot be freely administered when they are selected. For items rarely appearing, the associated exposure control parameters could be as high as 1.0, meaning that once these items are selected, they are almost always presented. The values of these parameters are determined from a series of iterative simulations using all items in the pool.

Investigations on the properties of the various exposure control strategies showed that the Stocking and Lewis conditional multinomial (SLC) procedure (Stocking & Lewis, 1998, 2000) was the most effective (Chang, 1998). The specific feature of the SLC algorithm that makes it the most appealing is that this method develops for each item an exposure control parameter with respect to a particular ability level. While the unconditional methods of exposure control only limit the item's overall appearance in reference to the examinee sample drawn from a target population, the SLC procedure directly controls the item exposure to examinees at the same or similar levels of proficiency. This procedure controls against an item being administered to almost all examinees at one particular ability level, even if the item's overall exposure rate is low for examinees across the entire ability continuum.

Since the exposure control parameters of an item are developed based on the existence of all other items in the pool, when there is any significant change in test structure of the pool, the exposure control parameters must be redeveloped. Stocking and Lewis (2000) pointed out that if an item pool is changed, even by the addition or deletion of a single item, the iterative simulations should be carried out again to guarantee continued exposure control for the new item pool. Maintaining pools of high quality is an ongoing process. Items are continuously under review and modifications of the pools are possible, even likely. It is conceivable that such iterative simulations must be repeated very frequently for the newly formed item pools. However, developing the exposure control parameters for the SLC method is especially tedious and time-consuming. If the modification is only to withhold a single item from use in the pool for some reason (for example, the item becomes dated for content reasons or it is found defective due to some flaw), is it still necessary to redevelop the exposure control parameters of the remaining items for this procedure? How much weight does one item carry in the values of the exposure control parameters of the other items in the pool? It is worthwhile to find out whether using the existing exposure control parameters in the modified pool could be warranted for the SLC algorithm. By doing so, not only could time savings accrue in the preparation of the exposure control parameters, but the interactions of these parameters could be further understood.

Research studies and practical experience have indicated that CAT administrations often lead to numbers of items that are seldom or never used, depending on the exposure control method and item pool employed (Chang, 1998; Chang, Ansley, & Lin, 2000; Veldkamp & van der Linden, 2000). According to the maximum information selection rule, the unused items are those with relatively low item information values. When compared with other items for a given ability level, they may not be competitive enough in contributing to the estimation of examinees' abilities. The unused items seem to make little contribution to the test. Therefore, it might be reasonable to suspect that removing the set of unused items would not affect the existing exposure control parameters of all other items, so that repeating the derivation process for the remaining items in such a reduced pool may not be necessary. This study will examine the effect of removing the unused items on the exposure control parameters for the SLC strategy.

Also, because the unused items are never administered to the examinees, will the CAT procedure still perform as well without these items in the pool, using either the existing or the redeveloped exposure control parameters? Is there still a sufficient number of items appropriate for selection in this reduced pool? These questions deserve a thorough investigation, so that the impact of unused items can be further understood, and values of the exposure control parameters as a result of the item interactions can be better described. At the same time, it is important to inspect more carefully the characteristics of the unused items, despite the fact that they possess very low information values for a given ability level.

Because unused items are also a consequence of utilizing a particular exposure control method for a specific item pool, this inspection will reveal information about the nature of the SLC algorithm in relation to the test structure of the pool. Findings of this study should be useful for constructing an even better pool, in terms of item usage, for the SLC method.

### **Purpose of the Study**

This study is designed to explore whether the removal of a single item or the unused items from the pool impacts the exposure control parameters of the remaining items in a CAT; that is, whether redeveloping the exposure control parameters is necessary when some items are eliminated from the pool. Specifically, this study attempts to achieve the following objectives:

1. to explore whether the exposure control parameters of the items are affected by removing a single item or the unused items from the pool;
2. to examine how the observed item exposure rates and measurement precision are affected by using the reduced pool where the unused items are eliminated;
3. to offer more information about the performance of the exposure control method in relation to the test structure of an item pool.

### **Method and Data**

Simulations were employed to carry out this research. The study design and methods for data analyses are described below.

#### **I. Decisions for the CAT Components**

The 3-PL model of IRT forms the basis for the investigations of this study. A single pool of 480 discrete items was used, in which the item parameters were calibrated from multiple forms of a large scale standardized paper-and-pencil test using BILOG (Mislevy & Bock, 1990) on a single ability scale.

The CAT administration process was initiated by assigning each examinee a common ability estimate of zero to simulate a situation where no a priori information was available about the individuals. Items were selected based on the maximum item information criterion with the SLC algorithm incorporated into the selection process. The target maximum exposure rate was specified to be .20. The content presented to the examinees was balanced according to the Kingsbury and Zara's mechanism (1989). The first item was administered following the algorithm of item exposure control, regardless of the item's content attribute. The percentage of items that had been administered in each content category was calculated and compared to the corresponding pre-specified percentage. The content area with the largest discrepancy between the empirical and the desired percentage was then identified; the

next item was selected from this content area and administered based on the algorithm of the SLC method. To estimate the examinees' abilities, Owen's Bayesian strategy (Owen, 1975) was employed for the provisional ability estimation and the maximum likelihood estimation method (Birnbaum, 1968) was utilized for the final ability estimation. Each examinee was administered 30 items.

## **II. Design of the Study**

Two explorations were designed for this study.

### **Study I: Removing a Single Item**

This study was designed to examine whether the removal of a single item caused the exposure control parameters of the remaining items to change. There are various reasons an item may become inappropriate and ought to be excluded from the current pool, and it may not necessarily be an item with very low information value for a given ability level. In addition, the information value an item possesses varies by ability level. An item with low information at one ability level may be associated with high information values for other ability levels. For the purpose of evaluating the removal effect of one item, this study considered three items of different information values at three selected ability levels. Arbitrarily, items of the highest, the middle, and the lowest information were chosen for removal at each of the theta points of -2.8, 0.0 and 2.8, respectively, with no considerations of the content area the item belonged to.

### **Study II: Removing All Unused Items**

For this part of the investigation, all of the unused items resulting from the operational CAT administrations were eliminated from the pool. *The reduced pool* is named hereafter for this condition to indicate the pool established based on the full pool without the unused items.

## **III. Procedures for Data Simulation**

Three parts of the simulations were involved in generating the data for this study. First, the exposure control parameters for each item in the pool were developed according to the SLC algorithm and the CAT design proposed for this study. Then, the simulations continued for the operational CAT administrations. The designated items were then removed from the pool and the exposure control parameters of the remaining items were redeveloped and applied to the operational testing situations.

### **Stage I: Developing Exposure Control Parameters**

The development of the SLC exposure control parameters was in reference to a particular level of proficiency. The adaptive tests were administered to a conditional sample of 3,000 examinees at each of the theta levels equally spaced over the interval of -3.2 and 3.2 with an increment of .40 (i.e., -3.2, -2.8,..., 3.2), totalling 17 ability points. This could be regarded as administering the adaptive tests to a

sample of 51,000 examinees (with 3,000 examinees at each of the 17 ability levels) whose abilities were uniformly distributed over the range of -3.2 and 3.2 on the theta scale.

The desirability of items for the SLC method was ordered based only on item information values, not on the weights as specified in the original algorithm for which the Stocking/Swanson weighted deviations model (WDM) (Stocking & Swanson, 1993; Swanson & Stocking, 1993) was employed to select items. The process was repeated until the observed maximum exposure rates were approximately equal to the desired level and the exposure control parameters were stabilized in the subsequent iterations. The stabilized parameters at the final round of iterations were the exposure control parameters to be used in operational adaptive testing. For each of the 480 items, a set of exposure control parameters was obtained for each discrete ability level. The parameters could be visualized as the elements in a matrix of 17 ability levels by the number of items in the pool. These were the parameters existing in the full pool of 480 items.

#### Stage II: Simulating Operational CAT Administrations

During the operational CAT administrations, the adaptive test was delivered to each examinee in a sample of 50,000 examinees drawn from a standard normal distribution,  $N(0,1)$ , being representative of the real examinee population. The exposure control parameters developed from the previous stage were utilized here to manage the administration frequencies of the selected items. The number of times an item was actually administered and the number of items in the pool never administered were recorded.

