312 787.6555

SIMULATING ADAPTIVE ADMINISTRATION
OF A NURSING LICENSURE EXAMINATION

Jennifer Bosma, Ph.D. and Eileen M. Dvorak, Ph.D., R.N. National Council of State Boards of Nursing Chicago, Illinois

Paper presented at the annual meeting of the American Educational Research Association Washington, D.C.

April 1987

Background

State boards of nursing are responsible for assessing the competence of graduates of nursing programs to practice registered or practical nursing. All state boards of nursing in the fifty states, the District of Columbia, the Virgin Islands, Guam, American Samoa, and Saipan require candidates for initial licensure as nurses to take the appropriate National Council of State Boards of Nursing licensure examination (NCLEX-RN or NCLEX-PN). The National Council is responsible to the state boards, which are its members, for the development of examinations which are psychometrically sound and legally defensible.

In order for the National Council to continue to meet the objective of psychometric soundness in an era of major changes in measurement theory and testing technology, staff at the National Council are investigating the applications of Item Response Theory and computerized administration to licensure examinations. One aspect of this investigation has been the use of existing data in simulation of an adaptive administration in order to assess (1) the number of Em 49 items that would be required for administration to each candidate before a pass/fail decision can be reached. (2) the match between that decision and the actual decision made after the full paper-and-pencil test, and (3) the degree to which the distribution of items in the simulated administration, up to the point of decision, conformed to the specified percentages for the test plan.

Theory and Research

The concept of tailoring tests, i.e. selecting items adaptively, to fit the examinee's ability level dates back to work in the early 1970s by Lord (Lord, 1970, 1971), and beyond that to work by Wald with sequential tests of statistical hypotheses (Wald, 1945) and Alfred Binet on the initial development of intelligence tests. However, these early paper-and-pencil adaptive tests were unreliable and inconvenient to administer because of the complicated directions examinees had to follow. With the general availability of microcomputer technology in the 1980s, these impediments to adaptive testing were removed: a computer program can quickly and unobtrusively perform the routing of the examinee through the tailored set of items comprising his or her test.

The development of Item Response Theory (see, for example, Lord, 1952; Rasch, 1960; Birnbaum, 1968) has allowed for great improvement in the process of adaptive item selection for use in tailored tests. Since the difficulties of items and the abilities of examinees may be calibrated on the same scale by application of IRT, precise formulae for the identification of the optimal item for selection as the "next item" have been developed (McBride, 1985). The scoring of adaptive tests using IRT models is also much more straightforward than scoring based on classical measurement models. As long as the item pool meets the requirement of unidimensionality and sufficient

care has been used in the calibration of the items, ability estimates obtained via adaptive testing are directly comparable from one examinee to the next (Reckase, 1981).

Recent literature on computerized adaptive testing (CAT) supports the improved accuracy and efficiency of the CAT measurement process (see, for example, Fortune, 1985; McBride, 1985; Weiss, 1982) for criterion-referenced testing. The expectation of greater accuracy is based on the way that testing terminates and a pass/fail decision is given only when the established confidence interval around each individual candidate's ability estimate no longer includes the passing criterion. Increased efficiency is expected to result because of the way that appropriate items are selected so that each response provides maximum information about a candidate's ability measure; in addition, test length may be further reduced for those candidates whose abilities are farther from the passing criterion, since their ability estimate range will exclude the pass score fairly early in the course of test administration.

Method

A computer program using the Rasch Model (see Rasch, 1960) for item calibration and ability estimation was developed by consultants at the University of Chicago to simulate the adaptive administration of NCLEX-RN Series 785.

A standard item of mid-range difficulty was selected as the first item in the simulated administration. After retrieving and scoring the candidate's actual response to that item, the program calculated a provisional ability estimate and selected as the next item that one of the remaining items whose difficulty came closest to the ability estimate. This process was repeated until the entire set of items on the original paper-and-pencil test was exhausted.

The data base for the simulation comprised individual item-level responses for 1762 candidates who sat for the July 1985 registered nurse licensure examination in New York. This was a subset of New York candidates consisting of those candidates who had baccalaureate-level educational preparation.

