

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 ~~the~~ ~~priority~~ ~~of~~ ~~psychometric~~ ~~soundness~~ ~~in~~ ~~an~~ ~~area~~ ~~of~~ ~~major~~

at the National Council are investigating the applications

Theory and Research

The concept of tailoring tests is 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

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

care has been used in the calibration of the items, ability estimates obtained via adaptive testing are directly

~~available from one examinee to the next (Beckage, 1981)~~

Recent literature on computerized adaptive testing (CAT) supports the improved accuracy and efficiency of the ~~QTE 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

~~is selected so that each response provides maximum~~

A standard item of mid-range difficulty was selected as the

that item, the program calculated a provisional ability estimate and selected as the next item that one of the

possibility. The case format was not accounted for in the adaptive simulation, i.e. items were selected individually,

the necessity of assuming that candidates would have

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

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

by the board of nursing for the full length paper and pencil

examination. A third assessment was also made comparing the

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,

~~patient and based on 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 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

"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/

considerably longer, thus maintaining the efficiency of the adaptive item selection process.

With respect to the ability estimation process for this

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

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

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

more than 100 items

(Insert Table 1 about here)

231

identive test (using a 12 standard seven band) all

investigated.

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

among a 300-item test which already met the test plan is almost certainly closer to the prescribed distribution than

that would be obtained in a real CPM administration in 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

statistics, immediate reporting of examination results to candidates and boards, year-round administration and

References

- Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord and M. R. Novick, Statistical Theories of Mental Test Scores. Reading, MA: Addison-Wesley, 1968.
- Fortune, J. C. and Associates. Understanding Testing in Occupational Licensing. San Francisco: Jossey-Bass, 1985.
- Hofer, P. J. and B. F. Green. The challenge of competence and creativity in computerized psychological testing. Journal of Consulting and Clinical Psychology, 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. Journal of Educational Measurement, 1971, 31, 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.
- _____. Theory of test scoring. Psychometric Monograph
- McBride, J. R. Computerized adaptive testing. Educational Leadership, 1985, 43, 25-28.
- Rasch, G. Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen: Danmarks Paedagogiske
- 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.

TABLE 1

ITEMS NEEDED FOR DECISION

Number of Items

	< 100	100-200	200-300	> 300	
--	-------	---------	---------	-------	--

Appendix

ST	ITEM	NBTAB			LDTAB			DIFF	RESP	PROB	MEAS	ERROR
1	BK1-32	0	0	1	0	0	0	0	75	-1.10	1.63	
2	BK4-20	0	0	1	1	0	1	0	67	-.37	1.15	
3	BK2-72	0	1	1	1	0	1	0	62	.14	1.03	
4	BK4-53	0	2	1	1	0	2	0	60	.58	.98	
5	BK4-87	0	2	1	2	0	3	0	59	.21	.83	
6	BK2-37	0	3	1	2	0	4	0	57	-.07	.76	
7	BK3-12	0	3	1	3	0	4	1	56	.18	.71	
8	BK3-9	0	4	1	3	0	4	1	55	.41	.68	
9	BK4-58	0	4	1	3	1	4	1	55	.20	.64	



