



CANADA'S TESTING COMPANY
ASSESSMENT STRATEGIES INC.

LA SOCIÉTÉ SPÉCIALISTE DES EXAMENS AU CANADA
STRATÉGIES EN ÉVALUATION INC.

Evaluation of speededness on the January 2012 Canadian Practical Nurse Registration Examination

27 January 2012

Executive summary

In January 2012 the Canadian Practical Nurse Registration Examination (CPNRE) was shortened to 180 items (plus 20 experimental items), administered over 4 hours. The purpose of this report is to determine if this shorter examination administered over a shorter time period fundamentally affected the score interpretations for the CPNRE.

In general, credentialing examinations are divided into two categories: 1) speeded tests and 2) power tests. The CPNRE is intended to be a power test where candidates are given time to finish the examination comfortably and demonstrate their true competence. It is expected that the majority of candidates will finish the examination and candidate ability is estimated as the total performance on the examination. The time limits for the CPNRE exist mainly to promote examination validity.

In considering the effect of the time in relation to completion, Assessment Strategies (ASI) made use of three criteria to determine if an examination is a speeded or a power test. Educational Testing Services (ETS) regards a test as speeded when 1) fewer than 80% of candidates complete the examination, 2) fewer than 100% of examinees reach 75% of the test and/or 3) the ratio of “not reached variance” to “total score variance” is greater than 0.15. Two refinements were made to these criteria to account for 1) candidates who run out of time and begin to skip some questions while responding to others and 2) candidates who begin to respond randomly when they run out of time. These three criteria were used in assessing all candidates. In addition, six subgroup comparisons were made: 1) Canadian-educated candidates writing the CPNRE for the first time in English, 2) Canadian-educated candidates writing the CPNRE for the first time in French, 3) Canadian-educated candidates writing the CPNRE for a second or third time in English, 4) Canadian-educated candidates writing the CPNRE for a second or third time in French, 5) internationally-educated candidates writing the CPNRE for the first time and 6) internationally-educated candidates writing the CPNRE for a second or third time.

Results indicate that 98.5% of candidates completed the scored items on the CPNRE. In addition, all candidates reached question 135 (75% of the examination). Finally, the ratio of “not reached variance” to “total score variance” for the January 2012 CPNRE is 0.009. This ratio is well below the established criteria of 0.15. The sum of the evidence indicates that the CPNRE remains a power test and candidates, as a group, were not disadvantaged by changes to the examination format. As a result, score interpretations remain valid.

1. Introduction

In January 2012, the Canadian Practical Nurse Registration Examination (CPNRE) was shortened from a 7.5-hour examination with 230-250 multiple-choice questions to a 4-hour examination with 180-200 multiple-choice questions. The January 2012 version of the CPNRE contained 180 operational multiple-choice items. The concern was whether candidates reached the last operational item.

The purpose of this report is to determine if this change in examination length and timing provided any systematic unfairness to candidates. Specifically, it is important to determine if the shorter CPNRE introduced an element of speededness that fundamentally modified the interpretation and validity of candidate results. Results are presented for all candidates and subgroup analyses are conducted by language, number of writings and whether the candidate was educated in Canada or internationally.

2. Literature Review

Background

Most standardized credentialing examinations are administered with a specific time limit. However, when time limits are set mainly for practical administration purposes such as administrative convenience and cost, it is important to allow candidates sufficient time to consider and complete all of the items. A large degree of speededness would undermine examination validity (Lu & Sireci, 2007).

Examinations are generally classified as being either speeded or power tests (Gulliksen, 1950). A pure speed test is a test composed of items that are so easy that examinees would rarely give a wrong answer. However, the tests are so long that no examinee would complete the entire test in the allotted time. A pure power test is a test in which all items are attempted and the examinee's performance is judged by the correctness of their responses. The CPNRE fits more closely to a power test. It is expected that most candidates will have time to contemplate each test item and complete the examination within the allotted time. Time limits are meant more for standardized administration.