The adaptive tests were also administered to conditional sample sizes of 3,000 examinees over 17 equally spaced ability points from -3.2 to 3.2. This step was carried out in order to compute the conditional maximum observed exposure rates and the error values conditional on each theta point.

#### Stage III: Removing the Single and Unused Items

At this stage, each of the chosen single items and the set of the unused items were, separately, removed from the pool. The items with the highest, moderate, and lowest information values at each of the theta points of -2.8, 0.0 and 2.8 were identified from an Info Table, which contained lists of the items, ordered by the amount of information they provided at the various ability points. The unused items were those never administered in Stage II, when the full pool of 480 items was applied in the operational CAT administrations. After each deletion, the exposure control parameters of the remaining items were redeveloped. The new set of exposure control parameters was compared with the existing parameters.

The process continued for the operational CAT administrations. For each modified pool condition, both sets of the existing and redeveloped exposure control parameters were respectively employed to control the administration frequencies of the selected items. The results for the various modified pools with the existing and redeveloped exposure control parameters were evaluated.

#### IV. Methods for Data Analyses

To analyze the extent to which the removed items affect the exposure control parameters of the remaining items, the descriptive statistics of the differences between the existing and redeveloped exposure control parameters were computed for each ability point. Also, the Pearson product-moment correlations of the parameters were obtained for the various ability levels. The results of applying the existing and redeveloped exposure control parameters in the operational testing situations were evaluated based on the maximum observed exposure rates, pool usage and the conditional standard errors of measurement (CSEMs). The maximum exposure rates observed were investigated in reference to the entire examinee group and also to the examinees at a particular ability point. The item pool usage was assessed using the numbers of items in the pool administered at least once. The CSEMs were the errors of the ability estimation as a result of introducing the SLC algorithm with the different sets of the exposure control parameters into the item selection process. Meanwhile, the results based on the full pool of 480 items were incorporated for evaluating the effects of removing the unused items from the pool.

The characteristics of the unused items were inspected with respect to the descriptive statistics of the three IRT item parameters, and the item information values and exposure control parameters for each ability point. These values were compared with those of the used items having high, moderate and low observed exposure rates, respectively. These used items were ordered based on the exposure rates observed in the operational CAT administrations and were classified in such a manner that the first third of the items were in the high exposure group, the second third were in the moderate exposure group and the last third were classified into the low exposure group.

#### Results and Discussion

The exposure control parameters of all items in the initial full pool were derived in reference to each ability level. These parameters were employed in the operational CAT administrations to control the exposure rates of the items. Forty-eight items in the full pool were never administered to examinees. These unused items and each of the designated single items were, separately, removed from the full pool and the exposure control parameters of the remaining items were redeveloped. For each single item deletion, the new parameters could be visualized as the elements in a matrix of 17 ability levels by 479 items. For the unused items deletion, the parameters could be visualized as the elements in a matrix of 17 ability levels by 432 items.

Figure 1 shows the results of the maximum exposure rates observed at each ability point during the iteration process for both full and reduced pools. A horizontal line is presented to indicate the pre-specified exposure rate of .20. Only the results of the full pool and the reduced pool of 432 items are



presented. It can be seen that the iteration results were very similar for both situations, even with the deletion of 48 items. The iteration results for the single item deletion conditions were virtually no different from the configurations displayed here.

Figure 1 also shows that the iteration curves did not approach the desired rate of .20. Also, the conditional observed maximum exposure rates for the various ability points seem to converge to different values. The closer the iterations were at the extreme ability levels, the higher the values the conditional observed maximum exposure rates converged to. These results reflected the nature of the item pools, in that there were fewer items appropriate for administration for the extreme ability levels than for the middle. The frequency distribution of the item difficulty values (i.e., the  $b$  parameters) for the full pool is displayed in Figure 2 to provide some idea of the numbers of items appropriate for the various ability levels.

Described below are the results of the relationships of the existing and redeveloped exposure control parameters, followed by the results of employing the existing and redeveloped exposure control parameters for the various modified pools in the operational testing situations. Then, the characteristics of the unused items are presented.

## **The Relationships of the Existing and Redeveloped Exposure Control Parameters**

### **The Differences of the Exposure Control Parameters**

Table 1 reports the results of the differences between the existing and redeveloped exposure control parameters for each theta point for the various removal conditions. The conditions for the single item removal are represented by the capital letters H, M, and L to indicate the highest, moderate and lowest informative items for removal at each of the three ability points of -2.8, 0.0 and 2.8, respectively. For the unused items removal condition, it is simply noted as UNUSED. The N indicated under each removal condition is the number of items remaining in the pool after the deletion. The N\_Unchanged line lists the number of items for which the redeveloped exposure control parameters remained the same as the existing parameters. The MDIFF line contains the average values of the differences between both sets of the parameters, with the redeveloped exposure parameters being subtracted from the existing ones (i.e., MDIFF = average of [existing parameters minus redeveloped parameters]). The MDIFFABS line reports the average values of the absolute differences between parameters.

Under the various conditions of removing the single items, the differences in the exposure control parameters seem insubstantial. Table 1 shows that about 66% to 75% of the exposure control parameters remained unchanged for the various ability levels, with fewer items having parameters unchanged at the upper middle part of the ability scale (around the theta points of -.80 to 1.6) than at both extremes. The findings indicate that removing one single item caused more items to have their

exposure control parameters changed for the upper middle part of the ability scale than at both extremes. This phenomenon might be explained by the fact that there were more items appropriate for administration for the upper middle part of the ability scale, as can be seen in Figures 1 and 2. Items appropriate for the upper middle part of the theta scale were thus used more often. When the pool was modified, a greater number of items were affected over this upper middle range than the two ends of the ability scale.

As for the comparisons among the various conditions of deleting items with the varying degrees of information values, there was no clear indication as to whether the effects were different. Removing items with the varying degrees of information values for the different ability levels did not seem to have noticeably differential effects on the results.

The values of MDIFF were negative or positive for the various single item removal conditions (see Table 1), as the redeveloped exposure control parameters could be larger or smaller than the existing parameters. However, it is somewhat interesting to see that the lower part of the ability scale tends to result in positive MDIFF values (i.e., on average, the redeveloped parameters were smaller than the original ones) while the middle and upper parts seem to produce more negative MDIFF values (i.e., on average, the redeveloped parameters were greater than the original ones). When the parameters were redeveloped to be smaller, the probability of administering an item once selected was reduced; in other words, the control for item exposure became stricter. When the parameters were redeveloped to be greater values, the probability of administering an item given selection was increased; the control for item exposure was lessened. The deletion of a single item might provide some opportunities for items appropriate for the middle and upper parts of the ability scale to be administered more freely.

It is expected that the average observed exposure rate of items would be increased when the pool size is reduced. That is, items would, in general, be used more often when there are fewer items in the pool. However, since there were more items appropriate for the middle and upper parts of the ability scale than for the extremes, the SLC algorithm opened up the opportunities for items appropriate at this middle and upper range so that their exposure control parameters were redeveloped to be higher to allow for more administrations given selection. The deletion of the items might just improve on the probability of administering items appropriate for the middle and upper parts of the theta scale.

The values of the average absolute parameter differences (see MDIFFABS in Table 1) seem to be small for the various conditions. The largest value was observed at the theta point of .80 for the various single item deletion conditions, except for H(0.0). On average, the exposure control parameters at the theta point of .80 suffered the biggest impact due to the single item removal. For the results with the H(0.0) condition, the largest MDIFFABS was observed at the theta point of -2.4, but the MDIFFABS values were in general greater than those of the other conditions. Notice that for the H(0.0) condition,

the removed item was the one possessing the greatest information value at the theta point of 0.0.

Removing this particular item led to the greatest impact among the various conditions, indicating that eliminating an item with the greatest information value for the middle ability point of 0.0 caused the greatest change on the exposure control parameters. This outcome was not surprising since the item being removed was extremely popular for most examinees, and was almost always selected.