The 300 test items on the July 1985 examination were all in multiple-choice format with four response options. All items had been previously field tested; and the assembled test met National Council guidelines for allocation to test plan categories, difficulty and discrimination statistics, and reading level. The items in the paper-and-pencil test were presented in a case format with four to nine items relating to each client situation. Although the items are carefully screened so that no item cues the answer to another, the case format may contribute to some lack of "local independence" among the items. No data are presently available to confirm or refute this

Juture

possibility. The case format was not accounted for in the adaptive simulation, i.e. items were selected individually, strictly on the basis of their Rasch difficulty.

Other limitations of the simulation procedure included the necessity of assuming that candidates would have responded similarly, even under the differing conditions of the CAT administration. One potential cause for different responses is the action of sitting at a computer terminal to take an examination. Research cited by Hofer and Green (1985) generally supports the equivalence of the testing modes when tests use multiple-choice items and do not deal with sensitive or personally-threatening issues. Research by Kiely, Zara, and Weiss (1986) has supported equivalence of computer-administered and paper-and-pencil administration for multiple-choice items not containing extremely long text passages.

A second possible cause of different examinee responses is the alteration in the order of presentation of items, and the lack of opportunity to review and change answers. No data are currently available to evaluate the effects of these factors, thus the conclusions below should be received tentatively pending further research with live subjects.

Results

By examining the computer output charting the course of the simulated adaptive administration (see Appendix), it was possible to determine at what point a pass/fail decision with a specified level of precision (±3 S.E.) would have been made. It was also possible to compare that simulated decision with the actual decision reported to the candidate and board of nursing for the full-length paper and pencil examination. A third assessment was also made comparing the distribution of the test plan categories of the items administered up to the decision point with the prescribed percentage for each category of the NCLEX-RN Test Plan.

The following brief explanation of the format of the computer printouts contained in the Appendix will serve to clarify the illustrations which follow:

-top line identifies data source, individual candidate, the raw score, the Rasch person measure (or ability estimate), the standard error of the measure, and two fit statistics

-the first column on the left gives each step in the simulated adaptive administration (i.e. each item "administered") a sequence number

-the column labeled "ITEM" identifies the original location of the item in the paper-and-pencil test -the five columns under "NBTAB" identify the coding of items with respect to the categories of the RN test plan addressing nursing behaviors (assessing, analyzing, planning, implementing, and evaluating);

figures in these columns give cumulative totals of codings of items administered to that point

-the three columns under "LDTAB" identify the coding of items with respect to the categories of the RN test plan addressing locus of decision making (nurse, client, and shared); figures are cumulative

-the column labeled "DIFF" gives the difficulty calibration of the item being "administered" at that step

- -the column labeled "RESP" gives the response of the individual to the item (0=wrong, 1=right)
- -the column labeled "PROB" gives the probability of the response, given the provisional person measure based on responses to all items administered thus far (including the present one)
- -the column labeled "MEAS" gives the provisional measure of the individual's ability, given responses to all items administered thus far (including the present one)
- -the column labeled "ERROR" gives the standard error for the measure in the preceding column
- -the diagram at the right portrays graphically the person measure (as \underline{B}) plus-and-minus the three standard error band (as dots to the right and left of \underline{B}); the vertical line down the center represents the criterion set for passing

The first page of the Appendix is part of the response record for the simulated adaptive administration of the examination to a candidate who ultimately failed the paper-and-pencil examination. With respect to item selection, it is interesting to note that for about the first 25 items, it was possible to very closely match the "next item" to the individual's current ability estimate even within the confines of this 300 item test. The item difficulty then becomes progressively divergent from the provisional estimate, indicating that appropriate items were no longer available for administration. With access to a larger pool of items, the appropriate item-difficulty/ examinee ability estimate match could be sustained considerably longer, thus maintaining the efficiency of the adaptive item selection process.

With respect to the ability estimation process for this candidate, the initial instability of the estimate (for about the first 20 items) can be seen readily in the diagram to the right. Thereafter, the estimate remains quite stable and the standard error band gradually decreases in width until approximately the 65th step of the administration, where it can be seen that the ability estimate surrounded by its ±3 standard error band completely excludes the passing criterion. These results indicate that a reliable decision for this candidate could have been reached after 65 items, rather than 300.