Whenever tests involve a time limit, the rate at which candidates work will affect their performance. A small percentage of candidates, no matter how much time is given, will not complete the examination. Stringent time limits will undermine the validity of a power test but overly-generous time limits will leave many examinees with extra time. The College Board Entrance Examination Board (CEEB; 1984) reported that generous time limits may have a troubling effect on candidates who have not yet finished including increased test anxiety. As a result, most examinations contain a mixture of speed and power components (Rindler, 1979).

There are several ways in which unintended test speededness may undermine test validity. For example, when speededness is unintended, it introduces construct-irrelevant variance and thus a construct which the examination does not intend to measure. Test speededness can also negatively impact content validity because some scored items are not attempted. This is problematic if the unattempted items fit into one or more content domains.

The effect of speededness is mitigated in the context of a criterion-based examination such as the CPNRE (Yu & Sireci, 2007). As long as the relative ranks of examinees in the score distribution are not

changed with the change in time limits, passmark adjustment processes (e.g., the Angoff method or statistical equating) can compensate for changes in examination difficulty. For example, if a shorter time limit renders the CPNRE more difficult, the passmark will be adjusted to compensate for this change (Kolen & Brennan, 2004). However, it is nevertheless important to demonstrate that a test is not overly affected by speededness. There are several statistical indices that are available to assess speededness using a single administration approach.

Methods of assessing speededness

Educational Testing Services (ETS) provides three guidelines to assess test speededness (Schaeffer, Reese, Steffen, McKinley & Mills, 1993; Swineford, 1973). A test is speeded when 1) fewer than 80% of candidates do not complete the examination, 2) fewer than 100% of examinees reach 75% of the test and/or 3) the ratio of “not reached variance¹” to “total score variance” is greater than 0.15. The “not reached variance” represents the variance of the number of items left unanswered following the last item for which the candidate responded. This statistic is divided by the “total score” variance in order to obtain the “not reached” to “total score” variance ratio. These three indices of speededness, in addition to the percentage of sample answering the last few questions (in this case, the last 25 questions), are typically used to determine relative speededness in an assessment. In this report, the results from the three indices are presented for all candidates. In addition, comparisons are made for six groups: 1) Canadian-educated candidates writing the CPNRE for the first time in English, 2) Canadian-educated candidates writing the CPNRE for the first time in French, 3) Canadian-educated candidates writing the CPNRE for a second or third time in English, 4) Canadian-educated candidates writing the CPNRE for a second or third time in French, 5) internationally-educated candidates writing the CPNRE for the first time and 6) internationally-educated candidates writing the CPNRE for a second or third time.

Refinement to these criteria

The three criteria presented above are based on the implicit assumption that a candidate will run out of time at the end of the examination and fail to respond to items that he or she has not reached. There are, however, two other scenarios that should be considered. Near the end of the examination, a candidate may realize that there is insufficient time to finish the test. In one scenario, the candidate would accelerate their work rate and would skip items that would take too much time. This candidate would have a sporadic response pattern near the end of the examination. In another scenario, upon recognizing that time is running out, the candidate would respond to all the remaining items, but in a random manner (Oshima, 1994).

To account for the last two scenarios, modifications were made when estimating the three criteria presented above. First, the “last reached” point was estimated as the point where the candidates responded to at least 3 questions in a row. It was not the point where they stopped submitting responses. To illustrate, using an example of a 180-item examination:

Original method

Question	170	171	172	173	174	175	176	177	178	179	180
Response	2	3	1	blank	blank	3	blank	blank	1	blank	blank

¹ Variance is a measure of distribution that is equal to the square of the standard deviation.

Under the original method, the candidate did not completely finish the examination. The candidate would have reached question 178 and then run out of time. Their “not reached” score would be 2.

Modified method

Question	170	171	172	173	174	175	176	177	178	179	180
Response	2	3	1	blank	blank	3	blank	blank	1	blank	blank

Under the modified method, the candidate also did not completely finish the examination. However, the candidate ran out of time and began to randomly respond to, and skip items following question 172. Under this modified method, the candidate would have reached question 172 before running out of time and their “not reached” score would be 8.