Other than the phenomenon observed with the removal of the most informative item for the middle ability point of 0.0, the values of the MDIFFABS do not seem to show a clear pattern as to what effects might be due to the different items' removal. That is, except for removing the most informative item for the middle ability level, the results in MDIFFABS did not show differences among the various removal conditions. Removing any other item seems to show similar effects on the change of the exposure control parameters. These outcomes might be attributed to the fact that items possess different information values at the various ability levels. The removed item may have low information value for a given ability level, but this item might be the one with high information at another ability level.

As presented in Table 1, the results of the differences in the parameters for the unused item removal condition were fairly similar to those observed with the single item removal, except with a smaller percentage of unchanged redeveloped exposure control parameters (around 64% to 70% for the various ability levels). The upper middle part of the ability scale still had a smaller number of unchanged items than the two ends. Removing all the unused items also caused more exposure control parameters to change for the upper middle part than for the extremes, in spite of the fact that these removed items were those never administered in the operational testing using the full pool.

Table 1 shows that the MDIFF values for the unused items deletion were negative for all ability levels, indicating that the redeveloped exposure control parameters were on average greater than the original parameters. It can also be noticed that the effects seem to be stronger at the middle and upper parts of the ability scale. The interpretations might be similar to those discussed with the single item removal conditions. With fewer items in the pool, the SLC algorithm increased the values of the exposure control parameters in general to allow for more administrations given selection. Because there were more items appropriate for the middle and upper parts of the ability scale than the other parts, the SLC algorithm tended to relax the exposure control for items more appropriate for these parts to a greater extent so that higher values of the exposure control parameters were redeveloped.

As can be seen in Table 1, the average absolute parameter differences (i.e., the MDIFFABS values) for this unused item deletion were in general greater than those observed with the single item deletion conditions. It was not surprising to see that removing more items from the pool caused the parameters to change to a greater extent. The largest MDIFFABS value for this condition was still observed at the theta point of .80. On average, the exposure control parameters for this theta point still

suffered the biggest impact due to the removal of the unused items.

### The Pearson Product-Moment Correlation

The Pearson product-moment correlations are contained in Table 2. The values were all significant at the alpha .0001 level, and were almost perfect. For all the removal conditions, there were very strong correlations between the original and redeveloped exposure control parameters at all ability levels. There seems to exist no specific patterns for these correlations.

### **Results of Employing the Existing and Redeveloped Exposure Control Parameters in Operational Testing**

The exposure control parameters resulting from the final round of the iterations were associated with the corresponding items to be used in the operational testing situations. Described below are the results of administering CATs based on the various modified pools using both existing and redeveloped exposure control parameters in operational testing. Since the results for the single item removal conditions were fairly parallel to each other, only those for the H(0.0) deletion condition are presented in the table and figures for discussion. The results based on the full pool of 480 items are included for the examination of the effect of removing all unused items.

Table 3 reports the summary statistics of the observed exposure rates as a result of applying the modified pools with both existing and redeveloped exposure control parameters to the entire examinee group. The N column lists the numbers of items that were administered to examinees at least once. Based on these items, the summary statistics were obtained. It can be seen in Table 3 that the observed maximum exposure rates resulting from using the two sets of the exposure control parameters were as low as the desired rate of .20 for either of the H(0.0) and unused items removal conditions. The pool utilization was also similar for using both sets of the exposure control parameters. For the reduced pool of 432 items, 427 items were used at least once, so there were still five items left unused. Also, compared with the results of the full pool (see the lower part of Table 3), removing all the unused items did not seem to cause differences on the observed exposure rates of the items. Table 3 presents also the results of the three exposure groups, which are discussed below in the examination of the characteristics of the unused items.

The conditional maximum observed exposure rates at each ability level are displayed in Figure 3. The values were higher than the desired rate of .20 at all ability levels, which was not unexpected given that the maximum observed exposure rates were stabilized, but to values higher than the desired rate in the derivation of the exposure control parameters.

For both H(0.0) and unused items deletion conditions, the results yielded by using the redeveloped exposure control parameters were almost identical to those of the full pool. However, the

curves were different when the existing exposure control parameters were applied. The maximum exposure rates conditionally observed at the lower part of the ability scale were especially higher than those of the full pool. Recall the results with the MDIFF values for the single item removal conditions where the redeveloped exposure control parameters tended to be smaller than the existing parameters for the lower part of the ability scale (see Table 1). The SLC algorithm seemed to exercise stricter exposure control for items at the lower part of the scale in the modified pools. When the existing exposure control parameters of larger values were utilized, the results were that the conditional maximum observed exposure rates were higher than expected. Although the differences between the existing and redeveloped exposure control parameters seemed insubstantial, as discussed in the earlier section, the existing exposure control parameters were in fact no longer appropriate for the modified item pools. These findings suggest that the exposure control parameters for a modified pool need to be redeveloped in order to ensure the continued exposure control for the items. Considering the effect of removing the unused items on the conditional maximum observed exposure rates, the performance of the reduced pool was competitive to that of the full pool as long as the redeveloped exposure control parameters were utilized.

The random errors conditional on each theta point are displayed in Figure 4. The CSEM curves yielded by using the existing and redeveloped exposure control parameters were almost identical for either of the modified pools. The employment of the existing or the redeveloped exposure control parameters appears not to affect the CSEMs. Applying the reduced pool also produced similar values of CSEMs to those of the full pool. Eliminating all unused items did not seem to affect the CSEMs either.

### **Examination of the Characteristics of the Unused Items**

The number of items that were never administered in the operational testing when the full pool of 480 items was employed was recorded. For the design considered in this study, 48 items were never used. The characteristics of the unused items were examined and compared with those of the used items having high, moderate, and low observed exposure rates. Table 3 reports the summary statistics of the observed exposure rates for these three groups of items. The average observed exposure rates of the high, moderate and low exposure groups were .13, .06 and .01, respectively. The maximum observed exposure rate was controlled to be as low as the desired rate of .20 in the high exposure group.

#### **With Respect to the Item Parameters**

Table 4 contains the summary statistics of the three item parameters for the various exposure groups. For the full pool, the mean and SD of the  $a$  parameters were 1.01 and .33, the mean and SD of the  $b$ s were .15 and 1.08, and the mean and SD of the  $c$ s were .18 and .08. As expected due to the maximum item information selection, the average value of the  $a$  parameters was the largest in the high

exposure group and the values decreased for less often exposed groups. The unused items had the lowest average of the  $a$  parameters. In terms of the  $b$  parameters, the high exposure items were those of average item difficulty, the moderate exposure items were those slightly more difficult and the low exposure items were those even more difficult. The unused items tended to be the easy ones. For the  $c$  parameters, the highest exposure items were associated with the lowest average guessing value while the unused items had the highest guessing value on average.

#### With Respect to the Item Information Values

Table 5 displays the descriptive statistics of the item information values at each theta level for the three exposure groups and the unused items. It can be detected that the maximum information values for the various ability levels appeared in different exposure groups. For the middle part of the ability scale (the theta points between -1.2 and .80), the maximum information values were observed in the high exposure group, where the average information values were higher than those for the other groups as well. For the slightly lower and upper parts (the theta points of -2.0 and -1.6, and of 1.2 and 1.6, respectively), the maximum information values were seen in the moderate exposure group, for which the average information values were the highest among the various groups. As for the very high and low ends of the scale, the maximum information values existed in the low exposure group. For the unused items, the item information values were low in general, although the average information values for some ability levels at the low end were the highest among the various groups.

#### With Respect to the Exposure Control Parameters

The descriptive statistics of the exposure control parameters are shown in Table 6 by ability level and exposure group. The results based on the item information and exposure control parameters were closely related, since the exposure control parameters were derived by the algorithm of the SLC procedure based on the maximum item information criterion for item selection. For the high exposure group, the average exposure control parameters were lower at the middle part of the ability scale than at both ends, indicating that items appropriate for the middle were being controlled for their exposure to a greater extent. Thus, the high exposure group contained items more appropriate for the middle than the two ends. Also, the average exposure control parameters were slightly higher at the upper than the lower end. For the moderate exposure group, the exposure control parameters at the upper end were somewhat lower than those in the high exposure group. This might be explained by the fact that items in the high exposure group were mostly appropriate for the middle part of the ability scale, but were in general not as appropriate for the upper end as those in the moderate exposure group. For the low exposure group, there existed items associated with lower exposure control parameters at the very high end than the moderate or high exposure group. The low exposure group contained some items especially appropriate for the very high end of the ability continuum. However, the fact is that except

for these few ability levels, items at all ability levels were associated with relatively high exposure control parameters. Most items in this low exposure group were not likely to be selected. Most items were not being administered, so the average exposure rate was still low.