The second page of the Appendix presents the response record of a candidate who passed the paper-and-pencil examination. For this candidate, the match between ability estimate and "next item" difficulty is not as good, most likely because the ability estimate reached a high range rather quickly and the items on the test tend to be of moderate to low difficulty. However, measurement of this candidate's ability is also efficient and accurate in that a passing decision can be made at about the 63-item mark.

The responses of a third candidate are illustrated on the third and fourth pages of the Appendix. Because this candidate's ability level is considerably closer to the passing criterion than either of the first two candidates, it takes about 132 items before the ability estimate with its standard error band excludes the passing point, and a decision to fail is reached.

The aggregate results for number of items required to make a decision are presented in Table 1. Overall, nearly half the decisions could be made in fewer than 100 items. Interestingly, the next largest group of candidates should have been administered more than the 300 items that were given in the paper-and-pencil test. The key to explanation of this phenomenon lies in the breakdown of score ranges within the table. Approximately two-thirds of those candidates with raw scores within 23 points (about one standard deviation) of the passing point (189) were the most-likely to require more than 300 items. Of candidates with

more extreme scores on the full test, virtually none required more than 200 items and over three-fourths required fewer than 100 items.

[Insert Table 1 about here]

When comparisons were made between the decisions reached for the paper-and-pencil test and the simulated adaptive test (using a ±3 standard error band), all decisions were identical. However, as was noted above, a substantial number of candidates would have required more than 300 items under these conditions.

A ±2 standard error band may also be quite acceptable, since it represents a 98% confidence interval. Using this band around the ability estimates and continuing the adaptive test until the band excludes the passing criterion, identical paper-and-pencil and CAT decisions were reached in all but sixteen cases (0.9% of the 1762). Of these sixteen discrepant decisions, two failed under CAT but passed the paper-and-pencil and fourteen passed CAT but failed the paper-and-pencil.

All sixteen decisions were reached in fewer than 60 items, thus it may be a situation of allowing "premature" decisions. Setting a minimum number of items is an obvious solution to this problem; however, the possibility that there might be other explanations for the discrepant results in terms of differences in the test-taking situation or in

the statistical properties of the measures should also be investigated.

The final area of interest in evaluating the simulated adaptive tests was the degree to which the distribution of items conformed to the specified ranges for the test plan categories. Table 2 presents the results for the three illustrative cases included in the appendix.

Table 2

Distribution of Adaptively Administered Items
Across Test Plan Categories

Category of Test Plan Nursing Behavio	Acceptable Range prs:	<u>#1</u>	Candidate <u>#2</u>	# <u>3</u>
Assessing	15-25%	14*	19	18
Analyzing	15-25%	28*	20	28*
Planning	15-25%	18	27*	20
Implementing	15-25%	23	19	19
Evaluating	15-25%	17	15	15
Locus of Decisi	lon:			
Nurse	20-30%	37*	29	33*
Client	10-20%	14	10	14
Shared	55-65%	49*	61	53*

The data in Table 2 indicate that reasonable conformity to prescribed test plan percentage ranges has been achieved even with unconstrained selection of items. However, two factors weigh against use of an unconstrained selection procedure for implementation of CAT for licensure examinations. First, a "reasonable approximation" is not

likely to be adequate to insure legal defensibility in the sense that the examination must be a demonstrable reflection of actual practice. Secondly, the distribution of items across categories in this simulated administration from among a 300-item test which already met the test plan is almost certainly closer to the prescribed distribution than what would be obtained in a real CAT administration in which items were selected from the entire 3000-plus item pool. Thus, the adaptive strategy used in future development of CAT software for licensure examinations will include a mechanism for assuring that each individual candidate's test meets the test plan.

Conclusions

The results of this study support the contention that licensure examinations administered in a computer adaptive mode will be more efficient, and will have accuracy at least as great as current paper-and-pencil examinations. This study demonstrates that the same pass/fail decision reached through a 300-item paper and pencil test can be made in fewer than 100 items on an adaptive test for about half the candidates. On the other hand, for those candidates not reaching clear pass/fail status within the 300-item test, CAT offers the possibility of extending testing by administering additional items until a clear determination can be made. Herein lies the improvement in accuracy, which

will better enable regulatory boards to protect the public from incompetent practitioners. Boards of nursing and candidates alike can benefit from the improved efficiency of the examination process resulting from a reduction in average administration times.

computer adaptive administration of licensure examinations offers other potential benefits not addressed in this study. Among these are monitoring of performance quality through application of item- and person-fit statistics, immediate reporting of examination results to candidates and boards, year-round administration and scheduling on demand, and a less stressful testing environment. The National Council plans to explore these benefits in future research.