The second refinement to the ETS method is to account for candidates who began to answer randomly near the end of the examination. In this case, the candidate may not have any omitted responses near the end of the examination. However, their score near the end of the examination would be much lower because they randomly responded without having time to fully contemplate the questions. To account for candidates who employed this strategy, their performance on the first 25 and the last 25 scored items of the examination was compared. Candidates who displayed a statistically significant ($p < 0.01$) drop in performance during the last 25 questions of the examination were deemed to have run out of time.

3. Results

The purpose of this analysis is to determine if recent changes in examination length changed the CPNRE from a power test to a speeded test. Analyses are provided for all candidates and for six subgroups: 1) Canadian-educated candidates writing the CPNRE for the first time in English, 2) Canadian-educated candidates writing the CPNRE for the first time in French, 3) Canadian-educated candidates writing the CPNRE for a second or third time in English, 4) Canadian-educated candidates writing the CPNRE for a second or third time in French, 5) internationally-educated candidates writing the CPNRE for the first time and 6) internationally-educated candidates writing the CPNRE for a second or third time.

To assess the relative speededness of the January 2012 CPNRE, the three indices of speededness discussed above were evaluated. In order to improve upon these indices, two enhancements were made to account for candidates who began to respond randomly when they ran out of time. The results from this analysis are presented next.

Index 1 – Percentage of candidates completing the examination

For a test to be considered a power test and not a speeded test, at least 80% of candidates must fully complete the examination. In the context of the January 2012 CPNRE, this means that at least 80% of candidates must have responded to all 180 operational items. Candidates were deemed to have not finished the examination if 1) they omitted one or more responses at the end of the examination, 2) they began to selectively skip and respond to items near the end of the examination or 3) if their performance on the last 25 items on the examination was significantly lower than their performance on the first 25 items ($p < 0.01$).

Table 1. Percentage of candidates who completed the January 2012 CPNRE

	Finished	Total	Percent completed
Canadian-educated English first time writer	1756	1780	98.7%
Canadian-educated French first time writer	21	22	95.5%
Canadian-educated English repeat writer	324	331	97.9%
Canadian-educated French repeat writer	36	38	94.7%
Internationally-educated first time writer	399	404	98.8%
Internationally-educated repeat writer	198	200	99.0%
Total	2734	2775	98.5%

Overall, 98.5% of candidates completed the CPNRE. There were differences across language groups. The lowest completion rates were found with Canadian-educated French writers. However, this group was relatively small and the numbers must be interpreted with caution. In addition, all numbers were well above the 80% criteria.

Index 2 – Percentage of candidates completing at least 75% of the examination

The second criterion is that 100% of candidates must complete at least 75% of the total examination. For the January 2012 CPNRE, this means that candidates must have reached at least question 135 before running out of time.

Table 2. Percentage of candidates who completed 75% of the January 2012 CPNRE

	Completed 75% of test	Total	Percent completed
Canadian-educated English first time writer	1780	1780	100.00%
Canadian-educated French first time writer	22	22	100.00%
Canadian-educated English repeat writer	331	331	100.00%
Canadian-educated French repeat writer	38	38	100.00%
Internationally-educated first time writer	404	404	100.00%
Internationally-educated repeat writer	200	200	100.00%
All	2775	2775	100.00%

As shown in Table 2, all candidates finished at least 75% of the January 2012 CPNRE.

Index 3 – Ratio of “not reached variance” divided by “total score variance”

The “not reached variance” represents the variance of items left unanswered for each candidate. The “total score variance” is the variance of the candidate’s total score on the January 2012 CPNRE. Results are presented in Table 3.

Table 3. Not reached variance divided by total score variance for the January 2012 CPNRE.

	Not reached variance	Total score variance	Ratio
Canadian-educated English first time writer	1.61	224.67	0.007
Canadian-educated French first time writer	0.41	318.66	0.001
Canadian-educated English repeat writer	3.66	141.20	0.026
Canadian-educated French repeat writer	28.67	113.79	0.252
Internationally-educated first time writer	2.17	348.87	0.006
Internationally-educated repeat writer	5.20	192.10	0.027
All	2.55	289.31	0.009

Using the ETS ratio of greater than or equal to 0.15 as the criterion for test speededness, the January 2012 CPNRE did not reach this threshold. For all candidates, the ratio of 0.009 was far below the stated criterion. In addition, although there were differences between subgroups, only one of the subgroups was above the 0.15 threshold. However, for the Canadian-educated French repeat writers, these numbers are skewed by 1) a relatively small sample size for this subgroup and 2) one candidate who did not reach the last 33 questions. As a result, using this criterion, the January 2012 CPNRE was not overly speeded.