It can be noticed that the value of the exposure control parameters could be 1.0 for items at any ability level or in any exposure group. Even for the high exposure group, there still existed rarely appearing items whose exposure control parameters were developed to be 1.0. Such phenomenon was not unexpected, since items could not be popular across all ability levels. As has been seen in Table 5, an item might be very informative at one ability level, but not informative at all at other levels. In terms of the minimum exposure control parameters, all values were as low as .20 for the high exposure group, suggesting that the high exposure group contained very popular items needing to be strictly controlled for their exposure to the desired rate. For the moderate exposure group, only for ability levels of 1.6 and above were the exposure control parameters derived to be .20 at the minimum. Apparently, items in the moderate exposure group were less popular than in the high exposure group. For the low exposure group, there still existed some popular items at the very high ability levels. But for the middle part of the ability scale, items were not as competitive for selection as those in the high or moderate exposure group, so the exposure control parameters were derived to be 1.0 in order to be administered once selected.

As for the unused items, the minimums of the exposure control parameters were all 1.0. They were rarely appearing items. Once the items were selected, they were always allowed to be administered to examinees.

## **Summary and Conclusions**

### **Summary of Results**

The current study explored whether removing a single item or the unused items had effects on the exposure control parameters of the remaining items; that is, whether redeveloping the exposure control parameters for the remaining items was necessary when some items were eliminated from the pool. For the purpose of evaluating the removal effect of one item, this study considered three items of different information values at three selected ability levels. Items with the highest, moderate and lowest information values were chosen for removal at each of the theta points of -2.8, 0.0 and 2.8, respectively. To examine the results of removing the unused items, this study employed a reduced pool by excluding the never administered items from the operational CAT administrations when the full pool was employed. The exposure control parameters were redeveloped for all items in the modified pools. Both existing and redeveloped exposure control parameters for the various pools were utilized in the operational testing situations and their performance was compared. Also, the characteristics of the

unused items were evaluated in detail. The results of the current study are summarized below.

### The Relationships of the Existing and Redeveloped Exposure Control Parameters

Under the conditions of removing both single and unused items, there were fewer items with the parameters remaining unchanged at the upper middle part of ability scale than at both extremes. Exposure control parameters at the upper middle part of the ability scale seemed to be affected by the item removal to a greater extent than the other parts of the scale. These outcomes reflected the nature of the item pools, in that there were more items appropriate for administration for the upper middle part than for the extremes of the ability scale. Also, the findings demonstrated no differential effects on the results by removing items with the varying degrees of information values for the different ability points.

An inspection of the average differences of the existing and redeveloped exposure control parameters indicated that under the single item removal conditions, the redeveloped exposure control parameters could be larger or smaller than the existing ones. However, for the middle and upper parts of the ability scale, the redeveloped parameters tended to be greater than the existing parameters, so the probability of administering items given selection for these parts was somewhat increased. For the unused item removal condition, the redeveloped parameters were on average greater than the existing ones and the results were stronger for the middle and upper parts of the ability scale. With fewer items in the pool, the SLC algorithm in general increased the values of the exposure control parameters to allow for more administrations given selection. But since more items were appropriate for the middle and upper parts of the ability scale than for the two ends, the SLC algorithm tended to loosen the exposure control for items appropriate for these parts to a greater extent.

The average absolute parameter differences for the unused items deletion were in general greater than those for the single item deletion. Removing a larger number of items from the pool led to a greater change on the exposure control parameters. Among the various conditions of the single item deletion, the findings did not suggest noticeable differences due to the removal of items with varying degrees of information values for the various ability levels, except for  $H(0.0)$ . These phenomena might be explained by the reason that an item possesses different information values, depending on the ability levels. Although the removed item was of low information at one ability level, this item might be of high information at some other ability levels.

The results regarding the Pearson product-moment correlations showed the redeveloped exposure control parameters obtained from the various removal conditions were all strongly correlated with the original parameters for all ability points.

### Results of Employing the Existing and Redeveloped Exposure Control Parameters in Operational



### Testing

For either the single item or the unused items deletion, the observed maximum exposure rates produced by using the existing and redeveloped exposure control parameters were as low as the desired rate of .20. Also, each of the modified pools was utilized to a similar degree by applying the two sets of the exposure control parameters. Both applications left similar numbers of items in the pool that were never administered in the operational testing situations.

However, as compared based on the conditional maximum observed exposure rates, the results of utilizing the existing exposure control parameters were different from those of the redeveloped ones. The maximum exposure rates conditionally observed at the lower part of the ability scale were especially higher than those of the full pool or the modified pools with the redeveloped exposure control parameters. The findings suggested that the existing exposure control parameters were no longer appropriate and the redevelopment of the parameters ought to be necessary when the pool was modified. In terms of the measurement precision, employing the existing or the redeveloped exposure control parameters did not yield differences in the CSEMs.

As to the effect of removing all the unused items, the findings indicated that as long as the redeveloped exposure control parameters were applied, the performance of the reduced pool was competitive to that of the full pool. Both overall and conditional maximum observed exposure rates for the reduced pool were as low as those for the full pool. Eliminating the unused items did not seem to cause effects on the CSEMs either.

### Examination of the Characteristics of the Unused Items

The characteristics of the unused items were compared with those of the high, moderate and low exposure groups, respectively. The results showed that the unused items had the lowest average discriminating value, were easier than the average difficulty level, and were those associated with the highest guessing value on average.

The item information values for the unused items were low in general, although their average information values for some ability levels were the highest among the groups. The outcomes also revealed that the maximum information values for the various ability levels did not always appear in the high exposure group. For the middle part of the ability scale, the maximum information values were seen in the high exposure group. But, for the slightly lower and upper parts, the maximum values were observed in the moderate exposure group and for the very high and low ends, the maximums were in the low exposure group.

All exposure control parameters of the unused items were 1.0. Obviously, these items were rarely appearing, so they were always allowed to be administered once selected. However, the results

were that these items were never selected in the operational testing, so they were never administered. For the high exposure group, the average exposure control parameters were lower in the middle than at both ends of the ability scale, saying that this group contained items more appropriate for the middle than for the two ends. For the moderate exposure group, the exposure control parameters at the upper end were somewhat lower than in the high exposure group. Similar to the observations with the item information values stated above, the items in the high exposure group were mostly appropriate for the middle part but were in general not as popular for the upper end as those in the moderate exposure group. Although the low exposure group contained some items especially appropriate for the very high ability levels, items at most parts of the ability scale were not likely selected since they were associated with relatively high exposure control parameters.

## Conclusions

This study explored how the deletion of a single item and the unused items might alter the exposure control parameters of the remaining items derived by the SLC algorithm. Although removing a single item or the unused items did not seem to have a substantial impact on the values of the exposure control parameters, the results of applying the existing exposure control parameters for the modified pools caused the conditional maximum observed exposure rates to be higher at the lower end of the ability scale. Thus, the original exposure control parameters were no longer appropriate for the modified pools. These findings provided evidence to support the view in Stocking and Lewis (2000) in that, in their opinion, the iterative simulations should be carried out again to guarantee continued exposure control for any modified item pools.

Employing the reduced pool with the redeveloped exposure control parameters yielded very similar results to the full pool in both overall and conditional maximum exposure rates, and in the CSEMs. The removal of all unused items seemed to have no effect on the performance of the SLC procedure in the operational CAT administrations. It might be that the items removed were those that would never be administered in the operational testing situations. These items were not influential enough in altering the results of the CAT procedure. Employing the full pool and the reduced pool with the redeveloped exposure parameters might provide similar enough conditions for the SLC algorithm to perform equally well. However, it should be noticed that this study took into account only the content matter of items during the item selection process. Contexts having nonstatistical constraints such as the overlap or the item set constraints were not considered in the current study. The effect of the unused items removal remains unknown under complex but more realistic test structures.