References

- Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord and M. R. Novick, <u>Statistical Theories of Mental Test Scores</u>. Reading, MA: Addison-Wesley, 1968.
- Fortune, J. C. and Associates. <u>Understanding Testing in Occupational Licensing</u>. San Francisco: Jossey-Bass, 1985.
- Hofer, P. J. and B. F. Green. The challenge of competence and creativity in computerized psychological testing.

 <u>Journal of Consulting and Clinical Psychology</u>, 1985, 53, 826-838.
- Kiely, G. L., A. R. Zara, and D. J. Weiss. Equivalence of computer-administered and paper-and-pencil ASVAB tests. (AFHRL-TP-86-13) Manpower and Personnel Division, Brooks Air Force Base, TX: Air Force Human Resources Laboratory, 1986.
- Lord, F. M. The self-scoring flexilevel test. <u>Journal of Educational Measurement</u>, 1971, <u>31</u>, 3-31.
- _____. Some test theory for tailored testing. In W. H.

 Holtzman (Ed.), Computer-assisted instruction, testing,
 and guidance. New York: Harper and Row, 1970.
- ____. A theory of test scores. <u>Psychometric Monograph</u>, 1952, No. 7.
- McBride, J. R. Computerized adaptive testing. <u>Educational</u> <u>Leadership</u>, 1985, <u>43</u>, 25-28.
- Rasch, G. <u>Probabilistic Models for Some Intelligence and Attainment Tests</u>. Copenhagen: Danmarks Paedagogiske Institut, 1960.
- Reckase, M. D. Tailored testing, measurement problems and latent trait theory. ERIC document ED207987, 1981.
- Wald, A. Sequential tests of statistical hypotheses.
 Annals of Mathematical Statistics, 1945, 16, 117-186.
- Weiss, D. J. Improving measurement quality and efficiency with adaptive testing. Applied Psychological Measurement, 1982, 6, 473-492.

TABLE 1

ITEMS NEEDED FOR DECISION

Number of Items

		< 100	100-200	200-300	> 300	
	under 165	38 (2.3%)	9 (0.5%)	2 (0.1%)	0 (0%)	49 (2.9%)
kange '	165-188	62 (3.7%)	29 (1.7%)	33 (2.0%)	145 (8.7%)	269 (16.1%)
	189-212	120 (7.2%)	48 (2.9%)	9 (0.5%)	452 (27.0%)	629 (37.6%)
MON	over 212	596 (35.6%)	108 (6.5%)	17 (1.0%)	4 (0.2%)	725 (43.3%)
		816 (48.8%)	194 (11.6%)	61 (3.6%)	601 (35.9%)	1,672 (100%)