Index 4 – Percentage of candidates responding to each of the last 25 questions.

The last index used is the number of candidates who responded to the last 25 questions. If an examination contains high levels of speededness, large numbers of candidates would not respond to the last questions. For this index, results are reported for all candidates in Table 4.

Table 4. Percentage of candidates responding to each of the last 25 items

Item #	Candidates responding	Total candidates	Percentage
156	2769	2775	99.8%
157	2767	2775	99.7%
158	2767	2775	99.7%
159	2764	2775	99.6%
160	2767	2775	99.7%
161	2767	2775	99.7%
162	2761	2775	99.5%
163	2767	2775	99.7%
164	2767	2775	99.7%
165	2764	2775	99.6%
166	2758	2775	99.4%
167	2758	2775	99.4%
168	2761	2775	99.5%
169	2761	2775	99.5%
170	2761	2775	99.5%
171	2761	2775	99.5%
172	2764	2775	99.6%
173	2761	2775	99.5%
174	2761	2775	99.5%
175	2758	2775	99.4%
176	2758	2775	99.4%
177	2756	2775	99.3%
178	2756	2775	99.3%
179	2747	2775	99.0%
180	2753	2775	99.2%

4. Conclusion

Score interpretations for the CPNRE are based on the assumption that the CPNRE is a power test and that time limits do not significantly affect candidate scores. As a result, it is important to assess if the recent shortening of the CPNRE to a 4-hour 180 operational item examination changed the fundamental nature of the assessment. Specifically, an investigation was done to determine if the January 2012 CPNRE contained excessive amounts of speededness that would affect the validity of the score interpretations. In order to conduct this analysis, three criteria were determined.

The results from all criteria indicate that the CPNRE does not contain excessive elements of speededness that would jeopardize score interpretations. Specifically, 98.5% of candidates completely finished the examination, 100% of candidates completed at least 75% of the examination and the ratio of “not reached variance” to “total score variance” was well below the threshold of 0.15 established by the empirical testing literature. This, in addition to the fact that approximately 99.0% to 99.5% of candidates

attempted the last 25 items on the examination indicates that the January 2012 CPNRE remains a power test and that the recent shortening of the examination did not introduce undue amounts of speededness that would jeopardize the validity of the examination.

References

- College Entrance Examination Board (1984). *The college board technical handbook for scholastic aptitude test and achievement tests*. New York: Author.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley.
- Kolen, M.J., & Brennan, R.L. (2004). *Test equating, scaling, and linking. Methods and practices (2nd ed.)*. New York: Springer-Verlag.
- Oshima, T.C. (1994). The effect of speededness on parameter estimation in item response theory. *Journal of Educational Measurement, 31* (3), 200-219.
- Rindler, S.E. (1979). Pitfalls in assessing test speededness. *Journal of Educational Measurement, 16* (4), 261-270.
- Schaeffer, G.A., Reese, C.M., Steffen, M. McKinley, R.L., & Mills, C.N. (1993). *Field test of a computer-based GRE General Test (ETS Report No. RR-93-07)*. Princeton, NJ: Educational Testing Service.
- Stafford, R.E. (1971). The speed quotient: A new descriptive statistic for tests. *Journal of Educational Measurement, 8*, 275-278.
- Swineford, F. (1973). *An assessment of the Kuder-Richardson Formula (20) reliability estimate for moderately speeded tests*. Paper presented at the annual meeting of the National Council of Measurement in Education (New Orleans, Louisiana, February 28, 1973).
- Yu, L., & Sireci, S.G. (2007). Validity issues in test speededness. *Educational Measurement: Issues and Practice, Winter 2007*.