The characteristics of the 48 unused items were inspected in detail. As expected, the information values of these unused items were low overall, but their average information values for the very low

ability levels were indeed the highest. However, for the ability distribution of  $N(0,1)$  employed in the current study, not many examinees were around this very low end, so items more appropriate for this part of the ability scale would not have been selected often. These unused items were not as competitive as those in the other exposure groups based on the maximum information selection rule. As a result of the item interactions, the exposure control parameters of these unused items were all derived to be 1.0, in spite that some of the unused items were somewhat popular for examinees at the very low end. It is conceivable that when there are more examinees with ability levels at this low end and items appropriate for this part of the ability scale are in greater demand, the unused items would have a better chance of being selected. This implies that the number of the unused items is a consequence of the ability distribution of the examinees' population, in addition to the characteristics of the items themselves. Considerations of the examinee population should not be ignored in the creation of a well-utilized item pool, as also emphasized in Guo et al. (2000) for constructing a CAT item pool. Further studies might be conducted to investigate the effects of the examinees' ability distributions on the pool utilization.

It is also important to attempt to determine the pool size for optimizing pool utilization. The results of the SLC algorithm were almost identical based on the full pool of 480 items and the reduced pool of 432 items. When the pool of 432 items was utilized in the operational testing, there were still five items never administered. Could the pool size be further reduced for optimal pool usage? This issue is, of course, very complicated since there exist many factors to consider, such as the desired maximums of the exposure rates, the characteristics of the items, the structures of the item pool, and the ability distribution of the real examinees' populations.

Both pool management and exposure control method implementation have been the focus of easing test security concerns in CAT. Maintaining pools of high quality is an ongoing process, and sometimes only a small modification may be needed. This study investigated how the deletion of a single item and the unused items might alter the exposure control parameters of the remaining items derived by the SLC algorithm. Results from this study have offered a better understanding of the effects of item interactions on the values of the exposure control parameters and also, on the outcomes of the observed exposure rates and measurement precision in operational testing. Along with a detailed examination of the characteristics of the unused items, this study has provided psychometric researchers and test practitioners valuable insights for a better pool construction design.

## References

- Birnbaum, A. (1968). Some latent trait models and their uses in inferring an examinee's ability. In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores* (pp. 395-479). Reading, MA: Addison-Wesley.
- Chang, S. W. (1998). *A comparative study of item exposure control methods in computerized adaptive testing*. Unpublished doctoral dissertation, The University of Iowa. Iowa City, IA.
- Chang, S. W., Ansley, T. N., Lin, S. H. (2000, April). *Performance of item exposure control methods in computerized adaptive testing: Further explorations*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- 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.
- Guo, F., Way, W. D., & Reshetar, R. (2000, April). *Test security and the development of computerized tests*. Paper presented at the NCME invited symposium: Maintaining test security in computerized programs--Implications for practice, New Orleans.
- Hetter, R. D., & Sympon, J. B. (1997). Item exposure control in CAT-ASVAB. In W. A. Sands, B. K. Waters, & J. R. McBride (Eds.), *Computerized adaptive testing: From inquiry to operation* (pp. 141-144). Washington, DC: American Psychological Association.
- Kingsbury, G. G., & Zara, A. R. (1989). Procedures for selecting items for computerized adaptive tests. *Applied Measurement in Education*, 2(4), 359-375.
- Mills, C. N., & Steffen, M. (2000). The GRE computer adaptive test: Operational issues. In W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: Theory and practice* (pp. 75-99). The Netherlands: Kluwer Academic Publishers.
- Mislevy, R. J., & Bock, R. D. (1990). Item analysis and test scoring with binary logistic models. *BILOG 3*. Chicago, IL: Scientific Software, Inc.
- 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(350), 351-356.
- Stocking, M. L., & Lewis, C. (1998). Controlling item exposure conditional on ability in computerized adaptive testing. *Journal of Educational and Behavioral Statistics*, 23(1), 57-75.
- Stocking, M. L., & Lewis, C. (2000). Methods of controlling the exposure of items in CAT. In W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: Theory and practice* (pp. 163-182). The Netherlands: Kluwer Academic Publishers.
- Stocking, M. L., & Swanson, L. (1993). A method for severely constrained item selection in adaptive testing. *Applied Psychological Measurement*, 17(3), 277-292.
- Swanson, L., & Stocking, M. L. (1993). A model and heuristic for solving very large item selection problems. *Applied Psychological Measurement*, 17(2), 151-166.
- Sympon, J. B., & Hetter, R. D. (1985, October). *Controlling item exposure rates in computerized adaptive testing*. Paper presented at the annual meeting of the Military Testing Association. San Diego, CA: Navy Personnel Research and Development Center.

- van der Linden, W. J. (2000). Constrained adaptive testing with shadow tests. In W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: Theory and practice* (pp. 27-52). The Netherlands: Kluwer Academic Publishers.
- Veldkamp, B. P., & van der Linden, W. J. (2000). Designing item pools for computerized adaptive testing. In W. J. van der Linden & C. A. W. Glas (Eds.), *Computerized adaptive testing: Theory and practice* (pp. 149-162). The Netherlands: Kluwer Academic Publishers.
- Way, W. D., Swanson, L., Steffen, M., & Stocking, M. (2001, April). *Refining a system for computerized adaptive testing pool creation*. Paper presented at the annual meeting of the American Educational Research Association, Seattle.