JB/mct

JEM 1986

SEM around pass see what happens within that band

Appendix

:W YORK MURSING TEST - 8.S. NAME 0357180316336	SCORE= 131 MEAS=	29 ERROR= .13 OUTFIT= 5.53 INFIT= 4.91
IIOR -2.00 .00 2.00 IST IYEM NOTAB LDTAB DIFF RESP PROB	·3.0	
1 8K1-32 0 0 1 0 0 0 0 1 .00 0 75 2 8K4-20 0 0 1 1 0 1 0 1 1 10 1 67	-1.10 1.63 37 1.15	B B
3 BK2-72 0 1 1 1 0 1 0 2 .37 1 62 4 BK4-53 0 2 1 1 0 2 0 2 .15 1 60	.14 1.03 .58 .98	·
5 BK4-87 0 2 1 2 0 3 0 2 .59 0 59 6 BK2-37 0 3 1 2 0 4 0 2 .21 0 57	.21 .83 ·.07 .76	B B
7 BK3-12 0 3 1 3 0 4 1 208 1 56 8 BK3-9 0 4 1 3 0 4 1 3 .18 1 55 9 BK4-58 0 4 1 3 1 4 1 4 .40 1 0 55	.18 .71 .41 .68	B B
9 8K4-58 0 4 1 3 1 4 1 4 .40 0 55 10 8K4-56 0 4 1 3 2 4 1 5 .20 1 54 1 8K4-47 0 4 1 4 2 5 1 5 .38 1 54	.20 .64 .38 .61 .56 .60	
12 BK4-44 1 4 1 4 2 6 1 5 .56 0 53 13 BK4-12 1 5 1 4 2 6 1 6 .40 0 53	.40 .56 .25 .54	
14 BK2-53 1 5 1 5 2 6 1 7 .25 0 53 15 BK2-92 1 5 2 5 2 6 1 8 .12 0 53	.12 .52 01 .50	B
16 8K4-89 1 6 2 5 2 6 1 9 .00 0 53 17 8K3-23 1 7 2 5 2 6 2 912 0 52	·.12 .49 ·.24 .48	. B
18 BK3-8 1 7 3 5 2 6 2 10 - 24 1 52 19 BK2-13 1 7 4 5 2 6 2 11 - 14 1 52	·.13 .46 ·.03 .45	. 8 .
10 BK2-18 1 7 5 5 2 6 2 1203 0 52 11 BK4-40 1 8 5 5 2 7 2 1212 1 52	· .13 .44 · .03 .43	. B .
12 8K4-13 1 8 6 5 2 7 2 1302 0 52 13 8K2-71 1 8 7 5 2 7 2 1411 1 51	12 .42 04 .41	. B .
!4 BK2-26 2 8 7 5 2 8 2 14 06 1 52 !5 BK4-91 2 9 7 5 2 8 3 14 .04 0 51 !6 BK4-90 3 9 7 5 2 8 4 14 02 0 52	.04 .40 03 .39	. B
16 BK4-90 3 9 7 5 2 8 4 1402 0 52 17 BK4-83 3 9 8 5 2 8 5 1410 1 51 18 BK4-78 3 9 8 6 2 9 5 1407 1 52	11 .39 04 .38 .03 .37	. B
29 BK2-42 3 9 8 7 2 10 5 14 .02 0 51 10 BK1-12 3 9 8 8 2 11 5 14 .01 1 50	03 .37 .03 .36	. B
11 BK1-28 3 9 9 8 2 11 5 15 .02 0 51 12 BK4-19 3 10 9 8 2 12 5 15 .01 1 50	03 .35 .03 .35	. B .
I3 BK1-51 3 10 9 8 3 13 5 15 .05 1 50 1 4 BK3-22 3 11 9 8 3 13 6 15 .09 0 51	.09 .34 .03 .34	. B .
I5 BK4-22 3 11 9 9 3 14 6 15 .06 0 51 I6 BK4-42 3 11 9 9 4 14 7 1509 0 49	02 .33 08 .33	. B .
17 BK4-51 3 12 9 9 4 15 7 1516 0 49 18 BK3-51 4 12 9 9 4 15 7 1617 1 52	13 .33 08 .32	. B .
39 BK4-29 5 12 9 9 4 16 7 1618 1 53 1 0 BK3-25 5 12 10 9 4 16 7 17 .08 1 48	04 .32 .02 .31	B B
11 BK4-21 5 12 10 9 5 17 7 17 .08 1 1 49 12 BK1-10 5 12 10 10 5 17 7 18 .10 0 52 13 BK4-1 5 12 11 10 5 18 7 18 .11 0 53	.07 .31 .02 .30 .02 .30	. B .