**Table 1 Results of Differences of the Exposure Control Parameters by Ability Level and Removal Condition**

Items Removed	Parameters	Theta																
		-3.2	-2.8	-2.4	-2	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2	2.4	2.8	3.2
H(-2.8) N=479	N_Unchanged	352	349	340	341	344	336	331	322	325	329	320	331	327	339	340	360	348
	MDIFF	0.0010	0.0002	0.0002	0.0005	0.0001	0.0005	-0.0011	-0.0029	-0.0018	-0.0024	-0.0016	-0.0016	-0.0011	-0.0014	-0.0010	-0.0010	-0.0020
	MDIFFABS	0.0039	0.0034	0.0047	0.0039	0.0027	0.0042	0.0037	0.0043	0.0048	0.0039	0.0072	0.0030	0.0042	0.0025	0.0040	0.0030	0.0040
H(0.0) N=479	N_Unchanged	352	348	335	335	337	337	326	323	325	323	324	331	324	340	340	346	350
	MDIFF	0.0023	0.0025	0.0011	0.0017	0.0005	0.0013	0.0005	-0.0020	-0.0010	-0.0006	-0.0016	-0.0003	0.0006	0.0010	0.0002	0.0008	-0.0006
	MDIFFABS	0.0064	0.0071	0.0078	0.0066	0.0060	0.0058	0.0053	0.0048	0.0063	0.0051	0.0068	0.0060	0.0055	0.0050	0.0059	0.0053	0.0075
H(2.8) N=479	N_Unchanged	351	351	337	336	341	340	329	324	325	328	320	330	330	339	342	351	345
	MDIFF	-0.0005	0.0004	0.0005	0.0000	-0.0010	0.0008	-0.0010	-0.0027	-0.0017	-0.0008	-0.0021	-0.0012	-0.0008	0.0000	0.0000	0.0000	0.0010
	MDIFFABS	0.0032	0.0036	0.0046	0.0050	0.0030	0.0028	0.0036	0.0046	0.0051	0.0029	0.0070	0.0029	0.0039	0.0030	0.0050	0.0040	0.0050
M(-2.8) N=479	N_Unchanged	353	347	337	334	336	338	328	321	320	329	319	328	327	339	342	352	350
	MDIFF	0.0020	0.0030	0.0019	0.0009	0.0022	0.0010	0.0002	-0.0017	-0.0008	0.0000	-0.0006	-0.0013	-0.0013	-0.0017	-0.0010	-0.0008	-0.0013
	MDIFFABS	0.0040	0.0040	0.0057	0.0059	0.0049	0.0035	0.0051	0.0061	0.0054	0.0042	0.0076	0.0036	0.0049	0.0025	0.0028	0.0026	0.0037
M(0.0) N=479	N_Unchanged	351	348	337	339	339	337	330	321	325	327	321	328	331	341	340	346	350
	MDIFF	0.0002	0.0020	0.0014	0.0004	0.0009	0.0016	-0.0011	-0.0037	-0.0019	-0.0026	-0.0012	-0.0010	-0.0011	-0.0016	-0.0010	-0.0012	-0.0022
	MDIFFABS	0.0026	0.0030	0.0046	0.0042	0.0029	0.0043	0.0045	0.0047	0.0040	0.0038	0.0066	0.0020	0.0027	0.0031	0.0030	0.0024	0.0036
M(2.8) N=479	N_Unchanged	354	350	341	337	344	339	330	323	325	330	319	330	325	342	339	350	352
	MDIFF	-0.0006	0.0001	-0.0009	-0.0010	-0.0010	-0.0003	-0.0014	-0.0027	-0.0021	-0.0022	-0.0022	-0.0008	-0.0026	-0.0023	-0.0020	-0.0020	-0.0018
	MDIFFABS	0.0032	0.0034	0.0044	0.0040	0.0030	0.0029	0.0037	0.0043	0.0040	0.0038	0.0053	0.0030	0.0045	0.0029	0.0050	0.0030	0.0037
L(-2.8) N=479	N_Unchanged	352	349	340	335	338	341	330	323	324	327	324	332	328	343	337	351	352
	MDIFF	0.0007	-0.0005	-0.0002	0.0000	-0.0009	-0.0001	-0.0027	-0.0034	-0.0020	-0.0012	-0.0008	-0.0004	0.0008	-0.0001	-0.0008	0.0001	-0.0003
	MDIFFABS	0.0033	0.0024	0.0039	0.0040	0.0021	0.0033	0.0037	0.0046	0.0050	0.0039	0.0074	0.0039	0.0056	0.0034	0.0054	0.0035	0.0044
L(0.0) N=479	N_Unchanged	354	352	338	334	340	339	331	321	322	331	323	333	328	341	339	346	350
	MDIFF	-0.0005	0.0001	-0.0002	-0.0017	-0.0008	-0.0003	-0.0009	-0.0039	-0.0018	-0.0023	-0.0025	-0.0019	-0.0005	0.0000	0.0000	0.0010	0.0000
	MDIFFABS	0.0029	0.0031	0.0041	0.0057	0.0036	0.0026	0.0035	0.0051	0.0052	0.0036	0.0063	0.0030	0.0040	0.0030	0.0050	0.0040	0.0050
L(2.8) N=479	N_Unchanged	351	348	339	337	339	337	329	323	323	329	319	328	328	341	342	347	351
	MDIFF	0.0007	0.0005	0.0008	0.0012	0.0008	0.0010	-0.0005	-0.0008	-0.0023	-0.0020	-0.0027	-0.0019	-0.0022	-0.0017	-0.0020	-0.0003	-0.0013
	MDIFFABS	0.0035	0.0033	0.0041	0.0041	0.0036	0.0038	0.0054	0.0031	0.0038	0.0052	0.0060	0.0026	0.0035	0.0028	0.0030	0.0023	0.0036
UNUSED N=432	N_Unchanged	304	302	291	291	296	290	282	276	275	282	275	283	282	295	292	303	299
	MDIFF	-0.0011	-0.0011	-0.0027	-0.0029	-0.0026	-0.0018	-0.0034	-0.0055	-0.0034	-0.0034	-0.0049	-0.0033	-0.0034	-0.0037	-0.0052	-0.0030	-0.0041
	MDIFFABS	0.0035	0.0038	0.0057	0.0065	0.0050	0.0047	0.0049	0.0066	0.0057	0.0044	0.0081	0.0042	0.0058	0.0040	0.0057	0.0040	0.0060

Note. N\_Unchanged indicates the number of items with the same existing and redeveloped exposure control parameters.

MDIFF represents the average value of the differences between the existing and redeveloped exposure control parameters, with the redeveloped parameters being subtracted from the existing parameters.

MDIFFABS represents the average value of the absolute differences between the existing and redeveloped exposure control parameters.

**Table 2 Correlations of the Exposure Control Parameters by Ability Level and Removal Condition**

Items Removed	Theta																
	-3.2	-2.8	-2.4	-2	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2	2.4	2.8	3.2
H(-2.8)	0.99859	0.99904	0.99847	0.99875	0.99906	0.99851	0.99882	0.99868	0.99817	0.99915	0.99738	0.99933	0.99868	0.99953	0.99758	0.99859	0.99821
H(0.0)	0.99674	0.99675	0.99623	0.99673	0.99739	0.99792	0.99798	0.99874	0.99705	0.99871	0.99721	0.99784	0.99772	0.99827	0.99762	0.99797	0.99593
H(2.8)	0.99892	0.99886	0.99841	0.99758	0.99813	0.99945	0.99888	0.99881	0.99792	0.99959	0.99763	0.99946	0.99883	0.99849	0.99718	0.99759	0.9972
M(-2.8)	0.99774	0.99712	0.99754	0.99687	0.99726	0.99887	0.99788	0.99794	0.99836	0.99887	0.99719	0.99932	0.9982	0.99954	0.99934	0.99945	0.99884
M(0.0)	0.99941	0.99838	0.99754	0.99836	0.99946	0.99862	0.99818	0.99876	0.999	0.99922	0.99792	0.99954	0.99963	0.99921	0.9988	0.99958	0.99903
M(2.8)	0.99917	0.99896	0.99831	0.99771	0.99869	0.99938	0.99889	0.99865	0.99898	0.99923	0.99822	0.99943	0.99877	0.99927	0.99708	0.99893	0.99911
L(-2.8)	0.99933	0.99958	0.99901	0.99804	0.99976	0.99932	0.99905	0.99881	0.99703	0.99917	0.99677	0.99916	0.99756	0.999	0.99692	0.99865	0.99814
L(0.0)	0.99905	0.99909	0.99884	0.99726	0.99818	0.99956	0.99889	0.9986	0.99795	0.99935	0.99778	0.99944	0.99881	0.99857	0.99653	0.99731	0.99746
L(2.8)	0.99922	0.99934	0.99867	0.99896	0.99908	0.9992	0.99796	0.99946	0.99912	0.9986	0.99827	0.99945	0.99946	0.99942	0.99871	0.99967	0.99914
UNUSED	0.99928	0.99928	0.99783	0.99634	0.99784	0.9988	0.99873	0.99809	0.99827	0.99917	0.99677	0.99888	0.99795	0.9989	0.99738	0.99856	0.99749

Note. N\_Unchanged indicates the number of items with the same existing and redeveloped exposure control parameters.

MDIFF represents the average value of the differences between the existing and redeveloped exposure control parameters, with the redeveloped parameters being subtracted from the existing parameters.

MDIFFABS represents the average value of the absolute differences between the existing and redeveloped exposure control parameters.

All the correlations were significant at .0001 level ( $p < .0001$ ).



**Table 3 Descriptive Statistics of the Observed Exposure Rates for the Various Conditions**

Pools	N	M	SD	Skew	Kurt	Min	Max
H(0.0) Removed						0.000	
Existing	430	0.06977	0.05139	0.35987	-1.00655	02	0.19950
Redeveloped	429	0.06993	0.05102	0.35122	-0.99623	02	0.19822
UNUSED Removed						0.000	
Existing	427	0.07026	0.05158	0.38329	-0.96678	02	0.20048
Redeveloped	427	0.07026	0.05152	0.36220	-0.98794	02	0.19866
The Full Pool	432	0.06944	0.05094	0.36144	-0.97549	02	0.19818
High Exposure	144	0.13031	0.02373	0.60987	-0.29616	06	0.19818
Moderate Exposure	144	0.06380	0.01735	0.08472	-1.12078	64	0.09502
Low Exposure	144	0.01423	0.01113	0.22776	-1.21673	02	0.03652

**Table 4 Descriptive Statistics of the Item Parameters  
for the Various Exposure Groups**

Exposure Groups	Item Parameters	M	SD	Skew	Kurt	Min	Max
The Full Pool N=480	a	1.00967	0.33146	0.80064	1.95247	0.19626	2.62976
	b	0.14984	1.07669	-0.25518	-0.32757	-3.44405	2.55902
	c	0.17638	0.08034	0.98093	1.52969	0.03408	0.50000
High Exposure N=144	a	1.18043	0.23482	0.48491	-0.14527	0.76140	1.86887
	b	-0.00962	0.46101	-0.33003	-0.35696	-1.08687	0.91304
	c	0.14555	0.07768	1.58430	4.38879	0.04022	0.49652
Moderate Exposure N=144	a	1.07575	0.33524	1.81559	5.45658	0.57622	2.62976
	b	0.21649	1.04485	-0.48977	-1.15781	-2.01056	1.70187
	c	0.17425	0.07127	0.53527	0.22022	0.03408	0.39780
Low Exposure N=144	a	0.87409	0.31919	1.11162	1.47129	0.40469	2.12031
	b	0.41493	1.41871	-0.51678	-0.84154	-3.44405	2.55902
	c	0.17934	0.06248	0.91955	1.66391	0.06020	0.42246
UNUSED N=48	a	0.70586	0.21605	-0.16699	-0.23427	0.19626	1.12616
	b	-0.36700	1.06907	-0.33251	-0.82387	-2.87023	1.31659
	c	0.26632	0.09419	0.40291	-0.20642	0.07508	0.50000

**Table 5 Descriptive Statistics of the Item Information values by Ability Level and Exposure Group**

Exposure Groups		Theta																
		-3.2	-2.8	-2.4	-2	-1.6	-1.2	-0.8	-0.4	0	0.4	0.8	1.2	1.6	2	2.4	2.8	3.2
High Exposure N=144	Mean	0.00090	0.00300	0.00960	0.02827	0.07302	0.16281	0.30732	0.48038	0.61274	0.62766	0.50244	0.31196	0.16283	0.07922	0.03799	0.01835	0.00898
	SD	0.00170	0.00580	0.01810	0.04840	0.10396	0.17823	0.24546	0.26978	0.27589	0.31734	0.35087	0.24817	0.12478	0.05947	0.03063	0.01673	0.00932
	Min	0	0	0	0	0.00002	0.00022	0.00265	0.02916	0.19418	0.16601	0.07989	0.03678	0.01530	0.00530	0.00183	0.00063	0.00022
	Max	0.01107	0.03487	0.10517	0.28545	0.57502	0.81845	1.38643	1.35625	1.73163	1.68827	1.96236	1.42610	0.61601	0.30088	0.16485	0.09864	0.05662
Moderate Exposure N=144	Mean	0.00790	0.01845	0.03888	0.06954	0.10166	0.12565	0.14371	0.16893	0.21840	0.30026	0.38927	0.44628	0.40170	0.24561	0.12547	0.06197	0.03075
	SD	0.01620	0.03509	0.07241	0.12688	0.16612	0.16967	0.15344	0.13530	0.14019	0.20300	0.30313	0.42344	0.52410	0.29551	0.13374	0.06639	0.03481
	Min	0	0	0	0	0	0	0	0.00001	0.00021	0.00455	0.01414	0.00584	0.00241	0.00099	0.00041	0.00017	0.00007
	Max	0.09751	0.19780	0.34945	0.70791	0.87109	0.67945	0.54413	0.47524	0.60704	0.87515	1.67293	2.65143	3.11494	2.01256	0.61300	0.30247	0.16700
Low Exposure N=144	Mean	0.01760	0.02670	0.03771	0.04967	0.06169	0.07367	0.08670	0.10243	0.12209	0.14672	0.17995	0.22875	0.28967	0.30889	0.24407	0.15377	0.08610
	SD	0.03770	0.05180	0.06502	0.07616	0.08374	0.08732	0.09065	0.10032	0.11648	0.12954	0.13779	0.17147	0.28303	0.38108	0.30772	0.19274	0.10402
	Min	0	0	0	0	0	0	0	0	0.00003	0.00030	0.00313	0.00566	0.00315	0.00175	0.00097	0.00054	0.00030
	Max	0.28183	0.36397	0.37741	0.32186	0.30327	0.28475	0.30786	0.38866	0.51041	0.56503	0.59481	0.73155	1.31923	2.39977	1.64078	1.25514	0.54550
UNUSED N=48	Mean	0.01772	0.02675	0.03895	0.05402	0.07088	0.08817	0.10569	0.12569	0.15010	0.17141	0.17284	0.14890	0.11224	0.07743	0.05090	0.03278	0.02105
	SD	0.02483	0.03424	0.04635	0.05975	0.07153	0.07823	0.07748	0.07106	0.07790	0.11056	0.13058	0.11829	0.08878	0.05993	0.03916	0.02615	0.01818
	Min	0	0.00002	0.00007	0.00026	0.00103	0.00386	0.01371	0.01645	0.01595	0.01532	0.01459	0.01378	0.01291	0.01110	0.00805	0.00471	0.00275
	Max	0.09683	0.10371	0.13882	0.18575	0.22393	0.25249	0.24907	0.25260	0.32820	0.41478	0.45997	0.43191	0.36110	0.24372	0.14350	0.09774	0.07873

Table 6 Descriptive Statistics of the Exposure Control Parameters by Ability Level and Exposure Group

[illegible]

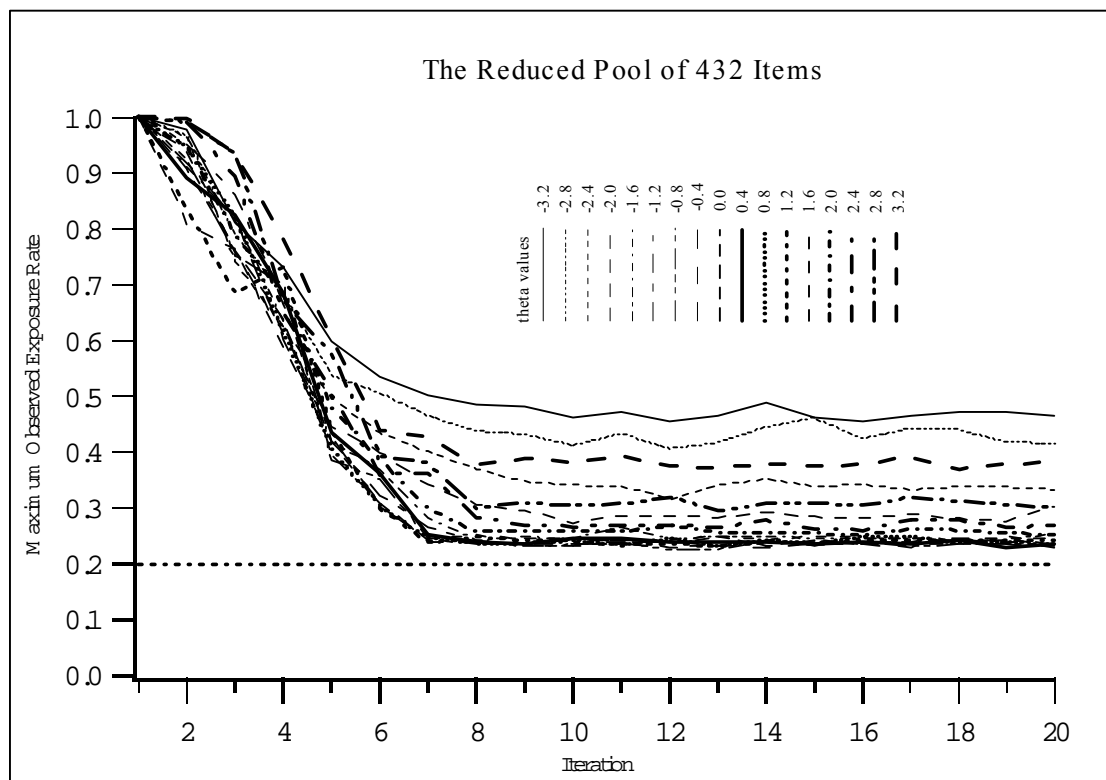
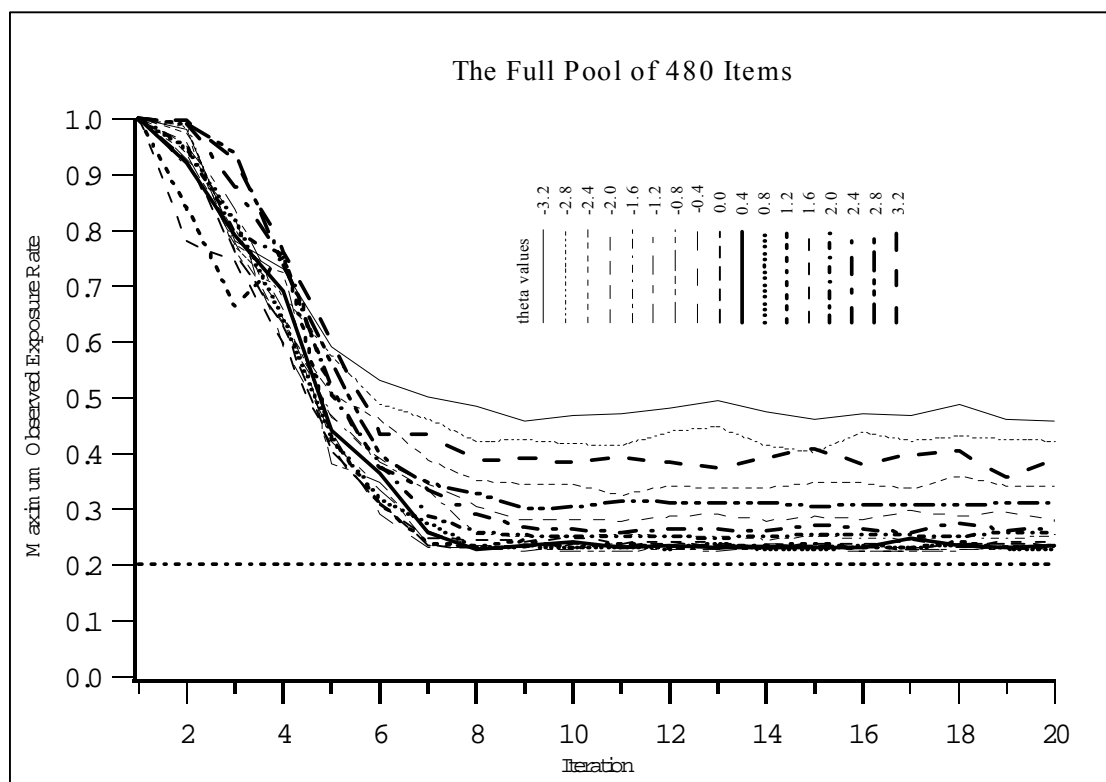