14 BK4-41 5 12 11 11 5 18 7 19 -18 0 47 15 BK4-7 5 12 11 11 6 19 7 19 -19 1 53	02 .30 07 .30 03 .29	. B .
16 BK2-2 5 12 11 11 7 19 7 20 - 19 1 55 17 BK2-81 5 12 11 11 8 20 7 20 .13 0 54	.01 .29	. B .
18 BK1-25 5 13 11 11 8 21 7 20 .13 0 55 19 BK1-2 5 13 11 11 9 22 7 2020 0 47	·.07 .29 11 .28	. B .
io 8K4-93 6 13 11 11 9 22 8 2021 1 53 ii 8K2-78 6 13 11 11 10 22 8 2122 1 54	07 .28 04 .28	. B .
i2 BK1-9 6 14 11 11 10 23 8 21 .14 0 55 i3 BK2-73 7 14 11 11 10 23 8 2222 0 47	·.07 .27 11 .27	. B .
64 BK2-93 7 15 11 11 10 23 8 2323 0 48 5 BK4-50 8 15 11 11 10 24 8 2325 0 48	15 .27 19 .27	. B . B
66 BK1-27 8 15 11 12 10 24 9 2328 1 53 67 BK4-30 8 15 11 13 10 24 9 2429 0 47 68 BK2-88 8 15 11 14 10 24 9 2529 0 48	15 .27 19 .26 23 .26	. B .
58 BK2-88 8 15 11 14 10 24 9 2529 0 48 59 BK1-54 8 15 12 14 10 24 9 2629 1 52 50 BK3-49 8 15 12 14 11 24 9 2730 0 48	23 .26 19 .26 23 .26	. B .
51 BK1-39 8 16 12 14 11 24 9 2831 1 52 52 BK2-32 9 16 12 14 11 24 9 2931 1 53	20 .25 17 .25	. B .
53 BK2-9 9 17 12 14 11 24 9 3031 1 54 4 BK5-44 9 18 12 14 11 24 9 3136 0 45	14 .25 17 .25	. B .
5 8K2-46 9 18 12 15 11 24 9 3239 0 45 6 8K2-56 9 18 12 16 11 24 10 3240 0 46	· .21 .25 · .24 .25	. B .
7 8K1-52 9 18 13 16 11 24 10 33 - 40 0 46 8 8K4-86 9 19 13 16 11 24 10 34 - 40 0 47	27 .24 30 .24	. B .
9 BK4-80 10 19 13 16 11 24 10 3540 0 48 0 BK2-61 10 19 14 16 11 24 10 3641 1 52	33 .24 31 .24	B
'1 8K1-35 10 19 15 16 11 24 10 37 ·.43 1 53 '2 8K4-74 10 19 16 16 11 24 11 37 ·.44 0 46	28 .24 31 .24	. B
73 8K3-57 10 19 17 16 11 24 12 3744 1 53 74 8K4-59 10 19 17 17 11 24 13 3747 1 55 75 8K3-7 11 19 17 17 11 24 13 3847 1 55	28 .23 26 .23 24 .23	. B .
75 BK3-7 11 19 17 17 11 24 13 3847 1 55 16 BK2-34 11 19 17 18 11 25 13 3848 0 44 17 BK4-92 12 19 17 18 11 25 14 3850 0 44	· .27 .23 29 .23	. B . B .
78 8K2-80 12 19 17 19 11 25 14 3950 1 55 79 8K1-40 12 19 17 20 11 25 14 4050 0 45	27 .23 30 .22	B .
10 BK2-75 12 19 17 21 11 25 14 4151 1 55 11 BK1-1 12 20 17 21 11 25 14 4252 0 44	28 .22 30 .22	. B .
12 BK4-38 12 20 18 21 11 25 14 4354 0 44 13 BK3-52 12 20 18 21 12 25 14 4456 0 45	· .33 .22 · .36 .22	. B B
34 BK4-61 12 20 18 22 12 25 14 45 56 0 45 35 BK1-3 12 21 18 22 12 26 14 45 56 0 46	·.38 .22 ·.41 .22	. B