Figure 1. Iterations of the Exposure Control Parameter Development

Figure 2. The Frequency Distribution of Item Difficulty Values for the Full Pool

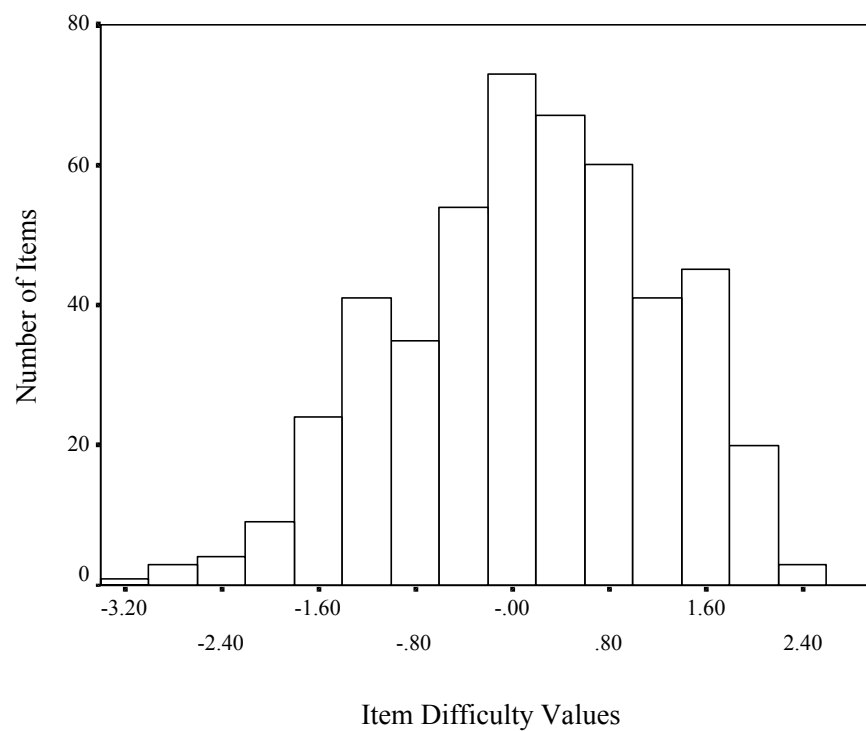


Figure 3. Conditional Maximum Observed Exposure Rates  
for the Various Conditions

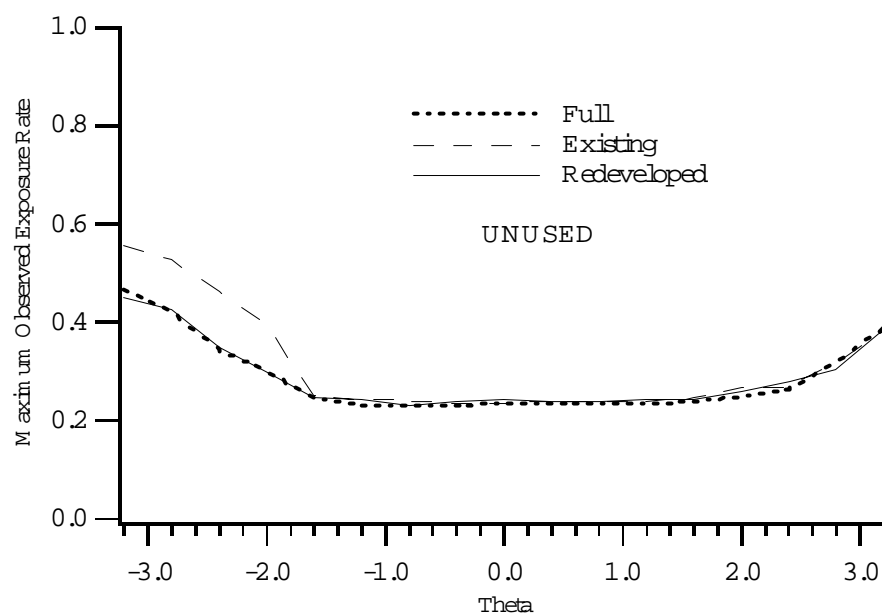
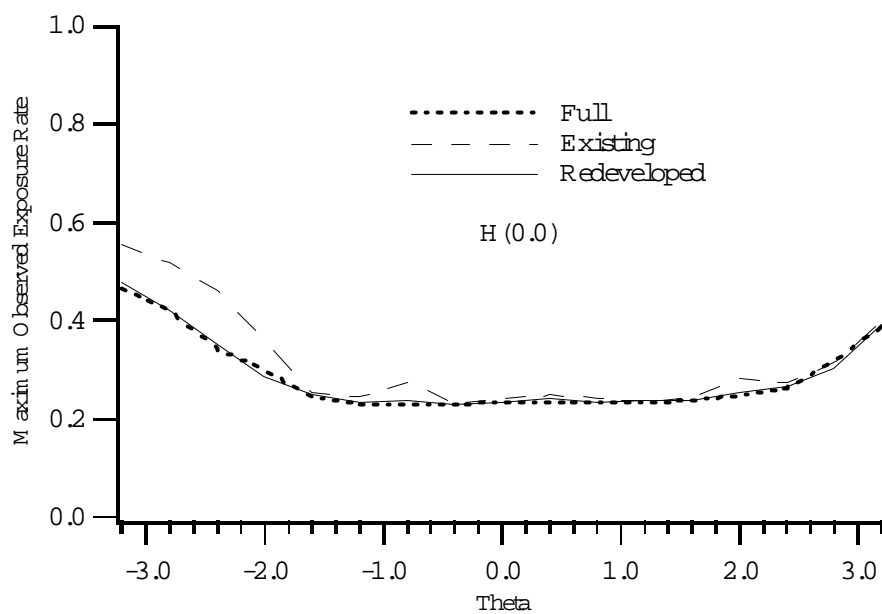


Figure 4. Conditional Standard Errors for the Various Conditions