W YORK MURSING TEST - B.S. NAME 0355420317229	SCORE= 207 MEAS=	
2 8K2-69 0 0 2 0 0 0 1 1 1.09 1 3 8K2-91 0 1 2 0 0 0 1 2 1.93 1 4 8K1-33 0 1 3 0 0 0 1 3 1.90 0 5 8K3-40 0 1 3 1 0 0 1 4 1.71 1 6 8K1-41 0 1 3 2 0 0 1 5 1.68 1	1.10 1.63 1.97 1.55 2.70 1.51 1.83 .98 2.20 .94 2.48 .92 2.71 .91 2.24 .74 1.93 .66 2.09 .65 1.86 .60 1.99 .58 1.80 .55 1.91 .54 2.01 .53 2.10 .52 2.18 .51 2.26 .51 2.29 .48 1.95 .46 2.01 .45 1.88 .43 1.77 .42 1.83 .41 1.89 .41 1.89 .41 1.89 .39 1.89 .39 1.89 .39 1.89 .39 1.89 .39 1.89 .39 1.85 .37	B B B B B B B B B B B B B B B B B B B

W YORK GLURSING TEST - B.S. 10R -2.00 .00 2.00 \$3 ITEM	SCORE 146 MEAS 1.3.0 MEAS ERROR 1.10 1.63 1.97 1.55 1.27 1.03 .86 .89 .52 .83 .22 .7905 .7631 .7407 .66 .13 .6105 .6021 .5805 .55 .08 .52 .21 .50 .09 .49 .02 .48 .09 .46 .19 .45 .09 .46 .19 .45 .09 .44 .18 .43 .09 .42 .01 .41 .09 .40 .17 .39 .10 .39 .02 .3805 .3711 .3718 .3718 .37	.05 ERROR= 2.0	.12 CUTFIT= 1.0 .0	.43 INFIT= .30 1.0 2.0 3 8 8 8	
8K4-42 5 9 9 6 4 14 7 12 -09 1 52 8K1-12 5 9 9 7 4 15 7 12 -01 1 51 8K4-89 5 10 9 8 4 17 7 12 -05 0 50 51 9 8 4 17 7 12 -14 0 51 9 8 4 17 7 12 -05 0 50 18 8K4-89 5 11 9 8 5 18 8 13 .04 1 51 51 31 0 8 5 19 8 14 .05 1 52 8 14 .05 1 52 8 14 .15 1 51 51 51 51 51 51 51 51 51 51 51 51 51 51 51<	.01 .34 .06 .34 .12 .33 .07 .33 .01 .33 .06 .32 .11 .32 .06 .31 .11 .31 .16 .30 .20 .30 .25 .30 .29 .30 .29 .29 .29 .29 .21 .28 .17 .28 .21 .28 .25 .27 .21 .27 .25 .27 .21 .27 .25 .27 .28 .26 .32 .26 .35 .26 .35 .26 .35 .25 .39 .23 .30 .22 .35 .22 .35 .22 .35 .22 .35 .22 .35 .22 .35 .22		. B	:	

	8 . 8 . 8 . 8 .
	. 6
	•
.35 .22 .33 .22 .33 .22 .33 .22 .33 .21 .28 .21 .29 .21 .31 .29 .21 .32 .20 .33 .20 .28 .20 .25 .20 .25 .20 .25 .20 .25 .20 .27 .29 .19 .20 .19 .19 .19 .19 .19 .19 .19 .19 .19 .19	
1 0 1 0 0 0 0 0 1 1 1 0 0 0 0 0 0 1 0 1	
78 .20 .46 .48 .48 .49 .50 .51 .55 .56 .56 .57 .58 .58 .58 .58 .58 .58 .58 .58 .58 .58	1
35 13 48 49 49 49 15 16 65 11 17 75 75 75 76 61 11 17 77 75 76 76 76 77 77 78 98 88 88 88 88 88 89 90 88 89 90 88 88 88 88 88 89 90 88 88 88 88 88 88 89 90 88 88 88 88 88 88 88 88 88 88 88 88 88	
1 22 20 13 1 22 20 13 2 22 20 13 2 22 20 13 2 23 20 13 3 23 20 13 3 23 20 13 3 23 22 14 3 23 22 16 3 23 22 16 3 24 22 18 3 24 22 18 4 25 25 21 18 26 25 19 2 7 27 19 2 7 27 19 2 7 27 19 2 7 27 19 2 8 29 27 19 3 1 28 19 3 1 32 20 3 2 3 2 20 3 3 3 3 2 2 2 1 3 3 3 3 2 3 2 3 2 2 2 3 3 3 3 3 3 3 3 3	34 36 22 34 36 22
121213131313131313131313131313131313131	23 28 6 77 29 6 50 29 6 50 29 6 52 29 6 53 29 6 55 29 6 56 29 6 57 29 6 77 29 6
\$	BK4- BK4- BK2- BK1- BK2- BK2- BK2- BK2- BK1- BK4- BK4- BK4-